BIOINFORMATICS APPLICATIONS NOTE Vol. 20 no. 15 2004, pages 2471–2472
doi:10.1093/bioinformatics/bth254

PerlPrimer: cross-platform, graphical primer design for standard, bisulphite and real-time PCR

Owen J. Marshall

Murdoch Childrens Research Institute, Royal Childrens Hospital, Flemington Road, Parkville, Vic., 3052, Australia

Received on December 17, 2003; revised on March 9, 2004; accepted on March 17, 2004
Advance Access publication April 8, 2004

ABSTRACT

Summary: PerlPrimer is a cross-platform graphical user interface application for the design of primers for standard, bisulphite and real-time PCR, and sequencing. The program incorporates highly accurate melting-temperature and primer–dimer prediction algorithms with powerful tools such as sequence retrieval from Ensembl and the ability to BLAST search primer pairs. It aims to automate and simplify the process of primer design.

Availability: Open-source and freely available from http://perlprimer.sourceforge.net

Contact: owen.marshall@mcri.edu.au

INTRODUCTION

Accurate design of primers is essential for effective PCR reactions. PerlPrimer is a cross-platform graphical user interface (GUI) application written in Perl that designs primers for standard PCR, bisulphite PCR, real-time PCR (quantitative PCR or QPCR) and sequencing. Whilst separate tools currently exist to find primers for all four methods, PerlPrimer provides several novel enhancements, allows the user to work on multiple projects at a time and provides interfaces to important tools such as gene sequence retrieval from the Ensembl database and the ability to BLAST search primer pairs.

OVERVIEW

PerlPrimer displays a tabbed interface, with each PCR method comprising a different tab as well as a separate tab for detailed primer information. A tab to design sequencing primers is also provided. Although the same algorithms are used for calculating primer melting temperature, amplicon size and primer–dimer stability, each PCR method provides a number of unique features.

For standard PCR, PerlPrimer includes an open reading frame (ORF) detection algorithm that finds the largest ORF in the input sequence, and will automatically set amplification boundaries to the gene found. The program allows the user to attach extra 5′ sequences to each primer. If these sequences include restriction enzyme cloning sites, PerlPrimer will automatically calculate the attached sequences to be in-frame (adding adenine bases where appropriate) for each primer pair if the frame of the cloning sites within the target vector are given.

Primer design for bisulphite PCR is based on the parameters given by Warnecke *et al.* (2002). Unlike an alternative program, Methprimer (Li and Dahiya, 2002), PerlPrimer by default selects for primers with converted cytosine bases at the 3′ end—allowing specific amplification of bisulphite-converted templates. Options are also provided to exclude CpG containing sequences, to eliminate primers containing repeats or runs either before or after bisulphite conversion, and to specify the minimum converted cytosine composition of the primers. PerlPrimer includes a CpG island finding algorithm based on the method of Gardiner-Garden and Frommer (1987) and will automatically set the amplification boundaries to encompass any CpG islands found. Users are advised to note that no attempt is made to automatically limit amplified sequences to small products of 300–600 bp, although the amplified range may be manually reduced if so desired.

For both standard and bisulphite PCR, functions are provided to automatically increase or decrease the amplification boundaries until primers are found that match the specified conditions. An option is also provided to set the amplified range from a selected primer pair, allowing an easy means to create nested primer sets.

Calculation of primers for QPCR is entirely automatic. PerlPrimer uses the application Spidey (Wheelan *et al.*, 2001) to find intron/exon boundaries, and screens all possible primer pairs against these data. To be selected, primer pairs must span at least one intron/exon boundary, and at least one primer must lie across an intron/exon boundary. Amplicon size is limited to between 100 and 300 bases by default. If a gene is requested from the Ensembl database, both genomic and cDNA sequences are retrieved automatically—making QPCR primer design both simple and extremely fast.

Primer pairs that are found by PerlPrimer are listed in a table displaying sequence, length, position, amplicon size and the $\Delta G^{\circ}37$ of the most stable extensible primer–dimer. Primers

are by default sorted by primer–dimer stability, but can also be sorted by any other category.

For all methods, a graphical display is provided showing the position of genes or CpG islands and the selected primer pair, with amplified ranges shown for standard and bisulphite PCR and intron/exon boundaries shown for QPCR. Users can also view a detailed alignment of the primers and sequence, showing (depending on the method used) the ORF with translated protein sequence and codons highlighted, CpG islands with highlighted CpG residues, or highlighted intron/exon boundaries. Selecting a primer pair in any method will bring up detailed thermodynamic parameters for the primer pair and a list of the most stable extensible and non-extensible primer–dimers. The primer pair can be BLAST searched using the NCBI server, and the results obtained can be further limited using a text search. Project files may be saved (including all DNA sequences and primer pairs) and a text report can be generated that lists the primer thermodynamic details, primer–dimers and a full sequence alignment. Selected primer pairs may also be copied in a tab-delimited format that can be pasted into most spreadsheet applications.

A tutorial demonstrating the cloning, sequencing and real-time expression analysis of the human progesterone receptor (*PGR*) gene is provided on the program website, including input and output files and relevant screenshots.

IMPLEMENTATION

PerlPrimer is written in Perl and requires Perl/Tk for the graphical interface. The program was developed and tested on both Microsoft Windows and GNU/Linux platforms, although it should run on any platform with Perl and Perl/Tk.

PerlPrimer calculates melting temperature via the nearest-neighbour thermodynamic approach, using the recent data from J. SantaLucia's group (Allawi and SantaLucia, 1997; SantaLucia, 1998). To adjust the calculated entropy for the salt conditions of the PCR, PerlPrimer uses the empirical formula derived by von Ahsen *et al.* (2001) and allows the user to specify the concentration of Mg^{2+}, dNTPs and primers, or use default, standard PCR conditions. The result is a highly accurate prediction of primer melting temperature, giving rise to a maximum yield of product when amplified.

Elimination of primer–dimer artefacts is important for reliable PCR amplification, and essential for accurate QPCR primer design. PerlPrimer calculates primer–dimers by creating a binding matrix between each primer and itself, and between the forward and reverse primers. The complementarity between the two primers is then read from the matrix, bonds between single, isolated base pairs are ignored (since these are generally unfavourable to stability) and the approximate $\Delta G^{\circ}37$ is calculated using the nearest-neighbour approach, incorporating mismatch data (Allawi and SantaLucia, 1997, 1998a,b,c; Peyret *et al.*, 1999) (currently no penalty is given to internal loops since these parameters are as yet unpublished, thus the value should be an over-estimate of stability). Since the ΔG° nearest-neighbour values are dependent upon entropy, which in turn is dependent upon the salt concentration of the PCR reaction, these values are recalculated whenever the salt conditions are changed by the user. Primer–dimer stability is calculated separately for extensible dimers that will produce amplifiable products, and for non-extensible dimers which will not cause visible artefacts but which will still reduce the free primer population in a reaction.

ACKNOWLEDGEMENTS

This work was supported and tested by the Chromosome Research group at the Murdoch Childrens Research Institute.

REFERENCES

Allawi,H.T. and SantaLucia,J.,Jr (1997) Thermodynamics and NMR of internal G·T mismatches in DNA. *Biochemistry*, **36**, 10581–10594.

Allawi,H.T. and SantaLucia,J.,Jr (1998a) Nearest-neighbor thermodynamics of internal G·C mismatches in DNA. *Biochemistry*, **37**, 9435–9444.

Allawi,H.T. and SantaLucia,J.,Jr (1998b) Nearest neighbor thermodynamic parameters for internal A·C. mismatches in DNA: sequence dependence and pH effects. *Biochemistry*, **37**, 2170–2179.

Allawi,H.T. and SantaLucia,J.,Jr (1998c) Thermodynamics of internal C·T mismatches in DNA. *Nucleic Acids Res.*, **26**, 2694–2701.

Gardiner-Garden,M. and Frommer,M. (1987) CpG islands in vertebrate genomes. *J. Mol. Biol.*, **196**, 261–282.

Li,L.C. and Dahiya,R. (2002) MethPrimer: designing primers for methylation PCRs. *Bioinformatics*, **18**, 1427–1431.

Peyret,N., Seneviratne,P.A., Allawi,H.T. and SantaLucia,J.,Jr (1999) Nearest-neighbor thermodynamics and NMR of DNA sequences with internal A·A, C·C, G·G, and T·T mismatches. *Biochemistry*, **38**, 3468–3477.

SantaLucia,J.,Jr (1998) A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proc. Natl Acad. Sci., USA*, **95**, 1460–1465.

von Ahsen,N., Wittwer,C.T. and Schutz,E. (2001) Oligonucleotide melting temperatures under PCR conditions: nearest-neighbor corrections for Mg(2+), deoxynucleotide triphosphate, and dimethyl sulfoxide concentrations with comparison to alternative empirical formulas. *Clin. Chem.*, **47**, 1956–1961.

Warnecke,P.M., Stirzaker,C., Song,J., Grunau,C., Melki,J.R. and Clark,S.J. (2002) Identification and resolution of artifacts in bisulfite sequencing. *Methods*, **27**, 101–107.

Wheelan,S.J., Church,D.M. and Ostell,J.M. (2001) Spidey: a tool for mRNA-to-genomic alignments. *Genome Res.*, **11**, 1952–1957.