Tropical Agricultural Research Vol. 24 (2): 193 – 201 (2013)
Short Communication

Comparative Study on Restriction Analysis and Primer Designing Programme

T.J. Koswatte, V.A. Sumanasinghe[1*] and P. Samaraweera[2]

Postgraduate Institute of Agriculture
University of Peradeniya
Sri Lanka

***ABSTRACT**. Many programmes are available in the internet to analyze restriction sites and design primers. The results of comparative evaluations of commonly deployed, freely available software for these using nucleotide sequences from the NCBI database are presented. Seven restriction analysis tools were compared for specificity, ability to detect the restriction site, graphical display of output, clarity, and user flexibility while the NEB cutter provides the highest user flexibility and the most clear graphical output. Just Bio-Cutter provides more details about the restriction enzymes than other tools. "Restriction Enzyme Digest of DNA" has the uniqueness of multiple sequence digestion. Seven primer designing programmes were evaluated for comprehensiveness, interactive design, user-friendliness, and the efficiency of designing primers. The Primer3 has higher capabilities and flexibility and, Primer Quest has comprehensiveness and better interactive designs. The Batchprimer3 has the unique ability to design primers for multiple sequences simultaneously.*

***Keywords:** Bioinformatics, primer designing, restriction analysis, software*

INTRODUCTION

Restriction analysis helps to identify restriction enzyme cleavage sites on a DNA sequence and is an important bioinformatics tool in genome mapping, gene cloning, and probe designing. In addition, restriction analysis is widely used in DNA finger printing techniques such as RFLP and AFLP and in software development related to molecular diagnosis. Type II Restriction endonucleases recognize a specific short DNA sequence and bind and cleave the DNA within the site of recognition (Cleaver, 2006). Most restriction sites are 4 to 6 base palindromes.

PCR (polymerase chain reaction) is one of the most popular, time saving and sensitive method for amplified synthesis of DNA sequences in biological and biomedical research today. However, there are several obstacles hindering the best performance of a PCR. False amplification of DNA is a serious problem in PCR. The most common causes of the artifact are the self-dimerization and unspecific annealing of primers. Optimal primer sequence and appropriate primer concentration are essential for maximal specificity and efficiency of PCR.

[1] Department of Agricultural Biology, Faculty of Agriculture, University of Peradeniya, Sri Lanka
[2] Department of Molecular Biology and Biotechnology, Faculty of Science, University of Peradeniya, Sri Lanka
[*] Corresponding author: sajanas@pdn.ac.lk

Several parameters such as the length of the primer, percentage GC content and the 3' sequence need to be optimized for successful PCR (Abd-Elsalam, 2003). If primers are too short they might hybridize to non-target sites giving undesired amplification products. If primers are too long, it reduces the efficiency of PCR. In practice, primers longer than 30 bases are rarely used (Brown, 2001). The ideal annealing temperature must be low enough to enable hybridization between primer and template but high enough to prevent mismatched hybrids from forming. This depends on the melting temperature (*Tm*), the point at which the paired primer-template hybrid precisely dissociates.

The GC percentage provides information about the strength of annealing. A 45%-60% GC content is recommended to avoid internal secondary structures and long stretches of any one base (Dieffenbach *et al.,* 1993). For primers with a GC content of less than 50%, it is necessary to extend the primer sequence beyond 18 bases to keep the melting temperature above the recommended lower limit of 50°C. The 3' terminal position in PCR primers is essential for the control of mis-priming (Kwok *et al.,* 1990). Software provides "Max 3" complementarities which bring the maximum stability for the last five bases at the 3' end. Default values are used in general primer selection. However, depending on the experimental conditions parameters have to be manipulated. The user flexibility and optimization are the most important factors.

A number of web-based free bioinformatics tools and commercial software are available for restriction analysis and PCR primer design. The objective of this study is to compare the freely available and easily accessible software for these two analyses.

MATERIALS AND METHODS

All programs were run on a computer with Intel dual core 2.0 GHz processor and 1GB RAM. The operating system was Windows XP.

Analysis of Restriction Site Detection Software

A 200 long nucleotide sequence (gi|153090190:301-500) randomly selected from the NCBI database was used for analysis (Fig. 1). The cleavage sites were identified using seven freely available restriction analysis tools: namely WatCut, Sequence extractor, JustBio-Cutter, NEB cutter, TAIR, RestrictionMapper and Restriction Enzyme Digest of DNA. The URLs and the features of these programs are listed in Table 1. The evaluation was carried out by comparing differences among recognized cleavage sites and specificity of each program.

```
>gi|153090190:301-500 Homo sapiens syntabulin (syntaxin-interacting) (SYBU),
transcript variant 8, mRNA

CACAGAGTGCAGCATCATGACAAGGAGATTTCTCGAAGCCGAATTCCCCGGTTGATTCTTCGGCCCCATA
TGCCCCAACAACAGCACAAAGTGTCCCCAGCCTCTGAGTCTCCTTTCTCTGAGGAAGAGAGCAGAGAGTT
CAACCCCAGCAGCTCTGGGCGCTCAGCGAGGACCGTTAGCAGCAACAGCTTCTGCTCAGA
```

Fig 1. The sequence used from NCBI for the identification of restriction sites

Table 1. Restriction Analysis programmes

Tool	URLs	Specificities
WatCut	http://watcut.uwaterloo.ca/watcut/watcut/template.php	Analyzes DNA sequences up to 50kbp in length Optimizes the display of results in a graphical format, as a plain table, or in a complete textual format along with the DNA sequence and its translated sequence
Sequence Extractor	http://www.bioinformatics.org/seqext/	Generates a restriction map and PCR primer map of a DNA sequence Gives translated sequence for submitted nucleic acid sequence. Manipulates six frame translation
JustBio-Cutter	http://www.justbio.com/index.php?page=cutter	Two alternating selections for restriction enzymes: all enzymes or only enzymes with recognition sites equal or greater than 6 bp Gives a full description about each enzyme and the commercial availability followed by REBASE
NEB cutter	http://tools.neb.com/NEBcutter2/index.php	The maximum sequence length is 300kbp Provides different type of outputs including restriction maps (Figure 3), theoretical digests and links to the restriction enzyme database REBASE, and gel view after digestion (Figure 4) Gives changeable digestion parameters
TAIR	http://www.arabidopsis. org/index.jsp	Provides flexibility in selecting enzymes whether it is cut by all enzymes, 3'overhang/ 5'overhang, blunt end, cut once or twice. Does not give full details about enzyme, restriction recognition site and the results fragment size. Gives cutting positions in graphical format
Restric-tion Mapper	http://www.restrictionmapper.org/	Output is expressed only as a table summary and does not provide pictorial output Provides information about sequence recognition position, cutting position, size of fragments, overhang, and the frequency of cutting
Restric-tion Enzyme Digest of DNA	http://www.biophp.org/minitools/restriction_diges/demo.php	Provides user flexibility in selecting restriction enzymes Provides the facility of submitting more than one sequence for digestion and all results are given in separate columns(Figure 5)

Analysis of Primer Designing Software

In order to evaluate the primer designing programs, a 1000 base long nucleotide sequence (gi|153090190 |ref|NM_001099747.1:1-1000) was randomly selected from the NCBI database. This sequence was submitted to seven different free, primer-designing tools without changing their default parameters. Table 2 shows these seven primer-designing software programs, URLs and their specificities. The initial comparison was based on comprehensiveness (the degree of manipulation), interactive design, and user-friendliness of servers. Then, the best software programs were selected and their ability to narrow down the primer binding sequences towards target regions was evaluated. Five trials were performed in this second stage by changing the target sequence, included region and excluded regions. Other parameters were kept constant in all five trials. The number of primers, product size, primer length, primer *Tm*, and percentage of GC were the constant parameters. The range of primer size was from 18 to 27 bp and the optimum was 20 bp. The annealing temperature was 57°C to 63°C with the optimum of 60°C. Percentage of GC was 30% to 80% and the optimum was 50%. All trials were preformed to obtain five primer pairs and product size ranged from 300 to 400 bp.

Table 2. Primer designing software programs and their URLs

Tool	URLs	Specificities
Primer3	http://frodo.wi.mit.edu/primer3/	Higher user flexibility for optimizing primer parameters
PDA	http://dbb.nhri.org.tw/primer/index.html.	Advanced parameter settings including dimer check, hairpin check, 5-GC content check, and 3-GC content check.
Primer-BLAST,	http://www.ncbi.nlm.nih.gov/tools/primer-blast	Combines primer design (using Primer3) and a specificity check via a BLAST search Designs primers that can amplify only a particular splice variant of a gene
PrimerQuest	http://eu.idtdna.com/Scitools/Applications/Primerquest/	Gives various kind of information clearly provides facility to purchase the primers
Genefisher	http://bibiserv.techfak.unibielefeld.de/genefisher2/	Does not provide flexibility to enter 'excluding and including regions' for primer designing
Batchprimer3	http://probes.pw.usda.gov/cgibin/batchprimer3/batchprimer3.cgi	Provides multiple sequence submission and gives primer for each sequence separately at once.
PRIDE	http://pride.molgen.mpg.de/pride.html	Facilitates designing primers and compares it with vector sequences.

RESULTS AND DISCUSSION

Restriction Site Identification

A total of 135 restriction enzymes were identified by all software programmes to the corresponding query sequence. Fig. 2 show the comparison of percentage of ability to identify the restriction enzymes by different programmes. The program Restriction Enzyme Digest of DNA exhibits the highest capability in identifying the relevant restriction enzyme for the target site (>70%) while more commonly used NEB Cutter registered only 40% level of capability.

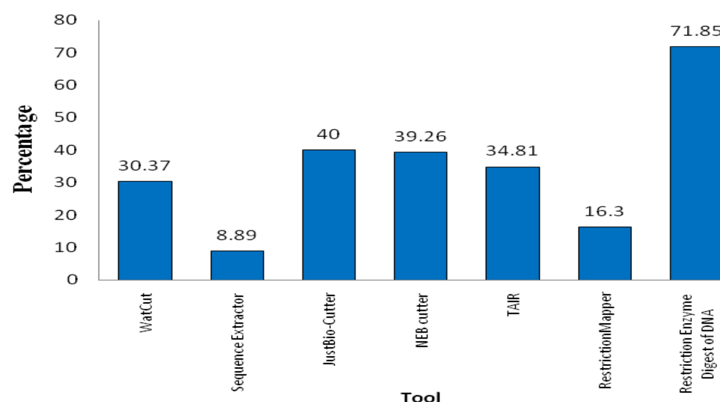

Fig. 2. Percentages of ability to identify the restriction enzymes

The study revealed that there are variations in the ability to recognize the restriction site, and in the pictorial expression of each computational restriction analysis tool. NEB cutter provides the highest user flexibility and clear pictorial expression (Fig. 3 and 4). JustBio-Cutter provides more comprehensive details about restriction enzymes than other tools. The program Restriction Enzyme Digest of DNA provides the multiple sequence digestion at the same time (Fig. 5). It has the highest (72%) ability to find the cleavage sites as well. NEB cutter and JustBio-Cutter have moderate ability (40%) of identification. Sequence Extractor and RestrictionMapper exhibited a low percentage of identification. For instance, JustBio-Cutter (at 21, 61,191bp), NEB cutter (21,161,191), and TAIR (22,162,192) could predict *Bbv*I restriction sites in template sequence but the other four software were not equipped for it.

The default setting of the restriction analysis programmes often identifies cleavage sites for all commercial enzymes that are present in query sequence. Although they have the same default guidelines, the results differ from one another. For an example, Watcut and RestrictionMapper have the same default settings but the ability of identification of cleavage sites differs. Consequently, it is important that the user should have great knowledge about the specificity of each program for identifying the best restriction analysis tool that suits the need.

#	Ends	Coordinates	Length (bp)
1	BsaJI-AluI	47-152	106
2	BbvI-AluI	162-188	27
3	BbvI-BsaJI	22-46	25
4	(LeftEnd)-BbvI	1-21	21
5	AluI-BbvI	153-161	9
6	BbvI-(RightEnd)	192-200	9
7	AluI-BbvI	189-191	3

Fig 3. **Graphical expression of NEB cutter programme**

Fig 4. **Theoretical gel view of NEB cutter Programme**

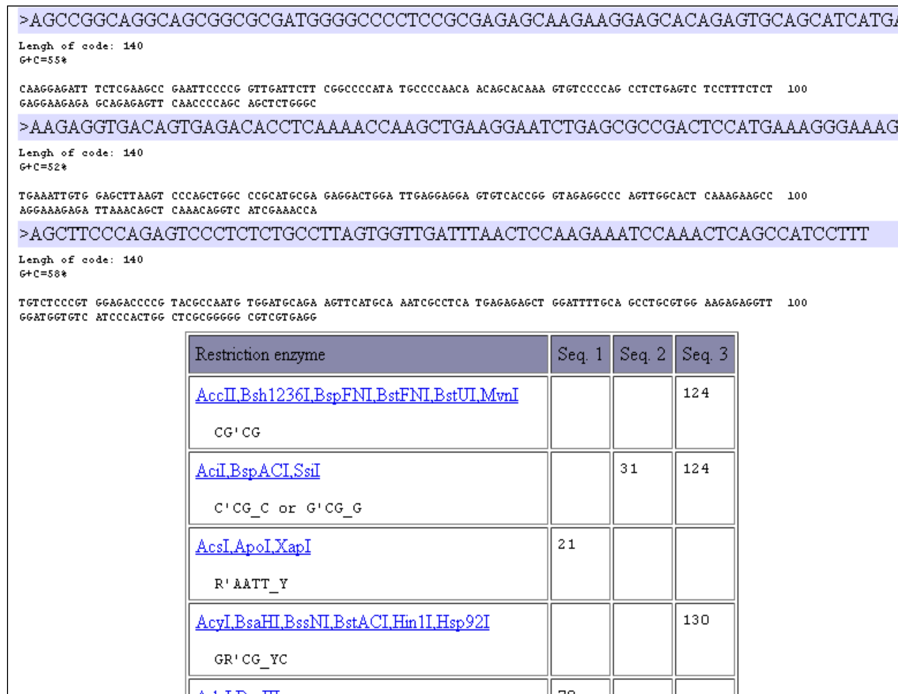

Fig. 5. Output of Restriction Enzyme Digest of DNA

Primer Designing

Based on the capability of narrowing towards the target region, the interactive designing, and the rapid accessibility of programs, the Primer3 and PrimerQuest were identified as the best primer designing software. PDA, Primer-BLAST, Genefisher and BatchPrimer3 tools do not provide the option to manipulate the excluding or including region on the template sequences. Manipulating the excluding and including region is important to design primers closer to the target region and it helps reduce unnecessary product formation in the PCR reaction. Therefore, PDA, Primer-BLAST, PRIDE, Genefisher and BatchPrimer3 showed less comprehensiveness in specifying the target region of the template sequence than Primer3 and PrimerQuest. Therefore, based on both comprehensiveness and user-friendliness shown by each program, it could be revealed that Primer3 and PrimerQuest are the best primer designing software. Therefore, further analyses with these two software programs were carried out. In this next stage, primer-designing parameters were equalized in Primer3 and PrimerQuest.

Five trials were conducted by changing the target, and the included and excluded regions. Other parameters such as number of primers, product size, primer length, primer *Tm*, and percentage of GC were kept constant in all five trials. Table 3 shows parameters deployed and the resulted primer sequence for each trial. According to the first trial, Primer3 and PrimerQuest have designed primers corresponding to template region of 518 to 899 bases. In addition, Primer3 identified two starting positions at 518 and 544 bp to design forward primers but the PrimerQuest identified only one position at 544 bp.

Table 3. Resulted primer pairs of Trial No: 1 to Trail No: 5

Trial	Database	Primer	Product	Start	Forward Primer	Tm	Start	Reverse Primer	Tm
Trial 1	Targets				Does not submit a target region				
	Excluded Regions				1-300				
	Included Region				500-400				
	PrimerQuest	1	301	544	ACACCCTCAGATGCTGGAAACA	58.4	844	GCTTGTGTGACGAAGCTCCATT	57.7
		2	355	544	ACACCCTCAGATGCTGGAAACA	58.4	898	TGCACAGCATGGACAGAAGGT	58.9
	Primer3	1	329	518	GCAGCCAGTCAGTGTCTCCT	60.6	846	AGGCTTGTGTGACGAAGCTC	60.6
		2	303	544	ACACCCTCAGATGCTGGAAA	60.7	846	AGGCTTGTGTGACGAAGCTC	60.6
Trial 2	Targets				346-238				
	Excluded Regions				1-300 &750-200				
	Included Region				300-370				
	PrimerQuest	1	350	309	TCACAGCTCTGGCTTGCAGAAT	58.8	658	AAGCGTGTCACCGTCGTGGAAA	61.2
		2	345	309	TCACAGCTCTGGCTTGCAGAAT	58.8	653	TGTCACCGTCGTGGAAAGCAC	59.8
	Primer3	1	342	313	AGCTCTGGCTTGCAGAATTT	59.2	654	GTGTCACCGTCGTGGAAAG	60.15
		2	330	325	CAGAATTTTCCACCCCAAAAT	60	654	GTGTCACCGTCGTGGAAAG	60.15
Trial 3	Targets				90-200				
	Excluded Regions				650-300				
	Included Region				1-550				
	PrimerQuest	1	365	18	TTCGGTTACTGAGTTGCTGCCT	58.5	382	TTCCAGGTCACCTTGGTTCTGT	58.1
		2	386	19	TCGGTTACTGAGTTGCTGCCTT	58.5	404	TGGCCGTTGTGGACATCTGGTATT	60.4
	Primer3	1	378	1	GGTCCCTTTGCCTTTTGTTC	60.8	378	AGGTCACCTTGGTTCTGTGG	60
		2	356	23	TTACTGAGTTGCTGCCTTGG	59.1	378	AGGTCACCTTGGTTCTGTGG	60

With respect to reversed primer pairs, both software identified two positions. In trial 2 PrimerQuest selected the same position (309) to design forward primers whereas Primer3 identified the same positions (654) to design reversed primers. In trials 3 and 4 also, the same pattern of identification was observed (Results not shown). In trial number5, both programmes could design primers including the same region (from 622 to 995 bp) with five base pair variation.

In all the trials, Primer3 and PrimerQuest did not predict identical primers for the same sequence. However, in the trial 4 and 5, both programs identified the same primer-starting positions with one to seven base pair variation. In trial 1, 2, and 3, Primer3 identified various starting positions compared to the PrimerQuest. Thus, it can be stated that Primer3 has higher capability of designing various primers corresponding to target region. The resulted primers had the same starting position with a possibility of changing the primer length and *Tm*. As a result, selection of the best primer pair that corresponds to the target region has limited probability. PrimerQuest was the most user-friendly among the seven software programmes, and it gave primer sequence expression with respect to query sequences in all

resulted primer pairs. Furthermore, it provided quick link to BLAST check on the primer pairs other than the input template. Primer3 had only one output sequence expression for the first primer or for the best primer only. However, it did not facilitate linking with BLAST to check untargeted annealing of primer sequences.

CONCLUSIONS

In restriction analysis, amongst the 7 software compared, the NEB cutter provides higher user flexibility to manipulate digestion parameters and it is the only package that gives simulated agarose gel profile predictions at 0.7, 1.0, 1.4, 2.0 and 3.0 concentrations after the digestion. “Restriction Enzyme Digest of DNA”, having the facility to digest multiple query sequences, is useful when comparing two or more sequence simultaneously. JustBio-Cutter is useful more in procurement as it provides more detailed description than other tools about enzymes such as supplier’s information, commercial availability, and variations of isozymers.

With respect to primer-designing tools, Primer3 and PrimerQuest are the best for optimizing parameters according to comprehensiveness, interactive design, and user-friendliness. Primer3 has higher capacity to design primers closest to a given target sequence minimizing regions of unwanted annealing. PrimerQuest exhibited higher user-flexibility and interactive designing ability. BatchPrimer3 is a multi-purpose package and is versatile with its ability to design primers for RAPDs, SSR, SNPs and DNA sequencing.

REFERENCES

Abd-Elsalam, K.A. (2003). Bioinformatic tools and guideline for PCR primer design."*African Journal of Biotechnology* 02, Academic Journals: *2*(5), 91-99.

Batchprimer3. http://probes.pw.usda.gov/cgi-bin/batchprimer3/batchprimer3.cgi. (Accessed on 18.10.2012).

Brown, T.A. (1995). Gene Cloning, 5th ed. Chapman and Hall, New York, 179-192.

Cleaver. (2006). http://cleaver.sourceforge.net/. (Accessed on 30.09.2010).

Dieffenbach, C.W., Lowe, T.M.J., and Dveksler, G.S. (1995). General Concepts for PCR Primer Design. In: PCR Primer, A Laboratory Manual, Dieffenbach CW, Dveksler GS Ed., Cold Spring Harbor Laboratory Press, New York, 133-155.

Genefisher2. University of Bielefeld. (2009). http:// bibiserv.techfak.uni-bielefeld.de/ genefisher2/ (Accessed on 19.09.2010).

IDT PrimerQuest. http://eu.idtdna.com/Scitools/ Applications/ Primerquest/. (Accessed on 19.09.2010).

JustBio-Cutter. http://www.justbio.com/index. php?page. (Accessed on 19.09.2010).

Kwok, S., Kellogg, D.E., McKinney, N., Spasic, D., Goda, L., Levenson, C.A., and Sninsky, J. (1990). Effects of primer-template mismatches on the polymerase chain reaction: Human Immunodeficiency Virus1 model studies. *Nucleic Acids Research*18, 999-1005.

Restriction enzyme digest of DNA http://www.biophp.org/minitools/restriction_digest/demo.php . (Accessed on 01.10.2010) .

Sequence Extractor. http://www.bioinformatics. org/seqext/. (Accessed on 01.10.2010).

TAIR http://www.arabidopsis.org/index.jsp. (Accessed on 01.10.2010).

University of Waterloo (2007). WatCut, An on-line tool for restriction analysis, silent mutation scanning, and SNP-RFLP analysis.http://watcut.uwaterloo.ca/watcut/watcut/template.php?act=help&cid=4da1904531f74e5d2d864ea8535e7cff . (Accessed on 17.09.2012).

http://tools.neb.com/NEBcutter2/index.php. (Accessed on 30.09.2010).

http://www.ncbi.nlm.nih.gov. (Accessed on 30.04.2012).

http://dbb.nhri.org.tw/primer/index.html. Accessed on 19.09.2010).

http://frodo.wi.mit.edu/primer3/. (Accessed on 19.09.2010).

http://www.ncbi.nlm.nih.gov/ tools/primer-blast. (Accessed on 19.09.2010).

http://pride.molgen.mpg.de/pride.html. (Accessed on 19.09.2010).

http://www.restrictionmapper.org. (Accessed on 01.10.2011).