PRISE Tutorial

Below is a detailed description of how PRISE was used to design sequence-selective PCR primers for an rRNA gene from the fungus *Pochonia chlamydosporia*. All of the files described below are available in the sub-folder "sample files" within your downloaded and unzipped folder "PRISE". 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.

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.

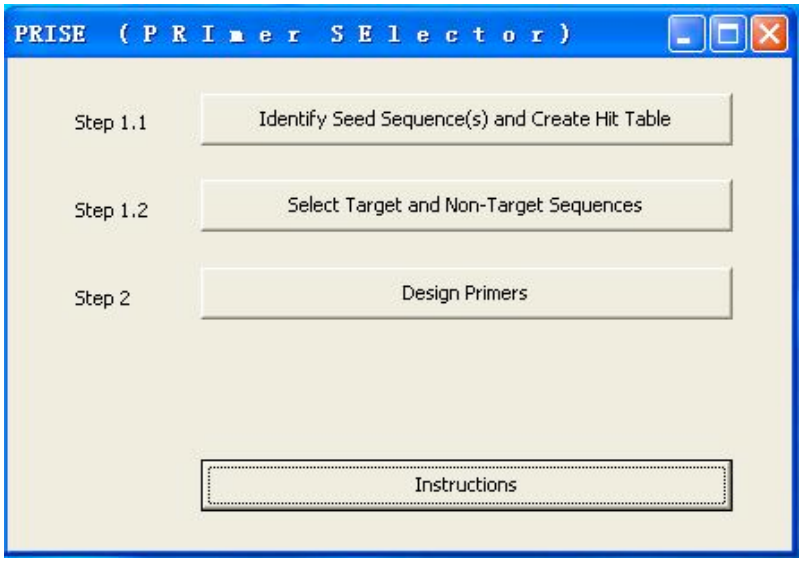


Figure 1. Opening window with links to the instructions and modules of the PRISE software.

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.

1. *Identify the Seed Sequence*: For our example, the Seed Sequence is an rRNA gene from the fungus *P. chlamydosporia* (see below). This sequence was saved in FASTA format as a plain text file labeled PochoniaSeed.

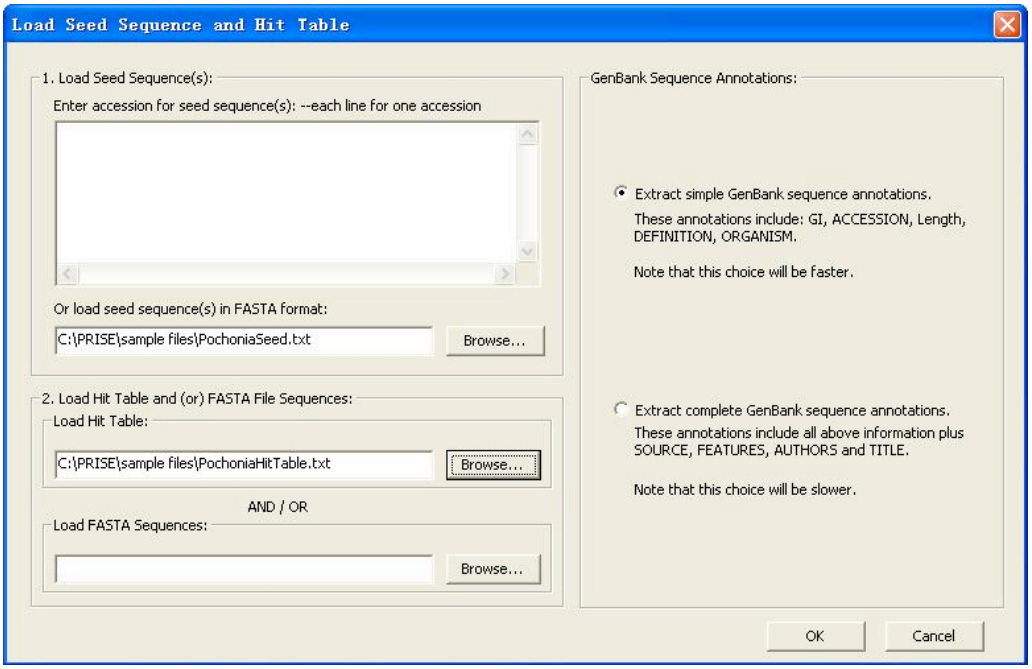
   CATGTGAACTTATACCATTTACAACCGTTGCTTCGGCGGGTTCTCGCCCCGGGC
   TTTACACCCCGGAACCAGGCGGCCCGCCGGGGGACCCAAACTCTAGATTTTTA
   TTTTGGCATGTCTGAGTGGAATCATTACAAAATGAATCAAAACTTTCAACAACGG
   ATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGCGATAAGTAATGTG
   AATTGCAGAACTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCGCCAG
   TATTCTGGCGGGCATGCCTGTTCGAGCGTCATTTCAACCCTCAAGCCCCAGCG
   GTTTGGTGTTGGGGACCGGCGAGTACAGAGGCTTTGGGGACTTGTCCCCCTTC
   CCTCGGCGCCGCCCCCGAAATGAATTGGCGGTCTCGTCGCGGCCTCCTCTGC
   GTAGTAGCACAACCTCGCATCAGGAGCGCGGCGCGGCCACTGCCGTAAAACG
   CCCAACTTTTTTTTAAGAG

2. *Create the Hit Table*: The Seed Sequence was subjected 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 will typically be a minimum of 500. For our example, the **Database** was **Others (nr etc.)** and number of **Max target sequences**, which is located in the **Algorithm parameters** section, was 500. 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. In our example, the Hit Table was labeled PochoniaHitTable. 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.

3. *Loading the sequences*. After opening the module, the user loads 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. For our example, in the **1. Load Seed Sequence(s)** section, we loaded the file PochoniaSeed from the “sample files” folder. In the **Load Hit Table** sub-section of **2. Load Hit Table and (or) FASTA File Sequences**, we loaded the file PochoniaHitTable from the “sample files” folder. Click **OK** and the Sequence Alignment Settings for Pairwise Identity Analysis window opens. Click **OK**.

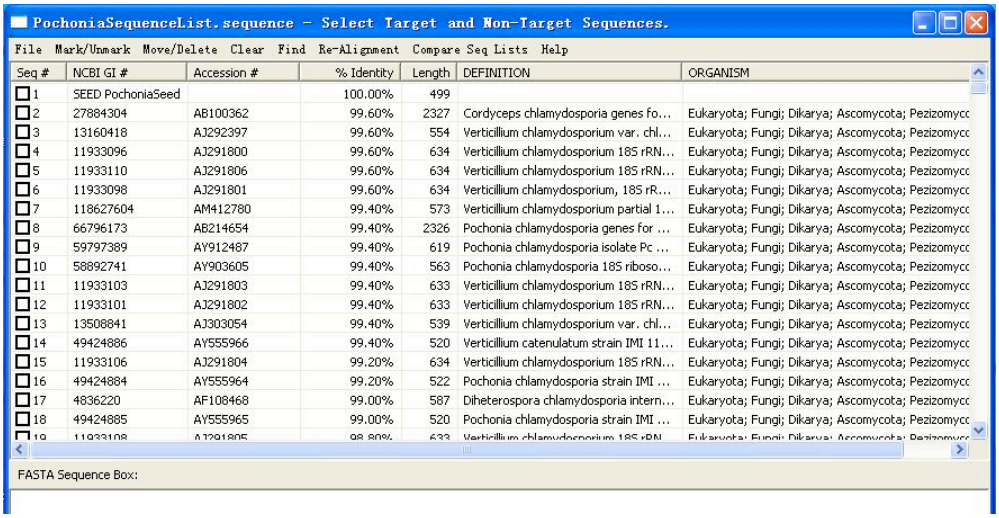


Collecting and parsing the sequence data. After the sequences are loaded, the software downloads all of the GenBank records associated with the Seed Sequence(s) and Hit Table sequences, parses the data contained within them into separate components, performs pairwise % identity analyses between the Seed Sequence(s) and Hit Table sequences, and displays these data in tabular form in the Select Target and Non-Target Sequences window. The title of the window will be the Hit Table file name followed by "- Select Target and Non-Target Sequences."

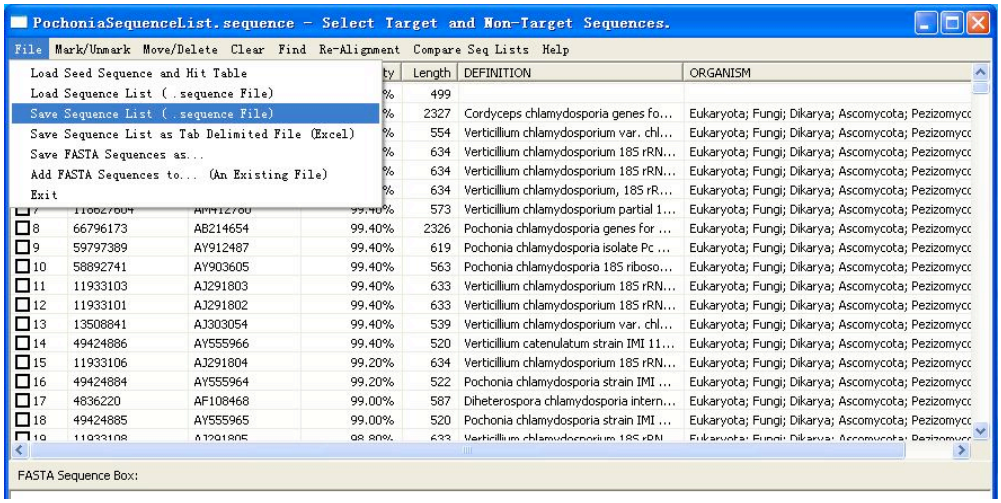
4. After the program finishes processing the data, which could take minutes to hours, depending on the number of sequences in the Hit Table, the speed of the internet connection and the capabilities of the computer, a sequence downloading report appears. This Report lists the accession number of sequence(s) 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. Click **OK**.

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 sequences to be selected by a variety of parameters such as sequence length, % sequence identity or GenBank parameters such as Definition or Source.

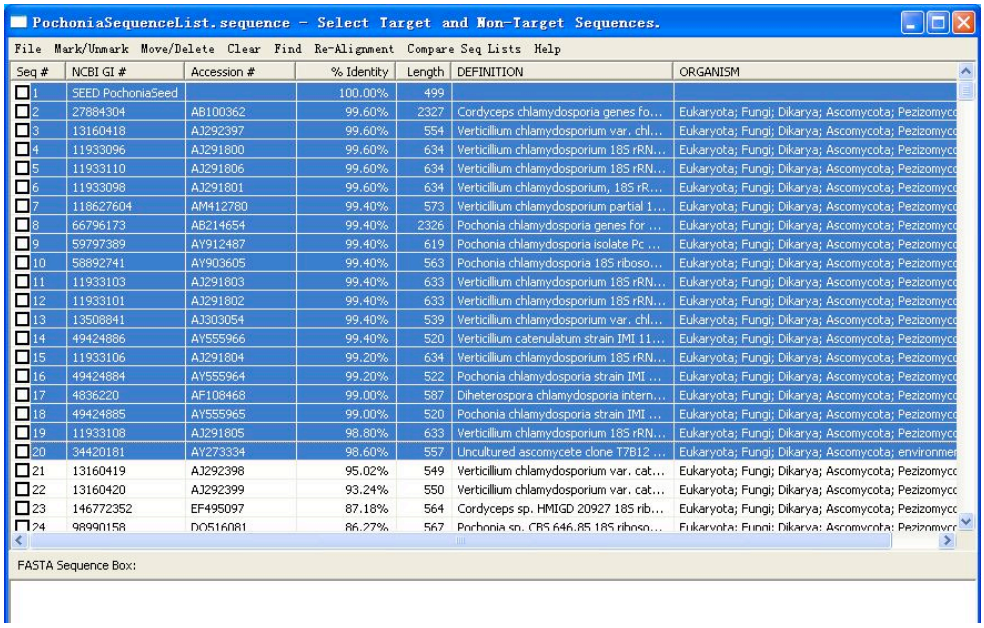
5. In our example, we clicked on the column heading **% Identity** to sort the sequences by this parameter.



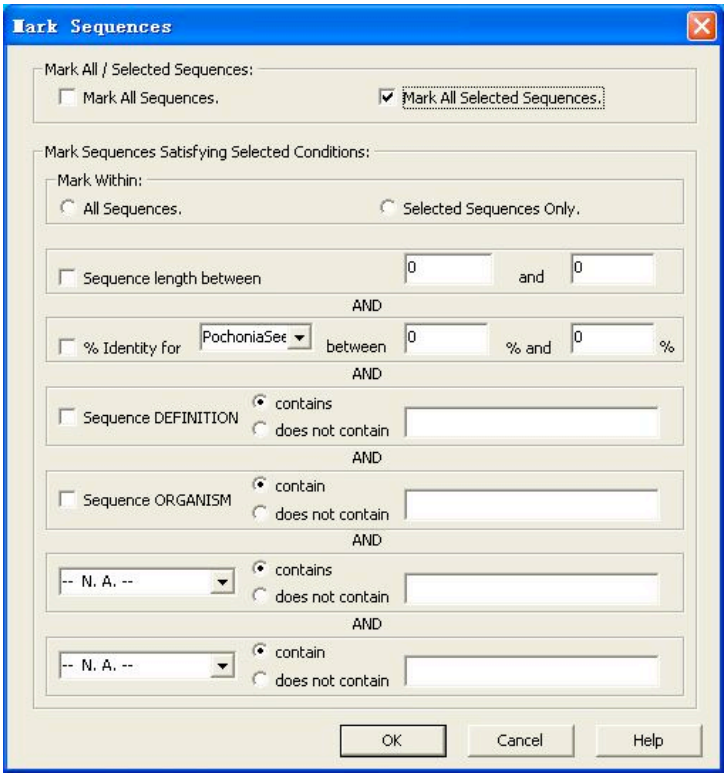
6. These data were saved as PochoniaSequenceList using the **Save Sequence List** option from the **File** menu.



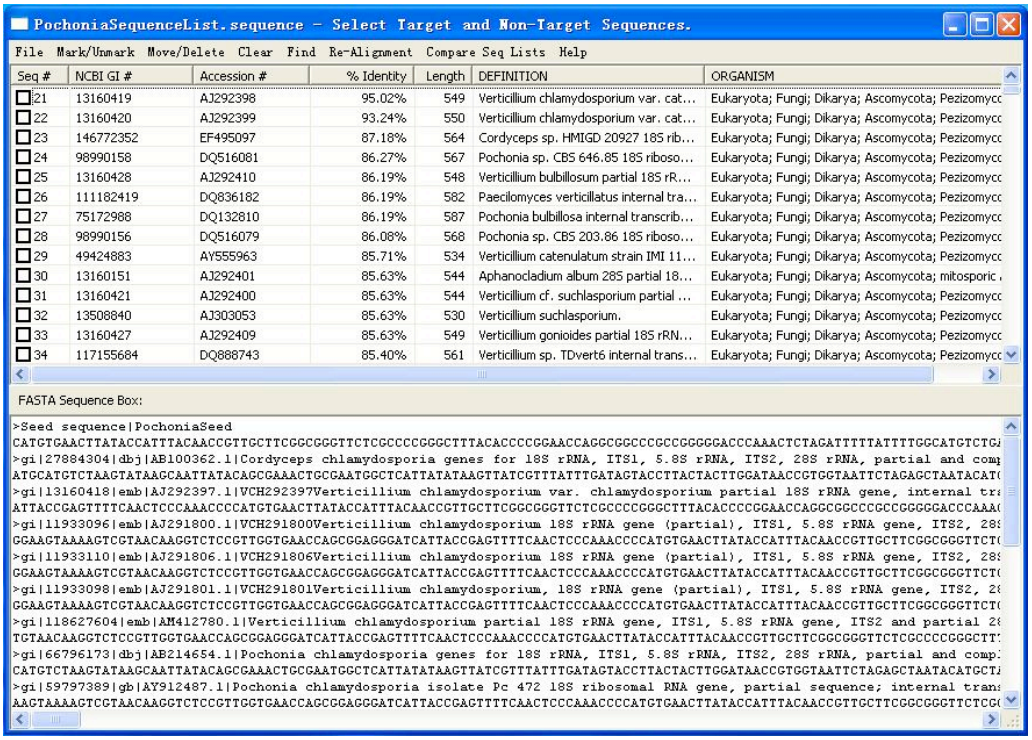
7. Sequences 1 through 20 were selected by clicking on any part (except the boxes in the column “Seq #”) of rows 1 and 20 while holding down the shift key.



8. These twenty sequences were “marked” by clicking the **Mark Sequences** option from the **Mark/Unmark** menu. This opens a new window where we selected the **Mark All Selected Sequences** option and then clicked **OK**.



9. These sequences were moved to the FASTA Sequence Box by using the **Move Marked Sequences to FASTA Sequence Box** option from the **Move/Delete** menu.



10. These sequences were saved as PochoniaTarget by using the **Save FASTA Sequences As** option from the **File** menu. These twenty sequences are the target sequences, which are those that the primers should amplify.

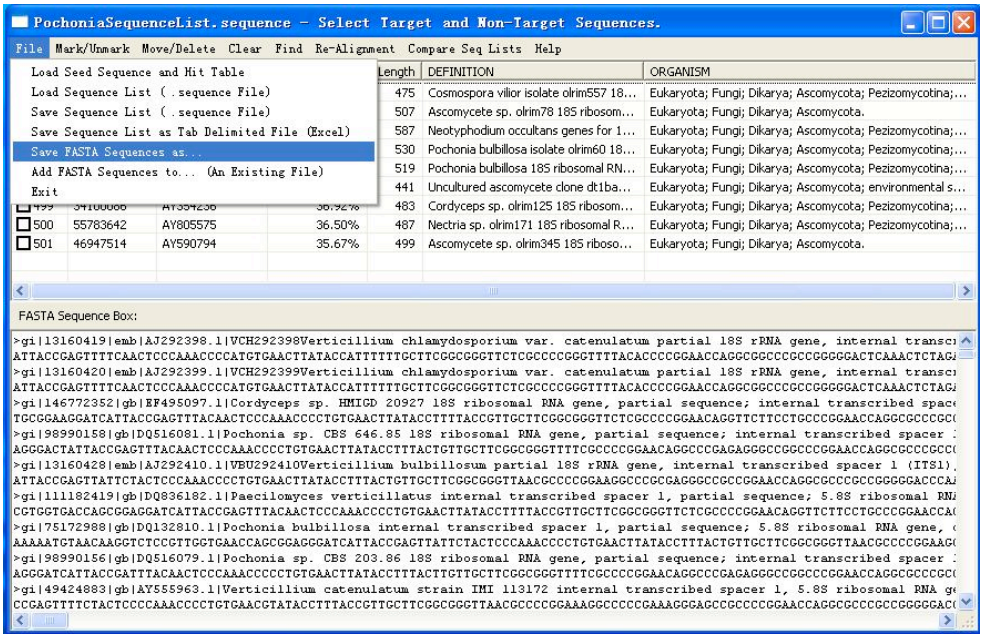
11. These sequences were removed from the FASTA Sequence Box by using the **Clear FASTA Sequence Box** option from the **Clear** menu.

12. Sequences 21 through 492 were selected by clicking on any part (except the box in the column “Seq #”) of rows 21 and 492 while holding down the shift key.

13. These 472 sequences were “marked” by clicking the **Mark Sequences** option from the **Mark/Unmark** menu. This opens a new window where we selected the **Mark All Selected Sequences** option and then clicked **OK**.

14. These sequences were moved to the FASTA Sequence Box by using the **Move Marked Sequences to FASTA Sequence Box** option from the **Move/Delete** menu.

15. These sequences were saved as PochoniaNonTarget by using the **Save FASTA Sequences As** option from the **File** menu. These 472 sequences are the non-target sequences, which are those that the primers should not amplify.



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. (See PRISE Manual for more detailed information on Primer Selectivity Settings.) 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 manual for details.)

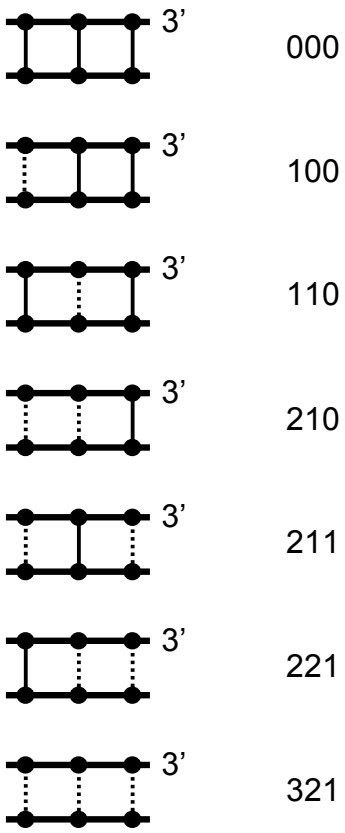
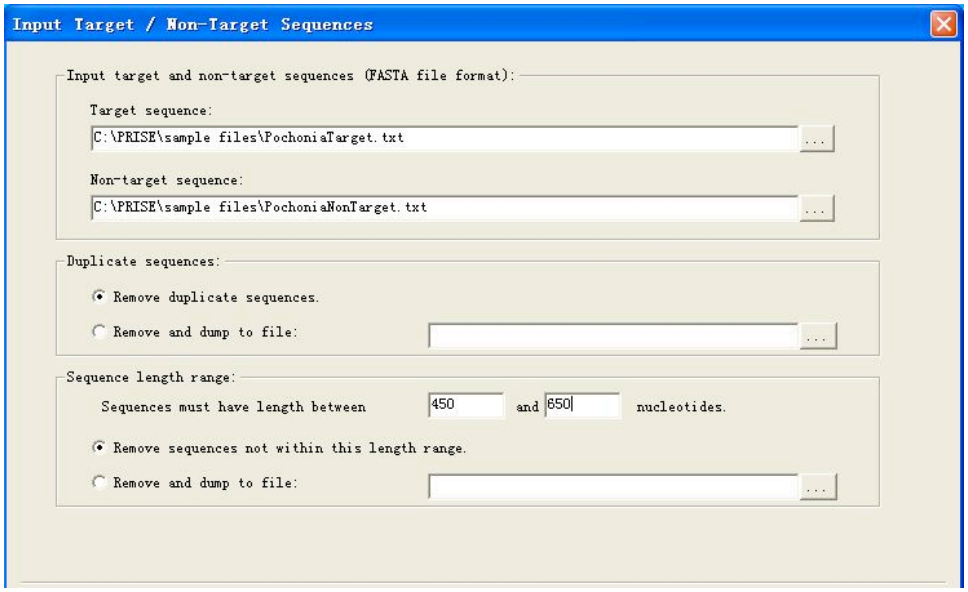


Figure 2. Scoring scheme for the sequence-selectivity component of the Design Primers module. On the left side are depictions of the last three 3’ nucleotides of a primer and its corresponding template. The primer is the top strand. Base-paired nucleotides are designated by solid lines. Non-based paired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

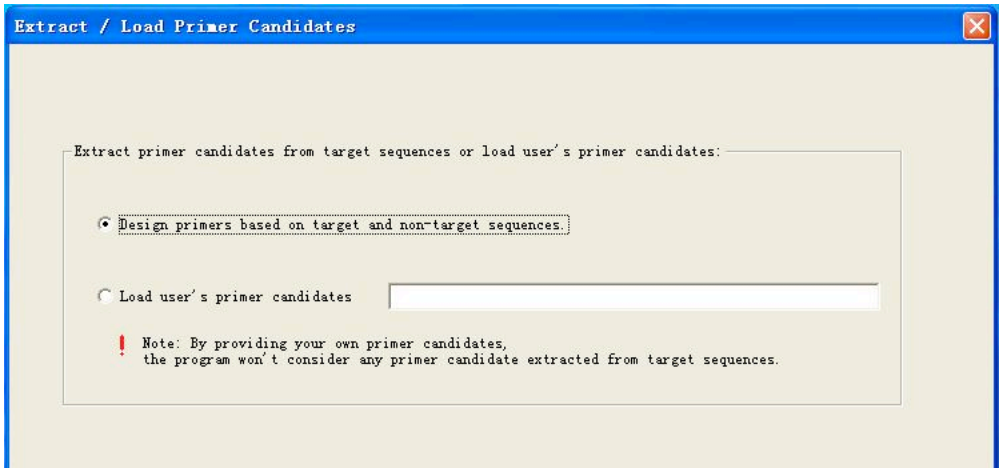
This scoring scheme can be set separately for target and non-target sequences. This useful feature gives a user the flexibility to define different stringency requirements for primer annealing within these two classes of sequences.

16. *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 pair list file or initiate a new primer design project. In our example, we selected the **Start designing new primers** option. Click **Next**. 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 as well as those that do not meet user-selected size criteria. In our example, we loaded files PochoniaTarget and PochoniaNonTarget from the "sample files" folder as the target and non-target files, respectively. We also chose the **Remove duplicate sequences** option and we set the **Sequences must have length between** option to 450 to 650 nucleotides. Click **Next**.



17. This opens the Extract / Load Primer Candidates window. Here, the user can choose to either (i) design new primers based on the target and non-target sequences and user defined primer criteria or (ii) assess the qualities of user-selected primers in relation to the target and non-target sequences and user-defined primer criteria.

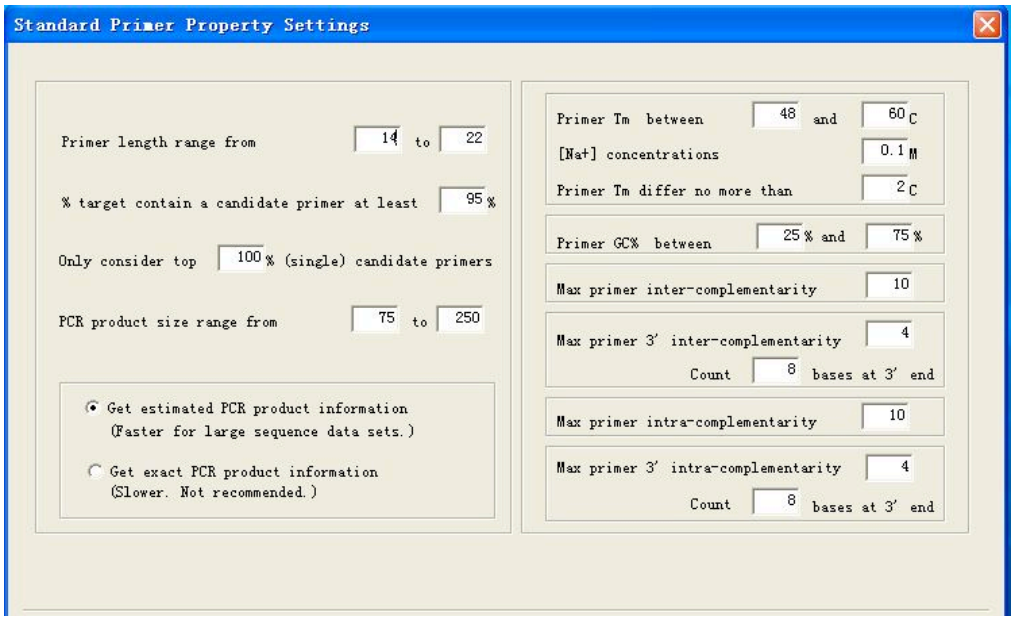
18. In our example, we chose the **Design primers based on target and non-target sequences** option. Click **Next**.



19. In the next window, titled 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 have the option of changing them.

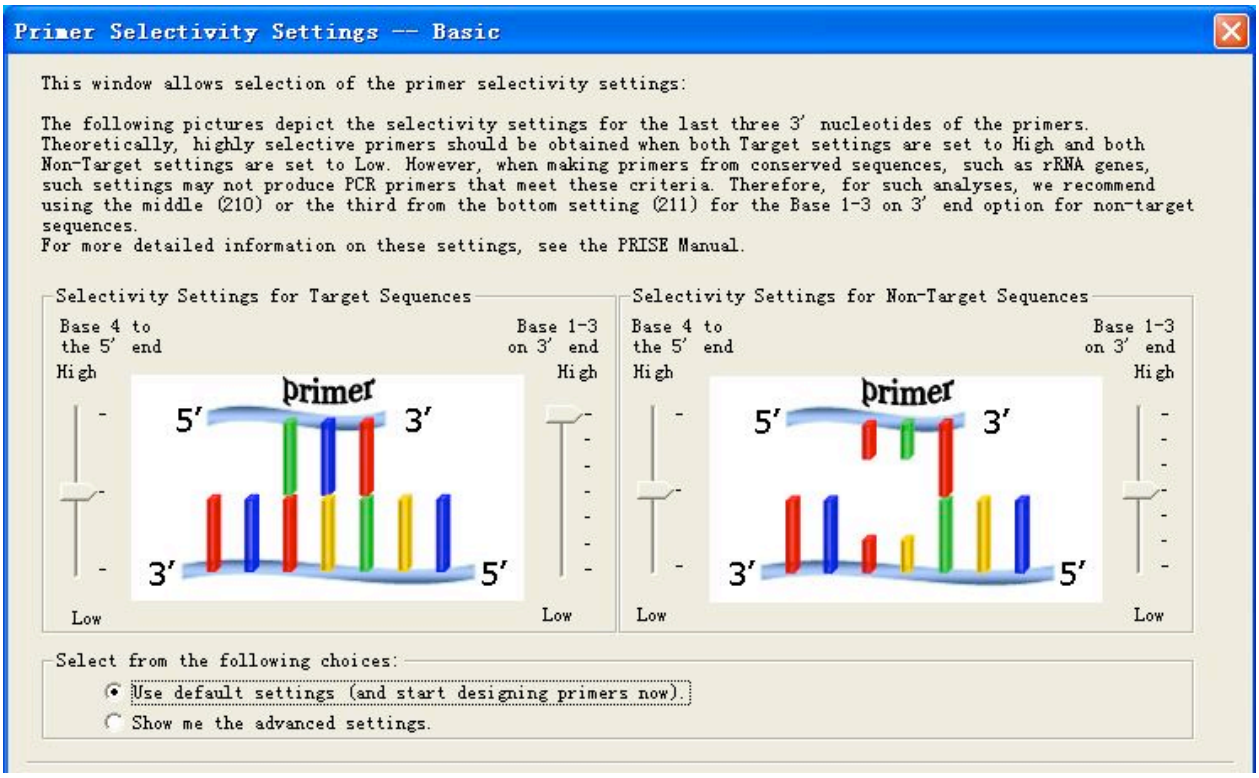
20. In our example, we chose the **Show me default settings** option. Click **Next**.

21. This opens the Standard Primer Properties Settings window. In our example, we used all of the default settings, except **PCR product size range from** was set to 75 to 250, and **Only consider top (single) candidate primers** was set to 100%. Click **Next**.

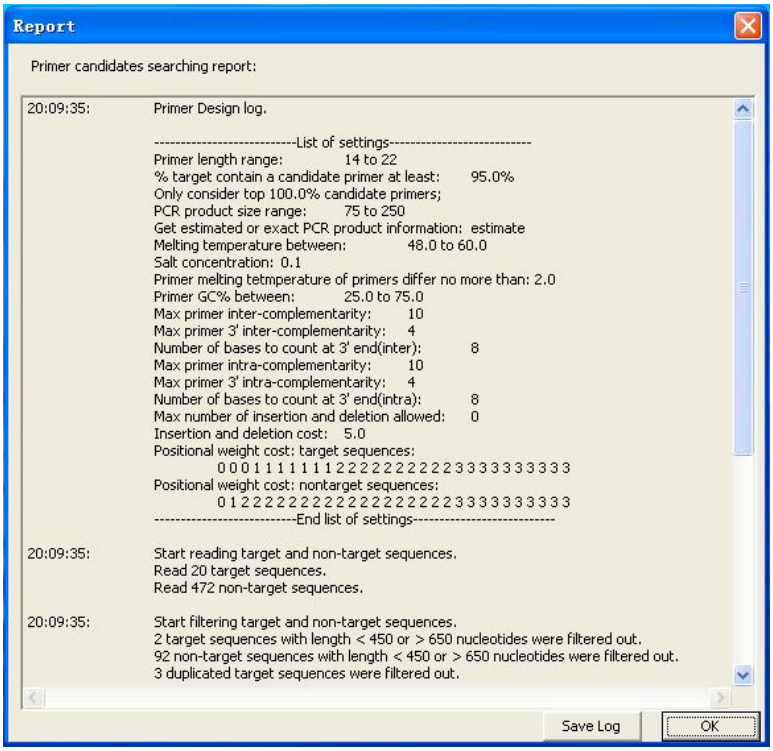


22. This opens the Primer Selectivity Settings window. We chose the medium and high settings for the target sequences **Base 4 to the 5' end** and **Base 1-3 on the 3' end**

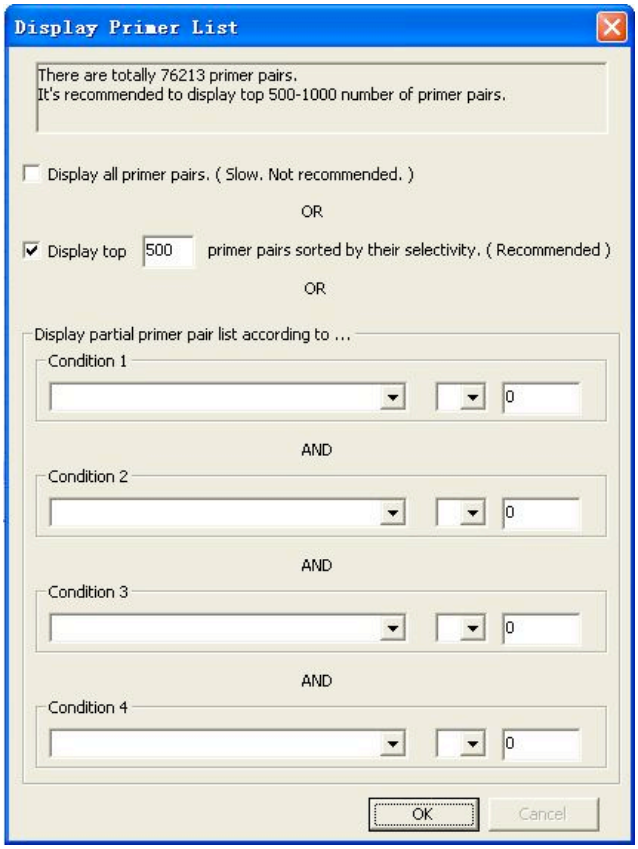
options, respectively, and medium settings for both options in non-target sequences. Click **Finish**.



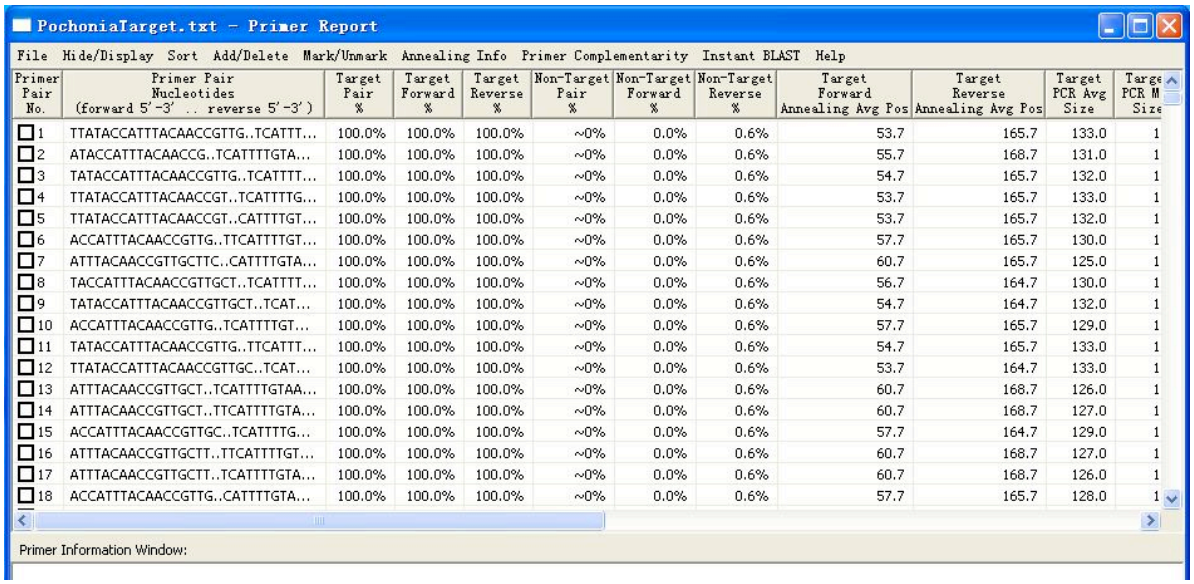
23. After the program finishes processing the data, which could take minutes to hours, depending on the size and complexity of the sequences in the target and non-target files, the user selected primer criteria and the capabilities of the computer, a primer design report appears. Information from this 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. Click **OK**.



24. The Display Primer List window appears. In our example, we set the **Display top primer pairs sorted by their selectivity** option to 500. Click **OK**.

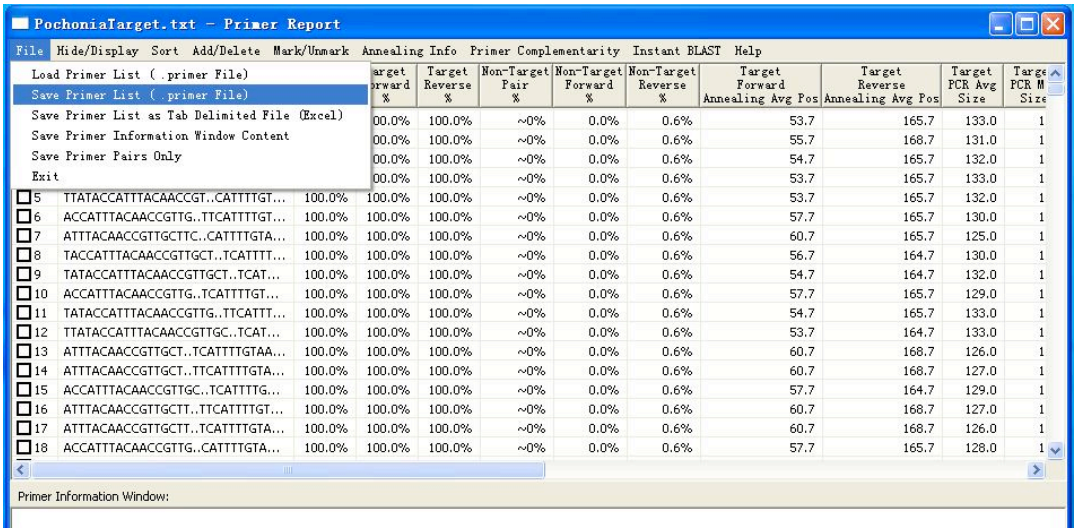


25. The Primer Report window containing the primer pairs appears. This window is named with the Target sequence file name followed by "- Primer Report."



Primer report. The Primer Report is a table that displays all of 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 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. This 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.

26. The primers and their properties can be saved by using the **Save Primer List** option from the **File** menu. In our example, we saved the primer list file as PochoniaPrimerList.



27. Primer lists can also be saved in a tab-delimited format, allowing the user to import the data into other programs such as spreadsheet software. This is accomplished by using the **Save Primer List as Tab Delimited File** option from the **File** menu. In our example, we saved this primer list file as PochoniaPrimerListTab.