BIOINFORMATICS APPLICATIONS NOTE Vol. 23 no. 17 2007, pages 2331–2333
doi:10.1093/bioinformatics/btm343

Genome analysis

QPRIMER: a quick web-based application for designing conserved PCR primers from multigenome alignments

Namshin Kim and Christopher Lee*

Center for Computational Biology, Molecular Biology Institute, Institute for Genomics and Proteomics, Department of Chemistry and Biochemistry, University of California, Los Angeles, 90095-1570, USA

Received and revised on April 11, 2007; accepted on June 23, 2007

Advance Access publication June 28, 2007

Associate Editor: Keith Crandall

ABSTRACT

Summary: We have developed a quick web-based application for designing conserved genomic PCR and RT-PCR primers from multigenome alignments targeting specific exons or introns. We used Pygr (The Python Graph Database Framework for Bioinformatics) to query intervals from multigenome alignments, which gives us less than a millisecond access to any intervals of any genome within multigenome alignments. PRIMER3 was used to extract optimal primers from a gene of interest. QPRIMER creates an electronic genomic PCR image from a set of conserved primers as well as summary pages for primer alignments and products. QPRIMER supports human, mouse, rat, chicken, dog, zebrafish and fruit fly.

Availability: http://www.bioinformatics.ucla.edu/QPRIMER/

Contact: leec@mbi.ucla.edu

1 INTRODUCTION

Polymerase chain reaction (PCR) is the most important experimental technique for comparative studies as well as gene cloning in order to amplify specific genes of interest from various organisms. There is a need for fast and flexible tool to design proper conserved primer candidates from multiple alignments. EC_oligos (Liu *et al.*, 2004) and DualPrime (Andersson *et al.*, 2005) can design conserved primers between two sequences. Software that can design conserved primers from multiple alignments also have been developed. Primaclade (Gadberry *et al.*, 2005) can handle many types of multiple alignments and Primer Premier 5 (PREMIER Biosoft International) is only available commercially.

Availability of multigenome alignments (Blanchette *et al.*, 2004) at UCSC Genome Browser (Karolchik *et al.*, 2003) provides an important new platform for designing conserved genomic PCR as well as RT-PCR primers of various organism. But, web-based applications utilizing multigenome alignments have not been reported, due in part to the challenges of fast database query from more than a billion alignment intervals. Pygr (A Python Graph Database Framework for Bioinformatics, http://bioinformatics.ucla.edu/pygr/) recently announced new features—NCList (Alekseyenko and Lee, 2007), which is designed for querying intervals of any genomic location of any genomes from multigenome alignments within 1 ms.

QPRIMER is a novel, automatic web-based application for designing conserved genomic PCR and RT-PCR primers from multigenome alignments on the web using Pygr and PRIMER3 (Rozen and Skaletsky, 2000). User-friendly and intuitive graphical interface enable one to browse any genomic location and design conserved PCR primers on the fly. Furthermore, it usually takes <5 s to get conserved genomic PCR primers.

2 FEATURES AND WEB INTERFACE FOR DESIGNING CONSERVED PRIMERS

Most software requires sequences or their multiple alignments as an input. Users need to do many time-consuming steps in order to get optimal primer sets. An important new feature of QPRIMER is its combination of genome browser and PRIMER3 on the web. As shown in Figure 1A and C, the user can browse a specific gene of interest using UCSC-like genome browser based on its genomic location. The user can select any region in the gene structure as a target for amplification, simply by clicking on it. Convenient zooming (700 bp) provides up to 1 bp to 1 pixel mapping for single nucleotide resolution.

Once the user has selected start and end positions for the primer target area (Fig. 1A), a PRIMER3 option window will be displayed at the bottom of page (Fig. 1B). The maximum product length is 1 kb by default for typical genomic PCR primers, but this can be increased in order to design RT-PCR primers covering a target exon. Total calculation time will be also increased. It should be noted that QPRIMER selects primers from only exonic regions. Number of PRIMER3 output is another important parameter because we need many combinations of forward and reverse primers within conserved blocks. QPRIMER chooses conserved primers by comparing synteny blocks retrieved by Pygr from multigenome alignment with optimal primer sets derived from PRIMER3. Total calculation time depends on total length of search area (product size). It will take <5 s for usual genomic PCR primer design, product size <1 kb.

*To whom correspondence should be addressed.

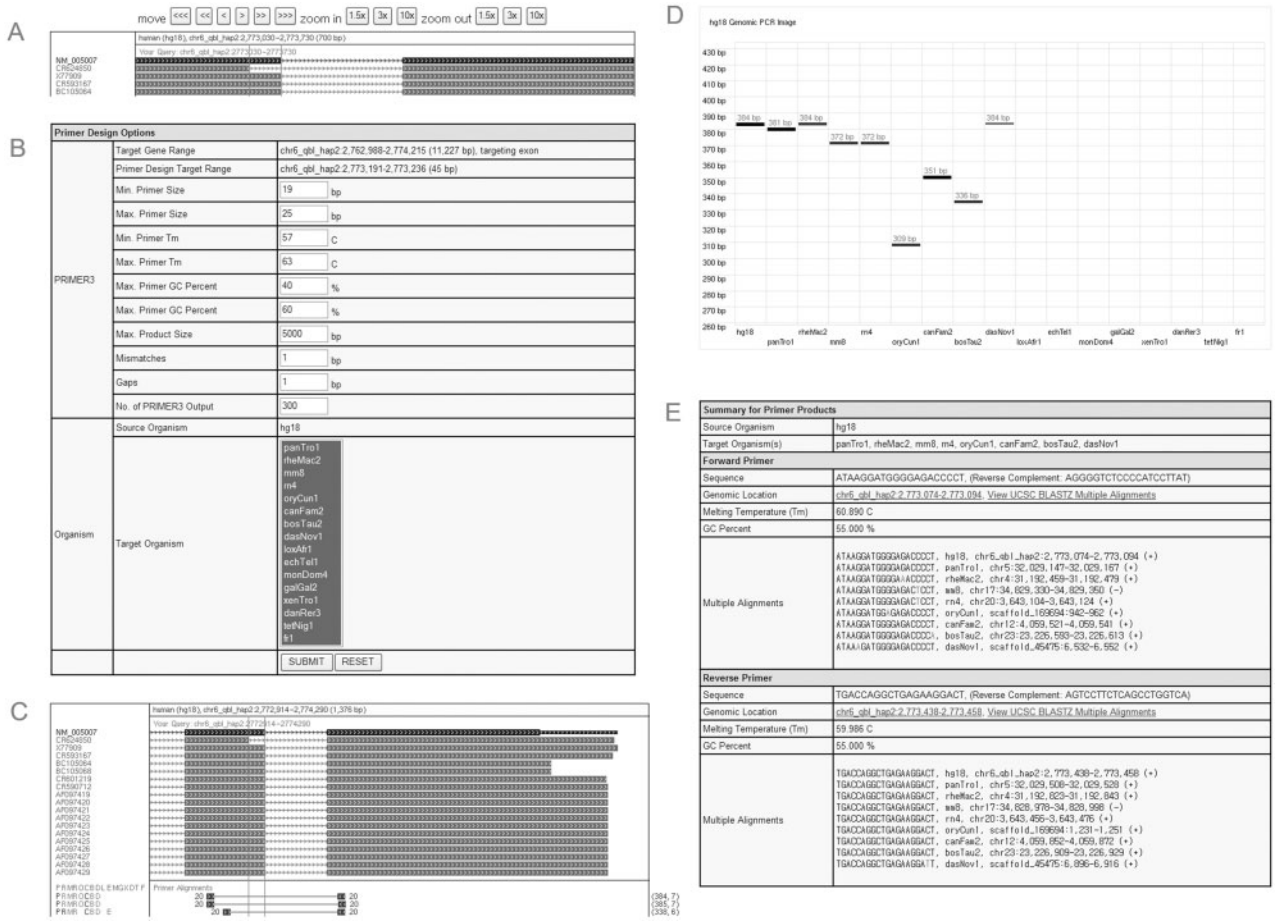

Fig. 1. QPRIMER output snapshots. (**A**) Selecting genomic region by clicking browser. (**B**) PRIMER3 options and target organisms. (**C**) Graphical view of output primer alignments. (**D**) Genomic PCR image. (**E**) Summary.

Genomic alignments of each primer pairs, organism information, product size and sum of mismatches as well as target area will be displayed as shown in Figure 1C (detailed information will be available at QPRIMER web site). Success of conserved primer design is solely dependent on whether the specific gene sequence is conserved over the target organisms or not. If the given gene is diverse or newly created, one may not get primer candidates. QPRIMER will give a short warning message noticed failure to retrieve conserved blocks from multigenome alignment within a second—Pygr performance—in this case.

If user clicks one of the primer sets in Figure 1C, electronic genomic PCR image will be displayed in a new window as shown in Figure 1D. Thickness and darkness of PCR bands represent total number of mismatches in that organism, better alignments will be thick and dark. Primer sequences, PRIMER3 summary, multiple alignments and products from various organisms will be displayed followed by genomic PCR image as shown in Figure 1E. QPRIMER does not check multiple hits for all genomes because it will require huge amount of calculation and not suitable as a web-based application. But, QPRIMER provides linkouts to UCSC In-Silico PCR (http://genome.ucsc.edu/cgi-bin/hgPcr). One can click the linkouts in order to check whether multiple hits exist or not. There would be no multiple hits for most cases because QPRIMER selects primers only within exonic region, not within Repeats.

In summary, one can easily design conserved PCR primers on the web via a simple, fast web interface.

ACKNOWLEDGEMENTS

We wish to thank Dr Meenakshi Roy and Min Jae Lee for valuable comments on this work and testing this application. This work is supported by DOE grant DE-FC02-02ER63421 and NIH grant U54 RR021813.

Conflict of Interest: none declared.

REFERENCES

Alekseyenko,A.V. and Lee,C.J. (2007) Nested Containment List (NCList): a new algorithm for accelerating interval query of genome alignment and interval databases. *Bioinformatics*, **23**, 1386–1393.

Andersson,A. *et al.* (2005) Dual-genome primer design for construction of DNA microarrays. *Bioinformatics*, **21**, 325–332.

Blanchette,M. *et al.* (2004) Aligning multiple genomic sequences with the threaded blockset aligner. *Genome Res.*, **14**, 708–715.

Gadberry,M.D. *et al.* (2005) Primaclade – a flexible tool to find conserved PCR primers across multiple species. *Bioinformatics*, **21**, 1263–1264.

Karolchik,D. *et al.* (2003) The UCSC Genome Browser Database. *Nucleic Acids Res.*, **31**, 51–54.

Liu,S. *et al.* (2004) EC_oligos: automated and whole-genome primer design for exons within one or between two genomes. *Bioinformatics*, **20**, 3668–3669.

Rozen,S. and Skaletsky,H. (2000) Primer3 on the WWW for general users and for biologist programmers. *Methods Mol. Biol.*, **132**, 365–386.