



An Application for the Simulation of Arbitrarily Primed PCR

Sandhya Dasu¹, Petri Urvil², Catherine Putonti^{1,3} and Yuriy Fofanov^{1,3}

¹ Department of Computer Science, University of Houston ² Department of Obstetrics & Gynecology, University of Texas Medical Branch ³ Department of Biology & Biochemistry, University of Houston



Abstract

Arbitrarily primed PCR (AP-PCR) is a procedure for typing and fingerprinting isolates of a species, by employing single, arbitrarily chosen primers. AP-PCR products can detect a large number of genetic polymorphism and can be used to identify or differentiate among species and sub groupings of species accurately. Although there has been substantial experimental application of this methodology has, so far, been limited.

Using the C# language, we have developed a Windows-based application to simulate AP-PCR fingerprinting. For a given DNA sequence, primer or set of primers, and experimental conditions, we can simulate the agarose gel result. Because many processes are involved during a PCR experiment, our application takes into consideration possible non-specific binding as well as interactions between primers (e.g. cross-hybridization). Although the time taken for execution is proportional to the length of the DNA, our application can consider genomes in seconds (microbial) to just a few minutes (eukaryotic).

Herein, we present the results produced by our application as well as the results produced in the lab for several single random primers and the *Escherichia coli* K12, *Drosophila melanogaster*, and Human genomes. Many of the bands predicted by our application were clearly visible in the real gel. The slight variations observed at a certain experimental temperature were found to disappear as small adjustments were made in the temperature during simulation. This holds with the fallibility of melting temperature calculation to estimate binding.

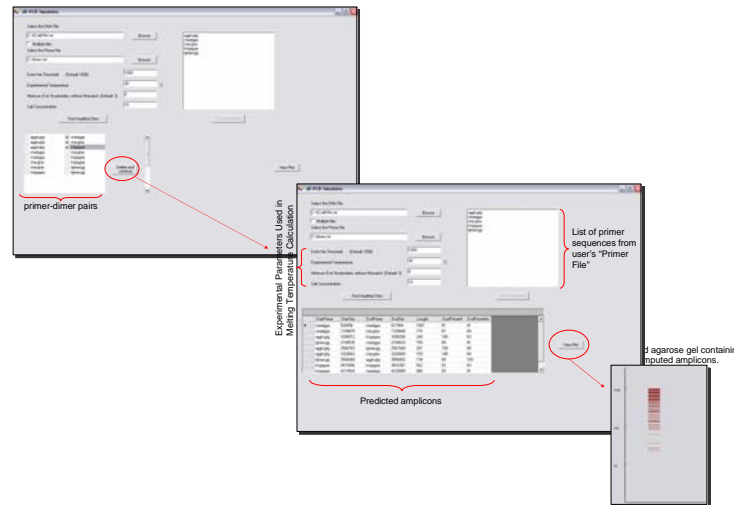
Overview

Detailed epidemiological studies of pathogens are hampered by difficulties in characterizing isolates. The ability to distinguish isolates would be useful in investigating the source of outbreaks of infection, the relatedness of isolates recovered from different patients, and the identities of multiple isolates from the same patient. Arbitrarily Primed Polymerase Chain Reaction (AP-PCR) is commonly used to identify strain specific variations in DNA. Arbitrarily chosen oligonucleotide sequences are selected as primers for the amplification of the target sequence. Because of variations in DNA, the number and sizes of amplicons will vary between different strains. Thus AP-PCR results in a DNA fingerprint that can be used to distinguish isolates. AP-PCR reactions are essentially PCR reactions but they amplify segments of DNA which are essentially unknown (random). Thus, AP-PCR is also known as Random PCR or Random Amplification of Polymorphic DNA (RAPD). AP-PCR products can detect a large number of genetic polymorphism and when linked to major genes can be potentially useful in identifying morphological traits. They can also be used in monitoring diversity within plant populations, for constructing linkage maps and for tracking hybrid species' origins.

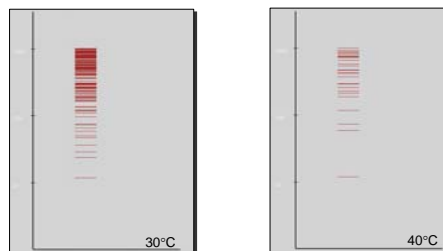
Despite the extensive usage of this technique, very little has been done in the form of theoretical and computational verification of the results it produces. Performing AP-PCR in the laboratory is an expensive process in terms of both time and resources. Our tool simulates the experiment providing a graphical representation of the expected output if the AP-PCR is run on agarose gel.

Application

Using the C# language, we have developed a Windows-based application to simulate AP-PCR. For a given DNA sequence and primer or set of primers, the program predicts where the primer and the target sample will bind. Furthermore, the application takes into consideration experimental conditions, e.g., melting temperature, salt concentration, etc., which may cause non-specific hybridization and subsequent amplification to occur. The present implementation of the application estimates the melting temperature of the primers using SantaLucia's Nearest-Neighbor method [1]. The primer binding locations and the resulting amplicons are presented to the user as is the agarose gel result. The application also assesses if primer interactions such as the formation of primer-dimers will occur. If such interactions are left unchecked, the predicted gel will differ from the actual experimental results. The time taken for execution is proportional to the length of the DNA.



The application permits the user to manipulate the experimental conditions thus influencing the simulated results. Of critical importance is the experimental temperature with respect to the calculated melting temperature for the primer(s). Our application takes into consideration the possibility for non-specific hybridization. The example below illustrates the potential impact on an agarose gel when non-specific hybridization occurs. The same primer was used with the *Drosophila melanogaster* genome at two different temperatures: 30°C (left) and 40°C (right). It is clear that non-specific hybridization is expected at 30°C resulting in many more amplicons.



Results

Several simulations were performed for experiments being conducted in the laboratory. We are presently testing just single primers for amplification. Comparing our predicted amplicons and the actual agarose gel has proven valuable for both verification of laboratory results as well as application design. There are many more factors influencing the success and results of a laboratory PCR reaction and gel. Not all of these factors have been represented in the current implementation of the application. The example below shows the results of an actual gel in which an 8-mer primer was used with an *E. coli* sample. Our prediction includes three amplicons, however there are just two bands visible in the agarose gel. It is possible that the bolder band in the gel is in fact the combination of two of these bands.



Comparison of agarose gel with predicted gel.

Conclusions & Future Work

AP-PCR is a useful tool which can be used to identify or differentiate among species and families of species accurately. Although there has been considerable experimental application of this technique, the theoretical and computational verification of this methodology has, so far, been limited. The software which we have developed can simulate any PCR reaction. Thus far we have found close correlation between the results of our application and the agarose gels generated in the laboratory. Our software has many potential uses including verification of laboratory results. Perhaps even more beneficial is the fact that any primer or group of candidate primers can be run against a sequence for a number of experimental conditions prior to performing the experiment. Future work includes the incorporation of more sophisticated methods of melting temperature estimation as well as further comparisons with experimental results using different primers, primer pairs and samples.

Acknowledgements: We would like to express our gratitude to the Texas Learning and Computation Center (TLCC) for partial support of this work. CP's work was supported by a training fellowship from the Keck Center for Computational and Structural Biology of the Gulf Coast Consortium (NLM Grant No. 5T15LM07093).

References: [1] SantaLucia, et al. (1996) *Biochemistry* 35: 3555-3562.