

BioParishodhana: A novel graphical interface integrating BLAST, ClustalW, primer3 and restriction digestion tools

Rajani Kanth Vangala*, Lucky Singh[§] & Ravi Prakash Gupta[§]

Institute of Bioinformatics and Applied Biotechnology; Electronics City, Bangalore, India; Head of Translational Research Thrombosis Research Institute 258/A, Bommasandra Industrial Area, Narayana Hrydayala Hospital, Anekal Taluk, Bangalore 560 099, India; Rajani Kanth Vangala – Email: rajani@triindia.org.in; Phone: +919880568987; *Corresponding author

§- authors contributed equally

Received May 25, 2012; Accepted June 16, 2012; Published July 06, 2012

Abstract:

Bioinformatics has emerged as an integral part of life sciences and biomedical research. The bioinformatics tools developed so far exist individually and do not cross talk leading biologists to spend more time in formatting the output from one tool as input for another tool. This leads to huge loss of time and cost. We herein have made platform which integrates the tools in a way that the output of one program can be directly used as input of another and does not need any modifications. Tools for similarity search, primer designing, and restriction enzyme digestion are required in almost all biological research; therefore we initially tried to integrate these tools. BioParisodhana platform optimizes the time spend in browsing and downloading applications and is an interactive, effective and user friendly.

Availability: <http://resource.ibab.ac.in/bioparishodhana.html>

Keywords: Integrative, BLAST, ClustalW, Restriction, Primer, BioParishodhana, Bioinformatics

Background:

The recent advances in life sciences and technology have started to produce a large amount of data in an extremely fast and cost-efficient way which requires the development of new algorithms, coupled with massively parallel computing. Besides, biologists are usually non-programmers, thus demanding intuitive computer applications that are easy to use by means of a friendly GUI. Several tools have been generated over the past decade in order to deal with the data generation but very little in integrating the tools and making biologist friendly interfaces. It is therefore of paramount importance to overcome such limitations, so that bioinformatics becomes much more widely used amongst biologists. Web based interfaces are good and popular as a biologist next to having a pipette and a pen, can access web browser and perform required analysis in a easy to use way. The main goal of our

ISSN 0973-2063 (online) 0973-8894 (print)
Bioinformation 8(13): 639-643 (2012)

project was to unify some of these existing bioinformatics applications and in one easy-to-use environment, independent of the computing platform, being a concentrator resource tool with a friendly interface allowing intuitive bioinformatics tool usage. Our platform BioParisodhana is a graphical interface integrating BLAST, ClustalW, Primer3 and restriction digestion tools allowing even a non-programmer laboratory scientists to chain different processes into workflows and customize them without code writing.

Implementation:

Architecture

We divide the task in three developmental phases and primarily use standalone versions of our chosen tools; BioPERL, PERL and CGI-PERL for achieving our goal **Table 1 (see supplementary material)**.

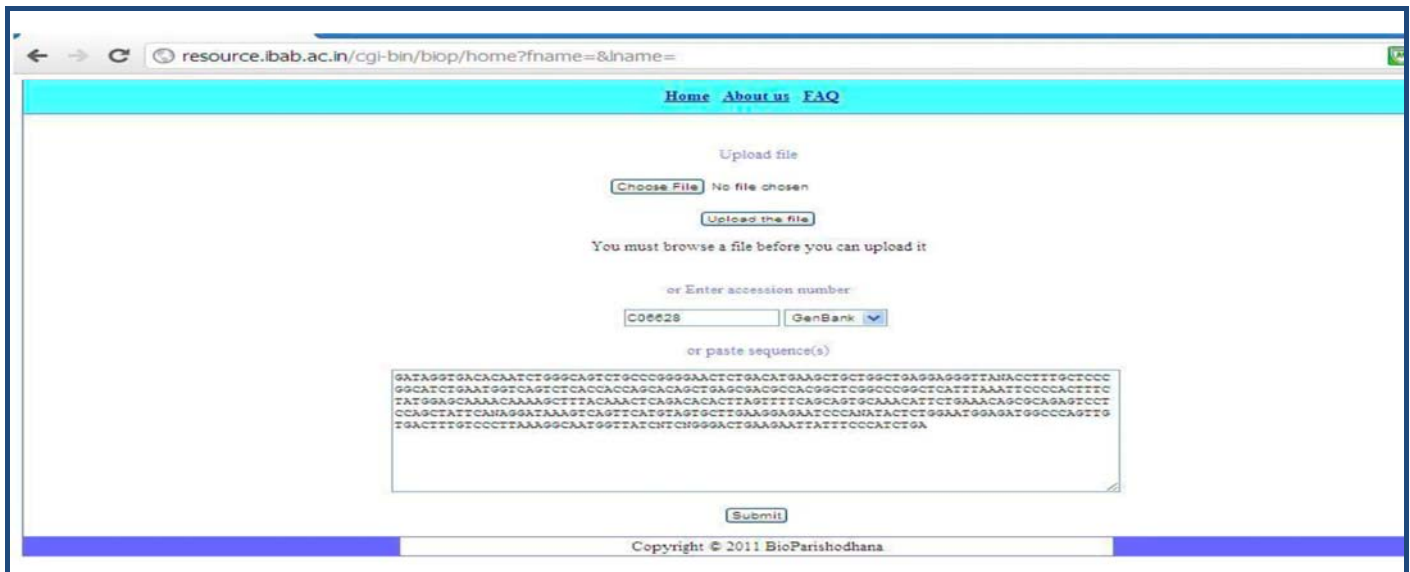


Figure 1: Front page for the user where a sequence data can be uploaded as a file or cut and paste or give the accession number/ID

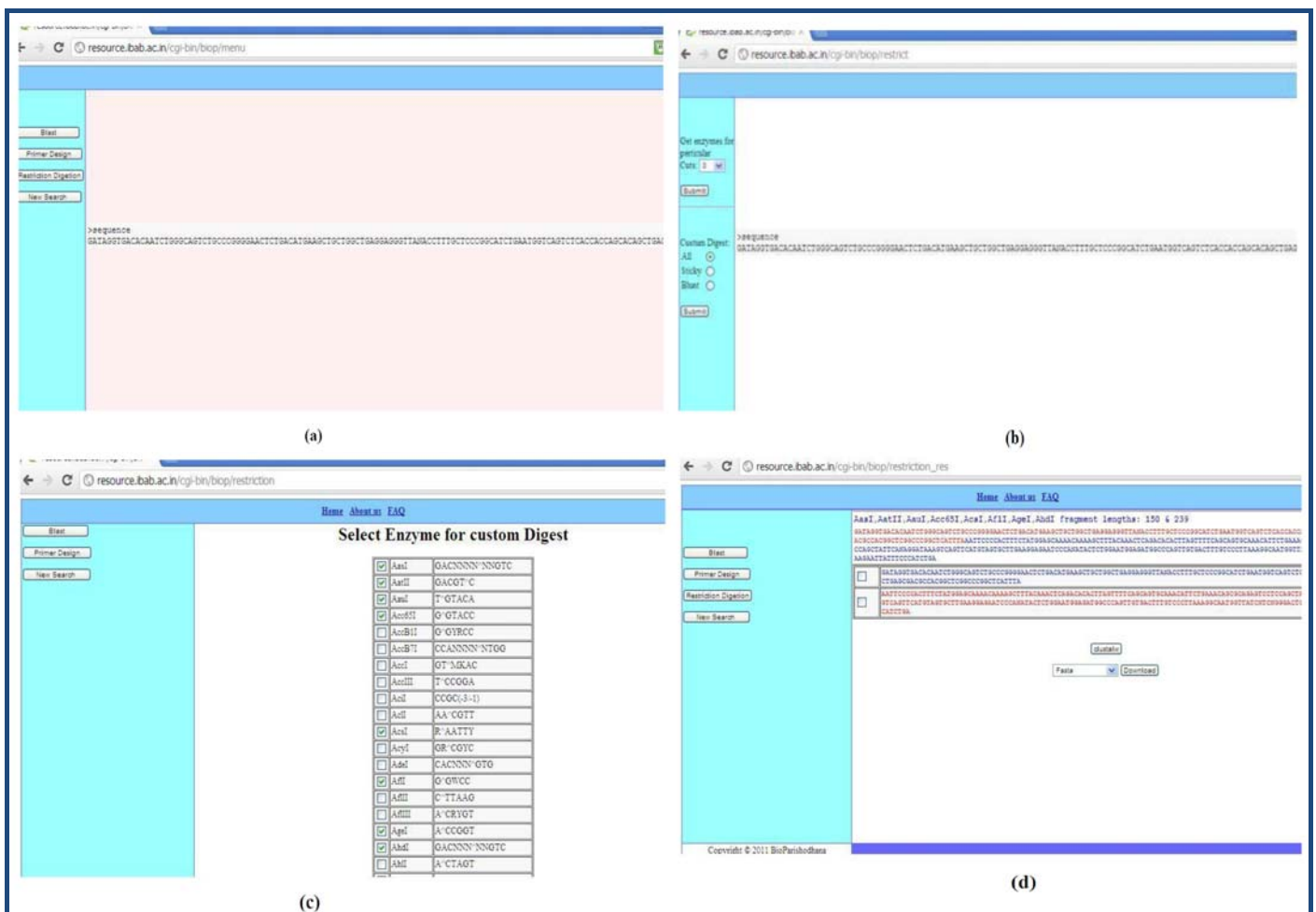


Figure 2: (a) After uploading the sequence, all the options will appear on the left side of the screen and main work area will have the input sequence. For example, restriction enzyme analysis of DNA; (b) The user will see a screen with all the required options on the left side and sequence in the work area; (c) The next screen will give all the restriction enzymes (depending on the selection made by the user: Sticky ends, Blunt ends or All), for example, sticky ends were selected. In this screen the user can select specific enzymes to cut with; (d) Depending on the selection of the enzymes, the digested fragments will be shown in the work area and all the options will be shown in the left part of the screen.

Input Phase

The first step in interlinking the tools is to make the input and output file formats compatible throughout. The input can be given in any of the following three formats: Upload, Paste and Accession or ID number. For any given input Bio::SeqIO module from BioPERL [1] was standardized to convert to FASTA format **Table 2** (see supplementary material).

Process Phase

BLAST [2], ClustalW [3], Primer3 [4] and a common restriction digestion tools were chosen for integration using perl coding. Of the 4 tools only BLAST requires database storage, to avoid data storage issues we use chose RemoteBlast **Table 3** (see supplementary material). For restriction digestion tool bioperl

module Bio:Restriction:Analysis is implemented **Table 4** (see supplementary material).

Integration Phase

A web interface (**Figure 1**) is created which serves to multiple users simultaneously, avoiding clashes in multiple queries and results. To achieve this, a random number is generated when the server receives the first request or home.cgi is accessed. This number becomes the unique id of the user till the session lasts and differentiates the individual users and creates a file with its reference for the user. The interlinking of the tools was done using PERL programs. We use CGI-PERL for achieving this task.

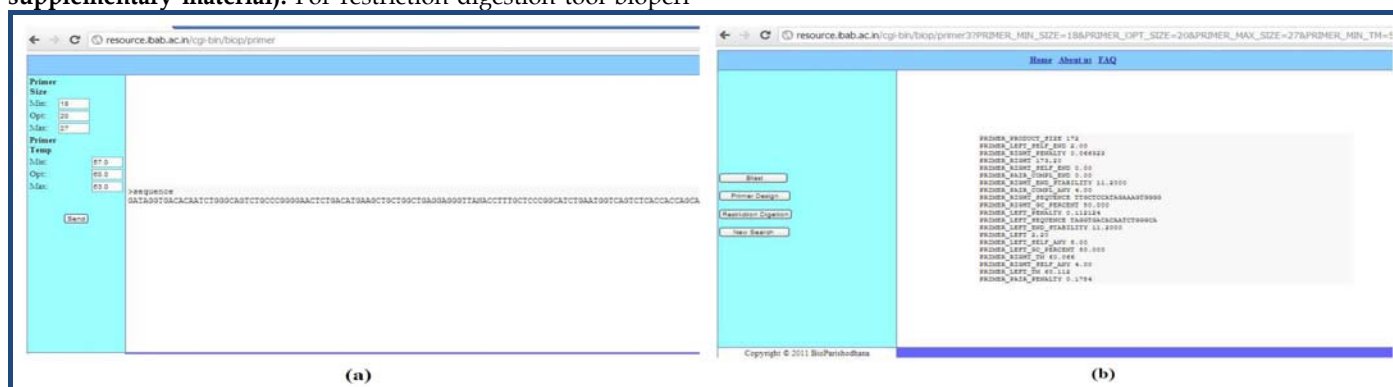


Figure 3: (a) In case of primer design, the required parameters will be shown after the input and sequence in the work area; (b) After identifying the primers the results will be displayed and all the other options which can be performed with the results.

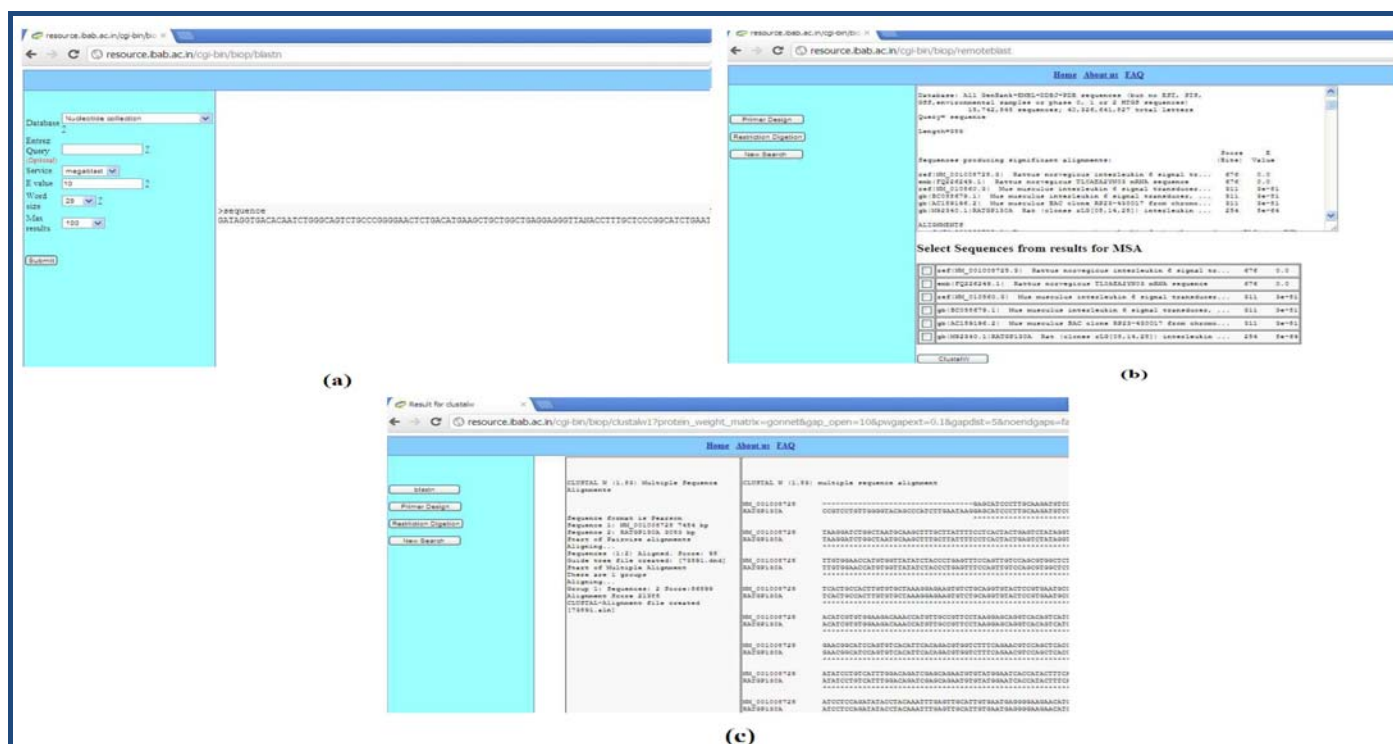


Figure 4: (a) The same sequence can be used for T-BLAST analysis and all the parameters will appear on the left hand side of the screen and sequence in the work area; (b) After selecting the parameters, the result of the BLAST will be shown and the user can perform the ClustalW to generate the family of the sequences or domain identification. Other options like primer design and restriction enzyme analysis can be performed for single sequence selected or multiple sequences from this output; (c) For further analysis, example, ClustalW, the results will be displayed as shown above and can be downloaded in different formats.

Discussion:

Restriction enzyme analysis and intuitional use of other tools: As seen in the (Figure 2b), the input sequence will be taken in the work area and required parameters which can be selected by the user are seen on the left side menu. For example, upon selecting the Sticky ends, a list of enzymes will be generated in the next screen (Figure 2c) and user can select the required enzymes and submit. This will generate a list fragments and their sizes with the sequence information (Figure 2d). This step can also give the user to go for specific fragment selection for primer design or further restriction digestion or perform BLAST analysis or also see if these two sequences can share any homology by performing ClustalW. The user can also download the sequences in the fasta format. Primer design and collaborative tools: The input sequence can be used for primer design and also a combinational analysis. Figure 3a shows the sequence and parameters which can be used for designing the primers (Figure 3b) and furthermore, all the sequences can be used to perform further BLAST, ClastalW and any other analysis needed. BLAST analysis with extended analytical capabilities: (Figures 4a, b, & c) show how BLAST tool can be used and in collaborative way with ClastalW. However, as it can be seen in (Figure 4b), apart from ClastalW, there are other tools like Primer design and restriction digestion can be performed on each of the hit and thus allowing the user to implement novel ideas and generate better hypothesis and analysis for their research. Some of the earlier tools like BioExtract [5] Discovery Net [6] and in GAP [7], where the tools exist as individual tools and work flows can be created by the

users have provided lot of help. We believe that BioParishodhana platform can give the much needed simpler and easier approaches for biologists to use the tools and analyze the data without having to download, learn and use the tools.

Requirements:

Project name: BioParishodhana; Project home page: <http://resource.ibab.ac.in/bioparishodhana.html>; Operating system(s): Linux; any restrictions to use by non-academics: none

Acknowledgment:

This work was supported by Department of Information Technology, Government of India to Institute of Bioinformatics and Applied Biotechnology (IBAB), Bangalore, India. Authors also thank faculty at IBAB who have supported and helped during the project.

References:

- [1] Stajich J *et al. Genome Res.* 2002 **12**: 1611 [PMID: 12368254]
- [2] Altschul SF *et al. J Mol Biol.* 1990 **215**: 403 [PMID: 2231712]
- [3] Chenna R *et al. Nucleic Acids Res.* 2003 **31**: 3497 [PMID: 12824352]
- [4] Steve Rozen & Helen J, *Humana Press.* 2000 pp 365-386
- [5] Lushbough CM *et al. Int J Comput Biol Drug Des.* 2008 **1**: 302 [PMID: 20054995]
- [6] Rowe A *et al. Bioinformatics.* 2003 **19**: 225 [PMID: 12855463]
- [7] Qi J *et al. Bioinformatics.* 2009 **26**: 127 [PMID: 19880367]

Edited by P Kanguane

Citation: Vangala *et al.* Bioinformation 8(13): 639-643 (2012)

License statement: This is an open-access article, which permits unrestricted use, distribution, and reproduction in any medium, for non-commercial purposes, provided the original author and source are credited

Supplementary material:

Table 1: Architecture of the work and integrating of the tools

Phases	Task	Source Code
Input Phase	Analyzing Input and its compatibility with all tools	Bio PERL
Process Phase	1. Blast	BioPERL
	2. ClustalW	ClustalW standalone
	3. Restriction Enzyme/Site	BioPERL
	4. Primer3	Primer3 standalone
Integration Phase	Interlinking of tools	PERL
Front End Phase	Interactive & User friendly platform	CGI-PERL

Table 2: Input phase: Sequence input

Module	Required Input	Output
Bio::SeqIO	File & File type	FASTA file

Table 3: For blast bioperl module Bio::Tools::Run::RemoteBlast is implemented.

Module	Required Input	Output
Bio::Tools::Run::RemoteBlast	All blast parameter respective to their type	Blast result in text file

Table 4: Process phase: Restriction analysis

Module	Required Input	Output
Bio::Restriction::Analysis	Bio::PrimarySeq object Name of Enzyme	Digested fragments