Multiplex PCR Assay Design by Hybrid Multiobjective Evolutionary Algorithm

In-Hee Lee, Soo-Yong Shin, and Byoung-Tak Zhang

Biointelligence Laboratory
School of Computer Science and Engineering
Seoul National University, Seoul 151-742, Korea
{ihlee,syshin,btzhang}@bi.snu.ac.kr

Abstract. Multiplex Polymerase Chain Reaction (PCR) assay is to amplify multiple target DNAs simultaneously using different primer pairs for each target DNA. Recently, it is widely used for various biology applications such as genotyping. For sucessful experiments, both the primer pairs for each target DNA and grouping of targets to be actually amplified in one tube should be optimized. This involves multiple conflicting objectives such as minimizing the interaction of primers in a group and minimizing the number of groups required for the assay. Therefore, a multiobjective evolutionary approach may be an appropriate approach. In this paper, a hybrid multiobjective evolutionary algorithm which combines ϵ-multiobjective evolutionary algorithm with local search is proposed for multiplex PCR assay design. The proposed approach was compared with another multiobjective method, called MuPlex, and showed comparative performance by covering all of the given target sequences.

1 Introduction

The Polymerase Chain Reaction (PCR) is a very powerful biological technique which is widely used to amplify DNA and plays a key role in biotechnology and biology research. In standard protocol, PCR can amplify only one target DNA at a time (Fig. 1(a)). But the biological or clinical assay usually involves multiple target DNAs, it is much more desirable to amplify these DNAs simultaneously. The multiplex PCR is an extension of PCR in which multiple target DNAs are amplified at the same time (Fig. 1(b)). It has a wide variety of applications in biology and is recently spotlighted as a core tool for high throughput single nucleotide polymorphism (SNP) genotyping [1,2,3].

For successful experimental results, a careful design of multiplex PCR assay is important. A multiplex PCR assay design is a complex problem composed of two optimization processes: optimizing primers for each target while minimizing the number of partition. First, the primers for each target DNA should be optimized so that the interactions between primers and non-target DNAs are minimized. Since the multiple targets are amplified in one tube at the same time, it is important that the primers for one target do not interact another targets or primers. If such an unwanted interaction happens, some of the target DNAs

S. Obayashi et al. (Eds.): EMO 2007, LNCS 4403, pp. 376–385, 2007.

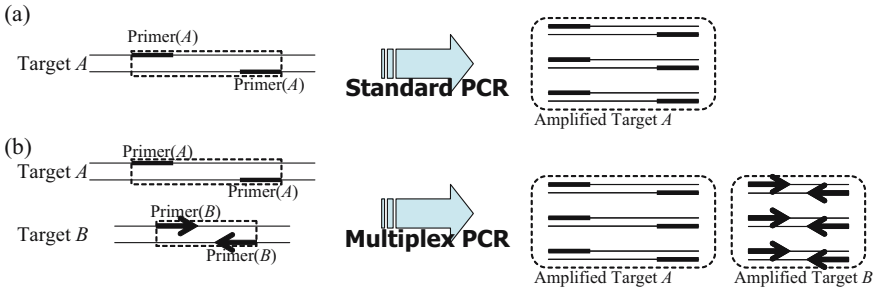

Fig. 1. (a) The concept of standard polymerase chain reaction (PCR). The region between primers in A (the dashed box) is amplified by PCR. (b) The concept of multiplex PCR. Multiple targets A and B are amplified simultaneously by primers A and B, respectively, in one experiment.

might not be amplified. Last, the grouping of target DNAs which will be amplified together should be decided. It would be ideal when all of the target DNAs can be amplified together. Unfortunately, it is very likely that the primers for some targets can not be chosen to prevent unwanted interactions. In such cases, the targets should be put to different groups for separate multiplex PCR runs. However, the number of such separate groups should be minimized.

There have been many studies to tackle this problem [4,5,6,7,8,9,10,11,12] . Most of these works assumed only one group and the targets that do not fit to be amplified together were discarded [5,6,7,8,11,12]. In [4,9,10], on the other hand, the partitioning of targets into multiple groups was handled. First, a set of primer candidates for each target is selected according to predefined conditions. Then, the targets are partitioned into appropriate groups in deterministic way while selecting optimal primers from the candidate sets. From the methodological point of view, most of the previous researches used a deterministic search and only a few evolutionary approaches are published [7,8,11].

Rachlin et al. formulated the design of multiplex PCR assay as finding cliques in graph to optimize both objectives [13]. According to their formulation, the nodes in graph G represent the target DNAs and edges connect two targets(nodes) which can be put into the same group. Each node has multiple states (candidate primers) and the state of two nodes determines whether they can be connected or not. They empirically showed that there is a tradeoff relationship between the specificity of each primer pair to their target and the overall degree of multiplexing. Moreover, it is well known that finding a clique in a given graph is a hard computational problem [14]. Considering these properties of multiplex PCR assay design, a multiobjective evolutionary approach with local search is suggested here.

The suggested multiobjective evolutionary algorithm is based on ϵ-MOEA which was originally suggested in [15]. The algorithm is modified to perform local search after the generation of every new offspring and a genotypical niching is adopted to keep the population diversity.