# DNAStar 中文使用说明书

# 编者:宋晨

—,	EditSeq	2
Ξ,	MapDraw	23
四、	MegAlign	32
五、	PrimerSelect	42
六、	Protean	54
七、	SeqMan II 开始	64

-, EditSeq

## 打开已有序列

我们从用苹果计算机打开"TETHIS21MA"和用 Windows 打开 "tethis21. seq"开始。

假设序列的末尾有载体序列污染。我们在用 EditSeq 打开序列的同时, 用 Set Ends 命令去除 5'和 3'污染序列。

- 从文件菜单 (FILE MENU),选择 Open。
- 打开文件夹"Demo Sequences"单击选定序列"TETHIS21"。
- 单击位于对话框右下角的 Set Ends 按
   钮。Set Ends 被打开(如右)。
- 在 5' 框和 3' 框中键入 50 和 850, 点击 0K。单击 0pen 打开序列。

当 EditSeq 窗口打开时,序列长度显示在右上角。通过"setting ends," 现在你只有最初序列中的 801 bp 的片段。Set Ends 选择在全部 Lasergene 应用程序中都可以使用。

TETHIS21MA(50)850)	
5. 20 th	850 3"
tccaa	CTCAT tcott
Length: 906 bp	Help
Range: 801 bp	Cancel OK

Triplet	Indicator Positio	on/Selection	Range	Length	
Protein/DNA Icon-		28	1921MA:SEQUENCE 5 = 612 30 40	50	801 bp
SEQUENCE	TTTHHHHATAHAT CAGATGCTTTTTT AGAAAAGGATAAT TAAAAAGAAAAAA	TATCCAATCAGA AATAGAAGATA( GCAGAAAAAAAA AGTAAAAAAATT(	HRCGCHGHHHHHGHHGH CAGATAACTGCAAGAGA TCATAATTTAATCAAAA CCATTTAAATGGCTCCC	ATTATCCHAT Atagaaraaa 1 Aaatataaat 1 Caagaargct 2	50 .00 .50 • 200 •
Windowpane Divider COMMENTS bp Range	LOCUS T DEFINITION T ACCESSION M	ETHIS21MA thermophile 31332	906 bp ds-DNA a macronuclear hi	l stone H2B-1	NV gen ▲
Feature Key		1835	551 1		
FEATURES	∽√note="histo mANA	ne H2B-1" 1341	745 2		
Proofread Aloud	/note='H2B-1	nRNA (major ified Search 💷	∽ alt., 3' end pu	at.)"	╡ ・ ・ ・ ・ ・
Find Previo	is Find Ne	xt Searc	sh		

Anatomy of an EditSeq Document

## 寻找开放读框

在这入门的一部分中,我们将确定序列中最大的 ORF,并翻译它。

- 从 SEARCH MENU 找到 ORF,点
   击打开会出现右边的对话框。
- 单击 Find Next 寻找第一个
   ORF 的位置。

Search		B
ORF Search		
🔾 Literal 🔹 Allow Ambiguity	ORF	🕕 Site
Find Previous Find Next	Options	Cancel

继续点击 Find Next 直到你把 ORF 的位置选定在位置 183-455。ORF
 的坐标会出现在 EditSeq 窗口的顶端附近。

## DNA 序列翻译

这一节中我们介绍如何翻译我们的 ORF, 不过任何序列中的读框内部分都可以用下面的方法进行翻译。如果你的选择是在三联码的读框内, 三



| 联码指示棒显示为实心黑线(如左图)。如果

你的选择是不在三联码的读框内, 左边的箭头和右面的箭头显示向左或 向右移动一个 bp, 以使所选序列成为三的倍数。

- 选定 ORF, 从 GOODIES MENU Untitled Pro #2: PROTEIN DB m.v. 3140 Position: 1 267 AA 20 30 \_\_\_\_\_\_ • FKWKLSNQNAEKEDVPLRCFFNRRVRQLEE JEKKKRJMQKKIJLQSKNJNQKEKKQKIPF KWLPRLPLLKRRSRRPDPKRRTRRRD QKPSPSTSSRSQSKSTLMSYFPRRLTLT PSLTTPSKESPQNPLSVSDSTREEPSHPGK SKPLSSSYVPYNSLDTPSPKVPRPSPSSLL 30 50 90 120 150 180 菜单中选择 翻 译 Ŧ (Translate). 4 Translate DNA Sequence TETHIS21MA(1,801) With Ciliate Macronuclear Code 翻译的蛋白 质序列出现在一 Molecular Weight 31408.08 Daltons 264 Amino Acids 💹 🏶 🔍 Unspecified Search 🖪 🔟 个新的未命 名窗口中 (如右
  - 图)。它是使用标准的遗传密码翻译的。

#### 使用其它遗传密码

根据你的序列的来源,你可以选择使用非标准的遗传密码进行翻译等操作。在这节中,我们将标准的遗传密码转换成 Ciliate Macronuclear 密码。



- 从 GOODIES MENU 菜单选择 Genetic Codes 打开,子菜单显示如左。
  - 单击"Ciliate Macronuclear"就实现了 遗传密码的转换, EditSeq 现在使用的就 是 Ciliate Macronuclear 的遗传密码。
  - 同样可以将遗传密码转换为其它类型。

### 遗传密码的编辑

这一节中我们修改 Ciliate Macronuclear 的遗传密码。

- 从 GOODIES MENU 菜单选择 Edit
   Selected Code。这将打开右面的窗
   口,窗口显示遗传密码是怎样翻译
   DNA和 RNA序列的。
- 如以 DNA 形式展示密码,点击 DNA 按钮。



- 编辑时,单击任何要编辑的密码,
   从其目前的位置拖到新氨基酸对应的位置则可。
- 如使用不同的启始密码子,单击 Set Starts 按钮。第二的遗传密码 窗就会被打开,可以进行起始密码子选择。

单击任何氨基酸(或者 codon 位置),该密码子就会变成绿色,而且
 旁边出现一个箭头,它就被设定为起始密码子了。如要去除,只需
 单击它即可。如不保存,则单击取消;如要保存,单击保存为。

## 序列的反向互补及反向转换

下面的步骤可以用于反向测定的序列的正确输入。

- 选定序列。
- 从 GOODIES MENU 菜单,选反向互补序列 (Reverse Complement),
   或者把序列颠倒过来 (Reverse Sequence)命令,则被选定的序列
   就被翻转互补或翻转过来了。

#### BLAST 检索

下面我们将在 NCBI 的 BLAST 服务器上对 TETHIS21 序列进行相似性比较。注意为了进行 BLAST 查找必须保证因特网的连接。如果你没有连接因特网,跳过这部分,继续下一部分。

- 选定序,或者从 EDIT 菜单中选择 Select All。
- 从网络检索菜单(NET SEARCH MENU),选择 BLAST 查找。BLAST 对话框就会出现。

Blast Query: Selection TETHIS21MA (1>801)	
Program blastn 💌 Database nr 💌	Help Cancel OK

- 程序默认为 blastn,数据库默认是 nr,参数转换请参照帮助。
- 单击 0K 开始查找。
- 寻找结果显示为两部

	blast	p of Selec	tion TET	HIS21MA	(54)551) vs	. nr 🗏 🖻	E
ELAS	Create Doo TP 2.0.8 [	ument La	sunch Brow ) 50 matcl	ser <b> </b>	Batoh Save ces reported	]	
V	Scare	Expected	i de	entifier	Descrip	otian	۲
	182	1e-45	3p P0899:	3   H2B 1_TE	TTH HISTONE	E H2B.1 ·	
	182	1e-45	pirl IA61:	301	histone	≥ H2B - 1	_
	192	1 <del>c 4</del> 5	prf  090:	9211A	histone	≥ H2B [Te	-
		Sc	ore -	182,	Identit	y = (94/	
		E×	pect =	1e-45,	Sinilar	= (94/	
	Query:	7Z RSETFA RSETFA	ATYTEKULK ATYTEKULK	DVHPDVG IS DVHPDVG IS	SKKAMNI NINSFI SKKAMNI NINSFI	NDSFERIA NDSFERIA	Ŧ
◀ Ⅲ						ł	0

分(如下)。上边的部分是按可能性的顺序显示检索到的序列的名字, 下面的部分显示上面部分选定序列与提交序列(上边序列)比较的 具体结果。有关"score"和"expectation"的详细的信息在 NCBI' s 网点——http://www.ncbi.nlm.nih.gov/BLAST.——可以找到。 一般来说,更主要的 score 和更低的 expectation 提示较好的相似 性。

 BLAST 结果窗最上边的 3 钮被用来打开,或者保存检索到的序列, 或者让你了解更多的有关信息。

下面我们从用"Create Document"钮来打开评分最高的5序列开始:

单击 Create Document。一个小
 的对话框出现。



- 单击 OK, EditSeq 自动查对多余序列。如果 EditSeq 提示至少 2 序列是同一个,请点击 OK。EditSeq 将从因特网数据库下在单一的序列,并分别打开一个单独的 EditSeq 窗口。

下面我们用"Batch Save"钮将 3-10 序列保存为 EditSeq 文件:

● 选定从顶端起第 3 个序列。单击 Batch Save。小的灰色的对话框出 现。

- 单击下拉菜单,选 Next。并在右面的文本框中写入 8。
- 点击 Set Location,显示文件夹对话框。
- 选定你要保存序列的位置。单击 0K 回到灰色的对话框。
- 单击 OK 保存序列,文件扩展为".seq"。在下载过程期间,EditSeq
   自动查对重复的序列。如果 EditSeq 提示至少 2 序列是同一个,请
   点击 OK。
- 除非你收到错误信息,否则可以认为你的序列已经成功下载。

最后我们可以使用"Launch Browser"钮查看序列的详细信息

- 选定序列。
- ●选择 Launch Browser。
- 你的网络浏览器
   将打开右面的窗口。



序列信息查看

现在我们要使用 EditSeq 菜单指令查看有关打开的 TETHIS21 序列的信息。 DNA Statistics Englishing Info about TETHIS21 M4(1,996)

选定序列的一部分。如果你倒是
 希望全选序列,从 EDIT MENU 菜
 单,选择 Select All。

DNA Statistics	
Sequence InFo about TETHIS21MA(1,906)	~
Total number of bases is 906	
1 A = 39.40 [357]	
<b>x</b> G = 12.47 [113]	
\$ T = 29.03 [263]	
\$ C = 19.09 [173]	
X Ambiguous = 0.08 [8]	
s A+T = 68,43 [629]	
\$ C+G = 31.57 [286]	
Dovis,Botstein,Roth Melting Temp C. 77.29	
Wallace Temp C 2730.08	
	-
4	• 4

 从 GOODIES MENU 菜单,选 DNA Statistics。就会出现右面的窗口, 显示序列信息。

## 序列校读

在我们学习保存和输出序列之前,下熟悉一下EditSeq的校读功能 这功能能帮助你校正测序胶中的错读。

- 选定序列。
- 单击校对发音图标(序列窗口底部张开的嘴),或者从SPEECH MENU,
   选Proof-Read Sequence。
- 电子的音声就会开始朗读所选的序列。(注:如果你听不见任何声音,检查你的计算机的喇叭是否已经打开。)
- 要改变音声read-back的速度,从SPEECH MENU菜单,选择Faster or Slower。
- 要停止校读,点击图标(手),或者从SPEECH MENU菜单,选择 Proof-Read Sequence。

## 序列的保存与输出

首先创建一个用于保存的序列。从文件菜单,选New中的New DNA,或者 New Protein。

9

- 将序列写入出现的窗口。
- 如果你输入非法字符,计算机会发出警告。
- 然后,我们将序列保存为EditSeq文件:

- 从文件菜单,选Save。
- 选定保存位置。
- 给序列命名。
- 单击保存则可。

以GenBank或GCG格式保存序列:

- 从文件菜单,选Export。
- 选定保存位置。
- 为sequence(s)选格式。
- 给sequence(s)命名。
- 单击保存则可。

以FASTA格式保存序列:

- 从文件菜单,选Export (1个序列),或者Export All As One (多 个序列)。当使用Export All As One的时候,如果DNA和蛋白质文 件同时存在,激活窗口的序列类型与你保存的类型是一致的。
   EditSeq仅仅将写入的序列保存为FASTA格式。
- 选定保存位置。
- 选FASTA格式。
- 给sequence(s)命名。
- 单击保存则可。

二、GeneQuest

GeneQuest可以帮助你发现和注释DNA序列中的基因,并帮助您操 作生物学所关心的DNA的其他feature:包括ORFS、拼接点连接,转录因 子结合为点、重复序列、限制性内且酶酶切位点等。通过应用"methods" 到序列,序列的feature可以以图形的形式展示出来。你可以在序列上 注释任何你发现的feature。和其它的Lasergene应用程序一样, GeneQuest也提供整合的BLAST和Entrez寻找功能。

GeneQuest能直接打开DNASTAR, ABI和GenBank文件。其他格式的 序列文件也可以使用EditSeq改为DNASTAR格式。如果你知道Genbank序 列的登录号或名称,你可以直接打开序列。另外,你还可以在Entrez 数据库进行序列查找和输入。

11

# 打开已有分析文件

在这一节中,我们将对已有的GeneQuest文件(也叫做GeneQuest分析) "Nematode R01H10."进行操作。

- 从文件菜单,选择Open打开一
   个和右边相似的窗口。
- 在苹果计算机上,从Show菜单 中选择GeneQuestDocument文 件。在Windows上,从文件类型 菜单(Files of Type)的中选择 GeneQuest Documents。

What Document?	
📋 Demo Sequences 🔹	<b>6</b> , <b>1</b> , <b>0</b> ,
Name	Date Modified 🚖
Remotode R01H10 Assay	11/5/97 🔺
🗢 💐 pF1753	1/30/98 💻
PFI 753-Pr10	10/4/93
🔁 pFI753-Pr12	10/4/93
😭 pFI 753-Pr16	10/4/93
PFI 753-Pr18	10/1/93 👻
Show: GeneQuest Document Files 🔹	SetEnds
2 Cancel	Open

用文件管理系统打开名为"Demo Sequences."的文件夹。双击
 Nematode R01H10,就可以打开下面的窗口。

Anatomy of a GeneQuest Assay Document



GeneQuest的DNA分析方法

打开GeneQuest文件后,下一步是选择应用方法。应用方法后,结果的 图形显示可以帮助你了解序列上感兴趣的features。打开序列后,你会 发现只有几种方法应用后的结果展示在窗口内。在下一部分中,我们将 学习如何把其它方法用于我们序列的分析。

- Title——给文件取名。
- Ruler——在文件中加入标尺。
- Sequence——显示文件中的序列。
- Patterns-Matrix——方法的运算参数。
- Patterns-Signal——转录因子结合位点数据库。
- Patterns-Type-In Patterns——使用键盘输入运算所需的Pattern参数。
- Repeats-Inverted Repeats——寻找反向重复序列。
- Repeats-Dyad Repeats——寻找Dyad重复和palindromes。
- Repeats-Direct Repeats——寻找正向重复序列。
- Gene Finding DNA Finder———在打开的DNA序列中寻找指定 DNA序列。分别显示正义连和反义连的寻找结果。
- Gene Finding Protein Finder——在打开的蛋白质序列中寻找指定
   DNA序列的翻译序列。显示结果为全部6个读框。
- Enzymes-Restriction Map——用DNASTAR酶目录中的酶分析打开 的序列,并以图形方式展示。
- Coding Prediction Borodovsky——用Borodovsky's Markov方法

来识别潜在的基因编码区,并以图形方式展示。

- Coding Prediction Starts Stops ORFs——根据指定的ORFs的最小 长度,寻找可能的开放读框,可以选择是否需要起始密码子。读框 的启始和中止点分别展示。
- Coding Prediction—Local Compositional Complexity——根据
   Shannon信息学原理寻找有基因编码提示信息的区域。
- Base Contents-Base Distribution——序列上4种碱基、A+T和G+C的频率、分布,以及AT和gc分布区域。
- Bent DNA Bending Index——DNA折叠预测。

## 用分析方法操作

调用新的GeneQuest方法的步骤是:从More Methods中选择方法,加入 方法帘(method curtain),待方法运行完毕后,选择性的拖取结果放 入分析界面

		File Edit Analysis Sites & Features Entrez Search Options Help
(assay surface)		Position: 1 Posit
即可(见右图)。		More Methods         Image: Constraint of the state of the st
在本节中,我们	More methods	Sequence         I<
使用Bent DNA -		Coding Prediction Map Coding Prediction Base Contents - Base Distribution Bent DNA - Bending Index Fs
Bending Index方	Method curtain ——	Coding Prediction - Local Composition.
法进行分析。		V NOR
• ~ ~	Assay surface	/

ANALYSIS

MENU选择Show Available Methods可以打开方法帘,也可以通过

拖动分析界面左上角的小环的打开方法帘。方法帘中包括已经用于分析的全部方法。

- 你将注意到方法帘没有Bent DNA Bending Index method。在方法 帘的顶端,点击More Methods打开一个下拉菜单,其中尤可以用于 分析的所有方法,点击Bent DNA - Bending Index method,该方法 就进入了方法帘。
- 若查看方法帘中的方法是否已经被应用,点击其右边的三角形。如
   果图标前有数字表明该方法已经使用,数字表示应用的次数。因为
   我们还没有应用Bent DNA Bending Index,所以点击三角形,会发
   现图标前没有数字。
- 点击白颜色的位置去除对图标的选择。
- 单击选定 "Bend Region,",将其拖到分析
   界面,释放鼠标。序列中可能会折叠的区域就会以小盒子的形式显示出来。

#### 方法参数改变

下面我们改变方法的参数,然后将分析结果与参数改变前的结果进行比较。

- 重复前面的操作,在方法帘中加入一个新的Bent DNA Bending
   Index,并把它拖如分析界面。现在你应该有2个完全一样的折叠区
   分析结果。
- 双击方法帘中任何一个结果,将打开一个参数对话框。
- 改变弧长度参数为30,单击OK。分析界面上就会出现根据此参数计



算得到的结果。

(注:如果你再次拖入新的方法时,参数的改变会自动应用到新的方法上)。

## 结果展示优化

- 组合或移动展示的结果:单击位于GeneQuest窗口左侧的调色板工
   具中的的选择器(手图标),在分析界面中可以随意拖动任何展示
   的结果到任何位置。
- 改变方法格式:单击目的选择器(手图标),可以选择分析界面上的任何方法。从OPTIONS MENU菜单选择Line Color,打开彩色选择子菜单。选定颜色后,方法题目和结果显示都将变成那个颜色。
   另外,你可以用类似的指令进行操作:Line Weight、Fill Color and Fill Pattern。
- 去除方法:单击选择方法帘中显示的方法,用退位或删除键去除应
   用的方法即可。

注意任何你除掉的方法都是能从Methods menu菜单再一次被使用。

<u>注释Features</u>\_

当你用Genbank输入序列时,序列中标注的Features会自动转换为图形 命令显示在方法帘中。这些Features 可以象任何别的方法那样拖到分 析界面上显示。

GeneQuest也允许你为你的DNA序列制作新Features:

- 单击位于GeneQuest窗口的左侧的调色板工具(箭图标)。
- 然后点击选定2641-4257的Informative Region。
- 点击调色板工具(铅笔图标),或者从SITES & FEATURES MENU

选择New Featur, 打开一个 Feature Editor对话框(如 右图)

● 你打开Feature Editor时, 首先进入Location设定。点

Untitled : Location					
OK Dancel Loca	tion 📔 Descriptio	n Style	]		
Title: Info region		Key: mise	_feature		
	Location —				
Segment:		Segment	t Name:		
<ul> <li>(2641 &gt; 4257)</li> </ul>	Ê Se	gmentA			
		Add Sec	gment		
		Delete S	egment 🛛		
		Set E	nds		
Show Feature Title	Above	Below			
☑ Show Segment Names	Above	Center	Below		

击"Info region"进入Title box,点击"Segment A"进入Segment Name box,接着,在对话框的底部选择标题和区段名字的位置。

- 点击Description按钮进入新的Feature Editor窗口。如果你愿意,可
   以在note中对Feature做标注,在key种选择Feature的种类。
- 点击Style按钮进入最后的Feature Editor窗口,选择字型,大小和颜
   色,以及你喜欢图型用于新建的feature。
- 点击OK关闭Feature Editor窗口。一个图形化展示的"Info region" 就会出现在分析窗口的底部。

下面我们把新建的feature与另外一个feature整合起来。

- 单击调色板工具(箭图标)。选定你刚做好的feature,单击工具调
   色板工具上的(链图标)。
- 然后点击名为 "R01H10.4——CDS." 的feature, 再点击连接(链图标)。
- 这样你的"Info region" featur将继续作为没有联系的方法而存在,
   但是,它的拷贝将作为R01H10.4featur的一部分出现。

BLAST检索

GeneQuest为你的序列在因特网或企业内部互联网数据库上进行相似 性检索提供2不同的方法。BLAST Selection指令允许你使用序列的任何 一部分进行检索。BLAST ORF需要你首先选择序列的ORF,然后他就 可以自动翻译,在蛋白质数据库中进行检索。在这一节中,我们用 BLAST在NCBI上检索与Nematode R01H10相似的序列。

注意必须保证您的计算机与因特网相连。否则,跳过这一部分。

- 单击范围选择器调色板工具(箭图标)。
- 如同在右边被显示的那样,单击任何"R01H10.5"的feature。注意 此时选定的仅仅 Nematode R01H10 Assay DE Selection: (12517<12478)+(12529<12802)+(13145<13287) = 579 (971 averall) 31543 Bp k 是外显子部分 1 (exons),内含 ð 8 子部分被自动去 BOHID.4 除。 从网络检索菜

Blast Query	Blast Query: Selection (13287<12317)				
Program Database	blastn v nr v	Help Cancel OK			

- 单,选BLAST Selection。会 出现左面的BLAST对话框。
- 不做参数改变,这样,程序

是blastn,数据库是nr。

● 单击同意开始查找。

寻找结果显示为两部分(如下)。上边的部分是按可能性的顺序显示检 索到的序列的名字,下面的部分显示上面部分选定序列与提交序列(上 边序列)比较的具体结果。有关"score"和"expectation"的详细的信息 在NCBI's网点--http://www.ncbi.nlm.nih.gov/BLAST.--可以找到。一般 来说,更主要的score和更低的expectation提示较好的相似性。BLAST

结果窗最上边的4钮被用 来打开,或者保存检索到 的序列,或者让你了解更 多的有关信息。"Put In Document"按钮相当于 Gene Finding - DNA

		blastn o	f Selection (13287<	12317) vs. nr 📃 🖽 🗄
BLAS	Put in Doci STN 2.0 8 (Je	ument [ 1 an-05-1999] 29	Launch Editor Laun matching sequences report	ch Browser 👔 🛛 Batch Save
<ul> <li></li> </ul>	Score	Expected	Identifler	Description 🔺
	547 325 283 40 40 40 40 40	le-154 5e-87 2e-74 0.45 0.46 0.46 0.46	enb [231590] DER01H10 enb [231590] DER01H10 enb [231590] DER01H10 enb [X13208] DP4 enb [X12200] DPPLCB enb [X5225] DPPLCB enb [X5225] DPLCUR dbj [D10246] DL0PLCA	Caenorhabditis eleg: Caenorhabditis eleg: Caenorhabditis eleg: Clostridiun perfrin; C. perfringens ple g Clostridiun perfrin;
		St Ex	oore = 547, ¢pect = 1e−154	Identity = (276/276) 🔺
	Query: Sbjct:	142 oga      12710 oga	ateeetgaaategaegtgg 	ogttetegatgoaagetggegete 
	1			▶ @

Finder,在打开序列中寻找检索到的序列。(如果我们使用blastp或 blastx,则相当于Gene Finding-Protein Finder)

- 在BLAST结果窗口中选择一个检索结果。
- 单击Put In Document。保存对话窗将打开。
- 选定保存位置,并命名。单击保存。
- 你选的序列将象应用方法那样以短的垂直的竖线自动出现在分析界
   面的底部。垂直的竖线显示BLAST结果的位置。

- 可以选择Zoom in/OUT来放大或缩小视野。直到你能清楚地查看序列。
- 其他按钮与EditSeq中介绍的功能相同。

## Entrez Database检索

GeneQuest允许你在因特网或企业内部互联网的Entez数据库上进行检 索。检索到的结果可以输入GeneQuest(或者EditSeq),或者保存为序 列文件。在这一节中,我们将检索与狨视觉色素(visual pigment in marmosets)序列有关的序列,然后学习GeneQuest或EditSeq是如何打 开其中的一个序列的。

● 从网络检索菜单(NET SEARCH MENU),选择新正文查找(New Text Search)来打

 开Entrez Query
 Or...
 And

窗口(如右)。

在文本框中敲入
 "visual



pigment." 。

- 从默认为"All Fields"的下拉菜单中选择"Text Word."。
- 用鼠标从右面再拖入一个检索框(New Term)。
- 在文本框中敲入"Callithrix",从"New Fields"菜单中选"Organism"。
- 单击查找。查找结果出现于Entrez Results窗口(如下)。
- 双击任何Entrez Results窗口里的序列,就可以查看其详细信息。

=			db=n:visual pigment[WORD] & Callithrix[ORGN]	Ð				
24 m	24 matching sequences reported							
Ý	D	Description						
	3290085 3290084 3290083 3290082 3290081 1103538 1103537	AF051595 AF051585 AF051584 AF051583 AF051582 138892 C 298891 C	Callithrix jacchus X-linked utswel pignent protein PSDI gene, exon 5 gil3200 Callithrix jacchus X-linked utswel pignent protein PSDI gene, exon 4 gil3200 Callithrix jacchus X-linked utswel pignent protein PSDI gene, exon 3 gil3200 Callithrix jacchus X-linked utswel pignent protein PSDI gene, exon 3 gil3200 Callithrix jacchus X-linked utswel pignent protein PSDI gene, exon 1 gil3200 Callithrix jacchus X-linked utswel pignent protein PSDI gene, exon 1 gil3200 jacchus intron 4 of utswel pignent gene (yelles elleles) gil1003538 jedb.308992	00000				
	1103535	X88893 C	jazohus intron 4 of visual pignent gene (green allele> gi 1103535 emb X88893					
4	1			•				

如果你想在GeneQuest或EditSeq里打开其中的一个序列:

- 从Entrez Results窗,单击你想打开的序列。
- 从网络检索菜单,选择用GeneQuest或者EditSeq打开(Open with GeneQuest/Open with EditSeq)。
- 选定文件保存的位置,并给文件取名。单击同意(视窗),或者保存(苹果计算机)。

如果你选择以GeneQuest打开文件,GeneQuest将打开文件并且将其默认的方法自动应用到序列上。如果你做选择以EditSeq打开文件,Entrez 序列将在主窗口中显示。

#### GeneQuest的其他特点

以RNA折叠形式查看序列的一部分:

- 选定目的序列,在ANALYSIS MENURNA菜单中选择Fold as RNA命令。
   序列酶切,模仿agarose凝胶电泳判定大小:
- 打开方法帘 (method curtain)。
- 从More Methods中选择Enzymes Restriction Map 加入方法帘。
- 单击图标左边的蓝色三角形,打开内切酶一览表。注意方法帘仅仅显示那些最少切割一次DNA序列的酶。刚打开酶一览表时,所有的酶都被选中了,在空白处单击一下去除选定。

- 选定特定的酶,拖入分析界面,即可
   以看见该酶的酶切位点在序列上的
   位置。
- 从SITES & FEATURES MENU菜单选泽
   Agarose Gel Simulation。
- 新窗口即显示酶切片段在
   electrophoretic的分离情况(如
   图)。

	LamBBX	Bgl II	Bgl I	Bgl II Bal I	Lam H3	
50000 -				bgi i	-50	0000
25000 -					-26	5000
10000 -					- 10	0000
5000 -					-50	000
2500 -					-26	:00
1000 -		2 2			- 10	000
500 -					-50	0
250 -					-26	;0
100 -					- 10	00
50 -	į				-50	L.

保存分析文件

- 从文件菜单,选保存。
- 选定文件的保存位置。给文件命名。
- 单击保存。
- 你所应用和展示的所有信息都会被保存。

三、 MapDraw

根据实验设计,分析和实验结果的展示需要的不同,MapDraw可 以制作6种类的酶切图。从简单的线性图到有注释的环形图,在展示限 制性酶切位点的同时,还可以同时展示序列的feature,六个读框及其翻 译结果。MapDraw也以自动展示从GenBank输入序列的features。

你可以按照位置、酶切频率等来排列你的酶切位点。另外, 你可以用手工选任何酶切位点的结合。酶切位点过滤器(filters)也可以使用Boolean operators联合使用。

MapDraw工具能使你规划酶切位点和克隆实验,产生详细,充分 地结果概括。和全部Lasergene应用程序一样,MapDraw也提供整合的 BLAST查找功能。 新酶切图制作

我们以组氨酸DNA序列"TETHIS21MA"(苹果计算机)和"tethis21.seq" (视窗)为例。首先我们寻找能够切割组氨酸DNA序列的酶切位点。

- 从文件菜单选择New打开右
   边的对话框。
- 打开"Demo Sequences."文 件夹。
- 双击打开TETHIS21(见下)。

What Sequence?							
🖏 Demo Sequences 🛛 🔹	) <b>6</b> . <b>1</b> . <b>0</b> .						
Nome	Date Modified 🔺						
SYNPUCI 94	5/4/92						
TETHIS21 MA	5/4/92						
TETHIS22MA	5/4/92						
Tetrahymene H2B1	5/4/92						
🚼 Tetrahymena H2B2	5/4/92 📃						
Yeast H2B1	5/4/92 👻						
Show: All available TETHI S21MA(1>906) aatttggtaaatacataaat Length: 906 bp Range: 906 bp							
Ø	Cancel Open						

Anatomy of a MapDraw Document						
View Selector	Enzymes Applied Position / Selection Length					
Mop Document Filter	16 of 516 5 zymes 906 bp					
Must Cut Here	MHind N SreQ I MCviS I Jsp509 I TacC I					
Don't Cut Here	Weil         Two I         Tappool           MCWF I         Tase C I         Trave I           McWF I         Tase C I         McWF I           McWF I         McWF I         McWF I           Tase C I         McWF I         McWF I           Tase C I         McWF I         McWF I					
New Feature						
Join	H2B-1 H2					
Six Translation	LysLysArgLys Giy CysArgLysLys Ser PheAsnGin Lyd II e I le LysLysLys					
Reading Frames	Leupho Prosor Phosor Lou (Ro Sor Dho Dho - Lou (ye Il è Leu PhotPre Leutan Dho Lou Dhe Sor Phot PhotPhot Leu II Le 10 Cryphe Phot Phot Leit Att Le Tyr PhotPhot Leu PhotPhot Tyr His Leu PhotPhot Rap Tyr Asn Leu - Phot Ibe Tyr II e Leu PhotPhot Phot					
Find Previous	Vinspecified Search 4 W					
F	Find Next Search Feature					

#### 过滤器类型

过滤器 (filter) 是你定义的可以使用的酶切位点的集合。在你自定 义过滤器之前,所有酶切位点都会用于序列分析。你可以用和/或来组 合过滤器。

MapDraw内置的过滤器如下:

- Overhang过滤器是根据一套你定义的Overhang准则归类的酶切位
   点。这些准则包括3'和5'端突出,与别的酶切位点互补以及兼并
   的末端突出。克隆试验中,可以使用这个过滤器寻找互补末端。
- 频率(Frequency)过滤器是指根据出现于指定的范围序列上的频率
   归类的酶切位点。我们将在本节的下一部分中使用这个过滤器型。
- 种类和复杂性过滤器(Class & Complexity)是根据酶切位点种类 归类的酶切位点。这些种类包括: I型(随机)把II型(明确),或 者I型+II型。这过滤器也可以根据位点复杂性、价钱和来源进行归 类。
- 手动选择过滤器(Manual Pick)允许你按需要选定酶切位点。在随 后的一部分中,我们学习使用手动选择过滤器的方法。
- Must Cut Here/Don't Cut Here调色板工具也是一种过滤器,随后进行阐述。

#### 频率过滤器应用

首先让我们用频率过滤器来删除任何可以两次切割我们序列的酶。

● 从ENZYME MENU,选New Filter,然后Frequency打开参数对话框。

25

- 在Min和Max对应的框中输入数字1和2,其他的参数不变。这个设定 将我们序列中切割多于两次的位点自动删除。
- 在过滤器名字框中, 输入"Two-cuts-max."。
- 单击App1y将过滤器用于序列分析——然后OK退出参数对话框。你将
   注意到在图上酶切位点的数目现在大大地减少了。

#### 应用手动过滤器

虽然减少了大约3分之2的位点,但是还有100以上的酶存在。我们还可以设定一些更严紧的命令进行限制。看一看我们的酶是否能整齐地用来切割TETHIS21序列的编码区。

- 从MAP MENU选Linear Minimapp。打开的窗口显示每个限制酶切割位 点的位置。
- 滚动窗口,直到你看到大的向右的箭头,上写"histone"。"Histone"
   是在最初的Genbank入口被注释的序列特点。当我们打开它的时候,



上单击种类调色板工具(Sort),接着选择Sort by Cuts Close to Selection。酶切位点即刻分类,这样,距特点最近而且超出选定范围的酶切位点将会排在最上面。

- 滚动到窗口的最顶端。你将注意到,在histone基因的左和右有2酶 切位点: Acs I和Apo I。两个酶切位点对我们来说都是理想的。
   现在我们将制作仅仅包括Apo I酶切位点的Manual Pick filter。
- 从ENZYME MENU,选New Filter——然后Manual Pick。打开酶切位 点过滤器编辑器。
- 把Apo I从右边的窗口拖进左边的窗口。给过滤器取名。在这我们把



过滤器叫做"Apo I."。

- 点击App1y——然后OK,就保存并应用
   "Apo I."过滤器了。
- 注意到现在线性的minimap仅仅由Apo I
   酶切位点和histone特点构成。

## 使用过滤器一览表

现在我们已经应用了2个过滤器:一个叫做"Two-cuts-max"的频率过 滤器和叫做"Apo I."的手动过滤器,MapDraw会自动把两个过滤器的 结果用"AND,"联合起来应用,这样,展示的结果需要满足两个标准: 切割一或两次,用Apo I切割。我们手动过滤器包括频率过滤器的单一 切点的内容,所以他自动的覆盖了频率过滤器的结果。下述过滤器一览 表允许我们随意编辑、组合各个过滤器。

TETHIS21	MA Map : Filter	B
Apo And Two-cuts-max And Or	Apo I	*
		4

从Map Document,单击 过滤器调色板工具(在 窗左上的漏斗图标)打 开过滤器一览表(如

下)。当前应用的所有过滤器都出现于窗口的左边。

- 为了除掉Apo I过滤器,点击选中它,从左窗口拖到右边的窗口,单 击应用即可。现在由于只有Two-cuts-max过滤器被应用,所以序列 上的酶切位点数目立即增加了。
- 把Apo I过滤器拖回,放在And旁,单击App1y再次应用Apo I过滤器, 序列上的酶切位点数目又立即减少了。(如果把Apo I过滤器拖回, 放在0r旁,单击App1y再次应用Apo I过滤器,序列上的酶切位点数 目不会改变,因为Apo I过滤器是Two-cuts-max过滤器的一部分)。
- 按上面的方法把两个过滤器都去除。

<u>使用Must Cut Here / Don't Cut Here调色板工具</u>

Must Cut Here / Don't Cut Here工具是另一种酶切位点过滤限制方法。我们将用这个方法再次重复上面的操作。

- 从MAP MENU,选Site & Sequence。
- 向下滚动,直到你看在大的向右指的箭头,上写"histone"。
   单击选中。点击Don't Cut Here调色板工具(红色园圈起的剪刀样图标),去除所有切割选定区的酶切位点。(点击Must Cut Here工具—
   一剪刀样图标——序列上将展示所有切割选定区的酶切位点。)

这样,仅仅剩下那些不切割选定区的位点。

#### 显示酶信息

内切酶编辑器(Enzyme Editor)使你能够查看内切酶的各种相关信息, 包括其名字,识别序列, isoschisomers,种类,价格,销售公司和有 效性等详细的信息。你对这些信息可以进行添加或删除操作。



#### 环形展示

你可以以环形图展示环形或线性序列,但是,如果序列是线性的, MapDraw会提示你。如果你想在环形序列上进行酶切位点限制,那末, 我们前面提到的Don't Cut Here filter完全可以做到这一点。

● 通过点击过滤器调色板工具检查当前是否只有Don't Cut Here过滤



器应用。如果已经应用,则 Don't Cut Here过滤器应 该在过滤器一览表的左边 的窗口里,而其他的过滤器 在右边。如果这不是这样, 请进行调整。

29

- 从MAP MENU,选择Circular Illustration,打开左边的窗口。(如
   果此时有太多的酶切位点,你可以加入其他的过滤器)。
- 通过OPTIONS MENU的Drawing Size调整图画的大小。2条红线相会的 角是你可以做的最小图画的范围。在窗口内点击任何位置,图形大 小将显示为那末大。
- 単击同意。

#### ORF图



止和开始codons可以使用指定的遗传密码。方法见EditSeq。

- 从OPTIONS MENU中选择ORF Criteria可以设定参数。你能调整最小的ORF长度,定义上游启动子和距上游启动子的距离,选定后单击同意即可。
- 放大ORF以查看翻译的序列。方法是点击调色板工具中Zoom In
   (有+的图标),接着,单击分析界面。重复这些2步直到你能看翻译。

## 显示选择

MapDraw's OPTIONS MENU提供各种指令允许你做下面的事情:

- 选择1个或3个字母的编码。
- 显示不同的读框组合 (Reading Frames 或 Line Layout)。
- 编辑用于翻译的遗传密码(Genetic Codes and Edit Selected Code)。酶显示可以选择垂直或者水平展示。
- 线性化环形序列 (Linearize Sequence)。
- 显示不确定位点。

保存退出

- 从编辑菜单(EDIT MENU),选保存地图(Save Map)。
- 选择保存位置,命名。
- 单击保存(苹果计算机)或同意(视窗)。
- 从文件菜单(FILE MEN),选择放弃(苹果计算机)或者退出(视窗),电脑提示保存或放弃;选择保存,则你操作的所有结果都会保存起来,否则结果会全部丢失。

31

四、MegAlign

MegAlign 提供 6 列队 (aligment) 方法,进行 DNA 和蛋白质序列的 配对和多序列比较 (multiple aligment)。多序列比较 (multiple aligment)可以在 MegAlign 的 worktable 进行查看和编辑。可以根据 队列 (aligment) 的结果制作进化树 (Phylogenetic trees),并且,有 关序列距离的数据和残基替代可以容易地作成表格。一般多序列比较 (multiple aligment) 的结果展示于队列 (aligment) 窗口,相似性和差 异用彩色的直方图展示。和全部 Lasergene 的应用程序一样, MegAlign 也提供整合的 BLAST 查找功能。

创建队列文件

MegAlign提供两种基本的队列方法: 配对比较和多序列比较。配对比较 可以比较任何2个选定的序列的相似性,而多序列比较对在Worktable 中所有序列进行比较。在我们入门的第一个部分中,我们将介绍使用2 不同的种类的配对比较方法。我们将从创建一个MegAlign文件,输入两 个histone序列开始。

- 从文件菜单,选Enter Sequences打开Enter Sequences对话框。
- 从你DNASTAR文件夹中的 "Demo MegAlign,"文件 夹,双击打开"Histone Sequences."文件夹,左 上两个序列是我们选用的 序列。

Enter Sequences						
😋 Histone seq 💠	1 of 2 Sequences Selected					
TETHIS21MA	D TETHIS21MA					
TETHIS22MA	D TETHISZZMA					
🙀 Tetrahymena H281						
Tetrahymena H2B2						
🖬 Yeast H2B1						
🖬 Yeast H2B2						
YSCH2B1 👻	*					
- Macintosh HD	Show All Types					
>> Add >> Eject	Remove Cancel					
Open Desktop	Help Done					

- 单击TETHIS21,点击Add钮。单击TETHIS22,点击Add钮。现在2序列
   将出现于右面的窗口。
- 单击Done把序列输入Worktable。



● 从OPTIONS MENU,使用Size 命令增加Worktable的字形大小直到可 以看清楚。

## 序列设置

下面我们练习subranging输入的histone序列。设置序列的末端在我们 这个例子中并不重要,但是当两个序列匹的全序列匹配不太好,或是蛋 白质和DNA的混合序列时就变得非常重要了。后者要求每个DNA序列都必 须是给出正确的翻译读框。序列设置可以通过给定末端位置手动操作设 置,也可以通过特定标注的feature进行。这里,我们将使用histone H2B-1的CDS进行操作。

- 点击选中(TETHIS21)。
- 从OPTIONS MENU,选

Set Sequence	e Ends
Feature List of TETHIS21MA	
histone H2B-1 CDS	-
H2B-1 mRNA mRNA	
H2B-1 mRNA mRNA	
	-
Segments of histone H2B-1	Change the Rest
<ul> <li>histone (232 &gt; 600)</li> </ul>	Next
	Help
	Cancel
	ОК

Set Sequence Limits, 接着,从Feature Table打开右面的序列末 端设置窗口。

- 顶端的histone H2B-1-- CDS feature已经选定。下面显示的是选定 序列的特点名字和范围。
- 单击Change the Rest用相同的feature设置第二histone序列。你将
   自动回到Worktable,这时,设置后的序列就出现了。

配对比较 (Pairwise Alignment)

MegAlign提供各种配对比较方法。我们序列是DNA,我们可以用的方法 是Wilbur-Lipman, Martinez-Needleman-Wunsch和Dotplot。在这入门 中,我们在使用Wilbur-Lipman的方法。

- 同时选中两个序列。
- 从ALIGN MENU中的One Pair选择Wilbur-Lipman方法。
- 使用默认参数进行比较,点击0K。
- MegAlign计算队列,然后在另一个窗口显示比较结果。

窗标题展示相似指数 (所右匹配残基的百 分比),缺口数目, 全的缺口长度和一致

] =		TETHIS21MA(232>6	00) vs TETHIS	22MA(231>5	99)	P
×	Yilbur-Lipman DNA A Kiupla : Ti Gao Republi	lignment 1975 : Mindeau (202				
7	Seq1(232)600) TETHIS2101A	Seq2(281>599) TETHIS22MA	Similarity Index	Gap Number	Gap Length	Consensus Length
	(238:600)	(237×599)	94.2	0	0	363
	v240 v2 CCCAAGAAAAGCTCC CCCAAGAAAAGCTCC CCCAAGAAAAGCTCC ^240	50 0260 CECTECCECTECTEAAA CECTEC CT CTEAAA CECTECTACTACTEAAA 250 ^260	v270 NASAASGTCAAS NASAASGTCAAS NASAASGTCAAS ^270	v280 AA66CCCCCCA AA66CCCCCCA AA66CCCCCCA ^280	v290 CAACCEAAA C ACCEAAA CCACCEAAA 2290	v300 Agaagaaca Agaagaaca Agaagaaca ^300
siti	ion : 1	4 11				•

序列长度。我们可以改变标志颜色来区别匹配和不匹配的序列。

单击队列调色板工具 (有"x"方框图标 ),打开下面的队列颜色编

- 辑读化框。 Match Color Mis match Color Consensus Color 单击深蓝色的方框,接着点击Match Color右面 ☑ Show Header 🗹 Show Ruler 的黑色方框,他就变蓝了。所有和一直序列一样 Show Names Show Context Show Identities As: 的碱基都立即变成深蓝色。如此重复可以改变不 Gesidue Characters Vertical Bars 匹配的序列的颜色。
- 另外,还可以通过选定或不选彩色选择方框下面的方框查看队列窗
   口中队列的变化。如你单击Vertical Bars,则匹配序列间的碱基变
   成了垂直的棒。

#### 使用点划分方法(Dot Plot Method)

点划分方法是先把要比较的序列叠加起来,然后计算匹配不当的数 目。

- 同时选定两个序列。
- 从ALIGN MENU中的One Pair1里选 择Dot Plot打开右面的窗口。
- 单击OK用默认的参数进行比较。
   MegAlign计算结果会在另一个窗口显示出来。



每个匹配都与特定的一组残基有特定的相似性(两个都设计在参数对话 框中), Dot Plot窗口中展示为蓝色。从左侧末端开始的红色斜线,表 示两个序列比较的位置。

	🗆 TETHIS21MA(232>600) vs TETHIS22MA(231>599) 巴 臣								
	Dotplot disgonal				-				
77	Seq1 TETHIS21MA	Seq2 TETHIS22MA	Similarity Index	Consensus Length					
	(469>498)	(267)296)	50.0	80					
	TETHIS21MA	0470 0480 ARCAAGAGAAGAACCCTCT	v490 Catccagesaa	1					
	TETH1S22MA	ARGAREGTERREAREGEEC	290 290	1	-				
Pasit	Position:-55 4 1								

● 双击任意一条蓝色的线会打

开左面的窗口,其中显示所选定的序列的比较结果。

## 多序列比较

为了在MegAlign中进行多序列比较,我们选择一个含有十四个相关的钙 调蛋白序列的文件进行操作。使用这个大的数据库我们可以制作进化 树,了解MegAlign的其他特性。

首先,我们要创建一个包括14 calmodulin序列的新文件:

- 从文件菜单,选Open打开DNASTAR文件夹的"Demo MegAlign"的文件夹。
- 双击打开"Calmodulin Alignment"。

从OPTIONS MENU,使用Size命令增加字体大小到看得清楚为止。
 现在我们需要选MegAlign's的两个多序列比较方法中的一个进行操作:
 Clustal或者Jotun Hein。如果已知序列有一定的同源性,我们推荐使用Jotun Hein,并且,如果有关序列相关性的背景未知,可以选择
 Clustal。我们使用的序列已知全部是calmodulins,所以,我们使用Jotun Hein方法。

在我们实行队列之前,我们应该选择一个权重表。MegAlign's残基权 重表用于对多序列比较进行评分,这样那些虽然残基不匹配,但残基化 学性质相似的序列的评分要比化学性质不相似的序列的评分要高。我们 的序列是蛋白质,并且我们将使用Jotun Hein方法,所以"Structural" 表是最好的选择。

36
Residue Weights of Calmodulin Alignment 🛛 🖻 🗎														lin /	lig	nm	ent				- 6	38
ii		ж	i i	C-er	least	ī	- Shen	Die	lan re	i.	Ē	štru	ctu	iral	-	a i	1					ai
-				-				-		-	-	_	-		-	-		-	_	_	+	-
- 1	¢.	5	Ţ			6		9	E.	9	н		ĸ	м		L	Y		Ŷ	*	1.	-
<u>e</u>	•																				LC.	
	4	2																			12	
1	2	÷.																			11	
51	2		1																		15.	
2	2	2	1	1	÷																12	
9	- 2		1	4		•	_														6	
"	2		-	2	- 2	2															L.	
D.			2	2	- 1	- 1															P.	
5	0	2	2	2	4	4	1	1													15.	
2		2		1	2	- 2	- 2	1	4	°.											12	
н	2		2	2	2		4	2	2	4	•										15	
•	2		2	2	2	- 2	3	2	3	- 2	4										P.,	
ĸ	0		4	2	- 2	2	4	2	9	4	- 2										K.	
м	2	2	2	2	2	1	1	2	2	2	2	2	2	•							P.	
	2	2	3	2	2	2	2	1	1	1	2	2	2	4	۰						Ľ.	
5	2	÷.	2	1	2	-	1	1	1	2	÷.	2	2	÷.	2	9					15.	
Y	2	4	1	1	5	4	2	1	4	2	1	2	1	4	2	2					Ľ	
5	3	3	2	1	3	2	2	1	2	1	2	1	1	3	4	4	4				Ľ.,	
Y	3	3	2	2	2	2	1	2	1	2	3	1	1	2	3	1	3	-	6		Ľ	
۳I	3	2	1	2	2	3	a	4	1	1	1	2	1	3	3	4	1	3	3	6	J۳.	
	c	s	T	P	٨	G	N	.0	5	9	н	R	ĸ	M	Г	L	Y	F	٧	*		
																					_	-
4																						16

- 从ALIGN MENU选择Jotun Hein Method。
- 队列进程窗显示比较完成的百分比。
- 当队列完毕,Worktable会显示队列的

结果。

- 从ALIGN MENU选择Set Residue
   Weight Table打开左面的窗口。
- 从上面的下拉菜单选择 Structural。单击同意。
   现在我们可以比较我们的

calmodulin序列了。



þ	-	-	-	-	-	-	-	Pa	ir D	lista	ince	s of	Calı	nod	ulin.	Alig	nn	nen	t ==		P	DE	3	)	通过VIEW MEN	UΦ
1									Pe	scew	: Ident	kγ –														- 1
Ш			1	2	3	4	5		6	7	8	9	10	11	12	2   1	3	14	1	1			11			
Ш		11	-	82.0	90.T	90.	1 90	0 19	0.7	9 <b>9</b> .J	86.7	192.0	190.	82.	7 99	7 9	1.37	99.7	1	7	Barley Colmodalin		Ш.			
Ш		2	20.6		84.7	83:	1 B4	.0 18	4.7	82.7	B1.1	80.	B4.	99.	3 82	0 8	2.0	81.3	2	-	Black Mold Calmodulin		Ш.		11.0	
Ш		3	10.0	17.2		9D.	109	4 1	00.0	0.00	88.7	100.3	100	0 85.	8 90.	0 90	1.D	90.7	3	- 1	Chicken Calmodulin		Ш.		for Sequence	
Ш		1	10.0	18.9	10.0	1	50	0 9	0.7	90.0	192.7	169.3	190.	84.	0 89.	3 9	1.0	90.7	1	1	Gilliste Calmodulin		Ш.		#\$209400m00	
Ш		5	10.8	18.0	0.7	10.8			9.1	89.1	88.0	100.0	109	84.	7 89.	3 8	1.3	90.D	5	1	Bectric Eel Calmodulin		Ш.			
Ш		6	10.0	117.2	0.0	110.0	Î D.	7 1	-	90.0	188.7	90.5	100	D 85.	3 90.	$\tilde{0}$	1.0	90.7	6	-	Human Calmodulin		Ш.			
Ш		7	0.7	19.8	110.8	10.8	1 TT	311	0.8		86.0	1912	190.	81	3 199.	3 9	17	99.3	7	- 1	Lilly Calmodulis		Ш.		<b>N</b> • . <b>LF</b> - 2	1
Ш	e.	8	14.7	215	12.3	17.7	113	111	2.1	15.5	1	88.	68	82	0 85	3 8	i.D	86.7	8	-	Paramecium Calmodulin		Ш.		Distanc合有户	トカ川
Ш	ž	9	8.5	22.4	110.0	11.2	10	a ji	0.0	9.2	12.0	1	50.	1 81.	3 91.	3 9	1.3	92.0	9	- 1	Parato Calmodulin		Ш.		21000000月八	14
Ш		10	10.0	117 2	0.0	110.0	I D.	T	0.0	10.8	12.3	10.0	1	85	3 90.	0 9	1.0	90.7	10		Bat Colmodulin		Ш.			
Ш		11	19.8	0.7	16.4	18.	117	2 11	6.4	18.9	20.6	121.5	16.	1	82	7 8	1.7	84.D	11	1	Red Bread Mold Calmodulin		Ш.			
Ш		12	1.3	120.0	10.8	110	r ti i	311	0.8	0.7	16.4	102	10.1	e 1 Ta.	8	<b>1</b> 8	1.D	98.7	12	-	Red Bryony Colmodulin		Ш.			
Ш		13	0.7	20.0	10.8	10.8	111	3 11	0.8	13	15.5	192	10	8119	8 Z J	1	-	98.D	13	1	Rice Calmodulin		Ш.		的丢别和相似。	灶
Ш		14	1.3	18.9	10.0	10.0	1 10	8 1	0.0	0.7	14.7	185	10	0 1 18.	0 1	1 1 2	Ĵ.		Ê 14	1	Soubean Calmodulin		Ш.			1-1-
Ш			11	2	3	4	15	-+-	6	7	1 8	÷ 0	10	11	11		3	14		- 1	,		Ш.			
Ш																				-1		14				
Ш	4	-															_			_	la la	10	1		<i>.</i>	
																					1	119			(++++)	
																									いがなり。	

- 为了继续,关上Residue

Substitutions and

Sequence Distances窗口。

<u>Phylogenetic Tree查看</u>

通过VIEW MENU中的Phylogenetic Tree打开下面的窗口。默认为
 Balanced Branches调色板(从顶端第3)。在phenogram中,距离长



Branches调色板工具(从顶端第4)。在cladogram中,枝长度(branch length) 是与祖先的节差异的评估。

● 为了继续这入门,关上Phylogenetic Tree窗口。

### 查看队列报告

队列报告以序列显示比较的结果。

- 通过VIEW MENU中的队列 报告命令(Alignment Report)打开报告窗口。
- MegAlign允许你改变队
   列报告的外观。从

Alignment Report Contents	
Show Consensus Disagreement Show Consensus Strength Show Consensus Show Ruler of the Consensus Show Se guences	Extra Space     between Residues     fit Breaks to Page     Break Alignment
<ul> <li>☑ Show Sequence Names</li> <li>☑ Show Sequence Positions</li> <li>□ Show Translated DNA</li> <li>☑ Use 3 Letter Amino Acids</li> </ul>	🗌 Exchange Names & Positions
Help Save as Defaults	Cancel OK

OPTIONS MENU中选择Alignment Report Contents打开右面的窗口, 选定第3-7和第9个方框,其它不变。

● 单击同意,更新报告。

### 创建Decorations和Consensi

另外,我们可以通过加入"decorations"和"consensi."来优化展示的效果。Decorations包括加框、加阴影和隐藏。在这节中,我们将给那些与一致序列不一致的氨基酸加阴影。



- 从OPTIONS MENU选择New
   Decoration来打开在左
   边显示的对话框。
- 在题目框中,输入类似 "Shade disagreements with consensus."
   的名字。
- 下一个排包括3下拉的菜单。选择第一个下拉菜单中的"Shade", 中间菜单的"residues differing from"。第3个菜单不变,默认 为Consensus。选择第一个下拉菜单中的"Shade"后,会激活其下 面的另外两个下拉式菜单。
- 从上边的菜单选择颜色,从下面的菜单选择阴影的样式。
- 距离单位框不变,为"0."。
- 单击同意,现在队列报告中所有与一致序列不一样的残基全都被加 上了阴影(如下)。

	1eg	Aliq	jn -	[A]	ligr	nm	enl	t R	epo	ort	of	Ca	ilme	odu	lin	Ali	gni	ner	it.n	neg	1]																
	File	E	dit	Ali	gn (	Vie	ew	C	Opti	ons	; I	Hel	p																								
	М	A	D	QI	Т	E	Е	Q	I	A	E	F	ΚE	A	F	s	L	FΙ	K	D	G	D	G	т	I	Т	т	K	ΕI	6	G	Т	V I	MI	R 9	5 L	Majority
									10								1	20									30									40	3
	-				-	-	-	-	-		-	-			-	~		1				-	-	2	-	-	-		-		-	-				-	
1	м	A	D	ųι	Т	Ð	D	ų	1	A	E	F	KE	A	F.	S	г	F. L	' K	D	G	D	G	U	T	Т	Т	ĸ	EI	ե	G	Т	۷.1	Mł	κ.	5 L	Barley Calmodulin
1	М	A	D	S I	Т	Ε	Ε	Q	A.	S	Е	Y	ΚE	A	F	s	L	FI	K	D	G	D	G	Q	Ι	Т	Т	K	ΕI	L	G	Т	V I	MI	R 9	5 L	Black Mold Calmodulin
1	М	A	D	QI	Т	Ε	Ε	Q	Ι	A	Ε	F	ΚE	A	F	s	L	FΓ	K	D	G	D	G	Т	I	т	т	K	E I	L	G	Т	V I	MI	2 9	5 L	Chicken Calmodulin
1	M	A	D	NI	Т	Ε	Ε	Q	I	A	E	F	ΚE	A	F	s	L	FL	K	D	G	D	G	Т	I	т	Т	K	ΕJ	L	G	Т	V I	MI	R :	5 L	Cilliate Calmodulin
1	М	A	D	QI	Т	Ε	Ε	Q	Ι	A	E	F	ΚE	A	F	s	L	FΓ	K	D	G	D	G	Т	Ι	Т	т	K	EJ	L	G	Т	V I	MI	2 9	5 L	Electric Eel Calmodulin
1	М	A	D	QI	Т	E	Е	Q	Ι	A	E	F	ΚE	A	F	s	L	FΓ	K	D	G	D	G	т	Ι	т	Т	K	ΕI	L	G	Т	V I	MI	R :	5 L	Human Calmodulin
1	М	A	D	QI	Т	D		Q	I	s	E	F	ΚE	A	F	S	L	FI	K	D	G	D	G	С	Ι	Т	Т	K	ΕJ	L	G	Т	V I	MI	R 9	5 L	Lilly Calmodulin
1	М	A	E	QI	Т	Е	Е	Q	Ι	A	E	F	ΚE	A	F	A	L	FI	K	D	G	D	G	Т	Ι	Т	т	K	ΕJ	L	G	Т	V I	MI	R :	5 L	Paramecium Calmodulin
1	М	A	E	QI	Т	Е	Е	Q	Ι	A	E	F	ΚE	A	F	s	L	FΓ	K	D	G	D	G	С	Ι	Т	Т	K	ΕJ	L	G	Т	V I	MI	R :	5 L	Potato Calmodulin
1	М	A	D	QI	Т	Е	Е	Q	Ι	A	E	F	ΚE	A	F	s	L	FΓ	K	D	G	D	G	Т	Ι	Т	т	K	ΕJ	L	G	Т	V I	MI	R 9	5 L	Rat Calmodulin
1	М	A	D	S I	Т	E	Ε	Q	V	s	E	F	ΚE	A	F	s	L	FI	K	D	G	D	G	Q	Ι	Т	Т	K	ΕJ	L	G	Т	V I	MI	R 9	5 L	Red Bread Mold Calmodulin
1	М	A	D	QI	Т	D	D	Q	I	s	E	F	ΚE	A	F	s	L	FI	K	D	G	D	G	C	I	Т	Т	K	ΕJ	L	G	Т	V I	MI	R 9	5 L	Red Bryony Calmodulin

Consensi是一致序列的另外一种图形展示方法,可以用来标志 ambiguous残基。另外,序列中一致和不一致的残疾可以用直方图在队 列报告中展示。当在某个位置上,任何一个序列的残基与其他序列不一 样时,一致序列就会以星号表示。并且用直方图的长短显示其中一致残 基的多少。

- 从OPTIONS MENU,选择New Consensus打开右边的对话框。
- 在题目框中,输入诸如 "A11 sequences matching potato."
   的名字。

Alignment Consensus All sequences match potato
Template: Potato Calmodulin 🔻
When all match 🔹 the template residue 🔻
show the template residue 🔻 otherwise show 🗴
A majority is 7 sequences 🗹 Show on Report
Show Histograms of: 🛛 Strength 🔲 Disagreement
Help Set Groups Cancel OK

- 下面有4个下拉菜单。在第一个菜单选择Potato Calmodulin。其他 的菜单不变: When all match; the template residue 和 show the template residue。
- 在右边框"otherwise show,"中输入星号(\*)。
- 选定"Strength."框。



方图可以显示新的一致序列的强度。

保存MegAlign文件

- 从文件菜单,选保存。
- 确定文件夹的保存位置。
- 在文件名框(视窗)或名字框(苹果计算机)中给序列命名。
- 单击保存(苹果计算机)或同意(视窗)。
- 对队列和队列报告的变化都被保存了。

五、PrimerSelect

PrimerSelect 能够帮助你设计 PCR, 测序和交杂实验所使用的引物 和探针。输入你的 DNA, RNA 或反向翻译的蛋白质模板序列后, PrimerSelect 就可以在 pentamer 窗口计算序列的融解温度, 自由能和 末端自由能。您可以通过控制包括引物浓度、盐分和.G 计算温度等参 数来限定计算结果。

在模板处理后, PrimerSelect 按照用户定义的参数确定引物的位 置,并给引物评分,然后筛选出模板序列上的最佳引物序列。引物 "workbenches"允许你修改你的引物,检查参数编辑对翻译读框,二 级结构,错误的引物位置和限制性酶切位点的影响。如果你选择并优化 了你的引物, PrimerSelect 可以让你建立有关模板,引物,扩增的文 件。和全部 Lasergene 的应用程序一样, PrimerSelect 也提供整合的 BLAST 查找功能。

42

# 创建 PrimerSelect 文件

我们将以克隆载体 pBR322 为模板 DNA 序列。在这一节中,我们选择引物,用来扩增位于 2013-2169 位的 H-strand Y effector。首先我们从打开模板序列开始。

- 从文件菜单选择 New 创建一个 PrimerSelect 文件。
- 从文件菜单,选择输入序列(Enter Sequence)。打开下面的对话
   框。
- 从"Demo Sequences"
   选中 pbr322.seq(视 窗口)的文件,或者
   PBR322(苹果计算机),



单击 Add 添加序列到右面的窗口。然后点击 Done 打开序列。



#### Anatomy of the PrimerSelect Document Window

### <u>定义引物特点和位置</u>

下面我们开始确定引物的类型。

● 从条件菜单(CONDITIONS MENU),选择 Primer Characteristics

	Minimum		Maximur	n
Primer Length	17 I	bp	24	bp
3' Pentamer Stability	8.5	-kc/M		
Melting Temperature	34.4	'n	66.7	۹C
Overall Stability	54.7	-kc/M	33.1	-kc/M
Unique 3' Sequence of	7 1	b <b>p</b>		
Accept Dimer Duplexing a	ส		2	եր
Accept Hairpin Duplexing	of		2	bp
Ignore Duplexing	8 1	bp fron	n 3' end	_
Ambiguous Residues			0	bp
Help D	efaults C	ancel	0	ĸ

样重叠示可以接受的。

- 不改变默认设定和单击 Cancel 退出。
- 从条件菜单,选择 Primer Locations 打开右边对话框。这个对话框

允许我们根据给的序 列的长度来限制引物 的位置。

●从下拉菜单的 Restrict Locations

Restrict Locations by:	Upper and Lower Prim	er Ranges 🛛 🔻
	Minimum	Maximum
Upper Primer Locations	1700	1950
Lower Primer Locations	2250	2500
Avoid Locations Contain	ing Repetitive Sequence	within 10 bp of 3'
	Help Default	s Cancel OK

打开左面的对话框。这对话框

可以用来设定各种参数诸如:

引物长度、最大的 3' pentamer

稳定性和可接受的重叠。注意

默认的引物长度为 17-24 bp,

并且1或2bp的二聚体和发卡

中选择 Upper and Lower Primer Ranges。当在全序列中设计编码 区的 PCR 引物时,这是最好的选择。

- 在 Upper Primer Locations 右面的方框中输入 1700 和 1950。在
   Lower Primer Locations 右面的方框中输入 2250 和 2500。
- 单击同意,会出现下面的窗口。成对的绿色和红色的三角形分别标

记所设定的正向和反向引物的范围。

### 寻找引物对

现在我们已经指定什么种类的引物对是



可接受的,剩下的工作就是让 PrimerSelect 为我们寻找理想的引物对。
从 LOCATE MENU 选择 PCR Primer Pairs。打开一个二分窗口,上面窗口中显示计算机寻找到的 28 对理想引物对,引物对按照评分值依次排列。因为调色板工具中的 Alternate Pairs 被默认激活 (在窗口左面工具中),所以底部窗口显示的是上面窗口选定的引物对的替代引物对一览表。





通过拖动在窗口的右侧上的小环拉开建议窗口,可以查看
 PrimerSelect 对选定引物的建议。

注意现在的"Best choice."是 PrimerSelect 根据在最初的条件中设定的参数计算得到的。看下面的窗口, PrimerSelect 列出了"Best choice"引物对的 25 替代选择。这些替代的引物对扩增的区域和上面选定的引物一样,所不同的是他们的序列有些不一样。

● 选择调色板工具中的 Alternate Products (左侧工具底部), 下面



窗口显示引物的位置和可能错配的位置,以及相应的产物长度。左面的那个单一的粗黑棒提示用我们

"Best choice"引物进行反应,会有大量产物被扩增出来。如果没有多个 5'或者 3'引物箭头也提示这一对引物没有错配点。

单击上面窗口里引物对,
 观察下面窗口里的产品棒
 变得越来越细表示预期的
 扩增物也越来越少。



注意有 8 个引物对有错误的配对点,用粉红色或淡绿色箭头指示出来。 第 13 对和第 22 对正向引物有错配点(右边显示),就会用第二条产品 棒的表示错误的产品大小。

查看引物的其他信息

	Ampli	ficatio	n Summar	y			E
	8		8				
Upper Primer: 2 Lower Primer: 2	8-men 5' 4-men 5'	60060. 60460	ATCCATACCGCC BASTCASTBASC	AGTTGT SAGBAA	. 8. 83.	1	
DNA 250 pH, Sat	t 50 mH		Upper Prime	ar Lou	er Primer	11	
Primer In Primer Overall Primer Location	Stablilty		65.5 °C −49.5 kc/t 1766 1788	- 78	63.3 °C 17.4 ke/m 1842361		
Product Tn - Pr Primers Tn Diff Optimal Annealt	lmer Tn erence ng Tenper	ature	-~0	7.9 °C 2.1 °C 1.0 °C			
Product Length Product Th (200 Product 80 Cont Product Th at 5	Nethod) ent xSSC		e B K	519 bp 1.3 °C 54.88 32.9 °C	;		
Product	Helting	Tenper	oture (SBC	Nethod	>	1	
Salt			Forma	ntde			
ni xssc	XSSPE	ΟĦ	108	208	508	11	
1 0.005 10 0.051 50 0.256 165 0.846 330 1.692 500 2.564 1000 5.128	0.006 0.052 0.312 1.031 2.052 3.125 6.250	53.1 69.7 81.3 89.9 94.9 94.9 102.0	46.6 63.2 74.8 83.4 88.4 91.4 96.4	40.1 55.7 68.3 76.9 81.9 84.9 89.9	20.6 37.2 48.8 57.4 52.4 55.4 70.4		
195 1.000	1.219	0.0	\$formanid	e = Tn	102.9 °C	1	Ŧ
4						F.	lą

PrimerSelect 能预测引物形成 self-dimers, pair-dimers and hairpins 的可能性。

- 单击选定 "Best choice" 引物。
- 从报告菜单(REPORT MENU),选
   Amplification Summary,打开左
   面的窗口。显示用所选个引物做
   PCR 扩增反应的报告。

- 选择 Composition Summary, 打开右面的窗口,描述扩增区 和引物对的碱基组成。
- 从报告菜单,选择 Self
   Dimers, Pair Dimers 和/或
   Hairpins 打开新窗口显示每
   种二级结构的预测结果。
   Self-Dimer 形成窗口如右。

Upper Primer: 29-men 51 00 Louer Primer: 24-men 51 00	Ition Summe DBCATCCATACCS AGCGAGTCAGTGA	CCAUTTUT GCCAUGGAGGAAG	3. 3.	
Primers	Uppen Prine	r Loue	- Priner	1
Single Strond Mr Extinction Coefficient I/E	7.0 k 4.78 nΠ/A2t 88.7 μg/A2t	0 4.03 0 30.6	16 k n11/4250 μg/4250	
Product	Composition	Quantity	Per Cent	1
Upper Strand Nr         191.3 k           Loven Strand Nr         191.8 k           Both Strands Nr         83.2 k           Length         619 bp           Tn (200, Rethod)         81.3 %	4 C O T I	152 179 161 128 0	24.6 28.9 20.0 20.7 0.0	
Th at 6x890 102.9 °C 80 Content 54.88	A+T G+C	280 340	45.2 54.9	



### 按照引物特点分类

PrimerSelect 在自动将引物按照从好到坏的顺序分类以前,已经计算 了引物的各个特点。Located Primers & Probes 窗口是我们能够根据 引物的不同特点,如引物位置、长度、溶解点或自由能等,来对引物进 行分类。这里让我们用溶解点来分类引物。

			Located	t Prim	ers		E
~	Start	End	Length	T1	<u>۵</u> ۵	∆Profile	_
Upp	er Pri	lers:	33 Locat/	ed			
~	1730	1753	24-ser	60.2	-46.0	79.2	
~	1731	1754	24-ser	58.1	-44.8	66.3	
✓	1752	1774	23-mer	66.6	-50.6	102.8	
¥	1760	1778	19-ser	59.6	-42.2	166.2	
~	1765	1787	22-ser	64.9	-48.Z	218. I	
~	1766	1788	23-mer	65.5	-49.5	218.4	-
Lor	er Pri	lers:	28 Locaty	ad			
~	2334	2351	18-ser	47.8	-35.5	49.1	4
~	2335	2351	17-sec	45.4	-33.9	40.8	⊨
~	2352	2375	24-mer	64.8	-49.1	62.1	
~	2354	2377	24-ser	62.7	-47.3	41.8	
~	2361	2384	24-ser	63.3	-47.4	116.7	H
~	2362	2384	23-mer	62.6	-45.8	108.6	
•						•	12

- 从报告菜单,选择 Located Primers
   & Probes 打开左面的窗口。窗口显示符合我们标准的引物的一览表。在没有特别指定的情况下,引物按照位置排列。
- 从LOCATE MENU选择 Sort Primers 打开下边对话框。
- 保留对正向和反向引物的选定,去除对 By Position 的选定。选定
   By Tm Nearest 方框,输入 65。

单击同意,现在两个窗
 口里的引物按照融解
 温度离 650C 最近的程
 度依次排列。

Sort Primers		
🗹 Sort Upper Primers	🔲 By Length nearest	24
🗹 Sort Lower Primers	🗹 By Tm nearest	65
🔲 By Position	🔲 By 🗚 Gnearest	0.0
3 5' end close to	🛄 By 🕹 G Profile nearest	1000.0
3" end close to	Help Cancel	ОК

#### 改变引物长度

PrimerSelect's工作台允许你在引物中引入限制性酶切位点,突变,改变使用的密码子,核查引物二级结构,查找错配点。在这一节中,我们将用正向引物工作台来扩展我们的引物长度。

- 双击选定正向"Best choice"引物。
- 从编辑菜单,选择 Work on Upper Primer 打开下面的窗口。
- 确认 Dimer, Hairpin 和 Priming Sites 调色板工具 (窗口左侧从顶 端起的 4-6) 已经激活。



让我们仔细查看一下工作台。在窗口 下边的第3排可以看见引物序列。序 列两端的三角形"handles"允许你 缩短或者延长序列。在窗口的左上角 的标题告诉我们引物的长度是 23 核 苷酸,溶解温度为 65.5 度;而底部 窗口的绿色棒显示引物在序列上的 大概位置。下面的信息告诉我们没有

大于 2 bp 的二聚体形成,并且引物只能形成一个并不理想的发卡样结

构。拖动窗口左边的滚动条之间的黑框调整上下窗口的比例。

拖动引物左上黑三角形直到长度延伸到 33 核苷酸。

现在引物的长度已经显示在窗口的标题中。引物的溶解温度也增加到 78.6度,将难以进行 PCR。看下面的窗口有二聚体和发卡形成,并标记 (坏!)。由于长引物容易形成稳定的二聚体和发卡样结构,所以我们 设计的引物经常较短。

再一次拖三角直到引物最初的23核苷酸。窗口又和我们开始打开它的时候一样了。

#### <u>引物中引入突变</u>

PrimerSelect 工作台可以很容易的在引物中引入突变。我们 "Best choice" 的正向引物序列只编码一个苯丙氨酸。在这一节中,我们把读



框1的苯丙氨酸通过改变读框中的密码子转 换成苏氨酸。转换完成后我们将可以观察到 我们变化的效果。

- 点击引物下面,读框 1 中绿色的苯丙氨酸,会出现一个小盒子(如右)。盒子中显示全部 4 种可能的苯丙氨酸密码子和词"0ther."标记的 GCC 是这个苯丙氨酸残基的密码子。
- 选择 Other 会自动跳出一个子菜单,其中显示所有的氨基酸及其代
   号,在 "T-Threonine."上点击一下,模板序列上的碱基组成没有

改变,但是引物序列和其编码的氨基酸(Thr-Ala)发生了改变。 Thr 显示为红色,表明它是从最初的引物序列上突变得到的。查看屏幕 底部的左角,可以看到引物形成了一个稳定的二聚体。我们可以通过给 我们的 Thr 残基做沉寂突变来删除这个不受欢迎的二聚体。我们试着把 苏氨酸的密码子改变为 ACC。这个密码子与苯丙氨酸密码子(GCC)只 有一个核苷酸不同。

● 单击红的 Thr 残基,从菜单中选 ACC 代替现在的 ACT。

下面的窗口显示从G到A的单一的核苷酸变化去除了二聚体。另外,正 向引物的融解点已经从65.5减少了到53.7。也许正向引物新的溶解温 度可以与反向引物匹配,但是我们可以找到与之更相匹配的反向引物。

- 在工作台的右上的白色文框输入突变引物的名字。
- 点击窗口左侧顶端的 0K 钮,保存突变的引物。
- 从 LOCATE MENU 中选择 PCR Primer Pairs 更新窗口。新的引物对会加入窗口中,突变的正向引物用淡绿色的箭头表示。
- 打开建议帘。没有警告的第一个突变引物对是从顶端起的第11个。
   这一引物对被评定为"OK",如果我们必须使用突变引物的话,虽然用它扩增的产品数量不多,但是它却是我们最好的选择。

#### 设计新引物

当你初次打开 PrimerSelect 的时候,引物目录是空的。如果你愿意,你可以用这个目录来储备任何在 PrimerSelect 中设计或者被修改的引物。LOCATE MENU 中的 Only Cataloged Primers 指令可以让你在已经

50

储存地引物中选择合适的引物供使用。在这一节中,我们学习如何设计引物并将他们输入目录表的方法。

- 从LOG MENU选择 Primer Catalog 打开引物目录表。注意到我们突 变的正向引物已经在目录表中。引物左面的检测标志表明引物处于 激活状态,V形臂章(>>)(视窗)或子弹(苹果计算机)显示它 已经通过了最初二级结构检测。
- 按回车键创建一个新的引物区。输入"Test Primer"。
- 按Tab键。
- 输入 "GIANTCATSNABMANYWARYRATS"。
- 按Tab键。
- 输入"Contains 46% ambiguous bases"。
- 按回车键结束。引物目录
   表应该如右图。

E		_		Primer	· Catalog		Η
E	/	•	None	Length	Sequence	Note	$\Box$
IF	~	٠	Ald to The mutant	23	BEEGEATECATACEACEAGTEGT		-
ŀ	r		Test Primer	20	GLANTCATSNABMANYRATS		
ŀ	~		Untitled	0			
1							-
	€	Ш				ł	47

如果我们的引物已经通过了最初二级结构检测,则引物左侧出现V形臂 章或枪弹。现在没有V形臂章或枪弹,说明我们引物测试失败。让我们 核查为什么:

- 从报告菜单,选 Primer Self Dimer,可以看出该引物可以形成12
   对稳定的二聚体。
- 从报告菜单,选 Primer Hairpins,可以看出该引物也可以形成 5
   个稳定的发卡。

如同我们预测的那样,我们序列里的模糊碱基的百分比过高会使得它那

以用来进行 PCR 反应。

### 使用寡核苷酸订购表

PrimerSelect 提供两种寡核苷酸订购表: Generic 和 IDT (Integrated Data Technologies, Inc.)。

- 点击选定正向"Best choice"引物。
- 从 LOG MENU 选择 01igo Request Form 中的 IDT 打 开 IDT 寡核苷酸订购表 (如右)。我们反向和正 向引物序列自动输入表 格。如果我们想定购引物, 我们只需填表打印,然后 传真或用电子信件邮给页 底部的地址即可。
- 从文件菜单,在返回附近
   到坐落的引物一对窗口口做选择。

PrimerSelect 文件保存

- 从文件菜单,选保存。
- 选定文件的保存位置,命名。



单击保存(苹果计算机)或同意(视窗)。任何对引物,或者引物
 目录表的操作都将和文件一起被保存。

六、Protean

Protean 可以使用多种方法分析、预测蛋白质结构,并以图形化的 方式展示出来。预测蛋白质结构。各类方法按照科学概念进行分类。几 个方法可以同时存在于一个概念群中,如用于分析疏水性概念的方法就 有好几种,不过有的概念只有一种方法可供选择,如柔韧性。你可以按 照任何顺序在 Protean 文件上展示各种方法计算的结果。另外,Protean 可以输入来自蛋白质数据库中标注序列特点,而且允许你注释新特点。 和其他 Lasergene 应用程序一样, Protean 也提供整合的 BLAST 查找功 能。

### 创建蛋白质分析文件

这一节,我们将为人的钙调蛋白序列创建一个新 Protean 分析文件。

- 从文件菜单选择 New 打开右面的对
   话框,然后打开名为"Demo Sequences."的文件夹。
- 双击"Human Calmodulin" (或 "Human Calmodulin. pro"),这样 就打开下面的分析文件窗口:

What Sequence?	
🔾 Demo Sequences 💠	👉, 🚺, 🔇,
Noma	Date Modified
I 10K_PLAKN	B/27/92 🔺
🛃 GALM\$HUMAN	5/4/92 -
hum-thistical polyphos4ptase	10/5/97
目 Human Calmodullin	5/4/92
🗢 💐 pFI 753	1/30/98
🧱 Tetrahymens H2B1	5/4/92 💌
Show: Protein Files 0	
Human Calmodulin(1>148)	
ADQLTEEQIA EEFVQMMTAK	Set Ends
Length: 148 AA Range: 148 AA	
Cancel	Open

Anatomy of a Protean Assay Document



### Protean's 蛋白质分析方法

打开序列后就是选择应用方法。方法结果的图形显示可以帮助你选择感 兴趣序列的特点。对于我们打开的序列,你会发现只有下面方法中的几 个被展示出来了,下列方法都可以使用:

- Title——给文件取名。
- Ruler——给文件加标尺。
- Sequence——显示序列。
- Protease Map——识别序列上的蛋白酶酶切位点,并且显示酶切图 谱。
- Patterns-Prosite 数据库——在 Prosite 数据库中检索你的序列。
- Patterns-Ariadne 文件——在序列上寻找用户指定的 Patterns。
- 电荷密度——电荷——预测电荷在特定的序列范围上的分布。
- 二级结构——Coiled Coil——预测跨膜区的阿尔法螺旋。
- 二级结构——Garnier-Robson——计算特定氨基酸残基在特定结构 内部的可能性。
- 二级结构——Deleage-Roux——蛋白质的二级结构类型预测。
- 二级结构——Chou-Fasman——通过序列氨基酸残基的晶体结构来 预测蛋白质二级结构。
- Hydropathy Goldman-Engleman-Steitz 预测可以跨过细胞膜的 非极性阿尔法螺旋。
- Hydropathy-Kyte-Doolittle——根据序列的氨基酸组成预测蛋白 质的疏水区和亲水区。
- Hydropathy Hopp-Woods 通过计算蛋白质序列上的最大局部亲水 性寻找蛋白质的抗原决定簇。
- Antigenicity Sette MHC Motifs——预测短肽上与老鼠 MHC II d
   型蛋白质相互作用的抗原位点。

- Antigenicity AMPHI 根据序列预测免疫优势辅助性 T 淋巴细胞 抗原位点。
- Antigenicity Rothbard-Taylor 预测含有特定基序(motif)的
   潜在T淋巴细胞抗原决定簇。
- Antigenicity Jameson-Wolf 通过联合现有的蛋白质结构预测方
   法预测潜在的蛋白质抗原决定簇。
- Amphiphilicity-Eisenberg-预测Eisenberg Moment。
- 表面可能性——Emini——预测特定区域位于蛋白质的表面的可能 性。
- 柔韧性-Karplus-Schulz-预测蛋白质骨架区的柔韧性。

#### 分析方法应用

应用新 Protean 方法的步骤是:把所用方法从 More Methods 菜单移入 方法帘,然后拖入分析窗口,基本上与前面介绍的 GeneQuest 方法相同。 在这一节中,我们将练习使用 Deleage & Roux 分析方法预测蛋白质结 构。

- 从分析菜单(ANALYSIS MENU)选择 Show Available, 或拖动分析 窗口左上的小环打开方法帘。方法帘中包括了全部已用于分析你的 序列的方法。
- 你会发现 Deleage & Roux 方法并不在其中。从方法帘的顶端,点击
   More Methods 打开下面的子菜单,里面有所有可以使用的方法。从
   Secondary Structure 的子菜单中单击 Deleage & Roux,这样

57



该方法已经被用于分析,否则没有被用于序列分析,数字的多少代

表方法正在被使用的次数。由于我们还没 有把 Deleage & Roux 用于序列分析,所 以其左侧没有数字。



● 点击方法左边的空白区域去除对方法的 \_\_\_\_\_\_
 选择。然后单击 "Alpha, Regions,",并将它拖入分析窗口。预测的蛋白质阿尔法螺旋区域即展示在分析窗口中。

### 改变方法参数

下面我们改变 Deleage & Roux 方法的参数,然后看一看参数改变前后 这个方法预测的结果有什末不同。

- 重复上述操作将另一个 Deleage & Roux 方法应用于分析窗口,并将
   两个 Deleage & Roux 放在一起。可以看出现在有两个完全相同的预测结果。
- 双击其中一个会打开一个参数对话框,选择 A + B,单击同意。此时在分析窗口上,相应的区域由于计算参数的改变立即发生相应变化。

现在你可以比较两个参数计算结果的差别。(注:此时,如果你把最初的方法再用于分析窗口,则改变的参数自动应用于新加入的方法。)

C Computer calculated)

优化结果显示(与 GeneQuest 相似,这里省略)

使用蛋白酶消化与 SDS PAGE 电泳

Protean 可以识别蛋白序列上的蛋白酶酶切位点,并将酶切片段的电泳 结果以图形展示出来。

- 从 More Methods 菜单中将 Protease-Protease Map 方法移动到方法 帘。
- 单击挨着方法名字的蓝色三角形查看蛋白酶一览表。点击方法左边的空白区域去除对蛋白酶的选择。
- 选择"Chymotrypsin" 和"CNBr.",将它们

拖入分析窗口。 两种蛋白酶的识别位点显 示如右。现在至少一种蛋 白酶被应用,我们可以模 拟凝胶分离。



单击调色板工具中的范围选择器(Range Selector)(箭图标)去
 除对应用的蛋白酶方法的选择。

 从 SITES & FEATURES 菜单,选择 SDS PAGE Ge1 Simulation。酶切片段的分 离结果会自动显示出来(如右)。在 每个上面凝胶柱对应蛋白酶的名字和 分子量。



- 当光标在凝胶上面移动时,你可以随
   时查看到相应位置的蛋白质分子量。移动光标到任何片段上。窗口
   最上边的标题显示片段的范围,大小,分子量,等电点和 HPLC 滞留
   时间。
- 为了继续一节,关上 SDS PAGE Ge1 Simulation 窗口。

Features 注释(与 GeneQuest 中讲解的相似,这里从略)

### 进行 BLAST 检索

Protean 允许你通过因特网或企业内部互联网进行序列数据库检索。 在这一节中,我们将在 NCBI 的 BLAST 服务器上检索与人的钙调蛋白序 列相似的序列。注意电脑必须与因特网相连接,否则,跳过这一的部分, 继续下一部分。

- 单击调色板工具中的范围选择器(箭图标),在序列上选择你想进行检索的部分。由于我们的序列比较短,所以我们选择全序列进行比较。
- 从 EDIT MENU 选择 Select All。

从网络检索菜单,选择 BLAST 查找。BLAST 对话框出现(如左),

Blast Query:	Human Calmod	lulin(1,148)	
Database [	nr	▼	Help
			Cancel
			ОК
		-	

默认的数据库是 nr。

单击同意开始查找。

BLASTP of Human Calmodulin(1,148) vs. nr

ADOLTEEDI AEFKEARSU-DKOSOGTI TITKELGTA/IRSUGOIPTEAELODHI NEUDADGMI 60 ADOLTEEDI AEFKEARSU-DKOSOGOTI TITKELGTA/IRSUGOIPTEAELODHI NEUDADDMI 61 ADOLTEEDI AEFKEARSU-DKOSOGOTI TITKELGTA/IRSUGOIPTEAELODHI NEUDADDMI 61

Launch Editor 📔 Launch Browser

298, 2e-80,

ing sequences repo

 Contracting Schwarting Sequences Physics

 Dipacted
 Identification

 Species
 Identification

 Species
 pir/INCE0

 Sa=80
 pir/INCE0

Score = Expect =

BLAST 结果显示为一个二分窗

Description colmodulin 1 Chosphorylose kinose, i colmodulin - boving gl[71664[ptr]]HCL CALMODULIN gi[71655[ptr][MCEE colmodu colmodulin - robbit gl[230824[pdb]31] Home septeme gi[640287[pdb]1001[8 Hom

identity = (148/148) 100.0# Similar = (148/148) 100.0#

미묘

٠

•

口(如下,结构、内容与相关操作参见 GeneQuest 的 BLAST 结果)。

1 15-1000

Expected

Create Documer

295 296

Query: 1

Shick: 2

现在我们选择三个 BLAST 序列, 在 Protean 中打开他们。

● 在 BLAST 结果窗

中选择一个结果。

Save Next 💌 1 Sequences in Macintosh HD:Getting Started:	Help Defaults
	Cancel
Set Location	ОК

单击 Create Document, 保存 对话框窗打开(如左)。从保 存的下拉菜单选择 Next (默认

为 Selected)。

- 在 "Sequences." 的左边文本框输入数字 "3"。
- 单击同意,打开三个新的选定序列的 Protean 的分析文件。 现在关上 BLAST 结果窗。

#### 二级结构模拟

Protean 能够展示诸如螺旋轮,螺旋网和 beta 片层等基本元件的二级 结构。另外, 它能以线性的 space-fill 样模型或者化学公式样模型展

示蛋白质序列。在这一节中,我们做一个阿尔法区域的螺旋轮二级结构。 二级结构 Garnier-Robson 方法已经应用于我们序列,阿尔法螺旋也已 展示出来。选定其中的一个阿尔法区域后,我们就能以螺旋轮展示其二 级结构。

- 单击调色板工具中的范围选择器(箭图标)。
- 点击Protean文件最上边的Garnier-Robson分析结果中的一个阿尔



法区域。

从分析菜单,选择 Model Structures 中的 Helical Wheel, 就会出现左边的二级结 构预测图。

为继续这一节,关上螺旋轮窗口。

### 展示滴定曲线

下面,我们将查看全蛋白质序列的滴定曲线。

- ▶ 单击 Protean's 调色板工具右边的白色空间。这使得选择回到起始 itration of Human C 位置,如果没有选中的序列, ъ. 2.1 11.1 在绘制滴定曲线时, Protean 1.1 ... £ 7. 就会默认是对全序列进行计 5.8 e, 算。 1.1 2.5
- 从分析菜单,选择 Titration Curve.



● 打开序列的滴定曲线窗口(如上)。Protean 会在等电点处加一个 蓝色的 "crosshairs",以显示 pH 和所带电荷。

保存分析的文件(略)。

七、SeqMan II 开始

SeqMan II 不仅仅能够将成千上万个序列装配成 contigs,而且在 装配前,可以修整质量差的序列以及从你的序列中清除污染数据,还提 供完善的编辑和输出功能。SeqMan II 可以使用 DNASTAR 独特的跟踪 (trace)品质评价策略,通过对自动测序仪产生的跟踪数据进行评价, 自动生成最精确的一致序列。

装配过程从序列输入开始,输入的序列可以是自动测序仪输出的、 也可以是文本格式或 DNASTAR 序列格式。另外, 你可以从因特网的 Entrez 或 BLAST 服务器输入序列。使用 Preassembly options 中的命 令可以进行包括修整质量差的数据,除掉特定的载体或污染序列等操 作。你甚至可以在装配过程中选择最后加入重复序列。

Contigs 裝配完毕后, Strategy View 可以用图形化的方式展示 contig 的范围和质量,并提示你什么地方需要补充序列。在需要的情 况下,你可以加入更多的序列,或其他的 SeqMan II 文件对现有的结果 进行再装配。Alignment View 允许你编辑,修整连续的一致序列,找 回以前修整过的数据,调整队列 (alignments)结果。另外, 你可以 查看每一个或所有序列的质量评分。SeqMan II 为你提供自动或手动工 具,帮助你选择是否加入多个 contigs。和其他 Lasergene 应用程序一 样, SeqMan II 也提供整合的 BLAST 查找功能。

64

## 输入序列片段

我们将从打开"Demo Seqman"文件夹中的 14 样品序列开始。每个 序列都是自动测序仪 (PE Applied Biosystems, Inc.) 输出的文件格 式 (\*.trace)。

- 从文件菜单,选择 New 打开 Unassembled Sequences 窗口和一个空的未命名 Project Summary 窗口。Project Summary 窗口在你装配序列之前始终是空的。
- 从 Sequences 菜单选择 Add 打开 Enter Sequences 对话框(如右)。
- 打开"Demo Seqman." 文件夹会出现 14 个序列。这些就是我们将要 发配的序列。
- 把片段序列通过把 Add
   A11 按钮移到右边的窗口。现在14序列名字应该出现于右边的窗口中。单击 Done 按钮。



● Unassembled Sequences 窗口现在会出现这 14 个序列。

	Unasser	mbled Sequences		E
Add Add Add Add Add Add Add Add	1 Sequences	Set Ends	Options SetVecto	
File	Limits	Vector	Type	
Sample 1.abi	(1)741)		Trace	-
Sample 10.abi	(1>702)		Trace	
Sample 11.abi	(1)837)		Trace	
Sample 12.abi	(1>792)		Trace	•
4				▶ 4/ <sub>2</sub>

● 通过拖拽其窗口右下角展开窗口,直到你能同时看全部样品名字。

#### Pre-Assembly Options 操作

在装配前面,你可以考虑进行两三个装配前的选择。你可以从你的片段 中除掉载体和污染序列,优装配顺序,设定片段末端和标记重复序列。 在这一节中,我们将从样品中自动除掉 Janus 载体序列并修整序列末 端。

- 选定全部 14 个序列。
- 在窗口右上的下拉菜单中选择 "Janus"作为 SeqMan I 要寻找的 载体。
- 点击窗口上边的 Trim Ends 按钮打 开右边的对话框。你将注意到在没

	Trim Ends	
Quality  Stringency  Low  Nedium  High  Other	⊖ Fixed	
Scan Selections	Scan All	Concel Scon Loter

有特别指定的情况下, SeqMan II 将用适中的严谨度进行末端修整。 不改变设定退出对话框。

单击窗口右上的 Options 按钮。打开下面的对话框,显示全部装配前的选择一览表。默认设定显示扫描载体,实行序列末端修整,使用最适当的序列装配顺序。

Pre-Assembly Options	
🗹 Trim sequence ends	
🗹 Scan for vector	
Remove contaminant sequences	
Optimize sequence assembly order	-
Don't add single sequence contigs	
🔾 Use sequence entry order	
🔲 Don't add new contigs	Cancel
Scan Selections Scan All	Scan Later

 单击 Scan All。报告窗口 将出现,但是现在可以放到 一旁。在扫描后,你将注意 到 Unassembled Sequences 窗口的 2 中心专栏发生了 变化。现在载体栏显示:载 体名字前都有一个检测通过的标志,说明 Janus 载体在全部 14 序列 中都已经检测到了。1 imits 栏展示新的序列末端,说明载体序列已 经去除,末端修整已经完成。

### 查看修整的数据

现在我们可以查看在扫描完成的末端修整和载体序列去除的细节报告。

	Trim Rep	ort 🚃		P	
SCAN FOR VECTOR					
TRIM ENDS					
NECIUM TRIM PARAMETERS					Ξ
Troce: Threshold	= 12				
Non-Trace: Window Si	ze = 50				
Maximum N	s = 2				
	AUERAGE	AHOUHT	TRINHED	PRE-TRIH	
NAME 1	QUALITY	TRIMMED	LENOTH	LENOTH	
"Eample 1.cb1" (9>538)	32	Z11	530	741	
"Sample 10.abi" (1:667)	29	35	667	702	
"Sample 11.abi" (15>690>	28	161	676	837	
"Sample 12.ab1" (1)570)	30	ZZ2	570	792	
"Sample 13.abi" (14>694)	25	127	6B1	808	
"Sample 14.abi" (44>615)	26	251	572	823	
"Sample 2.mbi" (24)523)	Z8	360	500	850	
"Sample 3.abi" (1>519)	32	263	519	782	
"Sample 4.sbi" (1>405)	34	412	405	817	
"Sample 5.abi" (12540)	30	253	540	793	
"Sample 6.abi" (2>548)	28	285	547	832	
"Sample 7.sbi" (19)403)	26	342	465	8D7	
"Sample 8.abi" (1>507)	30	345	507	852	
"Sample 9.abi" (16)686)	25	99	671	770	
14 Sequences Avenage	29	240	561	8D1	1
≪ ≫ Unspecified Search ◀ 🕕				۰.	4

从 PROJECT MENU,选 Trim Report 打开左 边的窗口,窗口中显 示:末端修整的参数、 每个序列的名字,平 均质量,被修整的数 量以及修整前和修整

的序列长度。

● 在你查看完报告之后,关上它。

要查看 SeqMan 的修整结果,让我们查看休整后序列的跟踪数据。

从 Unassembled
 Sequences 窗口,双击
 打开8号样本,打开右
 面的 chromatogram 窗口。



从 5'末端起变淡的部分将不会用于序列装配。注意大部分序列修整部

分优化后的质量。之所以从这里去除,是因为这些变淡的序列是 Janus 载体序列。

垂直的黑的棒出现于修整和未修整的序列之间。你可以根据需要拖动黑 棒调整用于装配的序列末端。

下面,让我们查看一下序列的 5' 末端:

- 单击一致序列的任何地方。
- 从编辑菜单,选择Go To Position。
- 在位置文本框输入"500"。
- 単击同意。

注意到最好的质量落到这区域上, SeqMan II 认为 507 位以外的序列质 量低于中等水平。

● 检查完跟踪数据后,关上那个窗口。

### 序列装配

现在我们准备调用序列装配程序。

● 单击 Unassembled Sequences 窗口左上的装配按钮。

	Untitle	d	
Name	Length	Seqs	
∷Contig 1	. 1431	14	4
			w
Sample 2.a	bi (53>523)		•
Sample 3.a	bi (56>519)		
Sample 4.a	bi(32>405)		=
Sample 5.a	bi(31>540)		
Sample 6.a	bi(54>548)		

Unassembled Sequences 窗口立即 消失,报告窗口显示装配的进度。 装配完整之后,出现左边的窗口, 窗口上边为装配好的一致序列,窗 口下边为所有连续的序列。在我们

例子中, SeqMan II 将 14 个序列装配成为单一的 contig。

## 查看覆盖面和构建结果

现在我们已经将所有的片段装配成一个序列,下面我们用 Strategy View 查看一下其中的矛盾碱基的数量和序列的覆盖程度。这个窗口图 形化的显示了每个片段在 contig 中位置和方向。

- 単击 Project window 中的 contig。
- 从 CONTIG 菜 Strategy of Contig 1 08 • Θ 1250 1500 250 500 750 1000 选择 单, Domentage Conflicts Sample 11.ab1(63)690; Sample 9.abi(58>685) Sample 1.abi(43>538) Strategy Sample 6.obi(54>548> Sample 13.abi(55>694) Sample 14.abi(6B>615) Sample 10.abi(33>667) View 打开右 Sample 7.abi(52>483) Sample 5.abi(31>540) Sample 8.ab1(45>507) Sample 2.abi(53>523) Sample 3.abi(56>519> \_\_\_\_ 边的窗口。注 Sample 4.abi(32>405> Sample 12.abi(48>570> €- $\neq$ ---意标尺下面

的那条带,显示学的覆盖程度。

条带的颜色和厚度代表覆盖序列的数量。你会看到在大致 200-900 之间 有绿色的较厚条带。这表示两条链都被测序,且高于最小覆盖面临界值 的区域。(覆盖面临界值可以通过 PROJECT MENU-Parameters-Strategy Viewing 来改变)。中间的蓝线表示只有一条链被测序的区域。。出现 于 contig 两边的细红线表示这个区域只测过一次序。

- 选定窗口左上的 Conflicts box 可以查看片段序列间的一致程度方型图。方型图大部分为黑色表示只有很少或没有矛盾碱基。蓝色条带表示矛盾碱基约为 10-20%的区域。
- 点击窗口左侧调色板工具中的 Zoom In (放大镜图标"+")。

- 双击135 位碱基调入该区域的 Alignment View。样品9中的矛盾碱
   基用红色表示出来。
- 关上 Alignment 窗口。

### 去除矛盾碱基和缺口

我们样品序列能够整齐地组装成单一的 contig 是因为 SeqMan II 除掉 了质量差的和载体数据。在一些情况下,为了获得满意的一致序列,必 须去除矛盾碱基和缺口。我们将尝试在 Alignment View 中编辑我们的 序列。

- 从 CONTIG 菜单,选 Alignment View,打开 Alignment View 窗口,
   上边为一致序列,下边为连续的序列。
- 从编辑菜单,选Find Conflict,然后点击Find Next。光标将迅速 跳到 102 位。单击底部的Find Next 按钮(>>),光标跳到 306 位。 样品 7 在这这个位置右一个"T",而全部其他的样品都是缺口。我 们可以通过查看样品 7 的跟踪数据去除这个矛盾碱基。
- 选定样品 7 的 302~310
   区域。
- 单击样品7名字左边的灰
   色的三角形,则显示出其
   跟踪数据,出现右面的
   Alignment View窗口。



● 点击 Zoom In 查看 chromatogram 细节。

跟踪数据表明样品 7 里的"T"与其左边的"A"一样可能是人工造成。 从一致序列可以看出 SeqMan II's 的 trace 品质评价标准自动忽略了 这些可疑的碱基。因此没有必要手工删除这些碱基。不过从学习的目的 出发,我们将用缺口替换"T",并从一致序列中除掉缺口栏。

- 选定"T",输入缺口(连字符)替换它。
- 单击一致序列的任何地方,接着从编辑菜单,选 Remove Consensus
   Gaps, 306 位的所有缺口都不见了。或者你也可以选定一致序列里的缺口,按退格键(视窗口)或删除键(苹果计算机)即可。

手动修改序列末端

在装配我们的 dcontig 之前, SeqManII 可以根据 trace 数据的质量和 Janus 载体序列自动的进行末端修整。虽然有足够的数据可以用来把序 列装配成单一的 contig, 但是也会存在这样的情况, 即当我们找回修 整过的数据时, SeqMan II 可以在单一的 contig 中加入多个 contigs。 在这一节中,我们将学习从已经打开着的 Alignment View 用手工找回 修整过的末端的方法。

为了区分修整过和没有修整过的数据,我们给修整过的数据加一个有颜

reasts annuming End Trimming Contaminant Screening Assembling Consensus Calling Stratogy Vlowing Editing & Color Blast Searching	All Editing Allowed     Only Gap Editing Allowed     No Editing Allowed     Match Color     Mismatch Color     Use Weight Color     Negated Weight Color     Trimmed Color
---	--

色的背景。

从 PROJECT MENU,选
 参数中的 Editing &
 Color 打开左边的对
 话框。

 确认 Use Weight Color 已经被选用,然后单击同意。你用手工找回 修整去掉的数据时,你会发现这些找回的序列有明亮的黄色的背景 出现。

如果修整完毕,则 Alignment View 中在序列的左边会有一个黑色的垂 直棒,右边有一个小的黑三角形。这一节,你可以选 14 个序列的某一 个。

 要找回修整去掉的序列末端,只需把垂直棒向序列的两端拖动即可, 可以看见以前修整去掉的序列有明亮的黄色背景。当你找回了修整 去掉的载体序列时,你会发现即使修整去掉的序列的峰型仍然很好, 但是其中也已经有很多碱基(红色显示)与您的一致序列不一样了。

在你的文件中不应 该留有这样的数 据。也许对组装 contigs 有用的数 据是那些由于序列



的峰型质量较差而被修整去掉的那些序列。

● 拖动竖线将序列恢复到原来修复的位置。

#### 文件保存与序列输出

保存 SeqMan II 文件:

- 从文件菜单,选保存。
- 选择文件保存的位置,命名。
- 单击保存,则以上所有操作的结果都保存进文件中了。
  只保存一致序列:
- 选定 Project Summary window, the Alignment View 或 the Strategy
  View 中一个或多个 contigs。
- 从 CONTIG 菜单,选 Save
  Consensus 中的 Single File
  打开右边的对话框。
- 选择文件保存的位置,命名, 确定保存类型(你可以选择 DNASTAR,GenBank或FastA格 式中的任何一个)。



- 选定 All and Add Features,不选 Include Gaps。这样就将完整的一致序列,以及 contig 构建的相关信息全都保存了,同时去除了一致序列中的所有缺口。
- 单击保存,保存序列。

输出编辑后的片段序列:

- 在 Project window 中选定你要输出的序列。
- 从 CONTIG 菜单,选输出序列。
- 这个命令与保存一致序列的命令一样,允许你以三种格式保存你的 序列。