Short Technical Reports

The Primer Generator: A Program that Facili- tates the Selection of Oligonucleotides for Site- Directed Mutagenesis

BioTechniques 26:672-676 (April 1999)

ABSTRACT

Site-directed mutagenesis is a powerful tool that has enabled molecular biologists to perform functional analysis of altered nucleic acids and proteins. Newer PCR-based mutagenesis techniques have reduced the process of mutagenesis to as little as one day. While each technique has its advantages, both require a strategy to isolate the desired clone from a population that contains mutagenized and wild-type genes. In this report, we describe a World Wide Web-based computer program that facilitates the design of mutagenic primers such that successfully mutagenized clones can be identified by the presence or absence of a unique restriction site.

INTRODUCTION

Many site-directed mutagenesis (SDM) protocols call for the use of two oligonucleotides (2,5). The mutagenic oligonucleotide introduces the desired mutation, and a selection oligonucleotide increases the fraction of successfully mutagenized clones and facilitates their subsequent identification. The selection oligonucleotide does improve the fraction of mutagenized clones but it is not 100% effective (3). Clones to which the selection oligonucleotide has annealed but are not annealed to the mutagenic oligonucleotide, are only identified with subsequent DNA sequencing.

If the mutagenic primer introduces or deletes a restriction site, then it can aid in the subsequent identification of mutagenized clones. In effect, the mutagenic primer can be carefully designed to also be a selective primer. The Primer Generator automates the selection of a mutagenic oligonucleotide that meets these criteria. This technique, of carefully selecting a muta-

genic oligonucleotide to facilitate subsequent mutant identification, is a useful and cost-effective method for optimizing the procedure of SDM.

MATERIALS AND METHODS

Software

The following software is used: Perl (10) with a CGI module (9) and National Center for Supercomputing Applications (NCSA) HTTPd server (6).

Algorithm

Figure 1 shows an outline of the program. Initially, the investigator inputs the nucleotide sequence that encodes the current protein fragment and the protein sequence desired following mutagenesis. The latter sequence is reverse-translated using the redundancy of the genetic code to produce a number of DNA sequences that could potentially encode the given protein sequence.

Following reverse-translation, the sequences are interrogated for restric-

Figure 1. Flowchart of the Primer Generator program.

tion enzyme sites using data from the Restriction Enzyme Database (REBASE) (<http://www.neb.com/rebase/rebase.html>). The results of these searches are independently compiled for the current and all possible target nucleotide sequences. Potential restriction sites in the existing sequence are subtracted from restriction sites in the mutagenized sequence, thereby generating a list of restriction sites unique to the mutagenized product. Similarly, restriction sites in the mutagenized sequence are subtracted from restriction sites in the existing sequence, which identifies sites that were deleted by the mutagenesis procedure.

From these operations, a list is compiled of unique restriction sites in the mutagenized sequence and of the

oligonucleotide probes needed to generate them. In addition, restriction sites that are deleted from the existing sequence through the mutagenesis procedure are also listed. The mutagenic primer output is ordered by an increasing number of base pairs that differ from the original sequence. These base pairs are boldfaced. A useful feature of the Primer Generator is the ability to limit the output by the maximal number of substitutions that will be tolerated in the mutagenic primer. It can be set for 1 through 10 or for no limit. This feature facilitates selection of a mutagenic primer that necessitates minimal alteration of the existing sequence. To decrease computation time, especially for longer sequences, it is also possible to limit the number of restriction endonu-

The Primer Generator

Original nucleotide sequence	GAGGCCCTCTCGCTTG
Original protein sequence	EASRL
Desired amino acid sequence	EARRL
Use of restriction enzymes with site size less than 5 bp	No
Use of only most common restriction enzymes	No
Maximum number of substitutions compared to original sequence	3

The following primers have been generated:

- Click on enzyme name to access corresponding entry in [The Restriction Enzyme Database \(REBASE\)](#).
- Positions of cut sites are calculated relative to the beginning of the primer.
- Differences between primer sequence and original nucleotide sequence are **bolded**

Primer Sequence	Deleted Restriction Sites	Cut Positions	New Restriction Sites	Cut Positions	Number of Substitutions
GAGGCCCGTCGCTTG	AatI	4	None		2
	Eco147I	4			
	Pme55I	4			
	SseBI	4			
	StuI	4			
GAGGCCCGGCGCTTG	AatI	4	AsuC2I	6, 7	3
	Eco147I	4	BcnI	6, 7	
	Pme55I	4	Bsp143II	12, 8	
	SseBI	4	BstH2I	12, 8	
	StuI	4	HaeII	12, 8	
			NciI	6, 7	

Figure 2. Partial representation of primers constructed by the Primer Generator program. For each primer, the following data are presented: (i) nucleotide sequence of the primer, (ii) list of restriction endonucleases that cleave the WT sequence but not the mutagenized sequence, (iii) number of residues from the beginning of the primer that the endonuclease cleaves, (iv) list of restriction endonucleases that cleave the mutagenized sequence but not the WT sequence, (v) number of residues from the beginning of the primer that the endonuclease cleaves and (vi) number of substitutions that the primer introduces into the WT sequence. The primers are sorted in order of decreasing number of changes introduced in the WT sequence. Underlined text represents hyperlinks.

Short Technical Reports

cleases to only the most commonly used enzymes. These enzymes are likely to be readily available in most laboratories.

Introducing or deleting a unique restriction site may necessitate mutagenizing more than the minimal number of residues required to elicit the desired change in the primary sequence. The mutagenesis procedure is however sufficiently robust to facilitate changing more than one base pair. A mutagenic oligonucleotide can also be chosen such that it introduces a unique restriction site in the mutagenized sequence, deletes a specific restriction site from the existing sequence or produces some combination of the two. A number of clones from the mutagenesis procedure can then be screened in a short amount of time with a restriction digest before being sequenced.

The restriction enzymes' names in the output are linked to the corresponding entries in the REBASE database,

which provides the user with additional information about the enzyme. Included among this information is the enzyme type, the microorganism from which it was isolated, the prototype of the enzyme (the isoschizomer that was discovered first), the recognition sequence, the sites at which the DNA strand is cleaved, commercial sources of the enzyme and references pertaining to the discovery and characterization of the enzyme.

RESULTS

To demonstrate the effectiveness of this algorithm, we programmed the software to develop primers that mutagenize a critical serine residue in the rapamycin and FKBP12 target protein 1 (RAFT1) to arginine. When administered to cells, rapamycin forms a complex with the immunophilin FKBP12.

This complex subsequently interacts with RAFT1 (8)/FRAP (1). This FKBP12-rapamycin-RAFT (1) interaction is dependent on the presence of a serine residue at position 2035 in RAFT1/FRAP (1). RAFT1/FRAP mutants that do not complex with FKBP12 in the presence of rapamycin are proving instrumental in elucidating the signaling capacity of this protein. Figure 2 summarizes the output of the program.

The output in Figure 2 is only a representative selection of the entire output of the program, which included 15 different mutagenic oligonucleotides. The mutagenic oligonucleotides (limited to the maximum of three substitutions compared to the original sequence) are capable of introducing 14 unique restriction sites, deleting the five existing sites and allowing the user to select from combinations thereof. Alternatively, with settings changed to limit restriction endonucleases used in analysis to

Short Technical Reports

only the most common enzymes and the maximum number of substitutions raised to four, we are still able to introduce six new and remove two old restriction sites (data not shown).

DISCUSSION

The Primer Generator has a number of uses. For example, an approach that has been used to define the binding specificities of modular protein-protein interaction domains (such as the phosphotyrosine-directed SH3 domain) is to independently vary the amino acid residue at a single position to the other 19 amino acids (7). This method has been successful in defining the binding characteristic of a single position (hydrophobic, basic, etc.) in the target peptide. By using the Primer Generator, one could order a single mutagenic primer that contains three N's in place of the codon to be mutated and contains an N in the wobble position of the adja-

cent invariant codons (over a thousand different sequences). A single mutagenesis reaction would be performed, and a few well-chosen restriction digests would rapidly identify the desired mutants, while preventing the sequencing of redundant mutants.

In addition to its use for mutagenesis, the Primer Generator can also be of use in clinical settings. A number of genetic diseases such as cystic fibrosis are characterized by a specific mutation that accounts for >70% of all cases (11). If one wished to screen polymerase chain reaction (PCR) products generated by primers flanking the region most commonly mutated, they need only input the wild-type (WT) and mutant sequences into the program to find restriction enzymes that can be used to distinguish between the two.

Computers are assuming an ever increasing role in molecular biology. The repertoire of computer-based tools and utilities available to investigators continues to grow. Here, we report an algorithm that allows one to choose a mutagenic oligonucleotide such that successfully mutated DNA can readily be identified by a restriction digest. The flexible, user-friendly interface offers the user extensive control of the program and its output. The hot links provide the user with additional information about the enzyme chosen. The availability of site-specific, hybrid restriction endonucleases is expected to increase the number of enzymes available for use by this program (4). A carefully chosen mutagenic oligonucleotide can save time and money by reducing the number of primers used and clones sequenced. The Primer Generator can be accessed at the Web site: (<http://www.med.jhu.edu/medcenter/primer/primer.cgi>).

ACKNOWLEDGMENTS

The authors wish to thank Dr. Robert Cottingham and Dr. Peter Li for constructive feedback and technical comments. We are also indebted to the Division of Biomedical Information Sciences of the Johns Hopkins University School of Medicine for allotting the project free space on the JHMI Affiliates Server.

REFERENCES

1. **Chen, J., X. Zheng, E.J. Brown and S.L. Schreiber.** 1995. Identification of an 11-kDa FKBP12- rapamycin-binding domain within the 289-kDa FKBP12-rapamycin-associated protein and characterization of a critical serine residue. *Proc. Natl. Acad. Sci. USA* 92:4947-4951.
2. **Hemsley, A., N. Arnheim, M.D. Toney, G. Cortopassi and D.J. Galas.** 1989. A simple method for site-directed mutagenesis using the polymerase chain reaction. *Nucleic Acids Res.* 17:6545-6551.
3. **Hutchinson, C.A., M. Phillips, M.H. Edgell, S. Gillam, P. Jahnke and M. Smith.** 1978. Mutagenesis at a specific position in a DNA sequence. *J. Biol. Chem.* 253:6551-6560.
4. **Kim, Y., J. Cha and S. Chandrasegaran.** 1996. Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proc. Natl. Acad. Sci. USA* 93:1156-1160.
5. **Kunkel, T.A.** 1985. Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proc. Natl. Acad. Sci. USA* 82:488-492.
6. **NCSA HTTPd Development Team.** 1998. NCSA HTTPd v. 1.5.2a. Available from: <http://hoohoo.ncsa.uiuc.edu/docs/Overview.html> [Internet].
7. **Pawson, T.** 1995. Protein modules and signalling networks. *Nature* 373:573-580.
8. **Sabatini, D.M., H. Erdjument-Bromage, M. Lui, P. Tempst and S.H. Snyder.** 1994. RAFT1: a mammalian protein that binds to FKBP12 in a rapamycin dependent fashion and is homologous to yeast TORs. *Cell* 78:35-43.
9. **Stein, L.D.** 1998. CGI.pm v. 2.42. Available from: <http://www-genome.wi.mit.edu/ftp/pub/software/WWW/> [Internet].
10. **Wall, L., T. Christiansen and R. Schwartz.** 1996. *Programming Perl*, 2nd ed. O'Reilly & Associates, Sebastopol, CA.
11. **Welsh M.J., L.C. Tsui, T.F. Boat and A.L. Beaudet.** 1993. Cystic fibrosis, p. 3799-3875. In C.R. Scriver, A.L. Beaudet, W.S. Sly and D. Valle (Eds.), *The Molecular and Metabolic Basis of Inherited Disease*, 7th ed. McGraw Hill, New York.

Address correspondence to Joseph F. Lawler, Department of Neuroscience, Johns Hopkins University School of Medicine, WBSB Room 814, 725 North Wolfe Street, Baltimore, MD 21205, USA. Internet: jlawler@welchlink.welch.jhu.edu

Received 30 June 1998; accepted 14 December 1998.

**Alexander Turchin and
Joseph F. Lawler, Jr.**
*Johns Hopkins University
School of Medicine
Baltimore, MD, USA*