BMC Bioinformatics

METHODOLOGY ARTICLE Open Access

Confronting two-pair primer design for enzyme-free SNP genotyping based on a genetic algorithm

Cheng-Hong Yang[1,2], Yu-Huei Cheng[1], Li-Yeh Chuang[3*], Hsueh-Wei Chang[4,5,6,7*]

Abstract

Background: Polymerase chain reaction with confronting two-pair primers (PCR-CTPP) method produces allele-specific DNA bands of different lengths by adding four designed primers and it achieves the single nucleotide polymorphism (SNP) genotyping by electrophoresis without further steps. It is a time- and cost-effective SNP genotyping method that has the advantage of simplicity. However, computation of feasible CTPP primers is still challenging.

Results: In this study, we propose a GA (genetic algorithm)-based method to design a feasible CTPP primer set to perform a reliable PCR experiment. The SLC6A4 gene was tested with 288 SNPs for dry dock experiments which indicated that the proposed algorithm provides CTPP primers satisfied most primer constraints. One SNP rs12449783 in the SLC6A4 gene was taken as an example for the genotyping experiments using electrophoresis which validated the GA-based design method as providing reliable CTPP primer sets for SNP genotyping.

Conclusions: The GA-based CTPP primer design method provides all forms of estimation for the common primer constraints of PCR-CTPP. The GA-CTPP program is implemented in JAVA and a user-friendly input interface is freely available at http://bio.kuas.edu.tw/ga-ctpp/.

Background

Genotyping is a common technique used in association studies of diseases and cancers. Although many high-throughput platforms of single nucleotide polymorphism (SNP) genotyping, such as SNP array [1] and real-time PCR using TaqMan probes [2], have been introduced, most laboratories still validate SNP or novel mutation by PCR-restriction fragment length polymorphism (RFLP) genotyping [3-6] because this method is inexpensive for small-scale genotyping. One shortcoming of PCR-RFLP is its long digestion time (usually in 2-3 hours) for restriction enzymes [7,8].

Recently, a restriction enzyme-free SNP genotyping technique called “PCR with confronting two-pair primers (PCR-CTPP)” was developed [9-12]. It has been applied successfully to at least 30 different SNP genotypings. For example, interleukin-1B (IL-1B) C-31T, interleukin-2 (IL-2) -330G, beta2-adrenergic receptor (beta2-AR) Gln27Glu, and aldehyde dehydrogenase 2 (ALDH2) were genotyped by PCR-CTPP for association studies with smoking behavior [13], pylori-induced gastric atrophy [14], severe coronary artery disease [15], and esophageal cancer risk [16], respectively. There is no doubt that the PCR-CTPP method is suitable and reliable for most cases of SNPs. This method considerably lowers the need to consume restriction enzymes. However, the criteria for the PCR-CTPP primers are only tolerant of a small difference in melting temperature ($T_{\text{m-diff}}$) between the four primers in the PCR-CTPP method [10]. Moreover, typical primer design constraints also need to be considered, such as primer length, primer length difference, PCR product length, GC proportion, melting temperature (T_m), GC clamp, dimer (including cross-dimers and self-dimers), hairpin structure, and specificity. Hence, the computational requirements needed to improve the primer design with PCR-CTTP are rather high.

* Correspondence: chuang@isu.edu.tw; changhw@kmu.edu.tw
[3]Department of Chemical Engineering & Institute of Biotechnology and Chemical Engineering, I-Shou University, Kaohsiung, Taiwan
[4]Department of Biomedical Science and Environmental Biology, Kaohsiung Medical University, Kaohsiung, Taiwan
Full list of author information is available at the end of the article

To design CTPP primers with many corresponding constraints, we introduce a genetic algorithm (GA) [17,18] to improve the design of CTPP primer sets. GA is a stochastic search algorithm modeled on the process of natural selection underlying biological evolution [19]. It constitutes a randomized search and an optimization technique that derives its working principle from natural genetics. Since GA computation is similar in nature to the evolution of the species, it best fits DNA behavior associated with SNP interaction [20] and general primer design [21]. The evolutionary computations involved, such as selection, crossover and mutation, are effective in achieving optimal solutions for many CTPP primer sets. After each run, chromosomes in a GA share information with each other and the superior solutions based on a fitness rule are refined from generation to generation. Therefore, CTPP primers obeying the typical primer design constraints described above can be mined.

Methods

Problem formulation

The CTPP primer design problem can be described as follows. Let T_D be a template DNA sequence, which is composed of nucleotide codes with an identified SNP. T_D is defined by:

$$\begin{aligned} T_D = \{B_i \mid\ & i \text{ is the index of DNA sequence}, \\ & 1 \leq\ i\ \leq \iota,\ \exists\ !\, B_i \in \text{IUPAC code of SNP}\} \end{aligned} \tag{1}$$

where B_i is the regular nucleotide code (A, T, C, or G) mixed with a single IUPAC code of SNP (M, R, W, S, Y, K, V, H, D, B or N) (is the existence and uniqueness). For the target SNP, we focused only on true SNPs described in dbSNP [22] of NCBI, i.e., deletion/insertion polymorphisms (DIPs) and multi-nucleotide polymorphisms (MNPs) are not included.

The CTPP primer design requires two pairs of short sequences which are constraining in T_D based on a defined SNP site as illustrated (Figure 1). The forward primer 1 (P_{f1}) is a short sense sequence in the upstream (5' end) of a defined SNP site for some distances; the reverse primer 1 (P_{r1}) is a short antisense sequence which contains a nucleotide (the minor allele of the defined SNP site) located at its 3' end; the forward primer 2 (P_{f2}) is a short sense sequence which contains a nucleotide (the major allele of the defined SNP site) located at its 3' end; and the reverse primer 2 (P_{r2}) is the antisense sequence in the upstream of a defined SNP site for some distances. These four primers are defined as follows:

$$P_{r1} = \{\overline{B_i} \mid i \text{ is the index of } T_D,\ R_{s1}\ \leq\ i\ \leq R_{e1}\} \tag{2}$$

$$P_{r1} = \{\overline{B_i} \mid i \text{ is the index of } T_D,\ R_{s1}\ \leq\ i\ \leq R_{e1}\} \tag{3}$$

$$P_{f2} = \{B_i \mid i \text{ is the index of } T_D,\ F_{s2}\ \leq\ i\ \leq F_{e2}\} \tag{4}$$

$$P_{r2} = \{\overline{B_i} \mid i \text{ is the index of } T_D,\ R_{s2}\ \leq\ i\ \leq R_{e2}\} \tag{5}$$

where both P_{f1}/P_{r1} and P_{f2}/P_{r2} are two primer pairs of PCR-CTPP. F_{s1} vs. F_{e1} and R_{s1} *vs.* R_{e1} indicate the start index *vs.* the end index of P_{f1} and P_{r1} in T_D, respectively. F_{s2} *vs.* F_{e2} and R_{s2} *vs.* R_{e2} indicate the start index *vs.* the end index of P_{f2} and P_{r2} in T_D, respectively. $\overline{B}_i$ is the complementary nucleotide of B_i, which is described in formula (1). For example, if B_i = 'A', then $\overline{B}_i$ = 'T'; if B_i = 'C', then $\overline{B}_i$ = 'G', and *vice versa*.

The target SNP site is defined at the end positions of P_{f2} and P_{r1}, which are indicated by the symbols F_{e2} and R_{e1}, respectively. As described in Figure 1, a vector (v) with F_{l1}, P_{l1}, R_{l1}, F_{l2}, P_{l2} and R_{l2} is essential to design the CTPP primer sets. This vector is defined as follows:

$$P_v = (\, F_{l1}, P_{l1}, R_{l1}, F_{l2}, P_{l2}, R_{l2} \,) \tag{6}$$

F_{l1}, P_{l1}, R_{l1}, F_{l2}, P_{l2} and R_{l2} represent the lengths of forward primer 1, PCR product between P_{f1} and P_{r1}, reverse primer 1, forward primer 2, PCR product between P_{f2} and P_{r2} and reverse primer 2, respectively. Consequently, the forward and reverse primers can be acquired from P_v, which is the prototype of a chromosome in GA and is used to perform evolutionary computations as described in the following sections.

Definition of the fitness function

The regular primer design constraints are used as values to design a fitness function to minimize the fitness value. The fitness function is defined as follows:

$$\begin{aligned} \mathit{Fitness}(P_v) = 3 * (& \mathit{Len}_{\mathit{diff}}(P_v) + GC_{\mathit{proportion}}(P_v) \\ & + GC_{\mathit{clamp}}(P_v)) + 10 * (\mathit{dimer}(P_v) \\ & + \mathit{hairpin}(\, P_v\,) + \mathit{specificity}(\, P_v\,)) \\ & + 50 * (\mathrm{Tm}(P_v)\ + \mathrm{Tm}_{\mathit{diff}}(P_v)) \\ & + 100 * \mathit{vg}\_\mathrm{Tm}_{\mathit{diff}}(P_v) \\ & + 60 * \mathit{PCR}\mathrm{len}(\, P_v\,) \end{aligned} \tag{7}$$

The weights (3, 10, 50, 60 and 100) of the fitness function are applied to estimate the importance of the primer constraints. These weights are set according to the experiential conditions for PCR-CTPP. They also accept adjustment based on the user experimental requirements.

Primer length

The feasible primer length for a PCR experiment is set between 16 and 28 bp. For longer primers, the T_m is

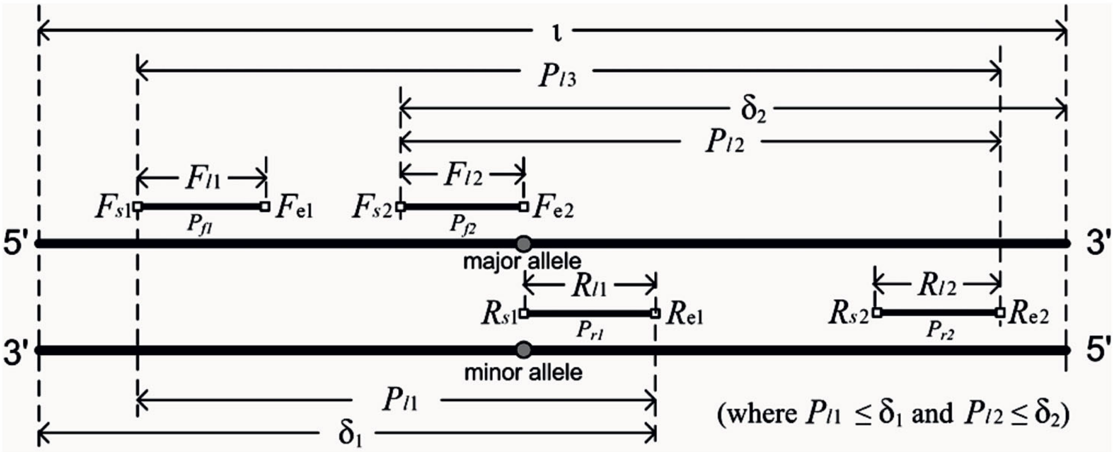

Figure 1 Parameters of the template DNA and the CTPP primer set. Symbols indicate: *F*: Forward primer; *R*: Reverse primer; *s*: Start nucleotide position; *e*: End nucleotide position; *P*: Length of PCR product using a primer set (*F/R*); *l*: Length of primer or product;ι: Length of DNA template; δ_1: Length from the R_{s1} end to downstream of template DNA; δ_2: Length from F_{s2} to the downstream end of template DNA.

increased whereas the T_m of relatively short primers is decreased. Accordingly, primers which are neither too long nor too short are suitable. We have limited the random values of F_{l1}, R_{l1}, F_{l2} and R_{l2} in an appropriate range; therefore, the primer length estimation is not considered to be joined to the fitness function.

A length difference (Len_{diff}) of less than or equal to 3 bp between the F_{l1}/R_{l1}, F_{l2}/R_{l2}, and F_{l1}/R_{l2} primer sets is considered optimal. The primer length difference function is defined as follows:

$$Len_{diff}(P_v) = \begin{cases} defect\_value = 3 \\ \text{if ABS } (F_{l1} - R_{l1}) \le 3, \\ \quad \text{then } defect\_value - 1 \\ \text{if ABS } (F_{l2} - R_{l2}) \le 3, \\ \quad \text{then } defect\_value - 1 \\ \text{if ABS } (F_{l1} - R_{l2}) \le 3, \\ \quad \text{then } defect\_value - 1 \\ \text{return } defect\_value \end{cases} \tag{8}$$

where $Len_{diff}(P_v)$ has a maximal fitness value of 3; the fitness value is decreased when the length difference between a primer pair is less than or equal to 3 bp. ABS represents the absolute value.

GC content and GC clamp

The function $GC\%(P)$ is proposed to represent the ratio of G and C nucleotides appearing in a primer:

$$GC\%(P) = \frac{G_{number}(P) + C_{number}(P)}{|P|} \tag{9}$$

where P represents a primer and $|P|$ represents the length of primer P; $G_{number}(P)$ and $C_{number}(P)$ represent the numbers of the nucleotides G and C, respectively.

In general primer design, the typical GC proportion constraint is set between 40% and 60%. However, the designed CTPP primers contain the target SNP to limit the range of the GC proportion. To relax this constraint, the constraint of GC proportion in a primer is adjusted to between 20% and 80%. Function $GC_{proportion}(P_v)$ is proposed with a maximal fitness value of 4 to lead the $GC\%(P)$ of CTPP primers corresponding to this constraint:

$$GC_{proportion}(P_v) = \begin{cases} defect\_value = 4 \\ \text{if } 20 \le GC\%(P_{f1}) \le 80, \\ \quad \text{then } defect\_value - 1 \\ \text{if } 20 \le GC\%(P_{r1}) \le 80, \\ \quad \text{then } defect\_value - 1 \\ \text{if } 20 \le GC\%(P_{f2}) \le 80, \\ \quad \text{then } defect\_value - 1 \\ \text{if } 20 \le GC\%(P_{r2}) \le 80, \\ \quad \text{then } defect\_value - 1 \\ \text{return } defect\_value \end{cases} \tag{10}$$

To meet the presence of G or C nucleotides at the 3' terminal of a primer to ensure a tight localized hybridization bond, the function $GC_{clamp}(P_v)$ is proposed with the maximal fitness value of 4 as follows:

$$GC_{clamp}(P_v) = \begin{cases} defect\_value = 4 \\ \text{if } 3' \text{ end of } P_{f1} \text{ is 'G' or 'C',} \\ \quad \text{then } defect\_value - 1 \\ \text{if } 3' \text{ end of } P_{r1} \text{ is 'G' or 'C',} \\ \quad \text{then } defect\_value - 1 \\ \text{if } 3' \text{ end of } P_{f2} \text{ is 'G' or 'C',} \\ \quad \text{then } defect\_value - 1 \\ \text{if } 3' \text{ end of } P_{r2} \text{ is 'G' or 'C',} \\ \quad \text{then } defect\_value - 1 \\ \text{return } defect\_value \end{cases} \tag{11}$$

Melting temperature

The melting temperature (T_m) for each CTPP primer must be considered carefully for PCR experiments. Here, we do not use the rough estimate 2 × (#A + #T) + 4 × (#G + #C), but a more elaborate equation containing the ionic strength, G and C content and length of the primer is concerned. The T_m calculation formula for a primer is described as follows:

$$\begin{aligned}\mathrm{Tm}_{BM}(P) &= 81.5 + 16.6 * (\log_{10}[\mathrm{Na}^+]) \\ &\quad + 0.41 * (\mathrm{GC}\%) - 675 / |P|\end{aligned} \tag{12}$$

where P represents a primer and $|P|$ represents the length of primer P; Na^+ is the molar salt concentration. The suffix BM represents the formula which was proposed by Bolton and McCarthy [23].

The function $Tm(P_v)$ is proposed to confine a CTPP primer set ranging from 45°C and 62°C with the maximal fitness value of 4:

$$\mathrm{Tm}(P_v) = \begin{cases} defect\_value = 4 \\ \text{if } 45 \le \mathrm{Tm}_{BM}(Pf1) \le 62, \\ \quad \text{then } defect\_value - 1 \\ \text{if } 45 \le \mathrm{Tm}_{BM}(Pr1) \le 62, \\ \quad \text{then } defect\_value - 1 \\ \text{if } 45 \le \mathrm{Tm}_{BM}(Pf2) \le 62, \\ \quad \text{then } defect\_value - 1 \\ \text{if } 45 \le \mathrm{Tm}_{BM}(Pr2) \le 62, \\ \quad \text{then } defect\_value - 1 \\ \text{return } defect\_value \end{cases} \tag{13}$$

Similar T_m between a primer pair is important to perform experiment in the same tube. The function $Tm_{diff}(P_v)$ is proposed with the maximal fitness value of 3 to guide the difference of the melting temperatures to less than or equal to 1°C:

$$\mathrm{Tm}_{diff}(P_v) = \begin{cases} defect\_value = 3 \\ \text{if ABS } (\mathrm{Tm}_{BM}(Pf1) - \mathrm{Tm}_{B}(Pr1)) \le 1, \\ \quad \text{then } defect\_value - 1 \\ \text{if ABS } (\mathrm{Tm}_{BM}(Pf2) - \mathrm{Tm}_{B}(Pr2)) \le 1, \\ \quad \text{then } defect\_value - 1 \\ \text{if ABS } (\mathrm{Tm}_{BM}(Pf1) - \mathrm{Tm}_{B}(Pr2)) \le 1, \\ \quad \text{then } defect\_value - 1 \\ \text{return } defect\_value \end{cases} \tag{14}$$

In order to balance the T_m values among a CTPP primers, function $Avg\_Tm_{diff}(P_v)$ is proposed to calculate the average T_m difference:

$$\begin{aligned}Avg\_\mathrm{Tm}_{diff}(P_v) = [&\mathrm{ABS}(\mathrm{Tm}_{BM}(Pf1) - \mathrm{Tm}_{BM}(Pr1)) \\ &+ \mathrm{ABS}(\mathrm{Tm}_{BM}(Pf2) - \mathrm{Tm}_{BM}(Pr2)) \\ &+ \mathrm{ABS}(\mathrm{Tm}_{BM}(Pf1) - \mathrm{Tm}_{BM}(Pr2))] / 3\end{aligned} \tag{15}$$

Dimer and hairpin

Primer dimers (annealing of two primers), such as cross-dimers (a forward primer and a reverse primer anneal to each other) and self-dimers (two forward primers or two reverse primers anneal to each other) must also be avoided. To check for the occurrence of primer dimers, the function $dimer(P_v)$ is proposed with the maximal fitness value of 10:

$$dimer(P_v) = \begin{cases} defect\_value \; = \; 10 \\ \text{if}(Pf1 \text{ and } Pr1) \text{ do not form a cross -dimer,} \\ \quad \text{then } defect\_value - 1 \\ \text{if}(Pf2 \text{ and } Pr2) \text{ do not form a cross -dimer,} \\ \quad \text{then } defect\_value - 1 \\ \text{if}(Pf2 \text{ and } Pr1) \text{ do not form a cross -dimer,} \\ \quad \text{then } defect\_value - 1 \\ \text{if}(Pf1 \text{ and } Pr2) \text{ do not form a cross-dimer,} \\ \quad \text{then } defect\_value - 1 \\ \text{if}(Pf1 \text{ and } Pf2) \text{ do not form a cross-dimer,} \\ \quad \text{then } defect\_value - 1 \\ \text{if}(Pr1 \text{ and } Pr1) \text{ do not form a cross-dimer,} \\ \quad \text{then } defect\_value - 1 \\ \text{if}(Pf1 \text{ and } Pf1) \text{ do not form a self-dimer,} \\ \quad \text{then } defect\_value - 1 \\ \text{if}(Pr1 \text{ and } Pr1) \text{ do not form a self-dimer,} \\ \quad \text{then } defect\_value - 1 \\ \text{if}(Pf2 \text{ and } Pf2) \text{ do not form a self-dimer,} \\ \quad \text{then } defect\_value - 1 \\ \text{if}(Pr2 \text{ and } Pr2) \text{ do not form a self-dimer,} \\ \quad \text{then } defect\_value - 1 \\ \text{return } defect\_value \end{cases} \tag{16}$$

The hairpin check is also implemented to avoid annealing due to the secondary structure of a primer. To check for the presence of a hairpin structure in CTPP primers, the function $hairpin(P_v)$ is proposed with the maximal fitness value of 4 as follows:

$$hairpin(P_v) = \begin{cases} defect\_value = 4 \\ \text{if } Pf1 \text{ do not form a hairpin,} \\ \quad \text{then } defect\_value - 1 \\ \text{if } Pr1 \text{ do not form a hairpin,} \\ \quad \text{then } defect\_value - 1 \\ \text{if } Pf2 \text{ do not form a hairpin,} \\ \quad \text{then } defect\_value - 1 \\ \text{if } Pr2 \text{ do not form a hairpin,} \\ \quad \text{then } defect\_value - 1 \\ \text{return } defect\_value \end{cases} \tag{17}$$

Specificity

Subsequently, the function $specificity(P_v)$ is proposed to check for repetition of each CTPP primer in the template

DNA sequence to ensure its specificity. The PCR experiment may fail when a designed primer is not sequence-specific (i.e. it reappears in the genome). The fitness value of the primers (P_{f1}, P_{r1}, P_{f2} or P_{r2}) appearing in T_D is evaluated using a binary value, i.e., when the primers repeatedly appear in T_D, the *specificity*(P_v) is defined as 1; or else the *specificity*(P_v) is defined as 0.

PCR product length

Finally, the function $PCR\mathrm{len}(P_v)$ is proposed with the maximal fitness value of 7 to calculate the appropriate lengths of the PCR products. Three ratios - i.e. ratio1, ratio2 and ratio3 - are introduced to the function $PCR\mathrm{len}(P_v)$ representing P_{l1}, P_{l2} and P_{l3}, respectively. The minimum length of PCR products needs to be greater than 100 bp.

$$PCR\mathrm{len}(P_v) = \begin{cases} defect\_value = 7 \\ \text{if } P_{l1} \ > \ 100\text{, then } defect\_value - 1 \\ \text{if } P_{l2} \ > \ 100\text{, then } defect\_value - 1 \\ \text{if } P_{l3} \ > \ 100\text{, then } defect\_value - 1 \\ \text{if } P_{l1} \text{ corresponds ratio1,} \\ \quad \text{then } defect\_value - 1 \\ \text{if } P_{l2} \text{ corresponds ratio2,} \\ \quad \text{then } defect\_value - 1 \\ \text{if } P_{l3} \text{ corresponds ratio3,} \\ \quad \text{then } defect\_value - 1 \\ \text{if all } P_{l1} \text{ and } P_{l2} \text{ and } P_{l3} \text{ correspond their ratios,} \\ \quad \text{then } defect\_value - 1 \\ \text{return } defect\_value \end{cases} \tag{18}$$

Algorithm

The proposed algorithm consists of five processes: (1) random initial population, (2) fitness evaluation, (3) selection, crossover, and mutation, (4) replacement, and (5) judgment on termination conditions. Figure 2 shows the flowchart of GA-based CTPP primer design. The five processes are described below:

(1) Random initial population

To start the algorithm, chromosomes P_v = (F$_{l1}$, P_{l1}, R_{l1}, F_{l2}, P_{l2}, R_{l2}) of particular number are randomly generated for an initial population without duplicates. F_{l1}, R_{l1}, F_{l2} and R_{l2} are randomly generated between the minimum and the maximum of the primer length constraint. The minimum and maximum lengths of the primer length constraint are set to 16 and 28 bp, respectively. The PCR product lengths, P_{l1} and P_{l2} are randomly generated between 100 bp and δ_1, and between 100 bp and δ_2, respectively. (δ_1 and δ_2 are the maximum tolerant PCR product lengths of P_{l1} and P_{l2} shown in Figure 1)

(2) Fitness evaluation

The fitness value in the fitness function is used to ascertain that an individual chromosome is either good or bad. We use formula (7) to evaluate the fitness values of all chromosomes in the population for related chromosomal operations later.

(3) Selection, crossover, and mutation

In GA, the processes for evolutionary computation include selection, crossover and mutation. Here, random selection is applied to select two chromosomes from the population. The two selected chromosomes are processed by the crossover operation. Uniform crossover is used to implement the crossover operation. The flowchart of the crossover process is shown in Figure 3, and an example of the crossover operation is shown in Figure 4. One-point mutation is applied in the proposed GA. The mutation process flowchart is shown in Figure 5, and an example of the mutation operation is shown in Figure 6.

(4) Replacement

After the evolutionary computation processes have been implemented, the two worst chromosomes in a population are replaced by the new offsprings, and the process is repeated in the next generation.

(5) Judgment on termination conditions

Once an optimal solution of chromosomes (i.e. the fitness value is 0) or the maximum number of iterations is achieved, the GA is terminated.

Other important criteria for CTPP primer design

There is already one alternative nucleotide in the defined SNP for the CTPP primers P_{f2} and P_{r1}. If a further SNP exists in any other CTPP primers, such as P_{f1} and P_{r2}, the T_{m} between all primers is more dynamic and is difficult to optimize. Accordingly, we avoid designing primers P_{f1} and P_{r2} with extra SNPs, i.e., all the designed primers for P_{f1} and P_{r2} including SNPs are eliminated without further processing.

Results

Dry dock experiments

The environment

The proposed algorithm was run using Xeon(TM) CPU 3.20 GHz × 2 and 2 GB RAM under the Microsoft Windows XP SP2 and JAVA 5.0 platforms.

Parameter settings

Four main parameters are set for the proposed algorithm, i.e. the number of iterations (generations), the population size, the probability of crossover and the probability of mutation. The respective values were 1000, 50, 0.6 and 0.001; the values are based on DeJong and Spears' parameter settings [24]. The population size was then set at 1000 and the other parameters were held constant; only the population size was increased (see Discussion for detail). The PCR product length was set to three ratios (ratios 1, 2, and 3) at 8, 13, and 20, respectively, which allowed for the distinct separation of PCR bands via gel electrophoresis. These ratios were chosen based on our previously conducted PCR experiments.

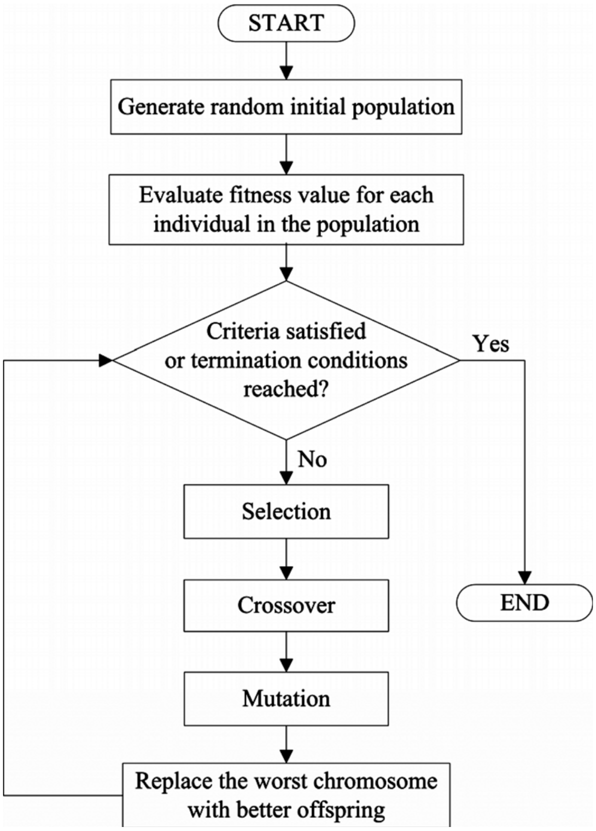
START
Generate random initial population
Evaluate fitness value for each individual in the population
Criteria satisfied or termination conditions reached?
Yes
No
Selection
Crossover
Mutation
Replace the worst chromosome with better offspring
END

Figure 2 Flowchart of the GA-based CTPP primer design. At first, a random initial population is generated and then all fitness values of all chromosomes in the population are calculated by the fitness function. A judgment on termination conditions is carried out, and if the termination conditions are reached then the algorithm will be finished, or else the algorithm proceeds with the following processes. Selection, crossover and mutation operations are performed and finally the worst chromosomes are replaced by the better chromosomes. The procedure is repeated in the next iteration until the termination conditions are reached.

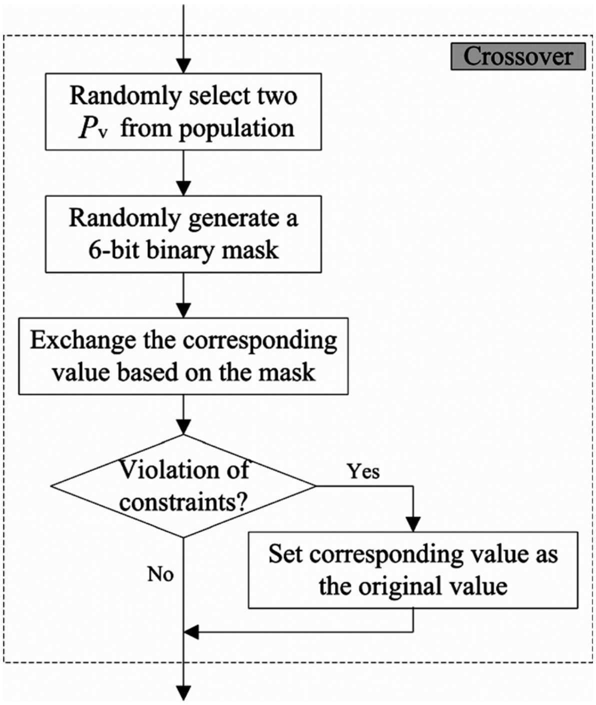
Crossover
Randomly select two P_v from population
Randomly generate a 6-bit binary mask
Exchange the corresponding value based on the mask
Violation of constraints?
Yes
No
Set corresponding value as the original value

Figure 3 Crossover flowchart for CTPP primer design. Two P_v from the population are randomly selected for crossover. At first, a six bit binary mask is generated and indicates which variables need to be exchanged. All exchanged variables are checked for violation of a constraint. If a constriction is violated, the exchanged variables will be restored, or else the process goes on to the next step.

Results for GA-based CTPP primer design in the example of the SLC6A4 gene

A point mutation in the SLC6A4 gene was recently identified and shown to be associated with autism spectrum disorders [25], psychosis [26], and bipolar [27] patients. The SLC6A4 gene is the member 4 for a solute carrier family 6 (neurotransmitter transporter, serotonin). The common constraints for CTPP primer design were used, including a flanking length of 500 bp, primer length of between 16~28 bp, GC% between 20~80%, primer T_m between 45 and 62°C, difference of CTPP primer T_m of less than 1°C, product length larger than 100 bp, and clearly separated PCR bands in gel electrophoresis. The SNPs for SLC6A4 gene (288 SNPs) are used as an example in this study which excluded the deletion/insertion polymorphisms (DIPs) and multi-nucleotide polymorphisms (MNPs). These SNPs were retrieved with 500 bp flanking length (at both sides of SNP) from SNP-Flankplus (http://bio.kuas.edu.tw/snp-flankplus/) [28]; the reference cluster IDs of these SNPs are shown in http://bio.kuas.edu.tw/ga-ctpp/dataset.jsp.

The entire CTPP primer set results are provided at http://bio.kuas.edu.tw/ga-ctpp/appendix.jsp and statistics of the results based on the common constraints of GA-based CTPP primers are shown in Table 1. For the 288 SNPs, there are 1152 representative primers for GC%, GC clamp, T_m, hairpin and specificity criteria (288 × 4 = 1152; a CTPP primer set contains four primers for a SNP). For the length difference, T_m difference and the product length criteria, there are only 864 (288 × 3 = 864; a CTPP primer set which contains four primers to lead three product lengths). The number of dimer is 2880 (each primer may form self-dimer and two different primers may form cross-dimer in a CTPP primer set). The primer lengths are all between 16 and 28 bp. In ID1 as shown in Table 1, the parameter settings are based on DeJong and Spears, the designed primer length difference violated the parameter settings for 215 primers (215/864). Most of the primer length differences were between 0 and 5 bp (data not shown). For GC%, 30 primers were less than 20%, 25 primers were more than 80% and the

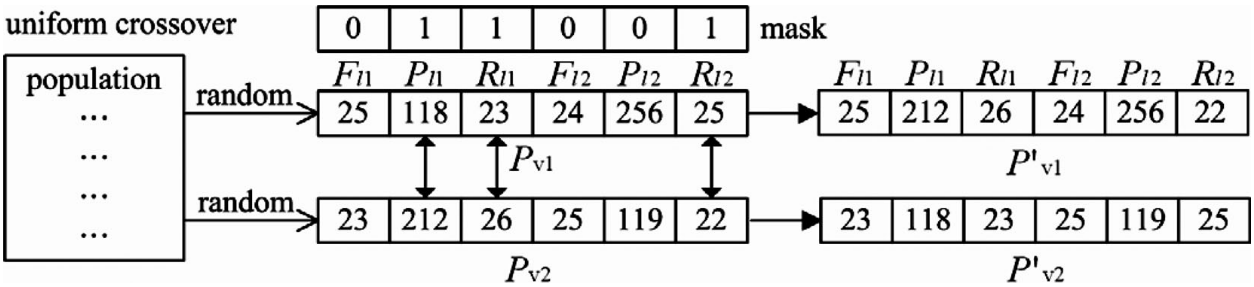

Figure 4 An example of a crossover operation for CTPP primer design. First, randomly select two P_v from the population, for example P_{v1} = (25, 118, 23, 24, 256, 25) and P_{v2} = (23, 212, 26, 25, 119, 22). Then randomly generate a mask of 6 binary bits i.e., 011001, and based on this mask, exchange the value of P_{l1}, R_{l1} and R_{l2}. Finally, the new offsprings P'_{v1} = (25, 212, 26, 24, 256, 22) and P'_{v2} = (23, 118, 23, 25, 119, 25) are generated.

GC% distribution was mainly between 30% and 70% (data not shown). Approximately half of the primers (645/1152) did not satisfy the GC clamp criteria. Most of the designed primers also satisfied T_m (998/1152); however, only approximately 23.6% (204/864) of the primer pairs satisfied the T_m difference criteria. The criteria for product length were satisfied in approximately 71.2% (615/864) of the designed primer pairs. For the criteria for primer dimer, hairpin and specificity, few primers were problematic (128/2880, 162/1152 and 35/1152, respectively).

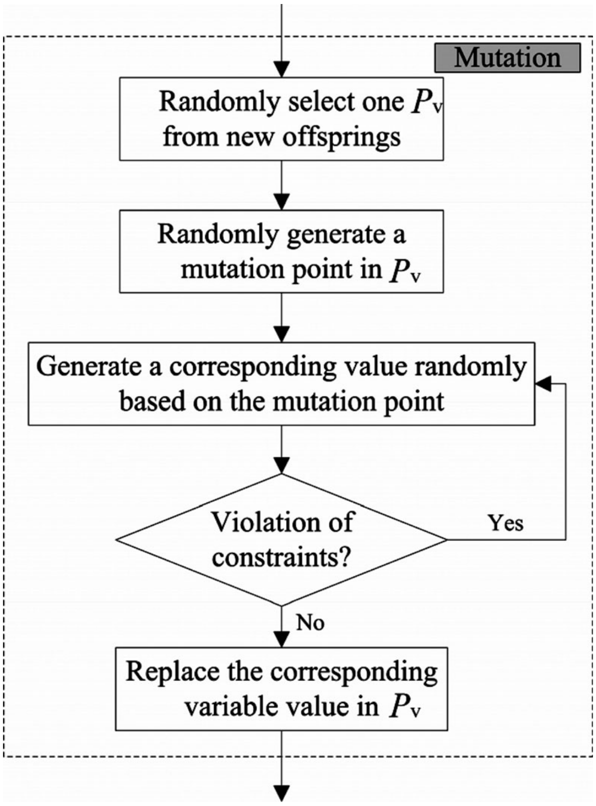
Mutation
Randomly select one Pv from new offsprings
Randomly generate a mutation point in Pv
Generate a corresponding value randomly based on the mutation point
Violation of constraints?
Yes
No
Replace the corresponding variable value in Pv

Figure 5 Mutation flowchart for CTPP primer design. When a mutation operation is performed, a mutation point in P_v will be randomly selected and a corresponding value generated. Then the variable will be checked for violation of a constriction. If a constriction is violated, a random point is reselected and the variable is regenerated; otherwise the original value is replaced by the variable and the process continues in the next step.

Genotyping experiment

Materials

One SNP rs12449783 in the SLC6A4 gene was taken as an example for the genotyping test. Three DNA samples with three known different SNP genotypes for rs12449783 were used to demonstrate the effectiveness of the GA-based CTPP primer design.

Validation of SNP genotyping by GA-based PCR-CTPP and TaqMan probe

The designed CTPP primer set for rs12449783 in the SLC6A4 gene is given in Table 2. DNA samples were added to the PCR reaction mixture (10 μl) containing 1 μl of 10× PCR buffer, 0.3 μl of 50 mM $MgCl_2$, 0.2 μl of 10 mM dNTPs, 0.6 μl of DMSO, 0.14 μl of 5 U Platinum Taq enzyme (Invitogen corp.), 0.12 μl of 350 μg/ml primer mix

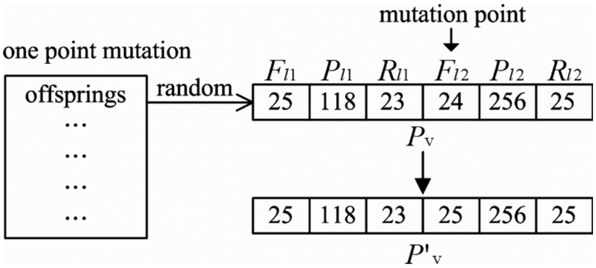

Figure 6 An example of a mutation operation for CTPP primer design. At the beginning, a P_v = (25, 118, 23, 24, 256, 25) is randomly selected from the offsprings. Then one mutation point of the position F_{l1}, P_{l1}, R_{l1}, F_{l2}, P_{l2} or R_{l2} is randomly generated. In our example the position the mutation point is F_{l2}. Then a random value between 16 and 28 of F_{l2} is generated. This value replaces the corresponding variable value in P_v. Finally, the new offspring P'_v = (25, 118, 23, 25, 256, 25) is generated.

Table 1 The statistics of the designed CTPP primers showing how many primers satisfied the common constraints for SNPs of the SLC6A4 gene*

		Constraints								
ID		primer length difference	GC%	GC clamp	T_m	T_m difference	product length	dimer	hairpin	specificity
1	Results	649/864	1107/1152	645/1152	998/1152	204/864	615/864	2752/2880	990/1152	1117/1152
2		684/864	1115/1152	653/1152	1103/1152	512/864	608/864	2760/2880	969/1152	1121/1152

* The results of ID 1 and ID2 are the parameter settings based on DeJong and Spears but ID2 is modified with an increased population size to 1000.

(1:1), and 7.64 μl of DNA in water. Primer mixtures of several combinations were used: P_{f1}/P_{r1} and P_{f2}/P_{r2}, and P_{f1}/P_{r2} (Figure 1). The used PCR program had the following paramaters: 94°C (4 min); 49 cycles at 94°C for (30 s), 50°C for (20 s), 72°C (20 s); and 72°C (5 min). PCR products were separated by electrophoresis on a 1.5% regular agarose gel followed by ethidium bromide staining.

In the principle of PCR-CTPP, two paired primers (four primers; P_{f1}, P_{r2}, P_{r1}, and P_{r2}) should be placed in one tube. Accordingly, when it is succeeded, two DNA bands are amplified for the heterozygotes and three DNA bands for the heterozygotes. As shown in Figure 7A, the samples were performed in PCR-CTTP using four CTPP primers (P_{f1}, P_{r2}, P_{r1}, and P_{r2}) within one tube (lanes 4, 8, and 12). Moreover, we also performed these PCR reactions separately for each set of CTPP primers ($P_{f2}P_{r2}$ for lanes 1, 5, and 9; $P_{f1}P_{r1}$ for lanes 2, 6, and 10; and $P_{f1}P_{r2}$ for lanes 3, 7, and 11, respectively), to clearly validate the performance for each combination of the CTPP primers ($P_{f2}P_{r2}$; $P_{f1}P_{r1}$; or $P_{f1}P_{r2}$) designed by our proposed computational algorithm.

As in Figure 7A for either the four CTPP primers or three different sets of two combined CTPP primers, the samples with AA genotype showed AA-negative for 228-bp ($P_{f1}P_{r1}$) and AA-positive for 105- ($P_{f2}P_{r2}$) and 294-bp ($P_{f1}P_{r2}$); the samples with CC genotype showed CC-negative for 105-bp and CC-positive for 228- and 294-bp; and the samples with AC genotype showed AC-positive for 228-, 105- and 294-bp. Accordingly, the bands of the A- and C-alleles-specific PCR amplifications were successfully demonstrated for AA/AC (105-bp) and CC/AC (228-bp), respectively. The internal positive PCR controls for all alleles (i.e., A and C) were confirmed. Therefore, it is clearly demonstrated that our proposed algorithm is able to provide the validated primers for PCR-CTPP.

Using the same samples in Figure 7A, the CTPP results were examined further using the TaqMan probes which were ABI no. hcv_7911133, VIC-/FAM-labels for ACACATAGAAAGTTACAGACTAGCA[A/C] GTCTGGTATTCATAAAGAATTGTGA, respectively. The TaqMan probe program was performed by a 2 step-protocol built-in the ABI real-time system (50°C, 2 min (stage 1, 1 cycle), 95°C, 10 min (stage 2, 1 cycle), 95°C, 15 sec (stage 3, 40 cycles), and 60°C, 1 min.). As shown in Figure 7B, the samples with AA, AC, and CC genotypes for rs12449783 in PCR-CTPP results (Figure 7A) were confirmed to be the same using the TaqMan probe assay. Therefore, the primer information for PCR-CTPP designed by our proposed algorithm was well proved.

Discussion

To date, many primer design approaches have been proposed, such as dynamic programming [29], parthenogenetic algorithm MG-PGA [30], greedy algorithm [31], heuristic algorithm [32], genetic algorithm [21,33], memetic algorithm [28] and any others. However, most of these methods do not provide the SNP genotyping function. In contrast, we reported the brief idea of the GA-CTPP method for primer design of SNP genotyping in the IEEE BIBE 2009 conference [34]. The differences between them and the improvements of the algorithm proposed in this study are described in the additional file 1. In this study, we present an improved GA-based algorithm which has been shown to be a robust search and optimization method for a number of practical problems, especially for highly complex problems for SNP genotyping with the CTPP primers design function. We had used electrophoresis to validate the reliability of the GA-based CTPP primer design method.

Table 2 The information for the designed CTPP primers for rs12449783 of the SLC6A4 gene

	CTPP primer set for rs12449783	Length (bp)	GC%	T_m	$T_{m\text{-diff}}$ (°C)	Product size (bp)
P_{f1}:	GATTATTAGTAGTTTCTGCA	20	30	50.96		
P_{r1}:	TTCTTTATGAATACCAGAC**G**	20	35	51.37	P_{f1}/P_{r1}: 0.41	P_{f1}/P_{r1}: 228
P_{f2}:	AGAAAGTTACAGACTAGCA**A**	20	30	51.37	P_{f2}/P_{r2}: 0.41	P_{f2}/P_{r2}: 105
P_{r2}:	ATGTTTAATCTCTGAGAAGA	20	35	51.37	P_{f1}/P_{r2}: 0	P_{f1}/P_{r2}: 294

The bold font represents a SNP.

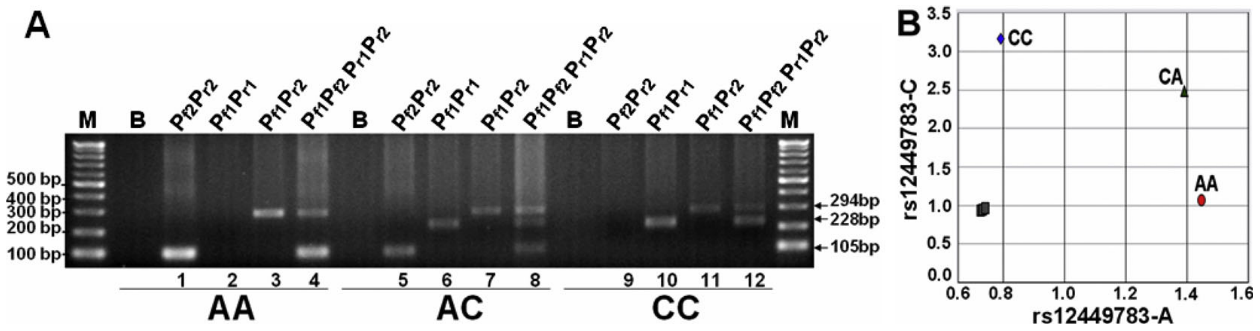

Figure 7 The SNP genotyping results of rs12449783 by (A) PCR-CTPP running in gel electrophoresis and (B) by TaqMan probe. (A) M and B indicate the 100 bp markers and blank, respectively. CC, AA, and AC indicate the control DNAs with known genotypes for CC, AA, and AC, respectively. The P_{f2}/P_{r2} (lanes 1, 5, and 9) and P_{f1}/P_{r1} (lanes 2, 6, and 10) primer pairs are designed to specifically amplify the A- and C-containing alleles for genotypes AA and AC *vs*. genotypes CC and AC, respectively. The P_{f1}/P_{r2} (lanes 3, 7, and 11) primer pair is designed to amplify the A- or C-containing alleles (CC, AA, and AC). For PCR-CTPP using four primers in one tube, it is marked with "P_{f1}, P_{f2}, P_{r1}, P_{r2}" for lanes 4, 8, and 12. The relative positions for P_{f1}, P_{f2}, P_{r1}, and P_{r2} are shown in Figure 1.

Influence of annealing temperatures

In PCR-CTPP, the designed annealing temperatures of primers are more important than in PCR-RFLP. When the T_m value is similar among the four primers of PCR-CTPP, the PCR competition between all possible DNA products is balanced [10]. When the annealing temperature is low, the PCR reactions are non-specific, leading to incorrect heterozygous genotyping. Therefore, a competitive or specific amplification is important to correctly genotype SNPs using PCR-CTPP. This problem is resolved by computationally finding similar T_m values for the four CTPP primers and by experimentally adjusting the annealing temperature for the PCR [10,35]. The GA used in this study to design the PCR-CTPP primers improves the efficiency by finding almost identical T_m values for the four primers. The *in silico* testing of the proposed GA-based PCR-CTPP primer design showed it to fit the T_m constraint to the primers reliably (Table 1).

Typical primer design constraints concerned

Since the T_m is important to our proposed GA-based PCR-CTPP method, further basic research is required to determine other factors to improve this automated PCR-CTPP system. This study is also concerned with the typical primer design constraints, such as primer length, primer length difference, GC proportion, GC clamp, dimer of primer pair (including cross-dimers and self-dimers), hairpin, PCR product length and specificity etc. as described in the Methods section.

Effect of population size

Dejong and Spears' parameter settings are the standard for most GAs, and for this reason, they were used in the present study. Typically, crossover is usually applied at more than or equal to the rate of 0.6, and the mutation rate is equal to 0.001 [24]. However, the population size 50 of Dejong and Spears's parameter settings is too small to provide the necessary sampling accuracy for the design of CTPP primer sets. Consequently, we tested the population size for 100, 200, 300, 400, 500, 600, 700, 800, 900, and 1000 to evaluate the primer design performance. When the population size was increased to 1000, it provides the more accurate sampling (as shown in the additional file 2). As shown in Table 1 ID2, the number of primers that satisfy the constraints was increased to 9.11% and 35.65% for the T_m and the T_m difference constraint, respectively. For other constraints, the numbers of satisfied constraints were almost similar. The results demonstrate that the increased population size can aid in the search for more feasible CTPP primer sets.

Available for GA-CTPP

The GA-CTPP can be accessed at http://bio.kuas.edu.tw/ga-ctpp/. GA-CTPP designs appropriate two-pair primers to genotype a defined SNP based on the parameter settings of DeJong and Spears. Parameter settings or the primer design conditions can be changed individually by users based on their requirements. When the input sequence contains multiple SNPs, the first SNP will be taken as the defined SNP to design CTPP primer sets. GA-CTPP reports an optimal set of confronting two-pair primers through a text file that records all information of the CTPP primer set for genotyping of the defined SNP.

Conclusions

PCR-CTPP may replace PCR-RFLP because the restriction enzyme digestion step can be omitted, resulting in lower costs and shorter genotyping times [10]; however, the PCR-CTPP is less developed for its computational tool providing PCR-CTPP primer design. A novel strategy for PCR-CTPP primer design has been introduced in this paper and the freely available web server implemented with this method was also constructed. With experimental validation, our proposed GA-based method is a useful algorithm to design feasible CTPP primers and it conforms to most of the PCR-CTPP constraints.

Availability and requirements

Project name: GA-CTPP: Confronting two-pair primer design using genetic algorithm.

Project home page: http://bio.kuas.edu.tw/ga-ctpp/.

Operating system(s): Operating systems free with web browser.

Programming language: Java.

Other requirements: Java 1.5.

License: none for academic users. For any restrictions regarding the use by non-academics please contact the corresponding author.

Additional material

Additional file 1: 'The differences between our previous publication in BIBE 2009 conference [34] and this study'.

Additional file 2: 'The performances for primer design using our proposed GA-CTPP algorithm between different population sizes of Dejong and Spears's parameter settings'.

Acknowledgements

This work is partly supported by the National Science Council in Taiwan under grant NSC97-2622-E-151-008-CC2, NSC96-2221-E-214-050-MY3, NSC96-2311-B037-002, NSC96-2622-E214-004-CC3, NSC98-2221-E-151-040-, NSC 98-2622-E-151-001-CC2 and the funds KMU-EM-99-1.4 and DOH99-TD-C-111-002. We also thank for the technical support from Mr. Ming-Tz Tsai.

Author details

[1]Department of Electronic Engineering, National Kaohsiung University of Applied Sciences, Kaohsiung, Taiwan. [2]Department of Network Systems, Toko University, Chiayi, Taiwan. [3]Department of Chemical Engineering & Institute of Biotechnology and Chemical Engineering, I-Shou University, Kaohsiung, Taiwan. [4]Department of Biomedical Science and Environmental Biology, Kaohsiung Medical University, Kaohsiung, Taiwan. [5]Graduate Institute of Natural Products, College of Pharmacy, Kaohsiung Medical University, Kaohsiung, Taiwan. [6]Center of Excellence for Environmental Medicine, Kaohsiung Medical University, Kaohsiung, Taiwan. [7]Cancer Center, Kaohsiung Medical University Hospital, Kaohsiung Medical University, Kaohsiung, Taiwan.

Authors' contributions

C-HY coordinated and oversaw this study. Y-HC participated in the design of the algorithm, and wrote the program and the manuscript. L-YC provided the biochemistry background and introduced the bioinformatics needed for primer design. H-WC performed and verified the PCR experiment, and modified the manuscript. All authors read and approved the final manuscript.

Received: 9 January 2010 Accepted: 13 October 2010
Published: 13 October 2010

References

1. Jasmine F, Ahsan H, Andrulis IL, John EM, Chang-Claude J, Kibriya MG: **Whole-genome amplification enables accurate genotyping for microarray-based high-density single nucleotide polymorphism array.** *Cancer Epidemiol Biomarkers Prev* 2008, **17(12)**:3499-3508.
2. Hui L, DelMonte T, Ranade K: **Genotyping using the TaqMan assay.** *Curr Protoc Hum Genet* 2008, **Chapter 2(Unit 2)**:10.
3. Chang HW, Yang CH, Chang PL, Cheng YH, Chuang LY: **SNP-RFLPing: restriction enzyme mining for SNPs in genomes.** *BMC Genomics* 2006, **7**:30.
4. Lin GT, Tseng HF, Yang CH, Hou MF, Chuang LY, Tai HT, Tai MH, Cheng YH, Wen CH, Liu CS, *et al*: **Combinational polymorphisms of seven CXCL12-related genes are protective against breast cancer in Taiwan.** *OMICS* 2009, **13(2)**:165-172.
5. Yen CY, Liu SY, Chen CH, Tseng HF, Chuang LY, Yang CH, Lin YC, Wen CH, Chiang WF, Ho CH, *et al*: **Combinational polymorphisms of four DNA repair genes XRCC1, XRCC2, XRCC3, and XRCC4 and their association with oral cancer in Taiwan.** *J Oral Pathol Med* 2008, **37(5)**:271-277.
6. Aomori T, Yamamoto K, Oguchi-Katayama A, Kawai Y, Ishidao T, Mitani Y, Kogo Y, Lezhava A, Fujita Y, Obayashi K, *et al*: **Rapid single-nucleotide polymorphism detection of cytochrome P450 (CYP2C9) and vitamin K epoxide reductase (VKORC1) genes for the warfarin dose adjustment by the SMart-amplification process version 2.** *Clin Chem* 2009, **55(4)**:804-812.
7. Chuang LY, Yang CH, Tsui KH, Cheng YH, Chang PL, Wen CH, Chang HW: **Restriction enzyme mining for SNPs in genomes.** *Anticancer Res* 2008, **28(4A)**:2001-2007.
8. NCBI: **Restriction Fragment Length Polymorphism.** [http://www.ncbi.nlm.nih.gov/genome/probe/doc/TechRFLP.shtml], (RFLP) (accessed September 2009).
9. Hamajima N, Saito T, Matsuo K, Kozaki K, Takahashi T, Tajima K: **Polymerase chain reaction with confronting two-pair primers for polymorphism genotyping.** *Jpn J Cancer Res* 2000, **91(9)**:865-868.
10. Hamajima N, Saito T, Matsuo K, Tajima K: **Competitive amplification and unspecific amplification in polymerase chain reaction with confronting two-pair primers.** *J Mol Diagn* 2002, **4(2)**:103-107.
11. Maruyama C, Suemizu H, Tamamushi S, Kimoto S, Tamaoki N, Ohnishi Y: **Genotyping the mouse severe combined immunodeficiency mutation using the polymerase chain reaction with confronting two-pair primers (PCR-CTPP).** *Exp Anim* 2002, **51(4)**:391-393.
12. Tamakoshi A, Hamajima N, Kawase H, Wakai K, Katsuda N, Saito T, Ito H, Hirose K, Takezaki T, Tajima K: **Duplex polymerase chain reaction with confronting two-pair primers (PCR-CTPP) for genotyping alcohol dehydrogenase beta subunit (ADH2) and aldehyde dehydrogenase 2 (ALDH2).** *Alcohol Alcohol* 2003, **38(5)**:407-410.
13. Katsuda N, Hamajima N, Tamakoshi A, Wakai K, Matsuo K, Saito T, Tajima K, Tominaga S: **Helicobacter pylori seropositivity and the myeloperoxidase G-463A polymorphism in combination with interleukin-1B C-31T in Japanese health checkup examinees.** *Jpn J Clin Oncol* 2003, **33(4)**:192-197.
14. Togawa S, Joh T, Itoh M, Katsuda N, Ito H, Matsuo K, Tajima K, Hamajima N: **Interleukin-2 gene polymorphisms associated with increased risk of gastric atrophy from Helicobacter pylori infection.** *Helicobacter* 2005, **10(3)**:172-178.
15. Abu-Amero KK, Al-Boudari OM, Mohamed GH, Dzimiri N: **The Glu27 genotypes of the beta2-adrenergic receptor are predictors for severe coronary artery disease.** *BMC Med Genet* 2006, **7**:31.
16. Yang SJ, Wang HY, Li XQ, Du HZ, Zheng CJ, Chen HG, Mu XY, Yang CX: **Genetic polymorphisms of ADH2 and ALDH2 association with esophageal cancer risk in southwest China.** *World J Gastroenterol* 2007, **13(43)**:5760-5764.
17. Goldberg DE: **Genetic algorithms in search, optimization, and machine learning.** New York: Addison-Wesley 1989.
18. Jong KD: **Learning with genetic algorithms: an overview.** *Mach Learning* 1988, **3**:121-138.
19. Holland JH: **Adaptation in nature and artificial systems: An introductory analysis with applications to biology, control, and artificial intelligence.** MIT Press 1992.
20. Chang HW, Chuang LY, Ho CH, Chang PL, Yang CH: **Odds ratio-based genetic algorithms for generating SNP barcodes of genotypes to predict disease susceptibility.** *OMICS* 2008, **12(1)**:71-81.
21. Wu JS, Lee C, Wu CC, Shiue YL: **Primer design using genetic algorithm.** *Bioinformatics* 2004, **20(11)**:1710-1717.

22. Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, Sirotkin K: **dbSNP: the NCBI database of genetic variation.** *Nucleic Acids Res* 2001, **29(1)**:308-311.
23. Sambrook J, Fritsch EF, Maniatis T: **Molecular cloning.** Cold Spring Harbor Laboratory Press Cold Spring Harbor, NY 1989.
24. De Jong KA, Spears WM: **An analysis of the interacting roles of population size and crossover in genetic algorithms.** Springer 1990, **1**:38-47.
25. Sakurai T, Reichert J, Hoffman EJ, Cai G, Jones HB, Faham M, Buxbaum JD: **A large-scale screen for coding variants in SERT/SLC6A4 in autism spectrum disorders.** *Autism Res* 2008, **1(4)**:251-257.
26. Goldberg TE, Kotov R, Lee AT, Gregersen PK, Lencz T, Bromet E, Malhotra AK: **The serotonin transporter gene and disease modification in psychosis: Evidence for systematic differences in allelic directionality at the 5-HTTLPR locus.** *Schizophr Res* 2009, **111(1-3)**:103-108.
27. Mandelli L, Mazza M, Martinotti G, Di Nicola M, Daniela T, Colombo E, Missaglia S, De Ronchi D, Colombo R, Janiri L, *et al*: **Harm avoidance moderates the influence of serotonin transporter gene variants on treatment outcome in bipolar patients.** *J Affect Disord* 2009, **119(1-3)**:205-209.
28. Yang CH, Cheng YH, Chuang LY, Chang HW: **SNP-Flankplus: SNP ID-centric retrieval for SNP flanking sequences.** *Bioinformation* 2008, **3(4)**:147-149.
29. Kampke T, Kieninger M, Mecklenburg M: **Efficient primer design algorithms.** *Bioinformatics* 2001, **17(3)**:214-225.
30. Wu J, Wang J, Chen J: **A practical algorithm for multiplex PCR primer set selection.** *Int J Bioinform Res Appl* 2009, **5(1)**:38-49.
31. Wang J, Li KB, Sung WK: **G-PRIMER: greedy algorithm for selecting minimal primer set.** *Bioinformatics* 2004, **20(15)**:2473-2475.
32. Chen YF, Chen RC, Chan YK, Pan RH, Hseu YC, Lin E: **Design of multiplex PCR primers using heuristic algorithm for sequential deletion applications.** *Comput Biol Chem* 2009, **33(2)**:181-188.
33. Chen SH, Lin CY, Cho CS, Lo CZ, Hsiung CA: **Primer Design Assistant (PDA): A web-based primer design tool.** *Nucleic Acids Res* 2003, **31(13)**:3751-3754.
34. Yang CH, Cheng YH, Chuang LY, Chang HW: **Genetic Algorithm for the Design of Confronting Two-Pair Primers.** *Ninth IEEE international Conference on BioInformatics and BioEngineering (BIBE)* 2009, 242-247.
35. Hamajima N: **PCR-CTPP: a new genotyping technique in the era of genetic epidemiology.** *Expert Rev Mol Diagn* 2001, **1(1)**:119-123.

doi:10.1186/1471-2105-11-509
Cite this article as: Yang *et al.*: Confronting two-pair primer design for enzyme-free SNP genotyping based on a genetic algorithm. *BMC Bioinformatics* 2010 **11**:509.