

Modelling Of Two Stages DNA Splicing Languages On De Bruijn Graph

Mohammad Hassan Mudaber^a, Yuhani Yusof^b, Mohd Sham Mohamad^b, Aizi Nor Mazila Ramli^b, Wen Li Lim^b

^aDepartment of Mathematics, Faculty of Natural Sciences, Kabul Education University, Afshar District, Kabul, Afghanistan

^bFaculty Industrial Sciences and Technology, Universiti Malaysia Pahang, Lebuhraya Tun Razak, 26300 Gambang, Kuantan Pahang Darul Makmur, Malaysia

ABSTRACT

Finding the sequence of the genome from its compositions as well as a mathematical graph is the most interesting topic in a field of DNA molecular. Since lack of technology is the big obstacle that biologists are facing to read a long sequence of the genome from beginning up to the end, therefore finding the compositions of the genome having very long sequence and also its description via de Bruijn graph is challenging or even impossible. In this paper, Yusof-Goode (Y-G) approach is used to generate the DNA splicing languages based on cutting sites of initial strings (one or two cutting sites) and crossing and contexts factors of restriction enzymes. The two short sequences of DNA (8bp) and two restriction enzymes are considered to create a connection between mathematics and DNA molecular. This relation will be presented as de Bruijn graph so that every edge of the de Bruijn graph gives a k-mer composition of DNA molecule and also each path of the de Bruijn graph gives a DNA sequence and vice-versa. Besides, the persistency and permanency of two stages DNA splicing languages can be predicted using this model.

KEYWORDS: de Bruijn graph, permanent, persistent, two stages splicing languages, Y-G splicing system

DOI: <http://dx.doi.org/10.11113/jt.v78.4236>