

BIOFFORC: Tool development for biological file format conversion

Chinnaiah Swaminathan Vinobha¹, Maruthamuthu Rajadurai^{2,*} and Ekambaram Rajasekaran³

¹P.G. and Research Department of Bioinformatics, PRIST University, Thanjavur; ²P.G. Department of Biotechnology and Bioinformatics, Bishop Heber College, Tiruchirappalli; ³Department of Biotechnology, Periyar Maniammai University, Vallam, Thanjavur; Maruthamuthu Rajadurai* - Email: mr.rajadurai@gmail.com; * Corresponding author

received July 23, 2008; revised September 02, 2008; accepted September 13, 2008; published October 25, 2008

Abstract:

The use of bioinformatics tools require different sequence formats at various instances. Every tool uses specific set of formats for processing. Sequence in one format is often required in another format. Thus, there is a need for sequence format conversion. A number of such tools are available in the public domain. Here, we describe BIOFFORC as a file format converter. The tool is developed with a graphical user interface in PERL.

Availability: <http://www.winningpath.com/biofforc/>

Keywords: bioinformatics; sequence; format; read; write; conversion; tools

Background:

The basic format underlying information in DDBJ/EMBL/GenBank is a flat file. The correspondence between individual flat file formats facilitated the exchange of data between each of these datasets [1]. GenBank describes each sequence entry with literature reference, functional data, location of mRNA coding regions and sequence [2]. Similarly, EMBL [3] and DDBJ [4] are also resources for biological and medical research data. The

sequence file format conversion with tools like READSEQ [5], FMTSEQ [6] and SeqVerter [7] are described in detail. Here, we describe BIOFFORC as a file format converter.

Methodology:

Process flow

A process flow for the tool is shown in Figure 1.

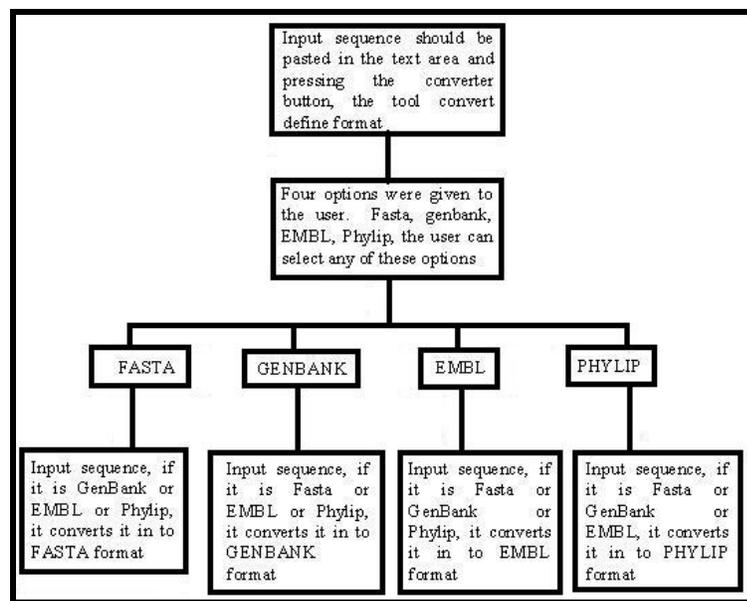


Figure 1: Process flow diagram for BIOFFORC tool.

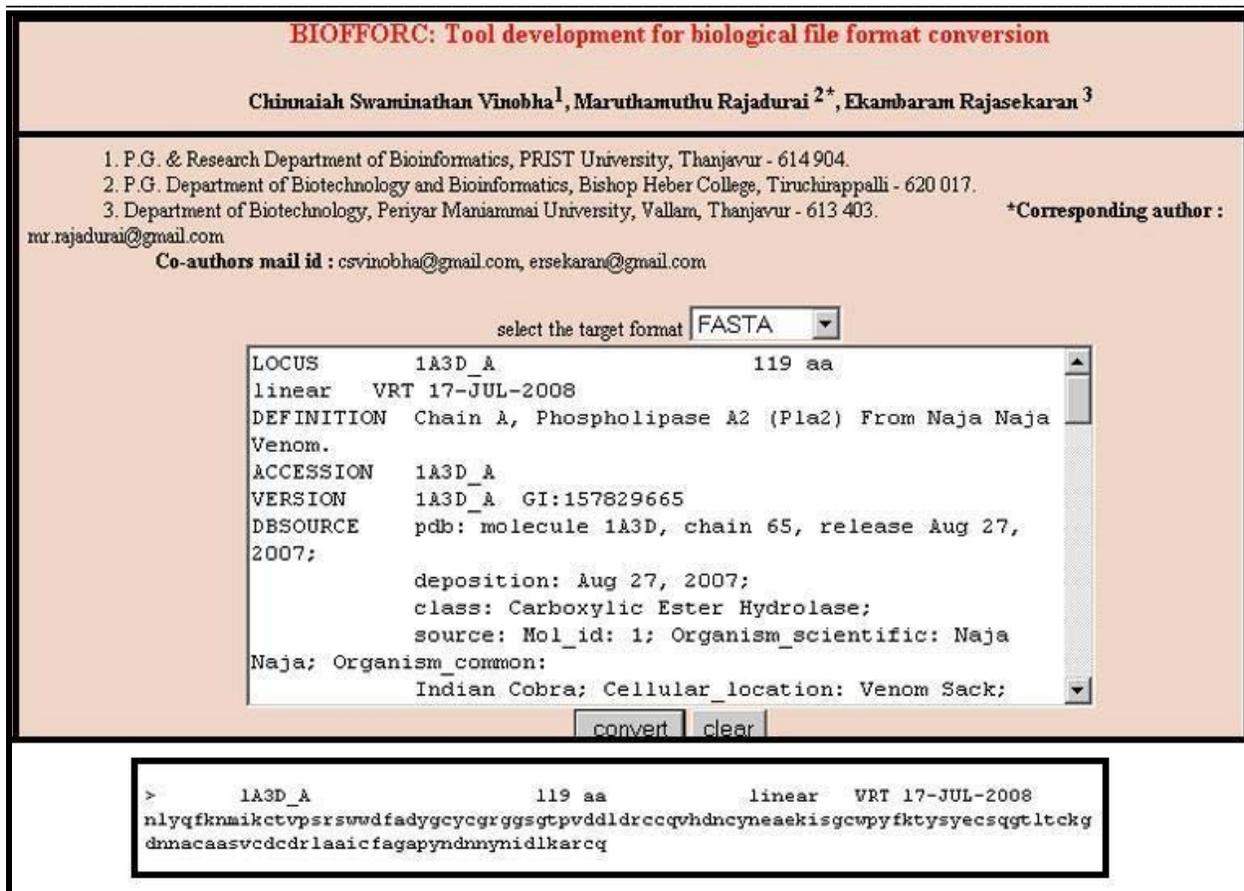


Figure 2: Sample input and output for and from BIOFFORC is shown.

Web interface:

The current version of BIOFFORC is a web based tool and it uses a common gateway interface (CGI) developed in PERL.

Input and output:

The use of the tool to convert a GenBank format to FASTA format is shown in Figure 2. The top panel in Figure 2 shows input in GenBank Format and the bottom panel shows output in FASTA format.

Caveats and future development:

The present version of BIOFFORC allows format conversion for four formats. We propose to expand conversion capability to several other required formats.

References:

- [01] <http://www.ncbi.nlm.nih.gov/>
- [02] D. A. Benson *et al.*, *Nucleic Acids Res.*, 36: D25 (2008) [PMID: 18073190]
- [03] P. Sterk *et al.*, *Methods Mol Biol.*, 406: 1 (2007) [PMID: 18287686]
- [04] <http://www.ddbj.nig.ac.jp/>
- [05] D. Gilbert, *Curr. Protocol Bioinformatics*, (2003) [PMID: 18428689]
- [06] <http://www.bioinformatics.org/JaMBW/1/2/>
- [07] <http://www.genestudio.com/seqverter.htm>

Edited by P. Kanguane

Citation: Vinobha *et al.*, *Bioinformatics* 3(2): 98-99 (2008)

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.