BIOINFORMATICS APPLICATIONS NOTE Vol. 21 no. 8 2005, pages 1701–1702
doi:10.1093/bioinformatics/bti219

*Genome analysis*

# MultiPLX: automatic grouping and evaluation of PCR primers

Lauris Kaplinski[1,2], Reidar Andreson[1], Tarmo Puurand[1] and Maido Remm[1,*]

[1]Department of Bioinformatics, University of Tartu, Tartu, Estonia and [2]Estonian Biocentre, Riia str. 23, Tartu 51010, Estonia

Received on October 11, 2004; revised on November 25, 2004; accepted on December 10, 2004
Advance Access publication December 14, 2004

## ABSTRACT

**Summary:** MultiPLX is a new program for automatic grouping of PCR primers. It can use many different parameters to estimate the compatibility of primers, such as primer–primer interactions, primer–product interactions, difference in melting temperatures, difference in product length and the risk of generating alternative products from the template. A unique feature of the MultiPLX is the ability to perform automatic grouping of large number (thousands) of primer pairs.

**Availability:** Binaries for Windows, Linux and Solaris are available from http://bioinfo.ebc.ee/download/. A graphical version with limited capabilities can be used through a web interface at http://bioinfo.ebc.ee/multiplx/. The source code of the program is available on request for academic users.

**Contact:** maido.remm@ut.ee

## INTRODUCTION

Polymerase chain reaction (PCR) is widely used in different areas of science and biotechnology. Grouping of PCR primer sets for amplification in a single tube (multiplexing) provides substantial savings in terms of time, chemicals and, most importantly sample materials. Thus, multiplexing is a powerful way for optimizing the cost of genetic analysis and some procedures of PCR are patented (Piepenbrock *et al*., 2004), which further demonstrates its importance. While a lot of effort has been put into the experimental optimization of multiplex PCR conditions (Henegariu *et al*., 1997; Zangenberg *et al*., 1999), little attention has been paid to the prediction of primer compatibility in multiplex reactions. Smaller multiplex groups are often composed manually and verified by trial and error, but this kind of approach is suboptimal for large datasets. Although there are software programs available for testing the compatibility of PCR primer pairs for multiplexing (Rychlik, 1995; Singh *et al*., 1998; Vallone and Butler, 2004), these do not perform automatic grouping and are unable to handle large datasets automatically. We have developed a program, MultiPLX, for automatic primer compatibility testing and grouping. It can also be used for the compatibility evaluation of the existing primer groups. The program is written in standardized C++ for maximum performance and compatibility. We have built binaries for Linux (×86 compatible processors), Windows and Solaris operating systems. Most of the functionality can be used through the web-based interface also.

*To whom correspondence should be addressed.

## ALGORITHM AND IMPLEMENTATION

The workflow of MultiPLX is divided into two main tasks: (1) calculation of compatibility scores for existing PCR primers and (2) grouping of PCR primers based on the cut-off scores. By scores, we are referring to all measurable parameters that may affect the compatibility of different PCR primer sets. Currently, MultiPLX is capable of calculating eight different pre-defined score types:

(1) maximum binding energy (deltaG) of two primers including 3′ ends of both primers
(2) maximum binding energy of 3′ end of one primer with any region of another primer
(3) maximum binding energy of any region of different primers
(4) maximum binding energy of 3′ end of one primer with any region of PCR product
(5) maximum binding energy of any region of primer with any region of PCR product
(6) maximum product length difference between compared PCR primer sets
(7) minimum product length difference between compared PCR primer sets
(8) maximum difference in primer melting temperatures between compared PCR primer sets.

The scores 1–5 are calculated by exhaustive evaluation of all possible gapless alignments and calculation of their deltaG values. The algorithm allows mismatches of any length. The thermodynamic calculations are performed using nearest-neighbor approximation. The default values of enthalpy and entropy of the dinucleotide pairs are the same as used in the program PROBESEL (Kaderali and Schliep, 2002), alternative set of enthalpy and entropy values from the Primer3 program is also available. The concentrations of monovalent salts, $Mg^{2+}$ and DNA can be changed from command line. In addition to the predefined score types, MultiPLX allows the use of a user-specified score, which can be imported into the program. For example, the ability of primers from different PCR primer sets to generate additional PCR products from the human genomic DNA, as calculated by auxiliary program GT4MULTIPLX (http://bioinfo.ebc.ee/gt4multiplx/) can be imported as a custom score into MultiPLX.

Compatibility scores are calculated for all possible pairwise combinations of PCR primer sets. MultiPLX allows in-depth examination of all default scores, including the listing of alignments between primers and their products from different PCR sets. Each score

**Table 1.** Running times of different stages of MultiPLX software with 10, 100, 1000 and 10 000 randomly generated primer pairs on a Pentium4 2.8 GHz processor on Mandrake Linux 9.2, 1 GB RAM

| Number of PCR primer sets to be grouped | Calculation of primer–primer interactions (h:min:s) | Calculation of primer–product interactions (h:min:s) | Calculation of groups (h:min:s) | Average multiplexing level (average number of primer pairs in groups) with 'normal' grouping stringency |
|---|---|---|---|---|
| 10 | <00:00:01 | <00:00:01 | <00:00:01 | 3.3 |
| 100 | 00:00:02 | 00:00:46 | <00:00:01 | 12.5 |
| 1000 | 00:04:39 | 01:14:26 | 00:00:13 | 29.4 |
| 10 000 | 07:51:27 | 129:16:31 | 00:22:25 | 49.3 |

Primer lengths are in the range of 18–22 nt, product lengths in the range of 200–300 bp, average primer GC% is 50. Grouping was performed using default options, without optimization.

type has a default cut-off value to determine whether two given PCR primer sets are compatible and thus can be amplified in the same group. Compatibility cut-off values for all scores can be adjusted by the user either as generic stringency criteria (low, normal, high) or as exact numerical values. Default compatibility cut-offs are generating approximately N3 plex (high stringency), 6–10 plex (normal stringency) or 15–20 plex (low stringency) groups of 100 PCR primer sets.

Automatic grouping of primers is based on compatibility scores and corresponding cut-off values, mentioned above. Two primer sets are considered incompatible if at least one of the compatibility scores exceeds the predefined cut-off value. Two grouping algorithms are implemented.

Algorithm 1. Number of friends (default).

(1) Sort PCR primer sets by the number of compatible primer pairs they have.
(2) In a sorted list of primers, try to insert each primer set into existing groups. If this is not possible, create a new group and move the primer set there.
(3) Repeat until all PCR primer sets have been grouped.
(4) Report the result—list of primer sets associated with their group numbers.

Algorithm 2. Random grouping.

The second algorithm is based on random grouping.

(1) Reorder PCR primer sets randomly.
(2) Try to group each primer set into existing groups. If this is not possible, create a new group and move the primer set there.
(3) Repeat until all PCR primer sets are grouped.
(4) Record the number of groups and group number for each primer.
(5) Repeat the steps 1–4 $N$ times (default value of $N$ is 10 000).
(6) Report the result that generated the lowest number of groups.

The random grouping approach typically generates groups of very different sizes. To make the number of primers in each group more uniform an extra optimization step can be invoked. Optimization tries to move PCR sets from larger groups into smaller ones and swap elements randomly between the groups, until all group sizes differ by no more than one primer pair. None of the grouping algorithms guarantees the optimal solution (lowest possible number of groups). However, the first algorithm typically gives smaller number of groups with shorter computation time. Randomized grouping may be desirable in situations where non-deterministic grouping solution is preferred.

MultiPLX can also be used for the evaluation of existing PCR groups against the calculated score values. Groups where some primer pairs break cut-off rules will be listed alongside the problematic primers. Using this list it is possible to examine, which primers do not fit together and need to be replaced. This option can be helpful for the evaluation of the existing PCR groups with problematic results.

The computation time required by the program for different numbers of primers is shown in Table 1.

## ACKNOWLEDGEMENT

This work was supported by the Estonian Ministry of Education and Research grant no. 0182649s04.

## REFERENCES

Henegariu,O., Heerema,N.A., Dlouhy,S.R., Vance,G.H. and Vogt,P.H. (1997) Multiplex PCR: critical parameters and step-by-step protocol. *Biotechniques*, **23**, 504–511.

Kaderali,L. and Schliep,A. (2002) Selecting signature oligonucleotides to identify organisms using DNA arrays. *Bioinformatics*, **18**, 1340–1349.

Piepenbrock,C., Adorjan,P., Rujan,T. and Schmitt,A. (2004) Method for amplification of nucleic acids of low complexity. European patent no. WO2004015139.

Rychlik,W. (1995) Selection of primers for polymerase chain reaction. *Mol. Biotechnol.*, **3**, 129–134.

Singh,V.K., Mangalam,A.K., Dwivedi,S. and Naik,S. (1998) Primer premier: program for design of degenerate primers from a protein sequence. *Biotechniques*, **24**, 318–319.

Vallone,P.M. and Butler,J.M. (2004) AutoDimer: a screening tool for primer-dimer and hairpin structures. *Biotechniques*, **37**, 226–231.

Zangenberg,G., Saiki,R.K. and Reynolds,R. (1999) Multiplex PCR: Optimization guidelines. In Innis,M., Gelfand,G.H. and Sninsky,J.J. (eds), *PCR Application: Protocols for Functional Genomics*. Academic Press, London, pp. 73–94.