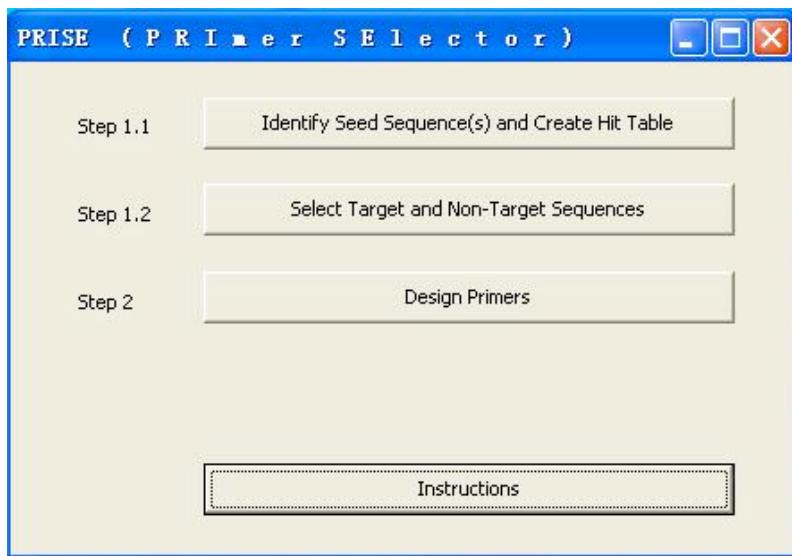


## PRISE Manual

**Computer and file requirements.** PRISE can be run on Windows 2000/NT/XP with a minimum of 512 MB of RAM (1 GB of RAM or more is recommended). Internet connectivity is also required. Note that files created in a non-Windows based environment may not function properly in this program.

**Overview of the design process.** Designing PCR primers using PRISE involves two general steps. Step 1, which is divided into two components (1.1 and 1.2), enables target and non-target DNA sequences to be identified and collected. Step 2 generates PCR primers designed to amplify target but not non-target sequences. A detailed step-by-step protocol (PRISE Tutorial), which demonstrates how the software was used to create sequence-selective PCR primers for a specific fungal rRNA gene, can be accessed via the Instructions or Help links.

**Running the program.** When the program is opened, a window with four selections appears. These link to instructions and modules for performing the various steps in the primer design process.



**Figure 1.** Opening window with links to the instructions and modules of the PRISE software.

### **Step 1.1: Identify Seed Sequence(s) and Create Hit Table**

**Overview.** The first step in the design process is to identify the Seed Sequence(s) and to create the Hit Table. The Seed Sequence(s) represent the DNA sequence(s) that the primers are designed to amplify. The Hit Table is a list of DNA sequences with various degrees of similarity to the Seed Sequence(s), from which the target and non-target sequences can be derived. The Hit Table is created by subjecting the Seed Sequence(s) to a BLAST (blastn) analysis.

*Note.* Although Steps 1.1 and 1.2 are designed to identify and collect target and non-target DNA sequences, there are certainly other strategies for accomplishing this task, which the user may decide to use instead of or in combination with our steps. The only requirement for using the primer design module of PRISE (Step 2) is that the target and non-target sequences be available in separate FASTA-formatted text files.

*Identify the Seed Sequence(s):* Identify the sequence(s) that the primers are intended to amplify and save them in FASTA format as a text file.

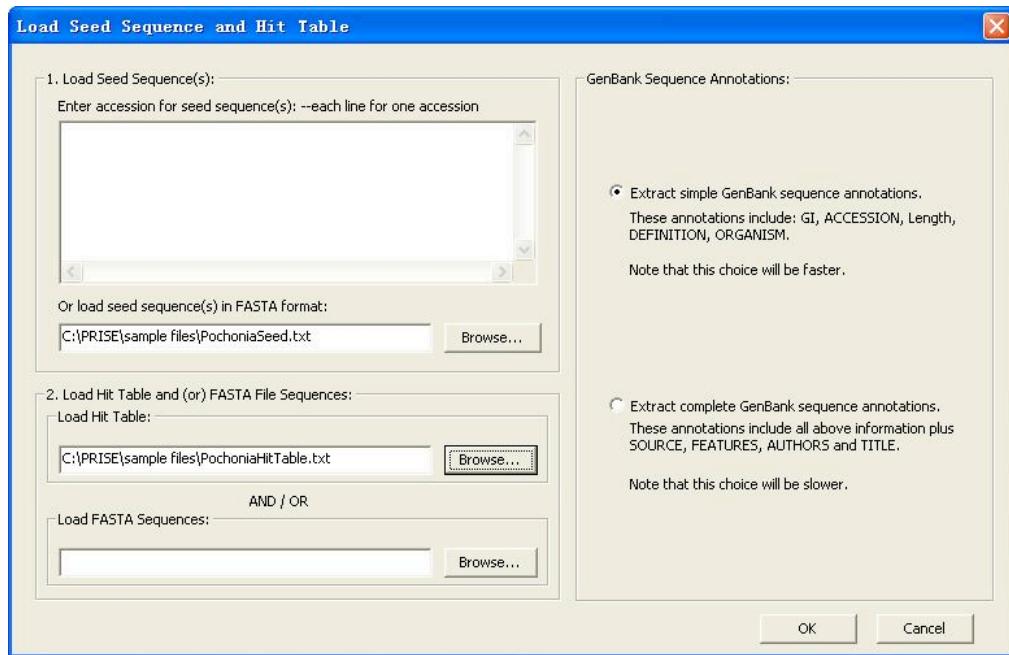
*Create the Hit Table:* Subject the Seed Sequence(s) to a nucleotide BLAST analysis using the program on the NCBI website (<http://www.ncbi.nlm.nih.gov/BLAST/>). In this analysis, the user needs to select the appropriate **Database** and number of **Max target sequences**, which, in our experience, will typically be a minimum of 500. The **Max target sequences** option is located in the **Algorithm parameters** section. In the current version of the program, after clicking on the **BLAST** button, click on **Formatting options**. Here, set **Show Alignment as Plain text** and **Alignment View as Hit Table**. In addition, set **Alignments** in the **Limit results** section to the value that was used for the **Max target sequences**. Click **View report** and save the output as a text file. This file is the Hit Table. Note that some web browsers do not allow output to be saved as text files.

## **Step 1.2: Select Target and Non-Target Sequences**

Once the Seed Sequence(s) and Hit Table have been created, the next step is to identify and collect the target and non-target sequences. This is accomplished with the **Select Target and Non-Target Sequences** module.

### **Using the module**

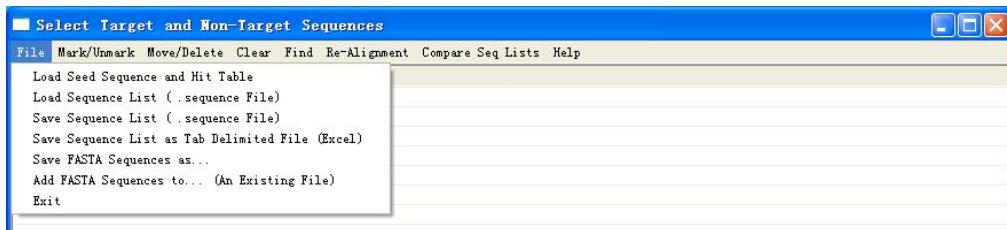
*Loading the Sequences.* After opening the module, the user inputs the Seed Sequence(s) and Hit Table files into the software. This is accomplished by selecting the **Load Seed Sequence and Hit Table** option from the **File** menu. This option opens a window titled Load Seed Sequence and Hit Table, where the appropriate files can be input. Note that this window also allows FASTA files to be input instead of or along with the Hit Table, allowing sequences other than those generated by a BLAST analysis to be utilized. In the next window, titled Sequence Alignment Settings for Pairwise Identity Analysis, the user can select settings for the pairwise identity analyses, which will be performed between the Seed Sequence(s) and the Hit Table sequences (and/or FASTA sequences if there are any).



**Collecting and Parsing the Sequence Data.** After the sequences are uploaded, the software downloads all of the GenBank records associated with the Seed Sequence(s) and Hit Table sequences, parses the data contained within them into separate components, performs pairwise % identity analyses between the Seed Sequence(s) and Hit Table sequences, and displays these data in tabular form in the Select Target and Non-Target Sequences window. (Note that degenerate nucleotides (ie. R, Y, etc) are considered mismatches in this pairwise analysis.) The title of the window will be the Hit Table file name followed by “- Select Target and Non-Target Sequences.” After the program finishes processing the data, which could take minutes to hours, depending on the number of sequences in the Hit Table, the speed of the internet connection and the capabilities of the computer, a sequence downloading report appears. This Report lists the accession number of sequence(s) from the Hit Table that are too large to be analyzed. The information in the Report can be saved as a text file for later.

**Sequence Selection.** Once these actions have been completed, the user can identify and collect the target and non-target sequences by applying sorting tools to the sequences assembled in the table. This is primarily done by using tools that allow the sequences to be selected by parameters such as sequence length, % sequence identity or GenBank parameters such as Definition or Source. Sequences can also be sorted by clicking on the column headings. Below is a description of all of the functions in this module, organized by the pull down menu they reside in. See the PRISE Tutorial for a few examples of how they can be used.

## File menu:



**Load Seed Sequence and Hit Table:** Allows Seed Sequence(s) and Hit Tables to be loaded. This window also allows the user to load a FASTA file instead of or along with a Hit Table, allowing sequences from sources other than a BLAST analysis to be utilized.

**Load Sequence List:** Allows previously created sequence lists to be loaded into the software.

**Save Sequence List:** Allows sequence lists to be saved in the format used by the PRISE software.

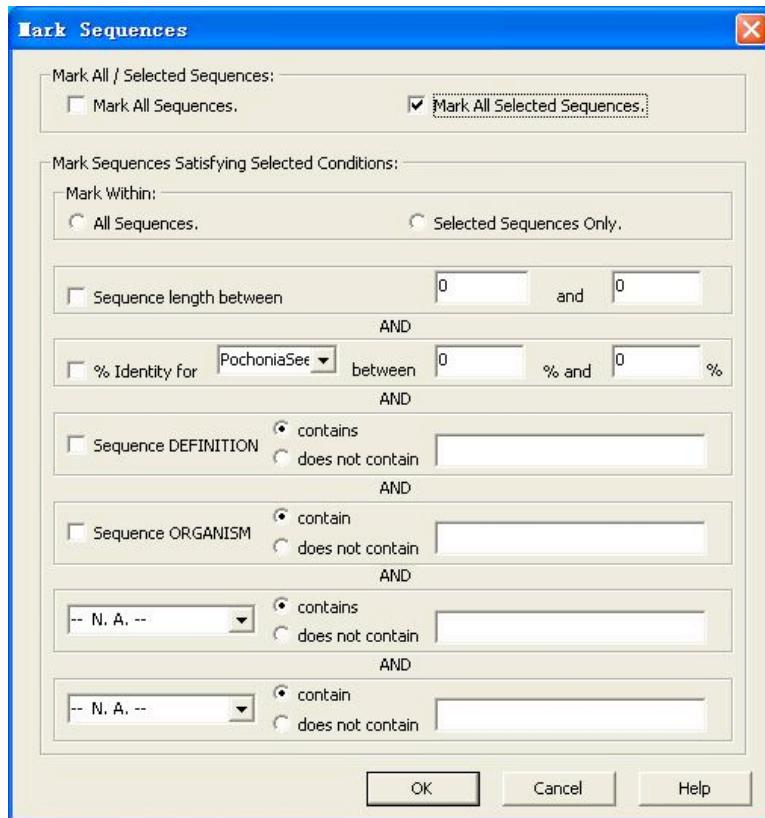
**Save Sequence List as Tab Delimited File:** Allows sequence lists to be saved in a tab-delimited format, which can be used in standard spread sheet software.

**Save FASTA Sequences As:** Saves the sequences in the FASTA Sequence Box in FASTA format as a text file.

**Add FASTA Sequences To:** Adds the sequences in the FASTA Sequence Box to another text file (typically one that contains other sequences in FASTA format).

## Mark/Unmark menu:

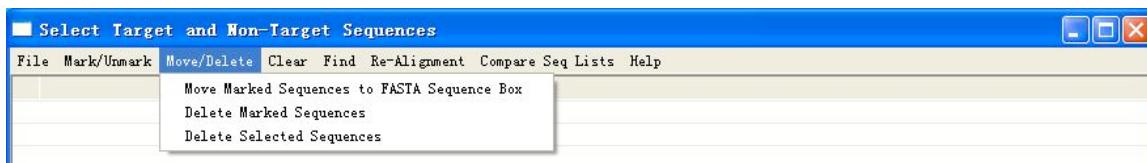
**Mark Sequences:** Allows sequences to be marked if they possess user-defined criteria. Marked sequences are designated by a check mark in the box in the "Seq #" column (and a yellow-highlighted row). Sequences that are marked can be moved to the FASTA Sequence Box, and then saved or merged with other FASTA files to make target and non-target sequence files.



**Unmark Sequences:** Allows sequences to be unmarked if they possess user-defined criteria.

**Reverse Marked and Unmarked Sequences:** Reverses the marked and unmarked designations.

#### Move/Delete menu:



**Move Marked Sequences to FASTA Sequence Box:** Moves marked sequences to the FASTA Sequence Box. Marked sequences are designated by a check mark in the box in the "Seq #" column. Sequences that are marked can be moved to the FASTA Sequence Box, and then saved or merged with other FASTA files to make target and non-target sequence files.

**Delete Marked Sequences:** Deletes marked sequences from the sequence list. Marked sequences are designated by a check mark in the box in the "Seq #" column (and a yellow-highlighted row).

**Delete Selected Sequences:** Deletes selected sequences from the sequence list. Selected sequences are designated by their rows being highlighted in blue. Sequences can be selected by clicking on any part of the row except the boxes in the column "Seq #." Standard key commands such as shift and control can be used with this function, allowing groups of sequences to be selected. Once sequences are selected, they can be marked or unmarked using the functions in the Mark/Unmark menu.

### Clear menu:

**Clear FASTA Sequence Box:** Deletes the sequences from the FASTA Sequence Box.

### Find menu:

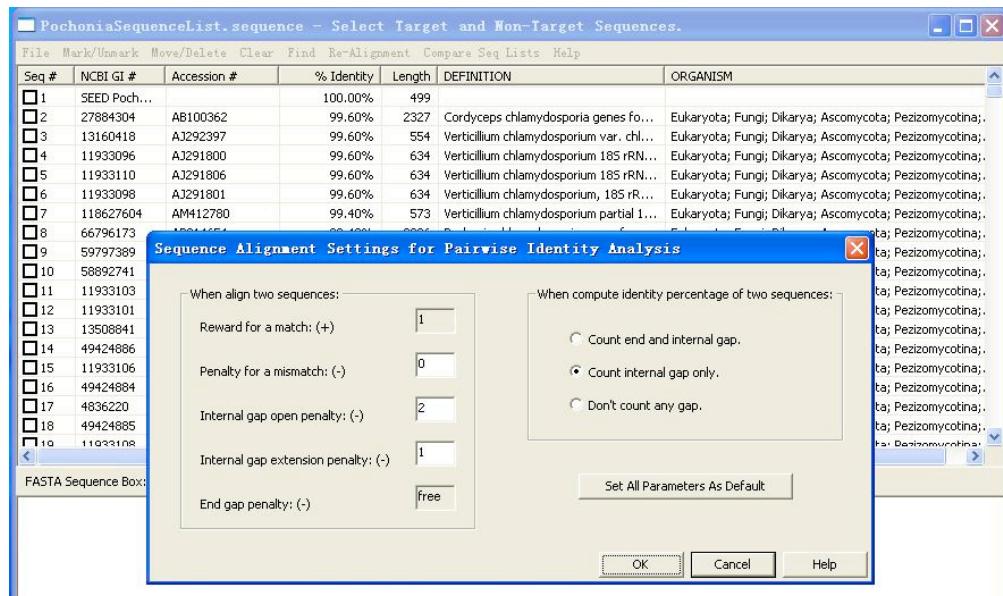
**Find Sequence:** Allows the user to search for sequences by user-defined criteria.

**Find Next:** Allows the user to search for sequences using the criteria that were input in the last *Find Sequence* search.

### Re-Alignment menu:

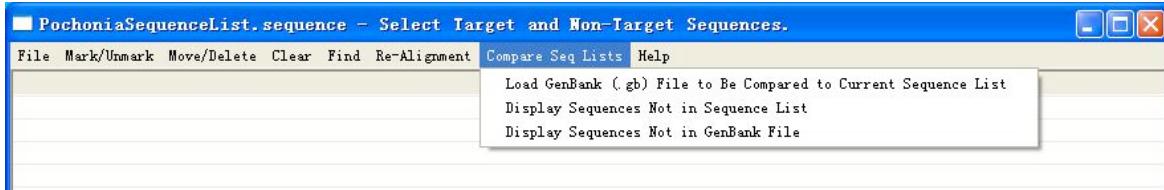


**Change Sequence Alignment Settings:** Allows the user to change the settings used for the pairwise identity analyses. The resulting changes in the alignment values for individual sequences can be viewed by using the **Display Pairwise Alignment** option, which is accessed via a right click. Note that these settings will not be saved unless the **Update % Identity for All Sequences** option is used (see immediately below).



**Update % Identity for All Sequences:** Allows the user to change the settings used for the pairwise identity analyses and then perform a new pairwise analysis on all sequences in the list. Note that any changes made with this option will be automatically saved in the Sequence List file.

### Compare Seq Lists menu:



These functions allow the user to compare sequences in the Sequence List, which is currently loaded in the PRISE software, to sequences in a GenBank file. Note that these sequences will be compared by their GenBank Accession number, not their nucleotide sequences.

**Load GenBank (.gb) File to be Compared to Current Sequence List:** Allows the GenBank file to be loaded into the software.

**Display Sequences Not in Sequence List:** Displays the sequences that are in the GenBank file but not in the Sequence List.

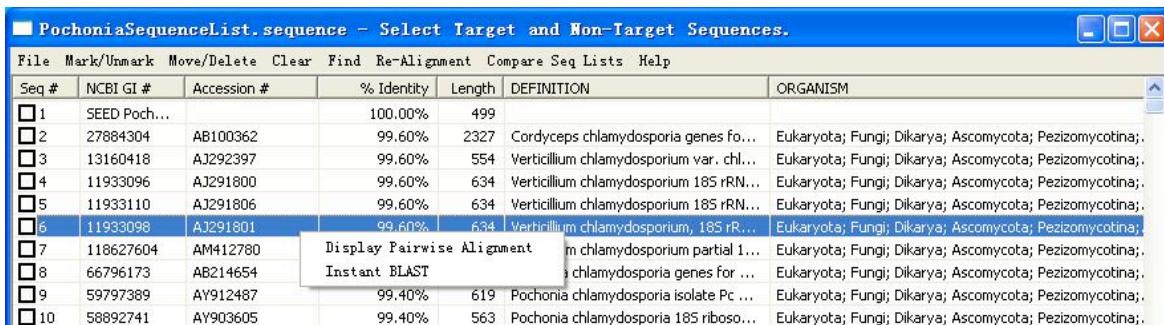
**Display Sequences Not in GenBank File:** Displays the sequences that are in the Sequence List but not in the GenBank file.

### Help menu:

**PRISE Manual:** Opens the PRISE Manual.

**PRISE Tutorial:** Opens the PRISE Tutorial, which provides a step-by-step protocol showing how the software was used to create sequence-selective PCR primers for a specific fungal rRNA gene.

### Right-click options:



**Display Pairwise Alignment:** Opens a window showing the alignment of the selected sequence and the Seed Sequence. Note that this function only works when one sequence is selected.

Pairwise Alignment		
Seed sequence PochoniaSeed V.S. subject sequence AJ291801; 99.60% identity. Identity: 497; Mismatch: 2; Internal gaps: 0; End gaps: 135.		
Sed 1	-----CATGTCAGCTTATACCAATTACA-----	23
Sbj 1	GCGAACTTAAAGCTCTAACAAAGGTCTCCGTTGGTGAACCCACGGGAGGGATCATTACCGACTTTCAACTCCAAACCCATGTCAGCTTATACCAATTACA-----	100
Sed 24	ACCCCTGCTTCGGCGGGGTTCTCCCGGCGGCTTACACCCCGGAAACCAGGGGGCCGCCGGGGACCCAAACTCTAGATTTTATTTGGCATGTCAG-----	123
Sbj 101	ACCCCTGCTTCGGCGGGGTTCTCCCGGCGGCTTACACCCCGGAAACCAGGGGGCCGCCGGGGACCCAAACTCTAGATTTTATTTGGCATGTCAG-----	200
Sed 124	TGGGAACTTACAAAATGAACTAAAACCTTCAACACGGATCTTGTGCTTCTGCATCGATGAAGAACCCAGCGAAATGCCATAACTATGTCATTGCA-----	223
Sbj 201	TGGGAACTTACAAAATGAACTAAAACCTTCAACACGGATCTTGTGCTTCTGCATCGATGAAGAACCCAGCGAAATGCCATAACTATGTCATTGCA-----	300
Sed 224	GAACTCAGTGATCGAACTTTGAACTCCACATTCGGCCCGGCACTATTCTGGCGGGATGCCCTGTCAGGGCTCATTTCAACCCCTCAAGGCCAGCG-----	323
Sbj 301	GAATTCACTGAAATCATCGAACTTTGAACTCCACATTGGCCCGGCACTATTCTGGCGGGATGCCCTGTCAGGGCTCATTTCAACCCCTCAAGGCCAGCG-----	400
Sed 324	GTTTGGTGTGGGGACCCGGCACTACAGCGCTTGGGACTCTCCCCCTCCCTCGGCCGCCGGGGAAATGAATTGGCGGTCTGTCGGCCCTCC-----	423
Sbj 401	GTTTGGTGTGGGGACCCGGCACTACAGCGCTTGGGACTCTCCCCCTCCCTCGGCCGCCGGGGAAATGAATTGGCGGTCTGTCGGCCCTCC-----	500
Sed 424	TCTCCGTACTAGCACACCTCCCATCAGGAGCCCCGGGGCACTGCCCTAAAAACGCCAACTTTTTTAACAG-----	499
Sbj 501	TCTCCGTACTAGCACACCTCCCATCAGGAGCCCCGGGGCACTGCCCTAAAAACGCCAACTTTTTTAACAGCTTGACCTCGAATCAGTAGGAAATA-----	600
Sed 500	-----500-----	
Sbj 601	CCCGCTGAACTTAAGCATATAAGCGGAGGA 634	

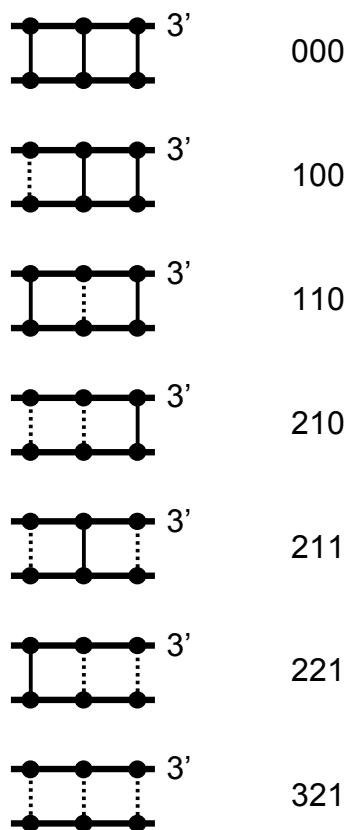
**Instant Blast:** Allows the sequence to be subjected to a BLAST analysis, by opening the Blast page at NCBI and loading the sequence. Note that this function only works when one sequence is selected.

## Step 2: Design Primers

The PRISE software allows selection of both standard PCR primer parameters, such as GC content, primer length, inter- and intra-complementarity, as well as criteria for sequence-selectivity. Selectivity is accomplished by identifying primers that should amplify target sequences but not non-target sequences. The prediction as to whether a PCR product will be made is based on a number of criteria that can be customized by the user to suit the application at hand.

One of the criteria used in this process is a scoring scheme that is used to define the likelihood that specific primer-template combinations will produce a PCR product. This scheme allows the user to set the design criteria for each position in the primer. Here, we describe only a simple version of this scheme that focuses on last three 3' positions. (For more detailed information on Primer Selectivity Settings, please refer to Appendix I.) Figure 2 shows various match-mismatch configurations and corresponding parameter settings. If the setting is set to xyz, then only primer-template pairs that satisfy this xyz match-mismatch configuration and those above will be considered as producing a PCR product. For example, for 000 setting, only exact matches at all three positions will be scored as creating a PCR product. If the setting is 210, then any primer-template pair with match-mismatch configurations of 000, 100, 110, and 210 will be counted as

producing a PCR product. (One match-mismatch setting does not appear in the figure for technical reason – see the Appendix I for details.)

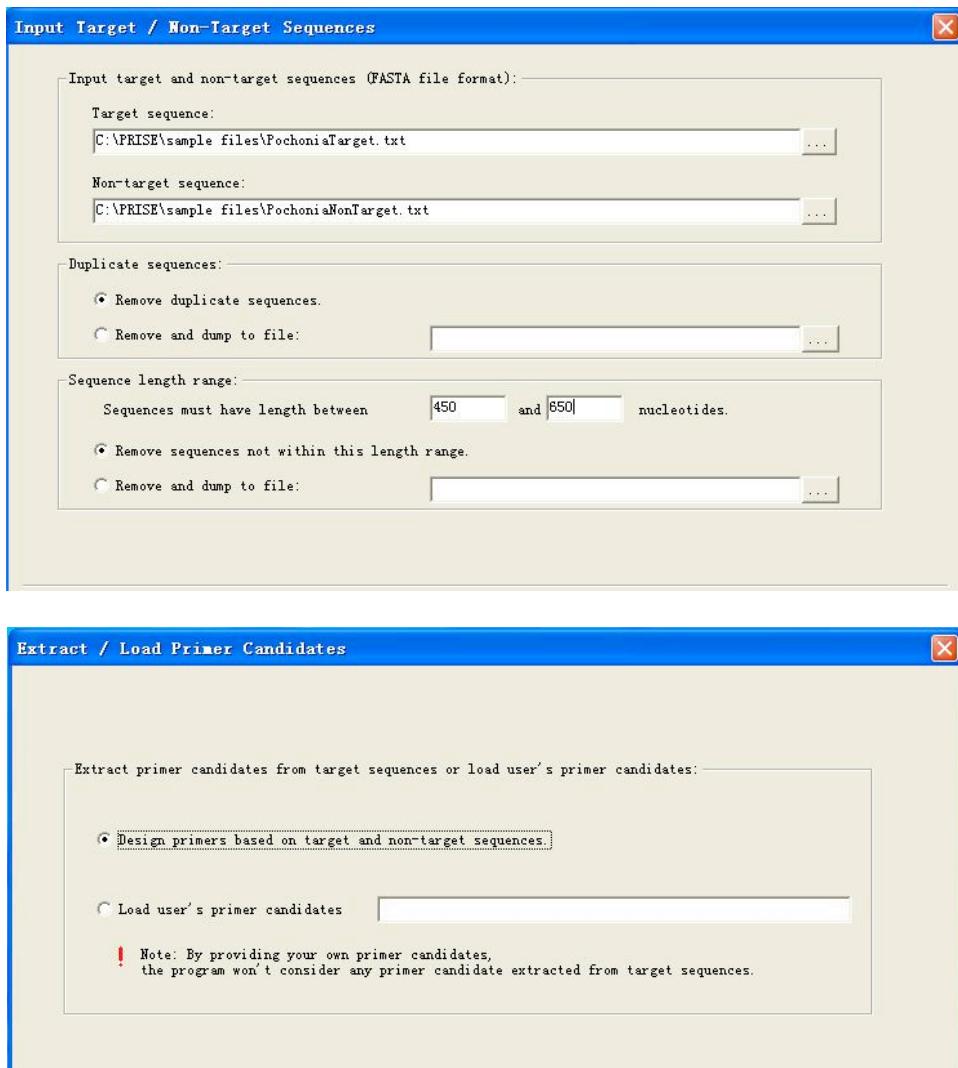


**Figure 2.** Scoring scheme for the sequence-selectivity component of the Design Primers module. On the left side are depictions of the last three 3' nucleotides of a primer and its corresponding template. The primer is the top strand. Base-paired nucleotides are designated by solid lines. Non-based paired nucleotides are designated by dashed lines. The score assigned to each type of template-primer pair is shown to the right.

This scoring scheme can be set separately for target and non-target sequences. This useful feature gives a user the flexibility to define different stringency requirements for primer annealing within these two classes of sequences.

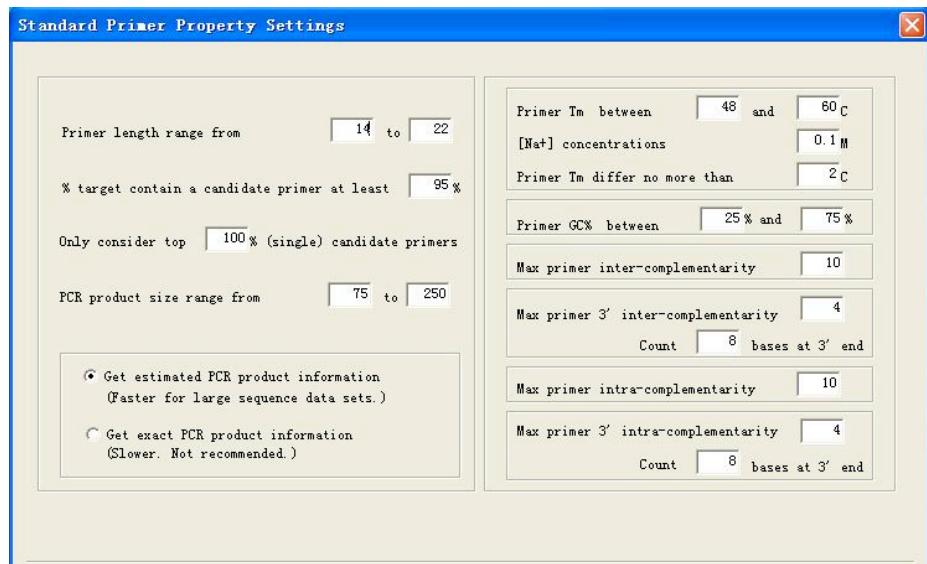
## Using the module

**Loading the Sequences.** After opening the **Design Primers** module, the Primer Design Page appears. In this window, the user can either load a previously created primer list pair file or initiate a new primer design project. In the following Input Target / Non-Target Sequences window, the user can load the target and non-target sequence files and select options to remove duplicate sequences and those that do not meet user-selected size criteria. In the following Extract / Load Primer Candidates window, the user can choose to either (i) **Design primers based on the target and non-target sequences** (and user defined primer criteria) or (ii) **Load user primer candidates** to assess their properties in relation to the target and non-target sequences and user-defined primer criteria.



**Standard Primer Property Settings.** In the next window, titled as Primer Design Settings, the user can select (i) **Use all default settings**, (ii) **Use previously saved settings**, or (iii) **Show me default settings**, the latter of which allows the user to review the default primer settings and to have the option of changing them. This opens the Standard

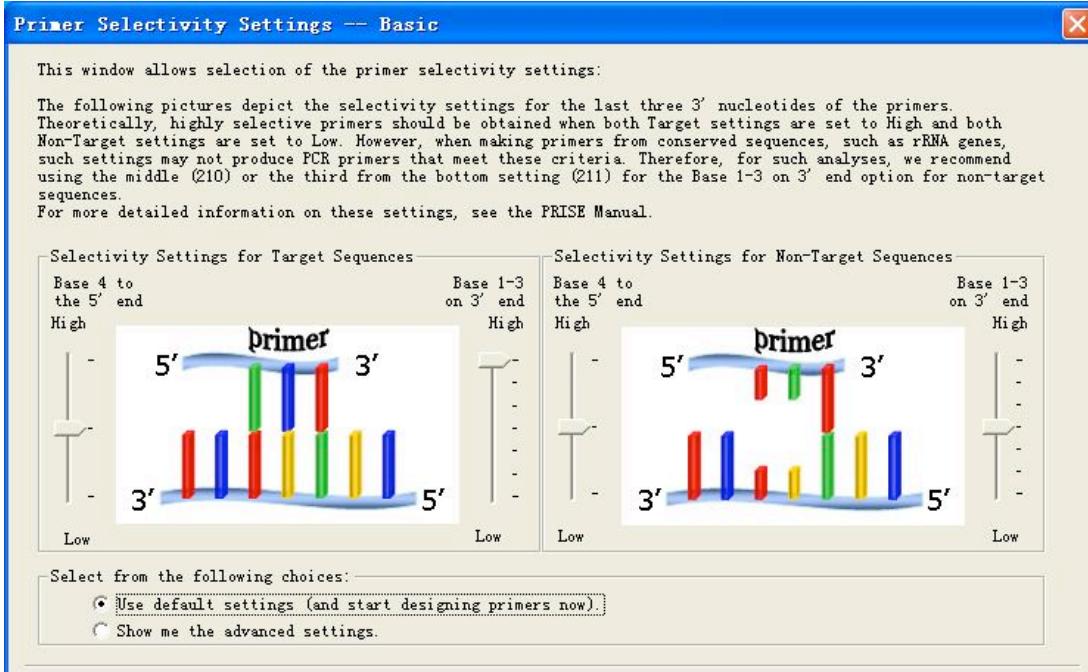
Primer Properties Settings window, which allows selection of standard primer properties such as GC content, primer length, and PCR product size.



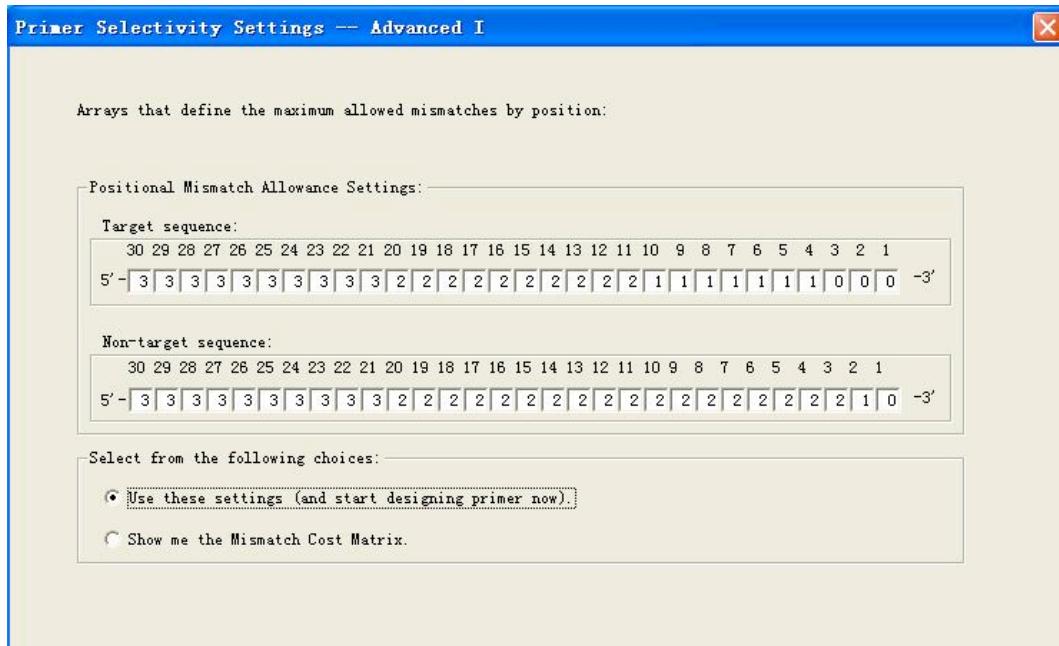
*Primer Selectivity Settings.* The primer selectivity settings are located in the next set of three windows. These three successive windows increase in both user complexity and control.

In the Primer Selectivity Settings – Basic window, the user can select to either use the default options or adjust the scoring scheme (described above and in Appendix I) for both target and non-target sequences. This window allows the user to set the selectivity settings for two separate regions of the primers: the last three 3' nucleotides and the other nucleotides.

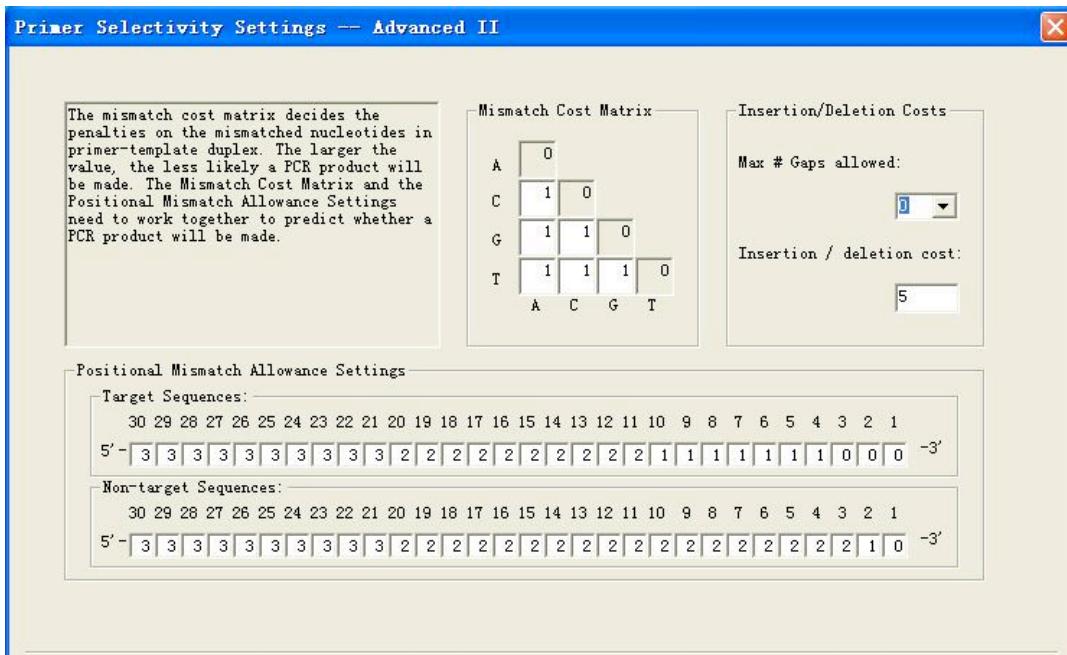
Theoretically, highly selective primers should be obtained when both Target settings are set to High and both Non-Target settings are set to Low. However, when making primers from conserved sequences, such as rRNA genes, such settings may not produce PCR primers that meet these criteria. Therefore, for such analyses, we recommend using the middle (210) or the third from the bottom setting (211) for the “Base 1-3 on 3’ end” option for non-target sequences.



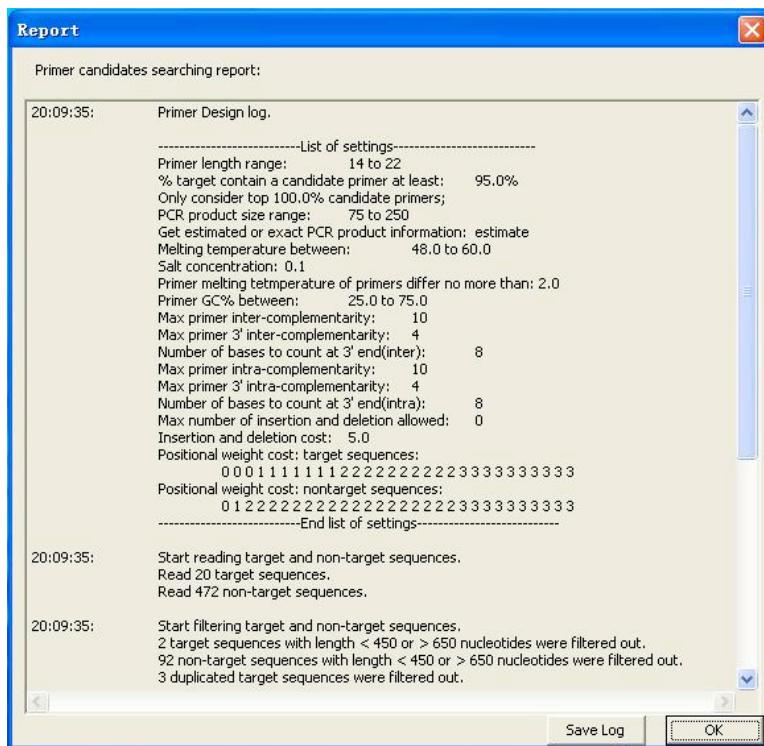
In the Primer Selectivity Settings – Advanced I window, the user can set the selectivity settings for every position in the primer. See above or Appendix I for more information on the selectivity settings.



In the Primer Selectivity Settings – Advanced II window, the user can adjust the Mismatch Cost Matrix and the Insertion/Deletion Costs. See above or Appendix I for more information on the selectivity settings.



After the program finishes processing the data, which could take minutes to hours, depending on the size and complexity of the sequences in the target and non-target files, the user selected primer criteria and the capabilities of the computer, a primer design report appears. Information from this Report can be saved as a text file. If optimal primers are not generated, the data in this file may help the user trouble shoot the design process.



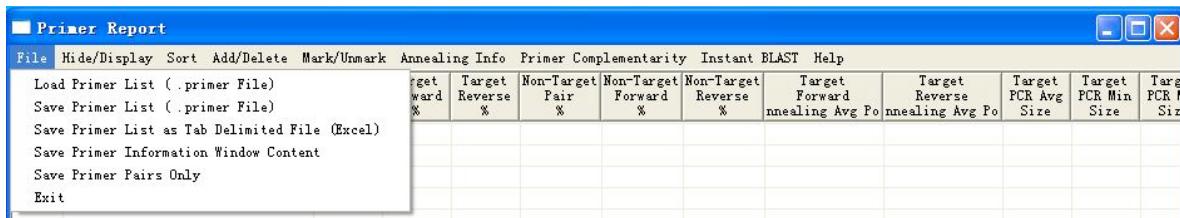
**Primer Report.** In the next window, titled Display Primer List, the user has the option of displaying all of the primers or a subset of these primers, which can be selected by their selectivity. The next window contains the primer pairs. The title of this window will be the Target sequence file name followed by “- Primer Report.” The Primer Report is a table that displays the primer pairs and their properties, including the % of target and non-target sequences predicted to be amplified, PCR product size, etc. To assist the process of selecting optimal primers, the primers in the table can be sorted by each of these parameters and by a formula that identifies primers that are most likely to amplify target but not non-target sequences (the “Selectivity Formula”). In addition, primers can also be sorted by clicking on the column headings. This software module also provides tools enabling the user to obtain detailed information concerning the selectivity of the primer pairs. These data include the percent of each nucleotide, at each position, in the target and non-target sequences in relation to the nucleotides in each position of the primers. In addition, the user can identify the target and non-target sequences that should or should not be amplified by each primer pair. This module also allows the user to load additional primer pairs, not necessarily created by the software, enabling the properties of these primers to be examined in relation to the target and non-target sequences. Finally, the primers and their properties can be saved in a tab-delimited format, allowing the user to import the data into other programs such as spreadsheet software.

PochoniaTarget.txt - Primer Report													
Primer Pair No.	Primer Pair Nucleotides (forward 5'-3' .. reverse 3'-5')	Target Pair %	Target Forward %	Target Reverse %	Non-Target Pair %	Non-Target Forward %	Non-Target Reverse %	Target Forward Annealing Avg Pos	Target Reverse Annealing Avg Pos	Target PCR Avg Size	Target PCR M Size	Targs	
□ 1	TTATACCATTTACAACCGTTG..TCATT... □ 2	100.0%	100.0%	100.0%	~0%	0.0%	0.6%	53.7	165.7	133.0	1		
□ 3	ATACCATTACAACCGTGT..TCATTGT... □ 4	100.0%	100.0%	100.0%	~0%	0.0%	0.6%	55.7	168.7	131.0	1		
□ 5	TATACCATTTACAACCGTTG..TCATT... □ 6	100.0%	100.0%	100.0%	~0%	0.0%	0.6%	54.7	165.7	132.0	1		
□ 7	TTATACCATTTACAACCGTGT..TCATTGT... □ 8	100.0%	100.0%	100.0%	~0%	0.0%	0.6%	53.7	165.7	133.0	1		
□ 9	TACCAATTACAACCGTTGCT..TCATT... □ 10	100.0%	100.0%	100.0%	~0%	0.0%	0.6%	57.7	165.7	132.0	1		
□ 11	TATACCATTTACAACCGTTG..TTCATT... □ 12	100.0%	100.0%	100.0%	~0%	0.0%	0.6%	54.7	165.7	133.0	1		
□ 13	TTATACCATTTACAACCGTTG..TCAT... □ 14	100.0%	100.0%	100.0%	~0%	0.0%	0.6%	53.7	164.7	133.0	1		
□ 15	TTTACCAACCGTTGCT..TCATTTGT... □ 16	100.0%	100.0%	100.0%	~0%	0.0%	0.6%	60.7	168.7	126.0	1		
□ 17	TTTACCAACCGTTGCT..TCATTTGT... □ 18	100.0%	100.0%	100.0%	~0%	0.0%	0.6%	60.7	168.7	127.0	1		
	ACCATTACAACCGTTG..TCATTGT... □ 19	100.0%	100.0%	100.0%	~0%	0.0%	0.6%	57.7	164.7	129.0	1		
	TATACCATTTACAACCGTTG..TTCATT... □ 20	100.0%	100.0%	100.0%	~0%	0.0%	0.6%	54.7	165.7	133.0	1		
	TTATACCATTTACAACCGTTG..TCAT... □ 21	100.0%	100.0%	100.0%	~0%	0.0%	0.6%	53.7	164.7	133.0	1		
	TTTACCAACCGTTGCT..TCATTTGT... □ 22	100.0%	100.0%	100.0%	~0%	0.0%	0.6%	60.7	168.7	126.0	1		
	TTTACCAACCGTTGCT..TCATTTGT... □ 23	100.0%	100.0%	100.0%	~0%	0.0%	0.6%	60.7	168.7	127.0	1		
	ACCATTACAACCGTTG..TCATTGT... □ 24	100.0%	100.0%	100.0%	~0%	0.0%	0.6%	57.7	164.7	129.0	1		
	TATACCATTTACAACCGTTG..TTCATT... □ 25	100.0%	100.0%	100.0%	~0%	0.0%	0.6%	60.7	168.7	127.0	1		
	TTATACCATTTACAACCGTTG..TCAT... □ 26	100.0%	100.0%	100.0%	~0%	0.0%	0.6%	53.7	164.7	130.0	1		
	TTTACCAACCGTTGCT..TCATTTGT... □ 27	100.0%	100.0%	100.0%	~0%	0.0%	0.6%	60.7	165.7	125.0	1		
	TTTACCAACCGTTGCT..TCATTTGT... □ 28	100.0%	100.0%	100.0%	~0%	0.0%	0.6%	56.7	164.7	130.0	1		
	TACCAATTACAACCGTTGCT..TCAT... □ 29	100.0%	100.0%	100.0%	~0%	0.0%	0.6%	54.7	164.7	132.0	1		
	TTTACCAACCGTTGCT..TCATTTGT... □ 30	100.0%	100.0%	100.0%	~0%	0.0%	0.6%	57.7	165.7	132.0	1		
	TTTACCAACCGTTGCT..TCATTTGT... □ 31	100.0%	100.0%	100.0%	~0%	0.0%	0.6%	57.7	165.7	130.0	1		
	TTTACCAACCGTTGCT..TCATTTGT... □ 32	100.0%	100.0%	100.0%	~0%	0.0%	0.6%	57.7	165.7	130.0	1		
	TTTACCAACCGTTGCT..TCATTTGT... □ 33	100.0%	100.0%	100.0%	~0%	0.0%	0.6%	57.7	165.7	130.0	1		
	TTTACCAACCGTTGCT..TCATTTGT... □ 34	100.0%	100.0%	100.0%	~0%	0.0%	0.6%	57.7	165.7	130.0	1		
	TTTACCAACCGTTGCT..TCATTTGT... □ 35	100.0%	100.0%	100.0%	~0%	0.0%	0.6%	57.7	165.7	130.0	1		
	TTTACCAACCGTTGCT..TCATTTGT... □ 36	100.0%	100.0%	100.0%	~0%	0.0%	0.6%	57.7	165.7	130.0	1		
	TTTACCAACCGTTGCT..TCATTTGT... □ 37	100.0%	100.0%	100.0%	~0%	0.0%	0.6%	57.7	165.7	130.0	1		
	TTTACCAACCGTTGCT..TCATTTGT... □ 38	100.0%	100.0%	100.0%	~0%	0.0%	0.6%	57.7	165.7	128.0	1		

Primer Information Window:

Below is a description of all of the functions in this module, organized by the pull down menu they reside in. Note that some of the functions are also available by right clicking on a row. See the PRISE Tutorial for a few examples of how they can be used.

## File menu:



**Load Primer List:** Allows previously created primer lists to be uploaded into the software.

**Save Primer List:** Allows primer lists to be saved in the format used by the PRISE software.

**Save Primer List as Tab Delimited File:** Allows primer lists to be saved in a tab-delimited format, which can be used in standard spread sheet software.

**Save Primer Information Window Content:** Saves information in the Primer Information Window as a text file.

**Save Primer Pairs Only:** Saves primer pairs as a text file. Such files can be used for a variety of purposes, including being loaded in the Extract / Load Primer Candidates window (see above) in future experiments.

**Exit:** Closes the Design Primer module.

## Hide/Display menu:

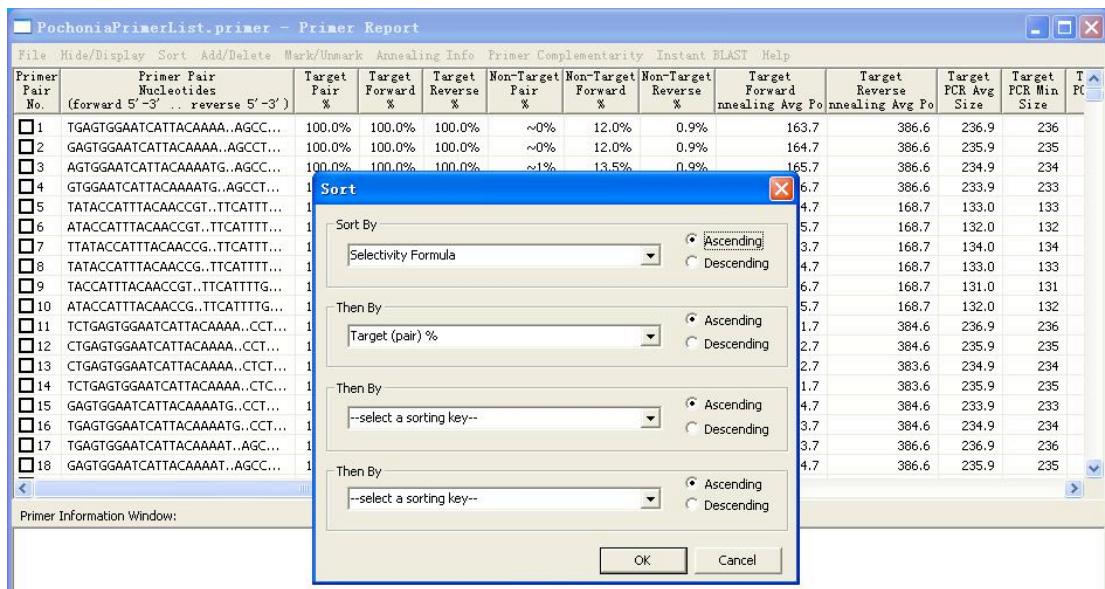
**Display All Columns:** Allows all data columns to be viewed. This function is only needed if the user had previously hidden columns.

**Hide/Display Columns:** Allows selected data columns to be hidden or displayed.

**Hide/Display Primer Pairs:** Allows selected primers to be hidden or displayed.

## Sort menu:

**Sort Primer List:** Allows the primers in the list to be sorted by a variety of user-selected criteria. One parameter that we find particularly useful is the Selectivity Formula, which is  $(1 - \% \text{ of target sequences estimated to be amplified})^2 + \frac{1}{2} [(\% \text{ of non-target sequences estimated to anneal with the forward primer})^2 + (\% \text{ of non-target sequences estimated to anneal with the reverse primer})^2]$ . The smaller the value generated by the Selectivity Formula, the more likely the primers will amplify target sequences and not amplify non-target sequences.

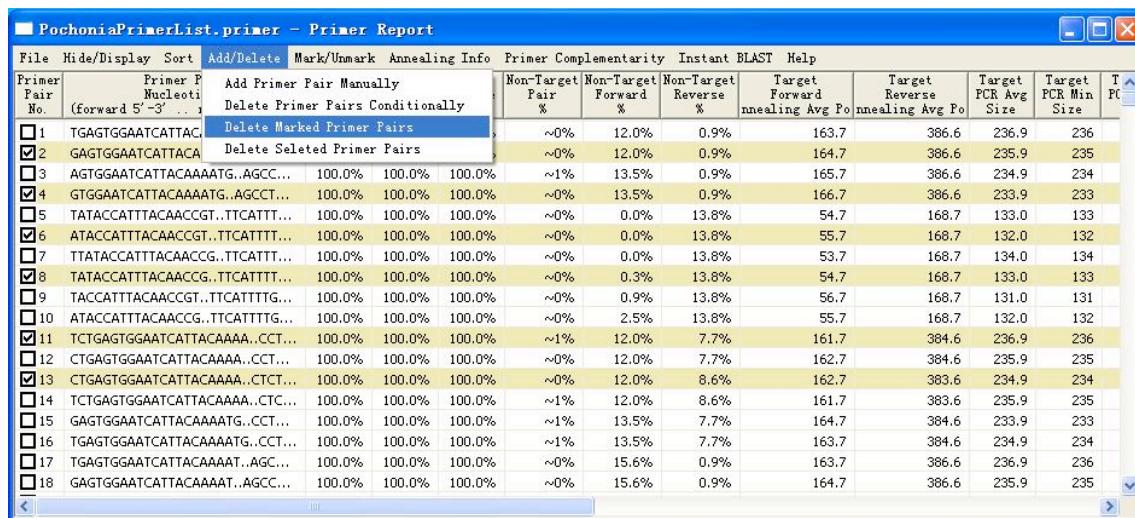


### Add/Delete menu:

**Add Primer Pair Manually:** Allows an individual primer pair to be added to the primer list, and its properties determined in relationship to the target and non-target sequence files and user-defined primer design settings.

**Delete Primer Pairs Conditionally:** Allows primer pairs to be deleted from the primer list by user-defined criteria.

**Delete Marked Primer Pairs:** Allows marked primers to be deleted. Marked primers are designated by a check mark in the box in the "Primer Pair No." column (and a yellow-highlighted row). Primers can be marked by clicking on the boxes in the "Primer Pair No." column or by using the Mark/Unmark functions below.



**Delete Selected Primer Pairs:** Allows selected primers to be deleted. Selected primers are designated by their rows being highlighted in blue. Primers can be selected by clicking on any part of the row except the boxes in the column titled “Primer Pair No.” Standard key commands such as shift and control can be used with this function, allowing groups of primer pairs to be selected.

PochoniaPrimerList.primer - Primer Report												
Primer Pair No.	Primer F Nucleotides (forward 5'-3')	Primer R Nucleotides (reverse 3'-5')	Add Primer Pair Manually		Primer Complementarity		Instant BLAST		Help			
			Add Primer Pair Conditionally	Delete Primer Pairs	Non-Target Pair %	Non-Target Forward %	Non-Target Reverse %	Target Forward Annealing Avg Pos	Target Reverse Annealing Avg Pos	Target PCR Avg Size	Target PCR Min Size	
1	TGAGTGGAAATCATTAC			Delete Marked Primer Pairs	~0%	12.0%	0.9%	163.7	386.6	236.9	236	
2	GAGTGGAAATCATTACA			Delete Selected Primer Pairs	~0%	12.0%	0.9%	164.7	386.6	235.9	235	
3	AGTGGAAATCATTACAAAATG..AGCC...		100.0%	100.0%	100.0%	~1%	13.5%	0.9%	165.7	386.6	234.9	234
4	GTGGAATCATTACAAAATG..AGCT...		100.0%	100.0%	100.0%	~0%	13.5%	0.9%	166.7	386.6	233.9	233
5	TATACCAATTACAAACCGT..TTCATTTT...		100.0%	100.0%	100.0%	~0%	0.0%	13.8%	54.7	168.7	133.0	133
6	ATACCATTTACAAACCGT..TTCATTTT...		100.0%	100.0%	100.0%	~0%	0.0%	13.8%	55.7	168.7	132.0	132
7	TATACCATTTACAAACCG..TTCATTTT...		100.0%	100.0%	100.0%	~0%	0.0%	13.8%	53.7	168.7	134.0	134
8	TATACCATTTACAAACCG..TTCATTTT...		100.0%	100.0%	100.0%	~0%	0.3%	13.8%	54.7	168.7	133.0	133
9	TACCATTTACAAACCGT..TTCATTTTG...		100.0%	100.0%	100.0%	~0%	0.9%	13.8%	56.7	168.7	131.0	131
10	ATACCATTTACAAACCG..TCATTTTG...		100.0%	100.0%	100.0%	~0%	2.5%	13.8%	55.7	168.7	132.0	132
11	TCTGAGTGGAAATCATTACAAAA..CTC...		100.0%	100.0%	100.0%	~1%	12.0%	7.7%	161.7	384.6	236.9	236
12	CTGAGTGGAAATCATTACAAAA..CCT...		100.0%	100.0%	100.0%	~0%	12.0%	7.7%	162.7	384.6	235.9	235
13	CTGAGTGGAAATCATTACAAAA..CTCT...		100.0%	100.0%	100.0%	~0%	12.0%	8.6%	162.7	383.6	234.9	234
14	TCTGAGTGGAAATCATTACAAAA..CTC...		100.0%	100.0%	100.0%	~1%	12.0%	8.6%	161.7	383.6	235.9	235
15	GAGTGGAAATCATTACAAAAATG..CCT...		100.0%	100.0%	100.0%	~1%	13.5%	7.7%	164.7	384.6	233.9	233
16	TGAGTGGAAATCATTACAAAAATG..CCT...		100.0%	100.0%	100.0%	~1%	13.5%	7.7%	163.7	384.6	234.9	234
17	TGAGTGGAAATCATTACAAAAATG..AGC...		100.0%	100.0%	100.0%	~0%	15.6%	0.9%	163.7	386.6	236.9	236
18	GAGTGGAAATCATTACAAAAATG..AGCC...		100.0%	100.0%	100.0%	~0%	15.6%	0.9%	164.7	386.6	235.9	235

### Mark/Unmark menu:

**Mark Selected Primer Pairs:** Allows selected primer pairs to be marked. Marked primer pairs are designated by a check mark in the box in the “Primer Pair No.” column (and a yellow-highlighted row). Marked primers can be saved in the PRISE program format or tab-delimited format using options in the File menu.

*Note that selected primers are designated by their rows being highlighted in blue. Primers can be selected by clicking on any part of the row except the boxes in the column titled “Primer Pair No.” Standard key commands such as shift and control can be used with the selection function, allowing groups of primer pairs to be selected.*

**Unmark Selected Primer Pairs:** Allows selected primer pairs to be unmarked.

### Annealing Info menu:

All of the functions below need to be performed on one primer pair. Before the function is performed, exactly one primer pair must be selected. Selected primers are designated by their rows being highlighted in blue. Primer pairs can be selected by clicking on any part of the row except the boxes in the column titled “Primer Pair No.”

**Primer Annealing Position Information:** Provides information on where the primers anneal to the target and non-target sequence.

**PochoniaPrimerList.primer - Primer Report**

File Hide/Display Sort Add/Delete Mark/Unmark Annealing Info Primer Complementarity Instant BLAST Help

Primer Pair No.	Primer Pair Nucleotides (forward 5'-3' .. reverse 5'-3')	Target Pair %	Primer Annealing Position Information									
1	TGAGTGGAAATCATTACAAAAA..AGCC...	100.0%	Percentage of Each Nucleotide in Target and Non-Target Sequences in Relation to Primer Sequences									
2	GAGTGGAATCATTACAAAAA..AGCCT...	100.0%	Target Sequences Annealing with Primer									
3	AGTGGAAATCATTACAAAATG..AGCC...	100.0%	Target Sequences Not Annealing with Primer									
4	GTGGAATCATTACAAAATG..AGCCT...	100.0%	Non-Target Sequences Annealing with Primer									
5	TATACCATTTAACACCGT..TTCATT...	100.0%	Non-Target Sequences Not Annealing with Primer									
6	ATACCATTTAACACCGT..TTCATT...	100.0%	100.0%	100.0%	0.0%	0.0%	13.8%	55.7	168.7	132.0	13:	
7	TTATACCATTTAACACCGT..TTCATT...	100.0%	100.0%	100.0%	~0%	0.0%	13.8%	53.7	168.7	134.0	13:	
8	TATACCATTTAACACCGT..TTCATT...	100.0%	100.0%	100.0%	~0%	0.3%	13.8%	54.7	168.7	133.0	13:	
9	TACCATTTAACACCGT..TTCATT...	100.0%	100.0%	100.0%	~0%	0.9%	13.8%	56.7	168.7	131.0	13:	
10	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	

Primer Information Window:  
Primer annealing position information for primer pair -- ATACCATTTAACACCGT..TTCATT...TGTAAATGATTCC

Target sequences:

```
>Seed sequence|PochoniaSeed
>gi|13160418|emb|AJZ92397.1|VCH292397Verticillium chlamydosporium var.
>gi|11933096|emb|AJZ91800.1|VCH291800Verticillium chlamydosporium 18S
>gi|118627604|emb|AM412780.1|Verticillium chlamydosporium partial 18S
>gi|59797389|gb|AY912487.1|Pochonia chlamydosporia isolate Pc 472 18S
>gi|58892741|gb|AY903605.1|Pochonia chlamydosporia 18S ribosomal RNA g
>gi|11933103|emb|AJZ91803.1|VCH291803Verticillium chlamydosporium 18S
>gi|1350841|emb|AJ303054.1|VCH303054Verticillium chlamydosporum var.
>gi|49424886|gb|AY555966.1|Verticillium catenulatum strain IMI 113078
>gi|11933106|emb|AJZ91804.1|VCH291804Verticillium chlamydosporum 18S
>gi|49424884|gb|AY555964.1|Pochonia chlamydosporia strain IMI 1161517 i
>gi|4836220|gb|AFLO8468.1|Diheterospora chlamydosporia internal transcr
>gi|49424885|gb|AY555965.1|Pochonia chlamydosporia strain IMI 113169 i
>gi|11933108|emb|AJZ91805.1|VCH291805Verticillium chlamydosporium 18S
>gi|34420181|gb|AY273334.1|Uncultured ascomycete clone T7B12 18S ribos
```

Non-target sequences:

**Percentage of Each Nucleotide in Target and Non-Target Sequences in Relation to Primer Sequences:** Provides the percentage of each nucleotide, at each position in the target and non-target sequences, in relation to the nucleotides in each position of the primers.

**PochoniaPrimerList.primer - Primer Report**

File Hide/Display Sort Add/Delete Mark/Unmark Annealing Info Primer Complementarity Instant BLAST Help

Primer Pair No.	Primer Pair Nucleotides (forward 5'-3' .. reverse 5'-3')	Target Pair %	Primer Annealing Position Information									
1	TGAGTGGAAATCATTACAAAAA..AGCC...	100.0%	Percentage of Each Nucleotide in Target and Non-Target Sequences in Relation to Primer Sequences									
2	GAGTGGAATCATTACAAAAA..AGCCT...	100.0%	Target Sequences Annealing with Primer									
3	AGTGGAAATCATTACAAAATG..AGCC...	100.0%	Target Sequences Not Annealing with Primer									
4	GTGGAATCATTACAAAATG..AGCCT...	100.0%	Non-Target Sequences Annealing with Primer									
5	TATACCATTTAACACCGT..TTCATT...	100.0%	Non-Target Sequences Not Annealing with Primer									
6	ATACCATTTAACACCGT..TTCATT...	100.0%	100.0%	100.0%	0.0%	0.0%	13.8%	55.7	168.7	132.0	13:	
7	TTATACCATTTAACACCGT..TTCATT...	100.0%	100.0%	100.0%	~0%	0.0%	13.8%	53.7	168.7	134.0	13:	
8	TATACCATTTAACACCGT..TTCATT...	100.0%	100.0%	100.0%	~0%	0.3%	13.8%	54.7	168.7	133.0	13:	
9	TACCATTTAACACCGT..TTCATT...	100.0%	100.0%	100.0%	~0%	0.9%	13.8%	56.7	168.7	131.0	13:	
10	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	

Primer Information Window:  
Percentage of each nucleotide in target and non-target sequences in relation to primer sequences information for primer pair -- AT: 15 target sequences and 326 non-target sequences.

Information for target sequences annealing with primer pair: (15 sequences)

A	T	A	C	C	A	T	T	A	C	A	A	C	C	G	T	..
A	100.0%	0.0%	100.0%	0.0%	0.0%	100.0%	0.0%	0.0%	100.0%	0.0%	100.0%	100.0%	0.0%	0.0%	0.0%	..
C	0.0%	0.0%	0.0%	100.0%	100.0%	0.0%	0.0%	0.0%	0.0%	100.0%	0.0%	0.0%	100.0%	0.0%	0.0%	..
G	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	100.0%	0.0%	0.0%	100.0%	0.0%	..
T	0.0%	100.0%	0.0%	0.0%	0.0%	100.0%	100.0%	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	100.0%	..

Information for target sequences not annealing with primer pair: (0 sequences)

A	T	A	C	C	A	T	T	A	C	A	A	C	C	G	T	..
A	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	..
C	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	..
G	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	..
T	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	..

**Target Sequences Annealing with Primer:** Shows the target sequences that anneal to the primer, using the user-selected primer design criteria.

**PochoniaPrimerList.primer - Primer Report**

File	Hide/Display	Sort	Add/Delete	Mark/Unmark	Annealing Info	Primer Complementarity	Instant BLAST	Help			
Primer Pair No.	Primer Pair Nucleotides (forward 5'-3' ... reverse 5'-3')	Target %	Target Forward %	Target Reverse %	Primer Annealing Position Information						
1	TGAGTGGAAATCATTACAAAAA..AGCC...	100.0%			Percentage of Each Nucleotide in Target and Non-Target Sequences in Relation to Primer Sequences						
2	GAGTGGAAATCATTACAAAAA..AGCT...	100.0%			Target Sequences Annealing with Primer						
3	AGTGGAAATCATTACAAAATG..AGCC...	100.0%			Target Sequences Not Annealing with Primer						
4	GTGGAATCATTACAAAATG..AGCT...	100.0%			Non-Target Sequences Annealing with Primer						
5	TATACCATTACAAACCGT..TTCATT...	100.0%			Non-Target Sequences Not Annealing with Primer						
6	ATACCATTACAAACCGT..TTCATT...	100.0%	100.0%	0.0%	0.0%	13.8%	55.7	168.7	132.0	13	
7	TTATACCATTTACAAACCG..TTCATT...	100.0%	100.0%	100.0%	~0%	0.0%	13.8%	53.7	168.7	134.0	13
8	TATACCATTTACAAACCG..TTCATT...	100.0%	100.0%	100.0%	~0%	0.3%	13.8%	54.7	168.7	133.0	13
9	TACCATTTACAAACCGT..TTCATT...	100.0%	100.0%	100.0%	~0%	0.9%	13.8%	56.7	168.7	131.0	13
					Update Estimated Information to Exact Information						

**Primer Information Window:**

```
>Seed sequence|PochoniaSeed
CATGTCATTTACACCGTACACCGTCTGCTCCGGGGTTCTCGCCCCGGGTTAACACCCCGAACCCAGGGGGGGGGGACCCAAACTCTAGATTTTATGGCATGCTGACTGGAATCJ
>gi|13160418|emb|IAJ292397.1|VCH292397Verticillium chlamydosporium var. chlamydosporium partial 18S rRNA gene, internal transcribed ATTACCGCTTCAACTCCAAACCCGATGTGAACCTATACCATTTACACCGTCTGCTCCGGGGGCTTACACCCCGAACAGCCGGCCGGGACCCAAACTCTAGATT;
>gi|11933096|emb|IAJ291800.1|VCH291800Verticillium chlamydosporium 18S rRNA gene (partial), ITS1, 5.8S rRNA gene, ITS2, 28S rRNA gen GGAACATAAAGCTGTAAACAAAGCTCTCGGTGTTGGTGAACAGCGGGAGGATCATTAACCGAGTTTCACTCCAAACCCCATGTGAACCTATACCATTTACACCGTCTGCTCCGGGGCTCTCGCCCGGGG
>gi|118627604|emb|IAJ412780.1|VCh291800Verticillium chlamydosporium partial gene, ITS1, 5.8S rRNA gene, ITS2 and partial 28S rRNA gen TGTAAACAAAGCTCTCGGTGTTGGTGAACACCGGAGGATCATTAACCGAGTTTCACTCCAAACCCCATGTGAACCTATACCATTTACACCGTCTGCTCCGGGGCTCTCGCCCGGGCTTACACCCGC
>gi|59797389|gb|AY912487.1|Pochonia chlamydosporia isolate Pc 472 18S ribosomal RNA gene, partial sequence; internal transcribed sp AACTAAAGCTGTAAACAAAGCTCTCGGTGTTGGTGAACACCGGAGGATCATTAACCGAGTTTCACTCCAAACCCCATGTGAACCTATACCATTTACACCGTCTGCTCCGGGGCTCTCGCCCGGGCT
>gi|58982741|gb|AY903605.1|Pochonia chlamydosporia 18S ribosomal RNA gene, partial sequence; internal transcribed spacer 1, 5.8S r: CGACCGAGGATATTGCTTCACTCCAAACCCCATGTGAACCTATACCATTTACACCGTCTGCTCCGGGGCTTACACCCCGAACAGCCGGCCGGGACCC
>gi|11933103|emb|IAJ291803.1|VCH291803Verticillium chlamydosporium 18S rRNA gene (partial), ITS1, 5.8S rRNA gene, ITS2, 28S rRNA gen GGAACATAAAGCTGTAAACAAAGCTCTCGGTGTTGGTGAACACCGGAGGATCATTAACCGAGTTTCACTCCAAACCCCATGTGAACCTATACCATTTACACCGTCTGCTCCGGGGCTCTCGCCCGGG
>gi|13508841|emb|IAJ303054.1|VCH303054Verticillium chlamydosporium var. chlamydosporium 5.8S rRNA gene and ITS 1 and 2, strain CB3e CATTACCGCTTCACTCCAAACCCCATGTGAACCTATACCATTTACACCGTCTGCTCCGGGGCTTACACCCCGAACAGCCGGCCGGGACCCAAACTCTAGAT;
>gi|49424886|gb|AY555966.1|Verticillium catenulatum strain IMI 113078 internal transcribed spacer 1, 5.8S ribosomal RNA gene, and : CCCAGTTTCACTCCAAACCCCATGTGAACCTATACCATTTACACCGTCTGCTCCGGGGCTTACACCCCGAACAGCCGGCCGGGGACCCAAACTCTAGATTTTA;
>gi|11933106|emb|IAJ291804Verticillium chlamydosporum 18S rRNA gene (partial), ITS1, 5.8S rRNA gene, ITS2, 28S rRNA gen GGAACATAAAGCTGTAAACAAAGCTCTCGGTGTTGGTGAACACCGGAGGATCATTAACCGAGTTTCACTCCAAACCCCATGTGAACCTATACCATTTACACCGTCTGCTCCGGGGCTCTCGCCCGGG
```

**Target Sequences Not Annealing with Primer:** Shows the target sequences that do not anneal to the primer, using the user-selected primer design criteria.

**Non-Target Sequences Annealing with Primer:** Shows the non-target sequences that anneal to the primer, using the user-selected primer design criteria.

**Non-Target Sequences Not Annealing with Primer:** Shows the non-target sequences that do not anneal to the primer, using the user-selected primer design criteria.

**Update Estimated Information to Exact Information:** If the “Get estimated PCR product information” option was selected in the Standard Primer Property Settings window when the primer design settings were chosen, this function recalculates the primer property values and reports the exact data. The estimated information option is typically used in the primer designing process, because the exact information option increases the computational time (see Appendix II for more information on this “Get estimated PCR product information” function).

### Primer Complementarity menu:

**Primer Report**

File	Hide/Display	Sort	Add/Delete	Mark/Unmark	Annealing Info	Primer Complementarity	Instant BLAST	Help
Primer Pair No.	Primer Pair Nucleotides (forward 5'-3' ... reverse 5'-3')	Target %	Target Forward %	Target Reverse %	Primer Inter-Complementarity			
					Primer Inter-Complementarity	Target Forward Avg Po	Target Reverse Annealing Avg Po	Target PCR Avg Size
					Primer 3' Inter-Complementarity			Target PCR Min Size
					Primer Intra-Complementarity			Target PCR Max Size
					Primer 3' Intra-Complementarity			

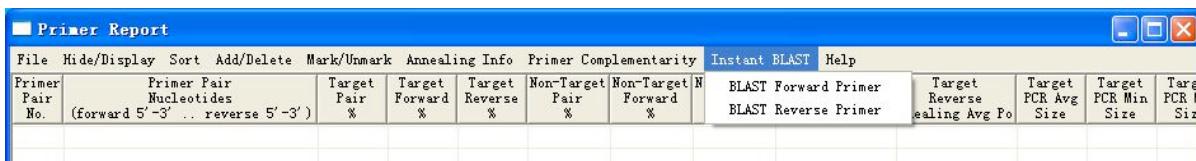
*Primer Inter-complementarity:* Provides information on the inter-complementarity of the entire primer.

*Primer 3' Inter-complementarity:* Provides information on the inter-complementarity of the last eight 3' primer nucleotides. Note that this value can be customized in the Standard Primer Property Settings window.

*Primer Intra-complementarity:* Provides information on the intra-complementarity of the entire primer.

*Primer 3' Intra-complementarity:* Provides information on the intra-complementarity of the last eight 3' primer nucleotides. Note that this value can be customized in the Standard Primer Property Settings window.

#### Instant BLAST menu:



*Blast Forward Primer:* Allows a single forward primer to be subjected to a BLAST analysis, by opening the Blast page at NCBI and loading the primer. Note that this function only works when one primer pair is selected.

*Blast Reverse Primer:* Allows a single reverse primer to be subjected to a BLAST analysis, by opening the Blast page at NCBI and loading the primer. Note that this function only works when one primer pair is selected.

#### Help menu:

*PRISE Manual:* Opens the PRISE Manual.

*PRISE Tutorial:* Opens the PRISE Tutorial, which provides a step-by-step protocol showing how the software was used to create sequence-selective PCR primers for a specific fungal rRNA gene.

## Appendix I: Primer Selectivity Settings.

Mispriming happens often in PCR experiments and it may or may not affect the PCR result. The efficiency of the polymerase to recognize and extend a mismatched duplex is not only sensitive to the number of mismatched nucleotide bases, but also to the nucleotide composition and location of the mismatches. Our Primer Selectivity Settings is composed of the *Mismatch Cost Matrix* and *Positional Mismatch Allowance Settings* to accurately evaluate the selectivity of a primer pair. Users can use default settings or customize the settings to suit their specific application. We now explain the fundamentals of our Primer Selectivity Settings.

1. *Mismatch Cost Matrix*: To capture various effects of mismatched nucleotides, the Primer Selectivity Settings allows the user to assign different penalties on the mismatched nucleotides in the Mismatch Cost Matrix. Each entry in the matrix specifies the penalty level of the corresponding mismatch in the primer-template duplex. Here the larger value of cost in the matrix, the more unlikely for a duplex with this mismatch to be predicted to be stable (and therefore a PCR to be made). The Mismatch Cost Matrix has entries for each nucleotide base A, C, G and T. The mismatch cost of ambiguous bases, such as N, R and Y, etc., will be obtained automatically by the average of mismatch cost between the non-ambiguous bases which the corresponding ambiguous bases represent. For example, since ambiguous base R denotes A and G, and base Y denotes T and C, the mismatch cost of R and Y  $mc(R,Y)$  can be calculated by the formula  $mc(R,Y) = ( mc(A,T) + mc(A,C) + mc(G,T) + mc(G,C) ) / 4$ .

2. *Positional Mismatch Allowance Settings*: This component captures the cost allowance of the insertion/deletion and mismatched nucleotides for position range in primer-template duplex. In the basic Primer Selectivity Settings, the exact Positional Mismatch Allowance for the three 3' end positions of primer can be specified for target and non-target sequences, respectively. If the setting is set to xyz, then only primer-template pairs that satisfy this xyz match-mismatch configuration and those above will be considered as producing a PCR product. xyz is the maximum allowed accumulated number of mismatches counting from right hand side (i.e., 3' end of primer).

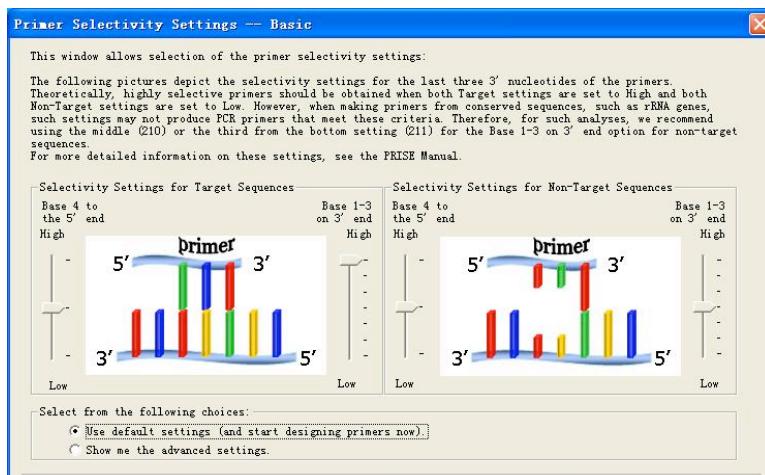
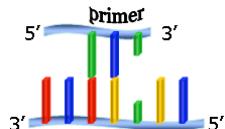


Figure 3. An example of basic Primer Selectivity Settings

An example of these settings is given in Figure 3, in which 000 setting is set for target sequences, and 210 setting is set for non-target sequences. This means that no mismatch is allowed on the three 3' end positions of primer for target sequences, while no mismatch on the first base on 3' end and at most two mismatches are allowed on the 2<sup>nd</sup> and 3<sup>rd</sup> bases on 3' end of primer is required for non-target sequences. Thus for target sequences, only exact matches at all these three positions will be scored as creating a PCR product, while any primer-template pair with match-mismatch configurations of 000, 100, 110, and 210 for non-target sequences will be counted as producing a PCR product. In this basic version of Primer Selectivity Settings, the approximative match-mismatch from the fourth base of primer's 3' end to 5' end can be specified, as well. This is illustrated by the example in Figure 3, in which high match percentage is required on the region from the fourth base to 5' end for target sequences, while medium match percentage is required on the segment from the 4th base to 5' end for non-target sequences. By moving the slider bars on the side of two pictures, these settings can be changed. Note that there are totally 8 different combinations of match-mismatch situation for the three 3' end positions, but only 7 pictures can be shown in this window and they represent the settings: 000, 100, 110, 210, 211, 221, 321. The picture below, which represents the 111 setting, is left out because the 111 setting and 210 setting are not compatible. More specifically, all of the above 7 settings are ordered strictly from more to less stringency in considering the likelihood of getting a PCR product. However, the 111 and 210 settings cannot be ordered by our Primer Selectivity Settings system.



111

Default Positional Mismatch Allowance Settings should be suitable for most applications, but they also can be customized using the advanced option. In this setting, the cumulative mismatch cost allowance for each primer position from 3' end can be specified. Each entry of positional mismatch array represents the maximum allowed cost for the region from 3' end to the corresponding point of the primer. We give an example to describe the use of these advanced settings. Consider the Mismatch Cost Matrix and the Positional Mismatch Allowance Settings for non-target sequences in Figure 4.

A	0			
C	1	0		
G	2	1	0	
T	1	3	1	0
	A	C	G	T

Non-target sequence:

30	29	28	27	26	25	24	23	22	21	20	19	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	
5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	4	4	4	4	4	4	3	3	1	1	0

Figure 4. An example of Mismatch Cost Matrix and Positional Mismatch Allowance Setting

This combination setting can be interpreted as:

- (1) No mismatch is allowed at the first base on 3' end;
- (2) At most one C-A, G-C, T-A or T-G mismatch and no G-A or T-C mismatch is allowed on the second to the third base on 3' end;
- (3) One T-C mismatch on the fourth base with no mismatch from the first to the third base on 3' end, or one G-A mismatch on the fourth base with at most one C-A, G-C, T-A or T-G mismatch is allowed on the second to the third base on 3' end.

Under this setting, the primer 5'-**ctaaactactgagaa**-3' will be predicted to amplify the sequence 5'-...**ctaaactactggaa**...-3' (more precisely, anneal to the reverse complement strand of this sequence), since the cumulative positional cost is 5'-...,2,2,2,2,0,0,0 -3', which satisfies the Positional Mismatch Allowance Settings. Note that in this example we didn't count the effect of insertion/deletion costs, and the calculation with these effect considered are more or less similar.

According to the fixed Primer Selectivity Settings, PRISE performs a local alignment for the primer against each sequence in target and non-target group, and predicts the position in the sequence where this primer anneals (or does not anneal at all).

## **Appendix II: Algorithms.**

1. *Speeding up the searching for a primer in a sequence:* One important step in PRISE computations is to predict the most possible annealing positions for every primer from a set of primer candidates in each target / non-target sequence according to the user specified Primer Selectivity Settings. We formulate it as a string matching (with mismatch and insertion / deletion) problem, in which we look for a segment in a sequence most similar to a given primer with possible insertion / deletions, or decide there is no such substring in this sequence. More specifically, for a given Primer Selectivity Settings, where the penalty for each type of mismatch, insertion / deletion penalty and maximum penalty allowance are specified, we perform a local string alignment of a primer versus a target / non-target sequence, and find a segment of the sequence matching the primer the best. This task is accomplished with dynamic programming. The naive solution of dynamic programming is rather slow since the number of candidate primers is in the hundreds of thousands. We speed-up this procedure by grouping the primer candidates into clusters, such that each of these clusters has a representative primer candidate and the other primer candidates in this cluster are proper-prefixes of the representative primer, and performing the dynamic programming computation once for each cluster versus each target / non-target sequence by computing string alignment table for the representative primer of each cluster and performing a table look-up for every primer in this cluster. This method reduces the total running time of primer searching three-fold.
2. *Get estimated PCR product information:* After we have the annealing position prediction of each single primer candidate in every target / non-target sequence, the next step is to pair each single primer in every possible way and compute the percentage of target / non-target sequences annealing (predicted) to each primer pair, and the average, maximum and minimum PCR product length of the primer pair in target / non-target sequences. To speed-up the computation, we use a random sampling. For each primer pair, we randomly pick 100 target / non-target sequences and compute the percentage and PCR length statistics, and then use these sample statistics to estimate the behavior of all the targets / non-targets.

### **Appendix III: Tm formula.**

$$T_m = 81.5 + 16.6 \log [Na^+] + 41(G + C)/length - 500/length$$