Creation, Innovation, Technology & Research Exposition (CITReX 2014), UMP, 05-06 March 2014

DNA SPLICING SYSTEM PREDICTOR

MOHAMMAD HASSAN MUDABER, YUHANI YUSOF AND MOHD SHAM MOHAMAD
Futures & Trends Research Group
Faculty of Industrial Sciences & Technology
Universiti Malaysia Pahang
Lebuhraya Tun Razak, 26300 Gambang,
Kuantan, Pahang Darul Makmur.

ABSTRACT

A user friendly interface by C# is constructed on showing the behavior of resulted DNA molecules. Up to two stages DNA splicing system, this system works based on the developed mathematical theorems by using Yusof-Goode (Y-G) approach. As a new innovation in DNA molecular, this system is able to predict the persistency and permanency of the DNA splicing system. On the biological side, this software replaced wet-lab, which optimize time and money.

INTRODUCTION

- Deoxyribonucleic acid (DNA) is a molecule constructing from nucleotides. Nucleotides consist of three main parts: nitrogenous bases (adenine, guanine, cytosine and thymine), deoxyribose sugar and phosphate group.
- The four bases found on DNA are adenine (*A*) guanine (*G*), cytosine (*C*) and thymine (*T*). There are two hydrogen bond between adenine and thymine and three hydrogen bond between guanine and cytosine. In the other word, adenine is only paired with thymine and guanine is paired by cytosine and vice-versa. This rule of pairing is written as [*A*/*T*],[*C*/*G*],[G/*C*] and [*T*/*A*].
- The structure of DNA is like a double helix, with sugar-phosphate (backbone) is on the outsides and the bases pairs are on the inside of the double helix (Weaver, 2005).
- Following the Watson-Crick base-pairing rules, and since the two strands of DNA are antiparallel, the double-stranded DNA looks like below.

  $5'...CGAGCTCG...3'$

  $3'...GCTCGAGC...5'$
- A restriction enzyme is an enzyme that recognizes specific DNA sequence usually 4-6 base pair in length and then cut it in a particular manner (sticky ends or blunt ends) . An important characteristic of sticky ends is that those produced from different molecules by the same enzyme are complementary. Then, the fragments of DNA re-join with their complementary ends by ligase and generate new DNA molecules (Walker and Rapley, 2009).

PRELIMINARIES

Definition 1: (Yusof, 2012) Yusof-Goode Splicing system

If $r \in R$, where $r = (a,x,b:c,x,d)$ and $s_1 = \alpha axb\beta$ and $s_2 = \gamma cxd\delta$ are elements of I, then splicing s_1 and s_2 using r produce the initial string I together with $\alpha axd\delta$ and $\gamma cxb\beta$, presented in either order where $\alpha,\beta,\gamma,\delta,a,b,c$ and $d \in A^*$ are free monoid generated by A with the concatenation operation and 1 as the identity element.□

Definition 2: (Head, 1987) Persistent

Let $S = (A,I,R)$ be a splicing system. Then S is persistent if for each pair of strings *ucxdv* and *pexfq*, in A^* with (c,x,d) and (e,x,f) patterns of the same hands: if y is a sub segment of *ucx* (respectively *xfq*) that is crossing of a site in *ucxdv* (respectively *pexfq*) then this same sub segment y of *ucxfq* contains an occurrence of a crossing of a site in *ucxfq*.□

Definition 3: (Gatterdam, 1989) Permanent

Let $S = (A,I,R)$ be a splicing system. Then S is permanent if for each pair of strings *ucxdv* and *pexfq*, in A^* with (c,x,d) and (e,x,f) patterns of the same hands: if y is a sub segment of *ucx* (respectively *xfq*) that is crossing of a site in *ucxdv* (respectively *pexfq*) then this same sub segment y of *ucxfq* is an occurrence of a crossing of a site in *ucxfq*.□

Definition 4: (Yusof, 2012) Palindromic

A string I of double stranded deoxyribonucleic acid (dsDNA) is said to be palindromic if the sequence from the left side of the upper single strand is equal with the sequence from the right side of the lower single strand.□

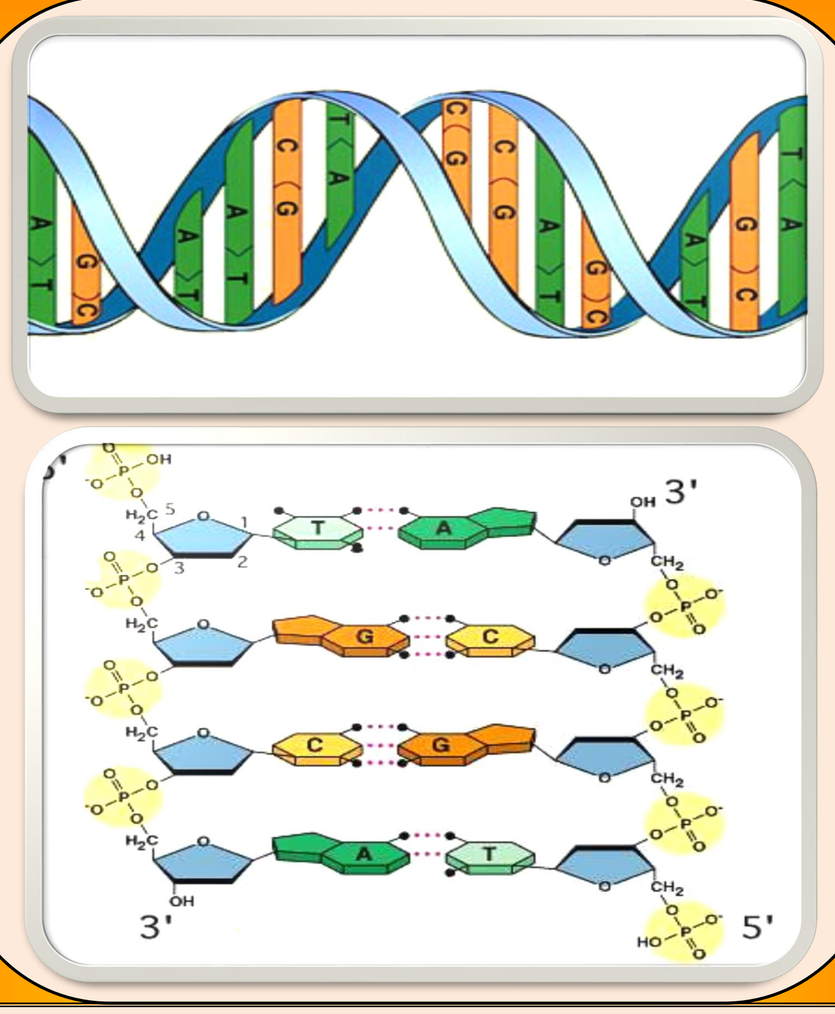

Figure 1: The Structure of Double-Stranded DNA

METHODOLOGY

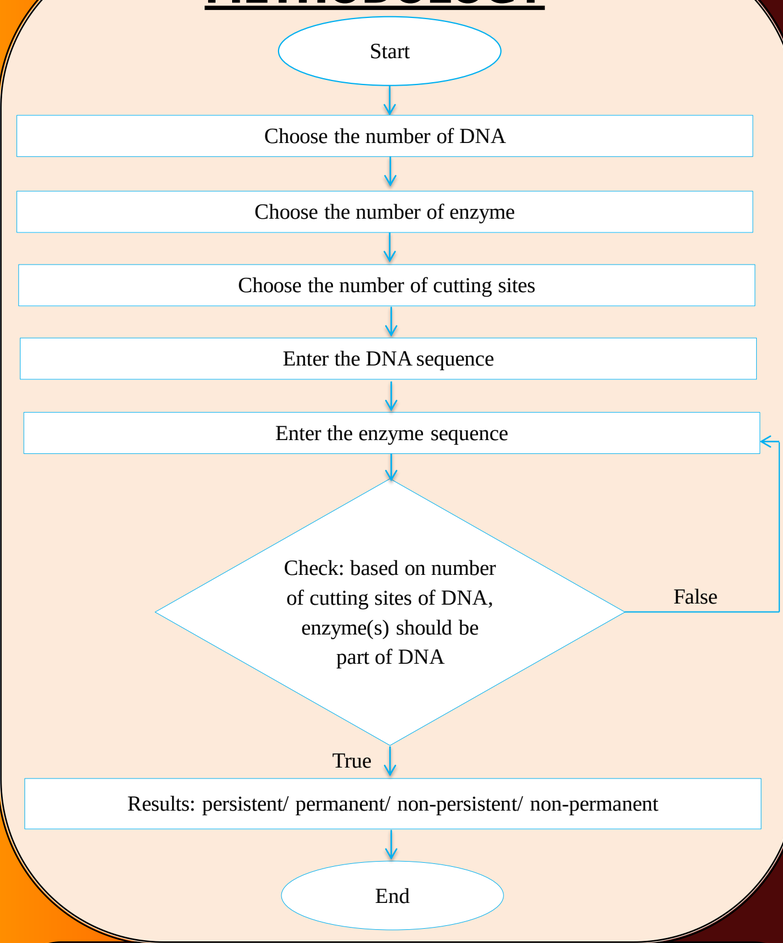

Figure 2: Flow Chart of Research Methodology

RESULTS

This program, shows the persistency and permanency of DNA splicing system based on the following mathematical theorems:

Theorem 1: The set of two stages splicing languages which is produced by Y-G splicing system consisting two initial strings (with one cutting sites) and two rules with disjoint crossing sites and palindromic sequences is persistent and permanent.

Theorem 2: The set of two stages splicing languages which is produced by Y-G splicing system consisting two initial strings (with one cutting sites) and two rules with disjoint non-palindromic crossing sites (or sequences) is persistent and permanent.

Theorem 3: The set of two stages splicing languages which is produced by Y-G splicing system consisting two initial strings (with one cutting sites) and two rules where the whole sequence of one rule is palindromic and the other is non-palindromic is persistent and permanent.

Theorem 4: The set of two stages splicing languages, that is produced by Y-G splicing system consisting two initial strings (with two cutting sites) and two rules with disjoint crossing sites and non-palindromic sequences), is persistent and permanent.

Theorem 5: The set of two stages splicing languages, which is generated by Y-G splicing system consisting two initial strings (with two cutting sites) and two rules with disjoint crossing sites and palindromic sequences, is persistent and permanent.

Theorem 6: The set of two stages splicing languages that is produced by Y-G splicing system consisting two initial strings (with two cutting sites) and two rules where the crossing of one rule is palindromic and the crossing of the other rule is non-palindromic, is persistent and permanent.

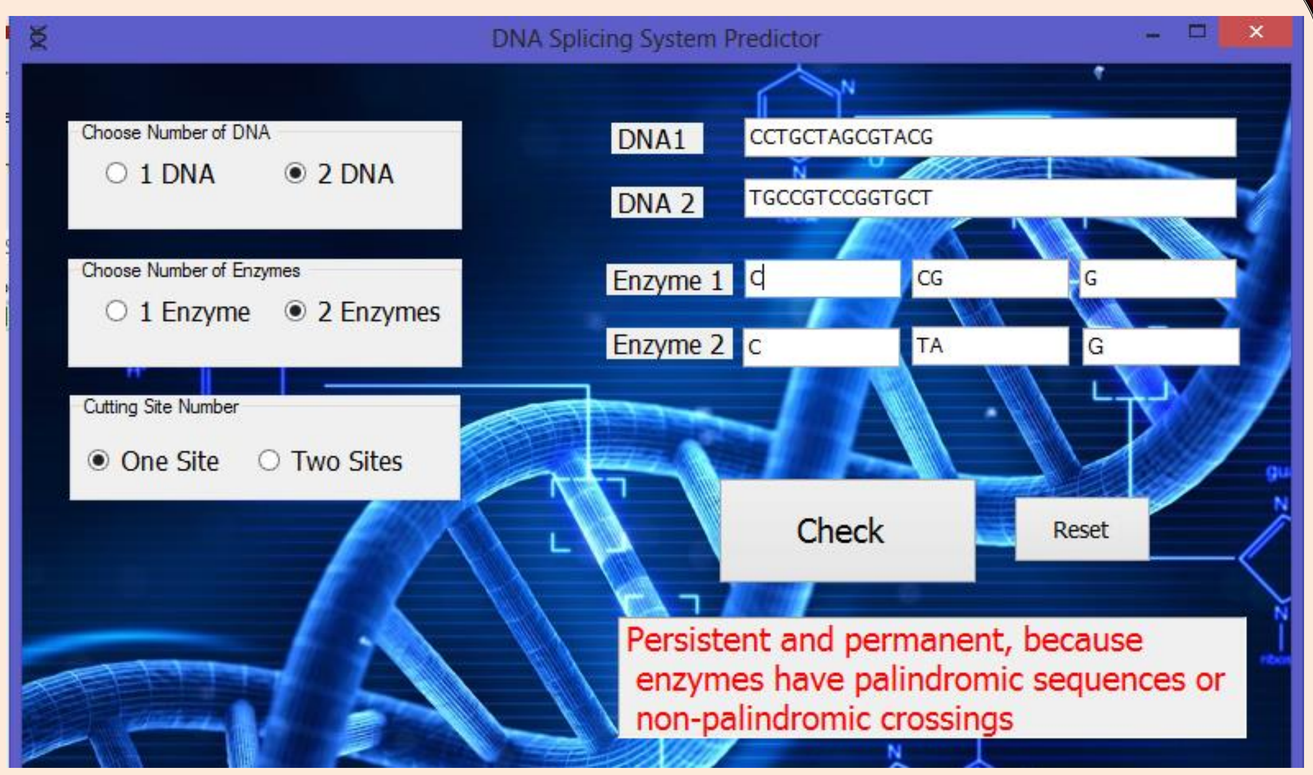

Figure 3: User Interface of DNA Splicing System Predictor

CONCLUSION

As a conclusion, this system works on predicting the persistency and permanency of DNA splicing system which benefits to certain organization such as : NEB

ACKNOWLEDGEMENT

All authors would like to thank Research and Innovation Department of UMP for funding this research through RDU Vote: 130354.

REFERENCES

Gatterdam, R.W.1989. Splicing System and Regularity. International Journal of ComputerMath. 31. 63-67.

Head, T. 1987. Formal Language Theory and DNA: An Analysis of the Generative Capacity of Specific Recombinant behaviors. *Bulletin of mathematical biology*, *49*(6), 737–759.

Weaver, R. F. 2005. *Molecular Biology*, New York: McGraw-Hill.

Walker, M.J and Rapley. R. 2009. Molecular Biology and Biotechnoloogy.5th edition. RSC Publishing.

Yusof, Y.2012. DNA Splicing System Inspired by Bio Molecular Operation. Ph.D Thesis. Universiti Teknologi Malaysia.