A Software Kit *SEP* for Designing Primer for SNP Genotyping and Other Purposes.

Project Information

Title	**A Software Kit SEP for Designing Primer for SNP Genotyping and Other Purposes.**
Author	**Name:** Mr. Chan Ying Wai **Email:** bc_cyw@stu.ust.hk **Name:** Mr. Cheng Wu Albert **Email:** bc_cwa@stu.ust.hk

A Software Kit *SEP* for Designing Primer for SNP Genotyping and Other Purposes.

Y. W. Chan
W. A. Cheng

Abstract

SNPs are single base pair mutations which occur with a high frequency (>=1% in population) throughout the genome. Many SNPs have been found to be highly linked to some human diseases. Optimal primer sequence is essential for SNP genotyping which is a process for identifying and measuring the presence of SNPs. The primers for genotyping need to bind the target region flanking the SNP specifically and efficiently. Bioinformatics has evolved to comprise of a set of essential tools for biological research and has enhanced research processes enormously due to the high speed computation and vast memory capacity of computers. This software kit (SEP) is a bioinformatics application used to aid in designing the suitable primers for SNP experiments.

Keywords: Bioinformatics, primer, SNP, genotyping, BLAST, specificity

Introduction

All humans differ in genomic DNA sequences by about 0.1%. The most common type of variations is SNPs which are single base variations between different individuals (one SNP per 1000 bases on average). Many SNPs have been shown to be highly correlated to certain human diseases. These disease associations may be explained by the fact that some SNPs are present in the coding regions directly affecting protein functions, while some SNPs are physically close to the disease locus. In SNP genotyping, primer design for amplifying the target region flanking the SNP (Fig1) for identifying SNP is an essential step. Not only should the pair of primers fulfill the general rules (*see appendix 1*) such as primer length, melting temperature, etc, but also should bind specifically in the particular region but not in other regions in human genome. This can be achieved by first finding the potential primers by a primer design program, e.g. Primer3, followed by checking the specificity of each primer through Blastn in National Center for Biotechnology Information (NCBI) (thereafter referred to as BLAST). If this process is done manually, it would be intensively time-consuming and yet researchers may not be able to find the most suitable primers.

The present software tool, called SEP (***S***pecific ***E***fficient ***P***rimer), acts as an adaptor for the two programs mentioned above and automatically perform the above processes and, most importantly, provides scoring methods for selecting the most suitable primers in terms of both

specificity and efficiency. Potential forward and backward primers are ranked based on the BLAST results.

Fig1. The primers need to be able to amplify the green region flanking the SNP(black spot). The green region is about 70 bases at 5' and 3' of the SNP

The program

Primer Designer - SEP (Specific Efficient Primer)

Main Flow Chart

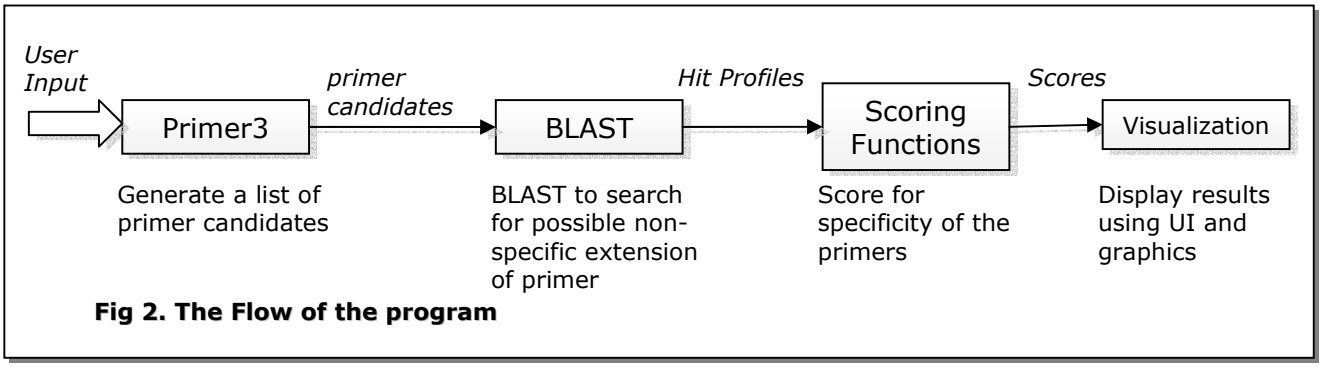

Fig 2. The Flow of the program

As illustrated in fig 2, SEP, upon receiving inputs from user about the locus of SNP and other parameters, call Primer3 to generate a list of efficient primer candidates. These primer candidates (or their corresponding binding sequence in template DNA) are BLASTed to search for occurrences of identical or similar sequences in the genome database. This process generates a hit profile for each primer which is passed through a scoring function and assigned a score. The results are then presented by visualization.

Main Specifications and Installation

System Requirements

- Linux Operating System
- GTK+ 1.2 or above
- X-Windows Environment
- PHP 4.0 or above(for the scoring plugins)

(The program has been successfully tested on Red Hat Linux 9.0. Running on other systems may require a rebuild of the program.)

Components

Primer Designer consists of mainly 4 components:

Main Components

- Scoring Plugins
- GUI Integration Component (GIC)

Scoring Plugins

The scoring plugins are written in PHP script which allows extension of new scoring functions. Read Appendix 4 for details.

External Dependencies/Components

- Primer3
- NCBI BLAST blastn

Primer3 and Blastn

Primer3 is a program for picking out primers for PCR reaction. It considers oligonucleotide melting temperature, size, GC content, and primer-dimer possibilities, PCR product size, positional constraints within the source sequence and miscellaneous other constraints. It is available at http://frodo.wi.mit.edu/primer3/primer3_code.html

Blastn is a program for comparing a query nucleotide sequence to all the sequences in a specified database. Each comparison is given a score reflecting the degree of similarity between the query and the sequence being compared. The higher the score, the greater the degree of similarity. If there are many sequences similar with the query, the query is lower in specificity. It is available at http://www.ncbi.nlm.nih.gov/BLAST

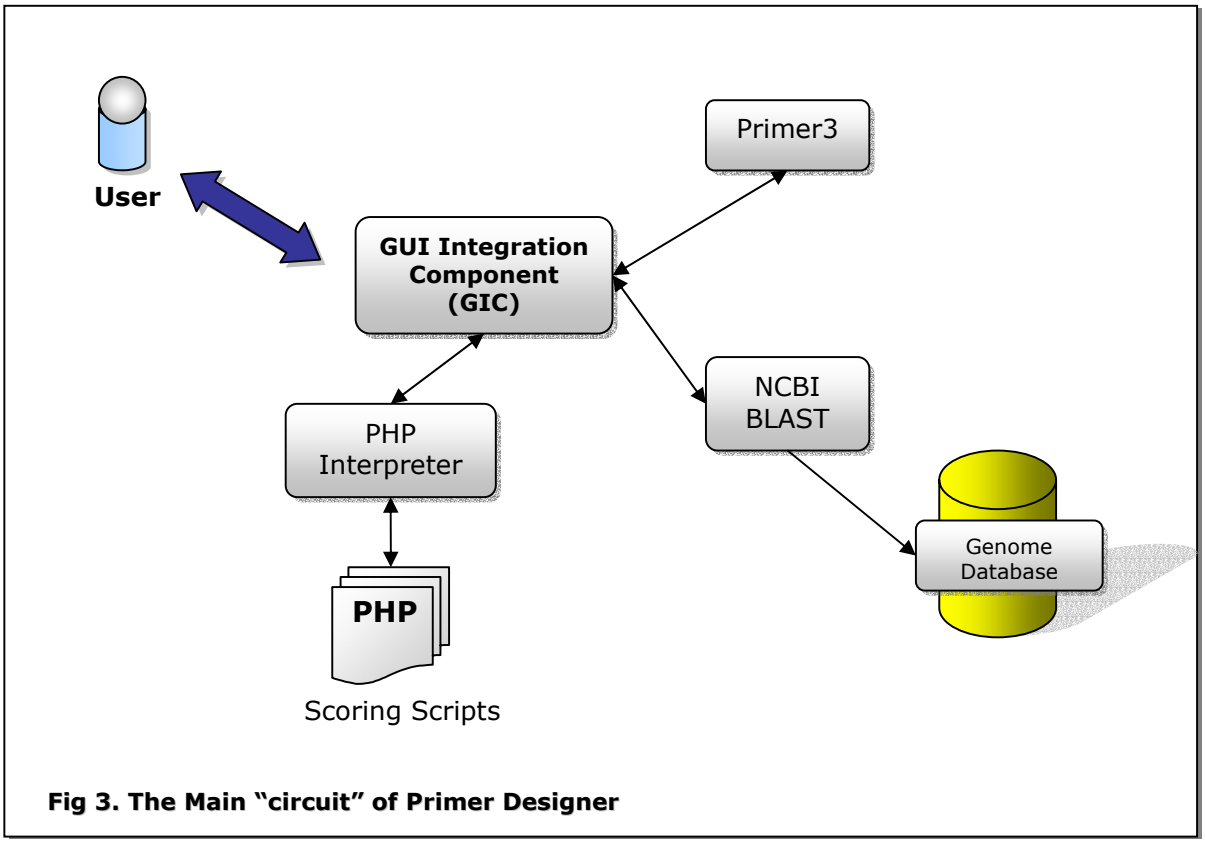

Fig 3. The Main "circuit" of Primer Designer

Distributed package

The distributed package is available in a zip file (file listing is attached as *appendix 2*.) containing the source files and binaries of The Scoring Plugins and GUI Integration Components. Primer3 and BLAST are external programs which are available online.

Sources of External Dependencies/Components

Primer3: http://frodo.wi.mit.edu/primer3/primer3_code.html [Available: 9 Jun 2004]
NCBI BLAST: http://www.ncbi.nlm.nih.gov/BLAST/ [Available: 9 Jun 2004]

For a listing of files in the package, *refer to Appendix 2*

Installation

The Binaries included in the package is built for Red Hat Linux 9.0. Other systems may require rebuild of the program prior to installation. (*see "Rebuilding the Program" below*). Proper running of the program depends on the correct configuration. The main configuration is primerDesigner.conf located in the same directory as the primerGUI executable.

Configuring the BLAST path and Primer3 path and configuration of Primer3 are discussed below in section "Configuring Primer Designer"

Rebuilding the Program

Go to the dev/integrate directory and run:

```
$ make
```

An executable primerGUI is generated.

The usage of program

Starting the program

The main program is initiated by running primerGUI executable.

User Interface

Main Window

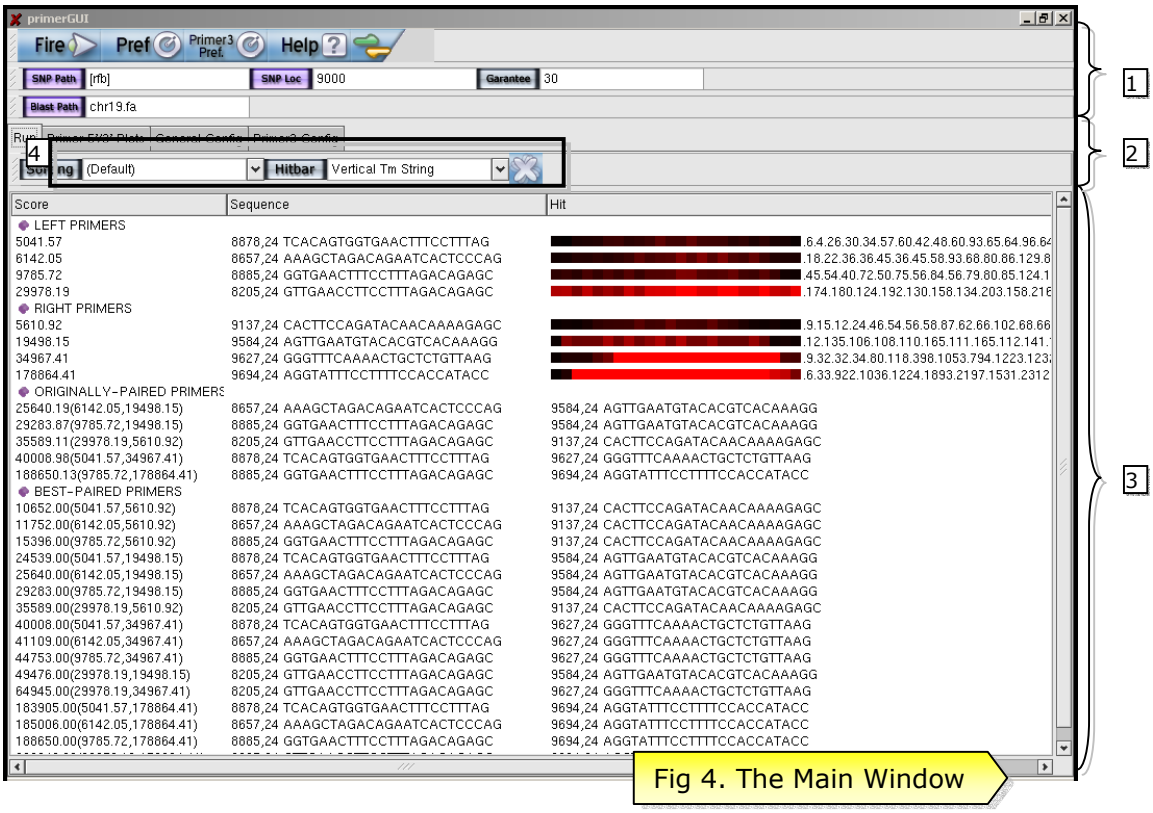

Fig 4. The Main Window

1	Main Toolbar	For the main manipulation of the program
2	Page Tabs	Changing between Pages

3	Page	The Main Workspace
4	Page Specific Toolbar	Manipulation of the Workspace

Main Toolbar

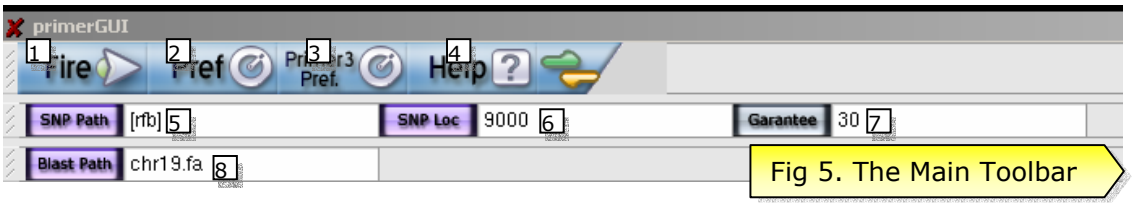

Fig 5. The Main Toolbar

1	Run the program
2	Change to "General Config" page
3	Change to "Primer3 Config" page
4	Open help documentations
5	The name of the file containing the SNP, [rfb] means the file is the same as the Blast path
6	Position of the SNP in the SNP file
7	The number of base before and after the SNP position that are needed to be amplified. If it is 30, it means 61 bases need to be flanked by two primers.
8	The name of the file containing the template DNA sequence.

Run Page

Fig 6. The Run Page

1	Choose 1 of the 5 different scoring algorithm (see the "Scoring Algorithm" part)
2	Choose 1 of the 2 Hitbar Mode (see the "Hitbar and score string" part)
3	Refresh the result
4	The potential left primers' score, position, length, sequence(5'->3'), Hitbar (3'->5') and the score string(3'->5')
5	The potential left primers' score, position, length, sequence(5'->3'), Hitbar (3'->5') and the score string(3'->5')
6	The originally-paired primers suggested by Primer3
7	Every left primer pairs with every right primer.
8	Hitbar (3'->5') and the score string(3'->5')

Primer 5'/3' Plots Page

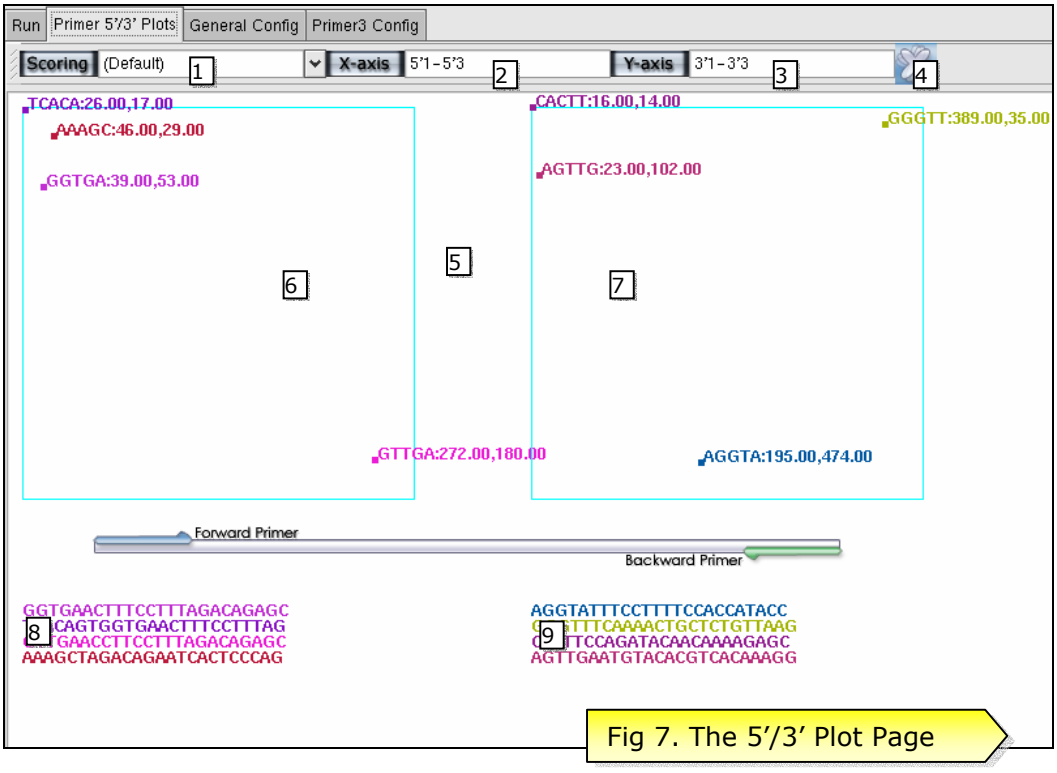

Fig 7. The 5'/3' Plot Page

1	Choose different scoring methods employed in the plots
2	Set the X-axis parameters of the plots In the format of [5/3]'x-[5/3]'y,…. e.g., 5'1-5'3 denotes a range of 1st base from 5' to 3rd base from 5'
3	Set the Y-axis parameters of the plots
4	Redraw the plots using new settings

5	Plot Area
6	Forward Primer Plot
7	Backward Primer Plot
8	Key for Forward Primers
9	Key for Backward Primers

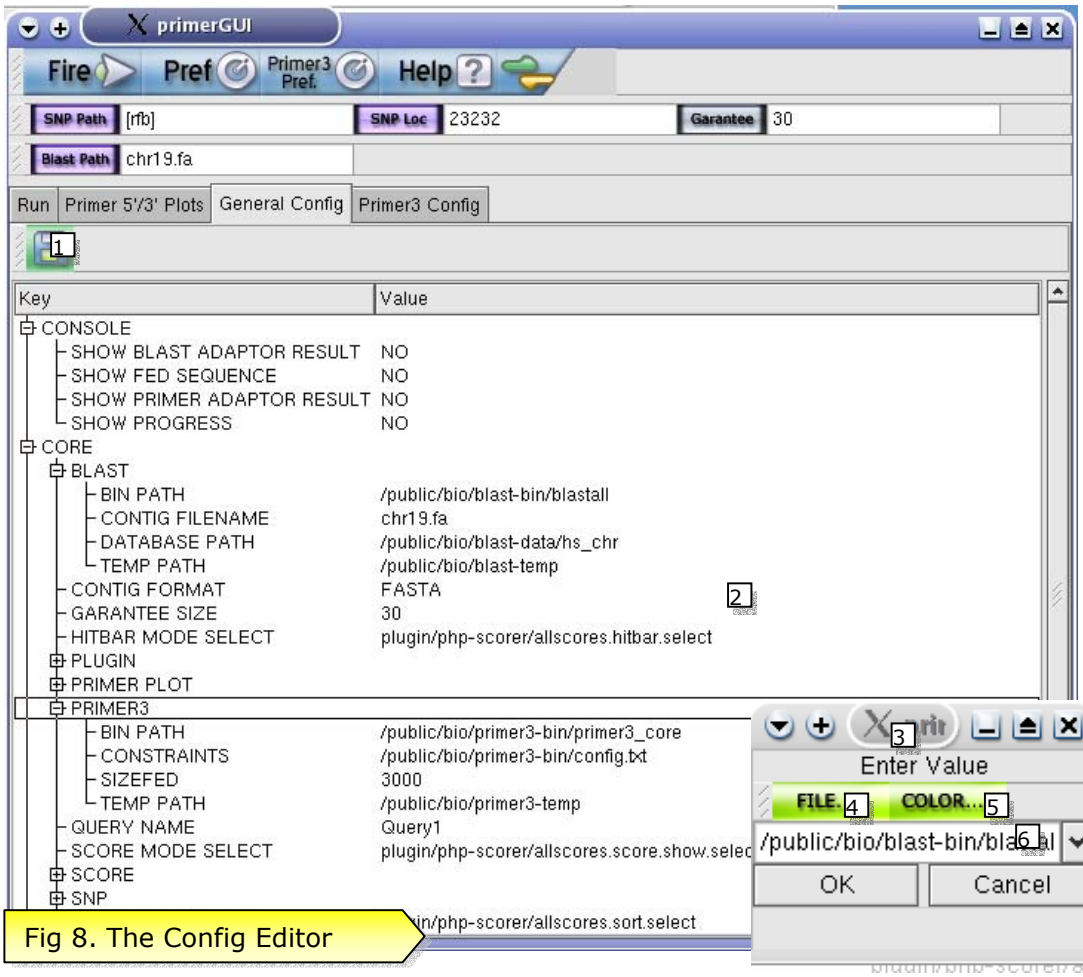

Fig 8. The Config Editor

1	Save the new settings
2	Setting Editor, Click entry to edit value
3	Value Edit Dialog
4	Open a file dialog
5	Open a color dialog
6	Edit Value directly

Interpretation of the Results

Scoring

The BLAST procedure generates a list of hits (similar sequence) for each primer. This hit profile is scored with different scoring method each focusing in different factors of primer annealing. Each forward and backward primer is given a score to indicate its specificity, the higher the score, the lower the specificity. The scoring methods, at least those distributed with the package, can be demarcated into 2 main types:

A One-score function assigns a real number score to a particular hit profile generated from BLAST. (fig 9)

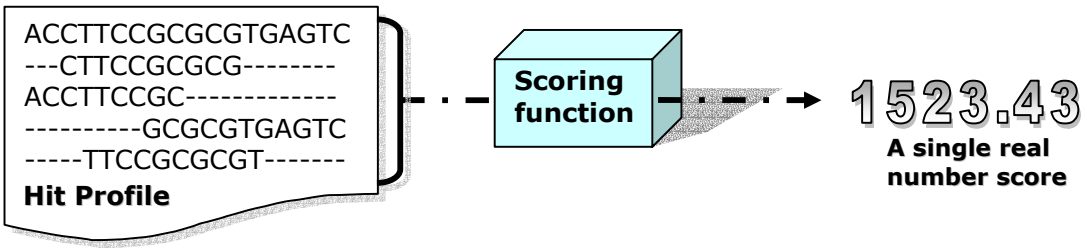

Fig 9, A One-score Function

A Sequence-score function assigns a real number to each position (base) of the primer and gives a sequence of real numbers as a score string for a particular hit profile generated from BLAST. The package includes several scoring algorithms. (fig 10)

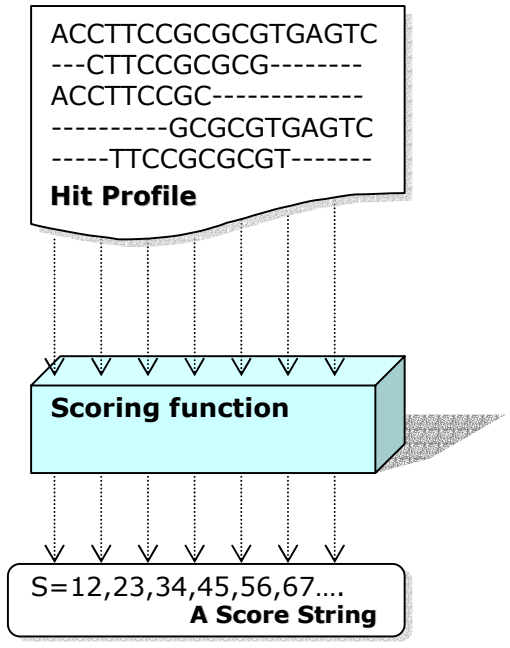

Fig 10, A Sequence-score Function

One-score Scoring Algorithms in the package

1. Scoring according to the number of hit sequences("Total Hit" option)
2. Scoring according to the number of total identical bases among the hits("Horizontal Hit Sum" option)
3. Scoring which give higher score to 3' identical bases than 5' identical bases among the hits("Horizontal Exp Score" option)
4. Scoring which similar to algorithm 3 but CG is two times higher score than AT("Horizontal Exp Tm Product Sum" option)
5. Scoring calculating the Tm between the primer sequence and the hits("Horizontal Tm Sum" option)

To illustrate the scoring algorithms, let's take an example:

Example 1

The hit profile (BLAST result) for a sequence 5'AATGCGT 3' is:

- - TGCGT
- - - GCGT
- - - GCG-

Algorithm 1 and 2

The score for example 1 by algorithm 1 is 3(no. of hits); by algorithm 2 is 12 (no. of identical bases).

Algorithm 3

For each hit, if the base is identical to the query,

Score := Score + B^p

Where B is a base number set by the user. P is the position of the identical base counting from 5' end.

So, if B=1.2, the score for the example 1 is

$= (1.2^3 + 1.2^4 + 1.2^5 + 1.2^6 + 1.2^7) + (1.2^4 + 1.2^5 + 1.2^6 + 1.2^7) + (1.2^4 + 1.2^5 + 1.2^6)$

=12.859

Algorithm 4

It is similar to algorithm 3. When the identical base is C or G, the B^{p_5} term in the equation will multiple by 2.

So, the score for the example 1 =

$(1.2^3 + 1.2^4 \times 2 + 1.2^5 \times 2 + 1.2^6 \times 2 + 1.2^7) + (1.2^4 \times 2 + 1.2^5 \times 2 + 1.2^6 \times 2 + 1.2^7) +$
$(1.2^4 \times 2 + 1.2^5 \times 2 + 1.2^6 \times 2)$

=20.407

Algorithm 5

T_m =2(A or T) +4(C or G), the score = score + T_m of each hit

-	-	T	G	C	G	T	0+0+2+4+4+4+2=16
-	-	-	G	C	G	T	0+0+0+4+4+4+2=14
-	-	-	G	C	G	-	0+0+0+4+4+4+0=12
					Total		**42**

Sequence-score Scoring Algorithms in the package

1. Scoring according to Tm Sum of all hits at each position of the primer("Vertical Tm String" option)
2. Scoring according to the number of total identical bases at each position of the primer among the hits("Vertical Hit String" option)

Referring back to example 1,

The hit profile (BLAST result) for a sequence 5'AATGCGT 3' is:

- - TGCGT
- - - GCGT
- - - GCG-

		-	-	T	G	C	G	T
		-	-	-	G	C	G	T
		-	-	-	G	C	G	-
Score Strings	**Algorithm 1**	0	0	1	3	3	3	2
	Algorithm 2	0	0	2	12	12	12	4

The score string for algorithm 1 is 0,0,1,3,3,3,2
The score string for algorithm 2 is 0,0,2,12,12,12,4

Hitbar

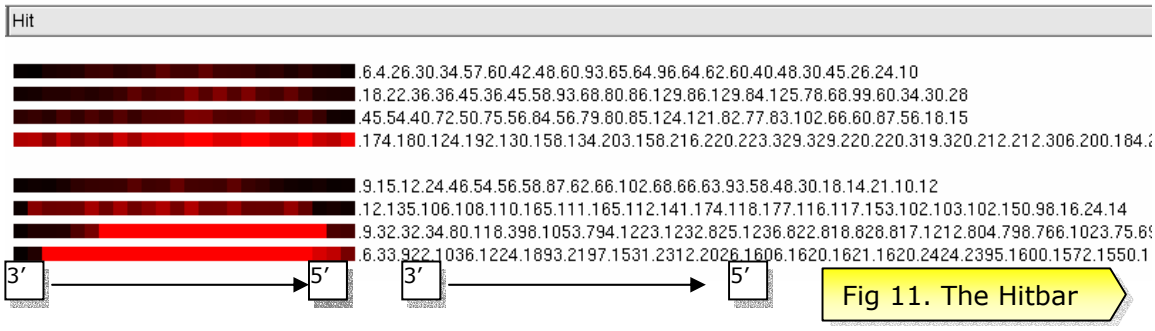

Fig 11. The Hitbar

The Hitbar is a visual representation of the score string generated by a sequence-score function. In the result in the run page, the hitbar appears as a bar with different intensity of red color across the length. Score of each base of the primers, from 3' to 5', are represented by different intensity of the color, the darker the color, the lower the score. Two modes of the Hitbar are distributed:

1. Vertical Hit String: The first number is the number of identical first 3' base between the primer and the hits, the second number is the number of identical second 3' base between

the primer and the hits and so on.

2. Vertical Tm String: Works similar to Hit String, but this time, when a A/T base is at a particular position, the score at that position is multiplied by 2, while G/C base multiply by 4

The numbers on the right of the bars are the corresponding score strings of the primer.

Primer 5'/3' Plot

The chance of extension of PCR primers at a particular location depends mainly on the ability of the 3' region of the primers to anneal to the DNA template at that location. However, the 5' region also contributes to the stability of the annealing by securing the primer to the template at some points. With the Primer 5'/3' plot, researches can elucidate these 2 factors at the same time and compare each primer candidates for their specificity. Each data point in the plot corresponds to a particular primer.

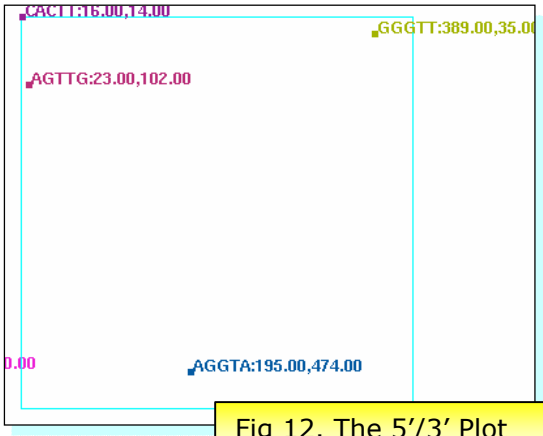

Fig 12. The 5'/3' Plot

The x-coordinate and the y-coordinate of each data point is calculated by summing the scores of the base in the range specified by the x-axis and y-axis parameters respectively.

$$\begin{cases} x = \sum_{i \in R_X} s_i \\ y = \sum_{j \in R_Y} s_j \end{cases}$$ **for score string S=$s_1s_2s_3.....s_n$; ranges R_x , R_y**

Configuring Primer Designer

There are two main sets of configuration of the program. One is primerDesigner.conf which defines settings for the GIC and config.txt (CORE_PRIMER3_CONSTRAINTS) for Primer3. The configuration can be edited in the text editor or in the preference pages.

(Refer to Apendix 3)

Setting up BLAST adaptor

CORE_BLAST_BIN PATH	The location of the blastall binary/executable e.g., /public/bio/blast-bin/blastall
CORE_BLAST_DATABASE PATH	The directory path of the BLAST databases

	e.g., /public/bio/blast-data/hs_chr
CORE_BLAST_TEMP PATH	A directory for the read/write of temporary file for BLAST: BLAST adaptor e.g., /public/bio/blast-temp

Setting up Primer3 adaptor

CORE_PRIMER3_BIN PATH	The location of the primer3 binary/executable e.g., /public/bio/primer3-bin/primer3_core
CORE_PRIMER3_CONSTRAINTS	The location of the config or constraints for Primer3 e.g., /public/bio/primer3-bin/config.txt
CORE_PRIMER3_TEMP PATH	A directory for the read/write of temporary file for Primer3: Primer3 adaptor e.g., /public/bio/primer3-temp

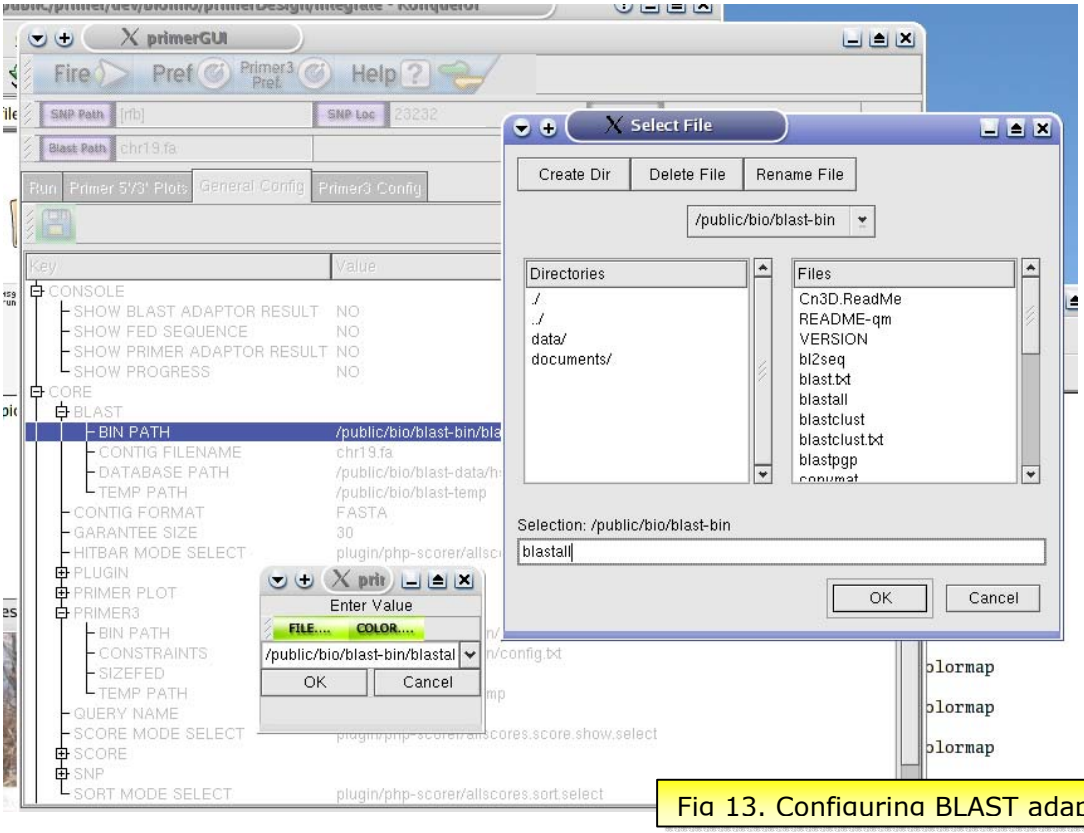

Fig 13. Configuring BLAST adaptor

References:

Abd-Elsalam. K. A. (2003) *Bioinformatics tools and guideline for PCR primer design*. African Journal of Biotechnology Vol.2 (5), pp.91-95

Duerinck K. F. (2001) *SNP – Single Nucleotide Polymorphism*
http://www.duerinck.com/snp.html [26 Mar 2004]

Ellipsis Biotherapeutics Corporation. (no date) *What is SNP Genotyping?* Available at:
http://www.ellipsisbio.com/what.is.html [2-June 2004]

National Center for Biotechnology Information (no data) *BLAST.* Available at:
http://www.ncbi.nlm.nih.gov/BLAST/ [2 June 2004]

Prezioso V. R.. (no date) *General Notes on Primer Design in PCR.* Available at:
http://www.brinkmann.com/PCR_appl_primer.asp [2 June 2004]

Robinson A. (1998) *An Introduction to the Use of COBRA for Molecular Biology* Available at:
http://industry.ebi.ac.uk/~alan/Presentations/SGM98/sld007.htm [26 Mar 2004]

Steve Rozen, Helen J. Skaletsky (1996,1997,1998) *Primer3*. Code available at http://www-genome.wi.mit.edu/genome_software/other/primer3.html [2-June 2004]

The International HapMap Consortium. (2003) *The International HapMap Project*. Nature Vol.426 pp. 789-796

Appendix 1

Criteria of selecting primer

1. Primer length

Both specificity and the temperature and time of annealing are at least partly dependent on the primer length. In general, the length is around 18-30 bases.

2. Melting temperature

The forward and backward primers should have similar melting temperature. The melting temperature should be around 55-60°C. The melting temperature can be calculated by following formula: Tm=2(A+T) + 4(C+G).

3. GC content

There should be around 45%-60% GC. PolyG or PolyC should be avoided to prevent non-specific annealing. PolyA and PolyT should also be avoided since it lowers the efficiency of amplification.

4. Specificity

Primers must be chosen so that they have a unique sequence within the template DNA.

5. Complementary Primer Sequence

Primers need to be designed with absolutely no intra-primer homology beyond 3 base pairs. If a primer has such a region of self-homology, "snap back" (partially double-stranded structures) can occur.

Inter-primer homology is also dangerous, it may cause primer dimer formation.

6. 3'-end Sequence

The 3' terminal position in PCR primers is essential for the control of mis-priming. A G or C residue at the 3' end of primers is needed. This "GC Clamp" helps to ensure correct binding at the 3' end due to the stronger hydrogen bonding of G/C residues. However, high GC content at 3' end (sticky end) will cause non-specific annealing.

Appendix 2

File Listing of the Package

bioinfo

bioinfo/primerDesign

bioinfo/primerDesign/integrate

bioinfo/primerDesign/integrate/basicgtk.cpp

bioinfo/primerDesign/integrate/gui-img

bioinfo/primerDesign/integrate/gui-img/.xvpics

bioinfo/primerDesign/integrate/gui-img/.xvpics/hitbar.xpm

bioinfo/primerDesign/integrate/gui-img/.xvpics/pcr.xpm

bioinfo/primerDesign/integrate/gui-img/.xvpics/reload.xpm

bioinfo/primerDesign/integrate/gui-img/.xvpics/tag-scoring.xpm

bioinfo/primerDesign/integrate/gui-img/.xvpics/tag-sorting.xpm

bioinfo/primerDesign/integrate/gui-img/.xvpics/tag-xaxis.xpm

bioinfo/primerDesign/integrate/gui-img/.xvpics/tag-yaxis.xpm

bioinfo/primerDesign/integrate/gui-img/btn_clear.bmp

bioinfo/primerDesign/integrate/gui-img/btn_clear.xpm

bioinfo/primerDesign/integrate/gui-img/btn_color.bmp

bioinfo/primerDesign/integrate/gui-img/btn_color.xpm

bioinfo/primerDesign/integrate/gui-img/btn_file.bmp

bioinfo/primerDesign/integrate/gui-img/btn_file.xpm

bioinfo/primerDesign/integrate/gui-img/btn_fire.bmp

bioinfo/primerDesign/integrate/gui-img/btn_fire.xpm

bioinfo/primerDesign/integrate/gui-img/btn_folder.bmp

bioinfo/primerDesign/integrate/gui-img/btn_folder.xpm

bioinfo/primerDesign/integrate/gui-img/btn_help.bmp

bioinfo/primerDesign/integrate/gui-img/btn_help.xpm

bioinfo/primerDesign/integrate/gui-img/btn_pr3pref.bmp

bioinfo/primerDesign/integrate/gui-img/btn_pr3pref.xpm

bioinfo/primerDesign/integrate/gui-img/btn_pref.bmp

bioinfo/primerDesign/integrate/gui-img/btn_pref.xpm

bioinfo/primerDesign/integrate/gui-img/bullet.xpm

bioinfo/primerDesign/integrate/gui-img/end_image.bmp

bioinfo/primerDesign/integrate/gui-img/end_image.xpm

bioinfo/primerDesign/integrate/gui-img/end_image-2.xpm

bioinfo/primerDesign/integrate/gui-img/end_mask.bmp

bioinfo/primerDesign/integrate/gui-img/end_mask.xpm

bioinfo/primerDesign/integrate/gui-img/hitbar.bmp

bioinfo/primerDesign/integrate/gui-img/hitbar.xpm

bioinfo/primerDesign/integrate/gui-img/pcr.bmp

bioinfo/primerDesign/integrate/gui-img/pcr.xpm

bioinfo/primerDesign/integrate/gui-img/reload.bmp

bioinfo/primerDesign/integrate/gui-img/reload.xpm

bioinfo/primerDesign/integrate/gui-img/tag_blast_path.bmp

bioinfo/primerDesign/integrate/gui-img/tag_blast_path.xpm

bioinfo/primerDesign/integrate/gui-img/tag_gar.bmp

bioinfo/primerDesign/integrate/gui-img/tag_gar.xpm

bioinfo/primerDesign/integrate/gui-img/tag_key.bmp

bioinfo/primerDesign/integrate/gui-img/tag_key.xpm

bioinfo/primerDesign/integrate/gui-img/tag_snp_loc.bmp

bioinfo/primerDesign/integrate/gui-img/tag_snp_loc.xpm

bioinfo/primerDesign/integrate/gui-img/tag_snp_path.bmp

bioinfo/primerDesign/integrate/gui-img/tag_snp_path.xpm

bioinfo/primerDesign/integrate/gui-img/tag_value.bmp

bioinfo/primerDesign/integrate/gui-img/tag_value.xpm

bioinfo/primerDesign/integrate/gui-img/tag-scoring.bmp

bioinfo/primerDesign/integrate/gui-img/tag-scoring.xpm

bioinfo/primerDesign/integrate/gui-img/tag-sorting.bmp

bioinfo/primerDesign/integrate/gui-img/tag-sorting.xpm

bioinfo/primerDesign/integrate/gui-img/tag-xaxis.bmp

bioinfo/primerDesign/integrate/gui-img/tag-xaxis.xpm

bioinfo/primerDesign/integrate/gui-img/tag-yaxis.bmp

bioinfo/primerDesign/integrate/gui-img/tag-yaxis.xpm

bioinfo/primerDesign/integrate/gui-img/tb_new.bmp

bioinfo/primerDesign/integrate/gui-img/tb_new.xpm

bioinfo/primerDesign/integrate/gui-img/tb_save.bmp

bioinfo/primerDesign/integrate/gui-img/tb_save.xpm

bioinfo/primerDesign/integrate/gui-img/tb_save_orig.bmp

bioinfo/primerDesign/integrate/gui-img/tb_save_orig.xpm

bioinfo/primerDesign/integrate/main.cpp

bioinfo/primerDesign/integrate/Makefile

bioinfo/primerDesign/integrate/pio.msg

bioinfo/primerDesign/integrate/pio.progress

bioinfo/primerDesign/integrate/pio.status

bioinfo/primerDesign/integrate/plugin

bioinfo/primerDesign/integrate/plugin/php-scorer

bioinfo/primerDesign/integrate/plugin/php-scorer/allscores.hitbar.select

bioinfo/primerDesign/integrate/plugin/php-scorer/allscores.plugin

bioinfo/primerDesign/integrate/plugin/php-scorer/allscores.score.show.select

bioinfo/primerDesign/integrate/plugin/php-scorer/allscores.sort.select

bioinfo/primerDesign/integrate/plugin/php-scorer/backup

bioinfo/primerDesign/integrate/plugin/php-scorer/backup/-all.php.backup

bioinfo/primerDesign/integrate/plugin/php-scorer/include

bioinfo/primerDesign/integrate/plugin/php-scorer/include/bas.inc

bioinfo/primerDesign/integrate/plugin/php-scorer/include/datio.inc

bioinfo/primerDesign/integrate/plugin/temp

bioinfo/primerDesign/integrate/plugin/temp/argv.in

bioinfo/primerDesign/integrate/plugin/temp/hits.in

bioinfo/primerDesign/integrate/plugin/temp/score.out

bioinfo/primerDesign/integrate/primerd

bioinfo/primerDesign/integrate/primerDesigner.conf

bioinfo/primerDesign/integrate/primerGUI

bioinfo/primerDesign/subPrograms

bioinfo/primerDesign/subPrograms/blast

bioinfo/primerDesign/subPrograms/blast/blast_main.h

bioinfo/primerDesign/subPrograms/interpreter

bioinfo/primerDesign/subPrograms/interpreter/bas.h

bioinfo/primerDesign/subPrograms/interpreter/configIO.h

bioinfo/primerDesign/subPrograms/interpreter/interpreter.h

bioinfo/primerDesign/subPrograms/interpreter/settingloader.h

bioinfo/primerDesign/subPrograms/primer3

bioinfo/primerDesign/subPrograms/primer3/biostring.h

bioinfo/primerDesign/subPrograms/primer3/config.txt

bioinfo/primerDesign/subPrograms/primer3/dictionary.h

bioinfo/primerDesign/subPrograms/primer3/input.txt

bioinfo/primerDesign/subPrograms/primer3/output.txt

bioinfo/primerDesign/subPrograms/primer3/primer_main.h

bioinfo/primerDesign/subPrograms/primer3/primertype.h

bioinfo/primerDesign/subPrograms/primer3/seq.txt

bioinfo/primerDesign/subPrograms/primer3/test-primer3.cpp

bioinfo/primerDesign/subPrograms/primer3/try.xpm

bioinfo/primerDesign/subPrograms/utilities

bioinfo/primerDesign/subPrograms/utilities/fastaLoader.h.

Appendix 3

Suggested Configurations of Primer3

Suggested Configurations of Primer3

PRIMER_SEQUENCE_ID=EXAMPLE
PRIMER_OPT_SIZE=22
PRIMER_MIN_SIZE=24
PRIMER_MAX_SIZE=27
PRIMER_MIN_TM=57.0
PRIMER_OPT_TM=60.0
PRIMER_MAX_TM=63.0
PRIMER_MAX_DIFF_TM=5.0
PRIMER_MIN_GC=40.0
PRIMER_OPT_GC_PERCENT=50.0
PRIMER_MAX_GC=60.0
PRIMER_MAX_POLY_X=4
PRIMER_SELF_END=3.00
PRIMER_SELF_ANY=6.00
PRIMER_GC_CLAMP=1
PRIMER_NUM_NS_ACCEPTED=0
PRIMER_PRODUCT_SIZE_RANGE=700-1000

(For the meaning of the input tags, please refer to the Readme file of the Primer3 available at http://frodo.wi.mit.edu/primer3/primer3_code.html
Two important input tags are missed in this configuration: they are SEQUENCE and the TARGET. They are generated when running the program. The value of SEQUENCE is the sequence in the SNP file. If the SNP location is X and guarantee is G, then the value of TARGER is X-G, 2*G+1.)

Appendix 4

Extending the Scoring Plugins.

To suit different needs of researchers, the scoring plugins components can be extended by editing the PHP scripts of the plugin. A plugin includes three files. E.g., for a plugin1, the 3 files are named:

- *plugin1*.plugin
- *plugin1*.hitbar.select
- *plugin1*.sort.select

plugin1.plugin contains the code for the scoring script. The two other files are the registry of the plugin for SEP to recognize.

For communicating with SEP, the PHP script requires including two files (bas.inc, datio.inc) located in the include directory.

```
require_once("include/bas.inc");
require_once("include/datio.inc");
```

The plugin first reads the hit profile from a file hits.in using a datio class. The hits are loaded into an array datio::arr by calling datio::loadArr("hits.in"). The hits are in 3' -> 5'. The plugin script then scores the hits by traversing the array scoring each hit and then storing the result into the map dat::map using the name of the score as the key. A one-score function should store a real number score while a sequence-score function should store the score string (from 3' to 5) as a string containing the real numbers separated by comma(,). The results are then passed to SEP by outputting the map by calling datio::saveMap("score.out")

A simple sample is listed below

```
<?php

//Lib for communicating with SEP
require_once("include/bas.inc");
require_once("include/datio.inc");

$dio=new datio;
$dio->loadArr("hits.in"); //load the hit profile
$dio->map["three_prime_score"]=0;
//initialize score three_prime_score to 0

foreach($dio->arr as $hit) //traverse hit array
{
  if($hit[0]=="a")
    $dio->map["three_prime_score"]++;
    //if the 3' first base is 'a', add one mark

}

$dio->saveMap("score.out");
//save result for SEP to read in
?>
```

This particular plugin script score the primer by adding one marks for a hit containing a 3' adenine (A).

To let SEP recognize the plugin and the scoring methods in the plugin, some values in primerDesigner should be changed:

CORE_SORT MODE SELECT=plugin/php-scorer/*plugin1*.sort.select
CORE_HITBAR MODE SELECT=plugin/php-scorer/ *plugin1*.hitbar.select
CORE_PLUGIN_SCORER_METHOD=plugin/php-scorer/ *plugin1*.plugin

The scoring method is then registered in *plugin1*.sort.select for one-score functions and *plugin1*.hitbar.select for sequence-score functions.

Each line in these files contain the information of one scoring function in the plugin.

(Default)=*three_primer_score*
Three Primer Score=*three_primer_score*

Where the string on the left of equal sign is the name of the scoring method displayed on the selection combos in the toolbar of the Run page and the Primer 5'/3' Plot page. (Fig *App4-1)* and the string on the right is the corresponding key of the score in the dat::map.

Fig App4-1. The scoring methods appearing in selection combos