Jurnal
Teknologi

Full Paper

PERSISTENCY AND PERMANENCY OF TWO STAGES SPLICING LANGUAGES VIA YUSOF-GOODE APPROACH: TWO INITIAL STRINGS AND TWO RULES

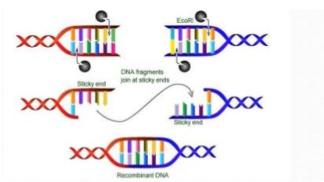
Article history
Received
21 November 2014
Received in revised form
7 July 2015
Accepted
1 October 2015

Mohammad Hassan Mudaber*, Yuhani Yusof, Mohd Sham Mohamad, Lim Wen Li

Faculty of Industrial Sciences & Technology, Universiti Malaysia Pahang,Lebuhraya Tun Razak, 26300 Gambang, Kuantan Pahang Darul Makmur, Malaysia

*Corresponding author
hmudaber@yahoo.com

Graphical abstract



Abstract

The notation of representing restriction enzymes in the form of double-triple in order to formulate Yusof-Goode (Y-G) splicing system was mathematically proposed by Yusof in 2012. The aim of introducing Y-G splicing system was to study the process of recombinant deoxyribonucleic acid or DNA strand in a translucent way. In real situation, when the recombination action occurs, the recombinant DNA strands which will arise often contain the patterns of the restriction enzymes. Persistency and permanency are two properties of splicing system, which show whether the recombinant DNA strands will be split by the existence of restrictions enzymes or not if the reaction goes to the second stage. In this research, the persistency and permanency of two stages splicing languages according to the number of cutting sites of initial strings as well as crossing sites and context factors of splicing rules are investigated. Therefore, a Y-G splicing system consisting of two initial strings (with two cutting sites) and two rules is used to present the above properties of two stages DNA splicing languages.

Keywords: Y-G splicing system, two stages splicing languages, persistent, permanent

Abstrak

Tatatanda untuk mewakili enzim pembatas dalam bentuk gandaan-tiga bagi memformulasikan sistem hiris-cantum Yusof-Goode (Y-G) secara matematik telah dicadangkan oleh Yusof pada tahun 2012. Tujuan memperkenalkan sistem hiris-cantum Y-G adalah untuk mengkaji proses rekombinan asid deoksiribonukleik atau rantaian DNA secara telus. Dalam situasi sebenar, apabila tindakan penggabungan berlaku, rantaian rekombinan DNA yang akan terhasil sering mengandungi bentuk enzim pembatas. Konsep berterusan dan kekal adalah dua sifat sistem hiris-cantum, yang menunjukkan sama ada rantaian rekombinan DNA akan berpisah dengan kewujudan enzim pembatas atau tidak jika tindakbalas menuju ke peringkat dua. Dalam kajian ini, konsep berterusan dan kekal bagi bahasa hiris-cantum dua peringkat mengikut bilangan bahagian pemotongan bagi jujukan permulaan serta lintasan dan faktor konteks bagi peraturan hiris-cantum dikaji. Oleh itu, sistem hiris-cantum Y-G yang terdiri daripada dua jujukan awal (dengan dua bahagian pemotongan) dan peraturan digunakan untuk mempersembahkan ciri-ciri dua peringkat bahasa hiris-cantum DNA di atas.

Kata kunci: Sistem hiris cantum Y-G, bahasa hiris cantum dua peringkat, berterusan, kekal

1.0 INTRODUCTION

Deoxyribonucleic acid (DNA) is a sequence of nucleotides consisting of three basic parts: phosphate group, deoxyribose sugar and nitrogenous base. The nitrogenous bases are categorized into purines (adenine, guanine) and pyrimidynes (cytosine and thymine). According to consensus, the notations *A*, *G*, *C* and *T* are used for representing the four bases, which are the abbreviations of adenine, guanine, cytosine and thymine, respectively. Double-stranded DNA (dsDNA) is formed by two types of chemical bonds, named phosphodiester bonds and hydrogen bonds. Nucleotides are connected by phosphodiester bonds, and the two strands of DNA are joined by hydrogen bonds. There are two hydrogen bonds between *A* and *T* and three hydrogen bonds between *G* and *C* and vice-versa [1]. Restriction enzymes (RE) are proteins, which cut the strand of DNA from specific sequences (restriction sites) in different ways, resulting in molecules with staggered or blunt ends [2]. Then the fragments of DNA with its complementary can combine by the existing ligase to produce new DNA strands. There are different types of restriction enzymes (RE) that have been isolated from bacteria and play a role as cell defence against invading viruses into the host's genome. In reality, the RE cleaves the genome site involving invading bacteriophages into fragments and destroys it before interpolating into host genome and indwelling the cell [3].

The mathematical modelling of Splicing system was first formulated by Head [4] under a framework of formal language theory. In addition, the concept of persistency in splicing system was discussed and its definition was presented. Besides, the equalities of persistent splicing languages and strictly locally testable (SLT) languages were shown. Some sufficient conditions for splicing languages to be strictly locally testable languages were provided [5]. It has been shown by Karimi that every finite language is an SLT language [6]. Then, another important property of splicing system, which is the permanency of splicing system was introduced by Gatterdam [7]. Despite the persistency and permanency of splicing system are two different properties, but the splicing languages that are produced by these two splicing languages are equal [8]. Some sufficient conditions for the persistency and permanency of splicing system were provided [9, 10]. Recently, Yusof [11] introduced a new variant of splicing system named Y-G splicing system. In addition, by using the notation of Y-G splicing rule, the non-semi simple splicing system was defined as well as the persistency and permanency of this splicing system were studied and discussed [12]. Mudaber [13] investigated on the persistency and permanency of two stages splicing languages and provided some sufficient conditions with respect to two initial strings (with one cutting site) and two rules for splicing languages to be persistent and permanent. In addition, the relations between stage one and stage two splicing languages were investigated based on the number of cutting sites of initial strings and the crossing sites properties of splicing rules using the Y-G approach [14].

In this study, some properties of two stages splicing languages are investigated based on two initial strings (with two cutting sites) and two rules (the crossing site of the first rule is palindromic and the crossing site of the second rule is non-palindromic) in terms of persistency and permanency presented as lemmas, theorem and corollary.

In the next section, the preliminaries related to this research are given.

2.0 PRELIMINARIES

This section contains definitions of Y-G splicing system, persistent and permanent. In this investigation, the persistent and permanent properties of splicing system are studied using Y-G model in order to ensure the results from this study are applicable and beneficial in molecular biology. Therefore, since Y-G approach is used in this research, its definition is first given.

Definition 2.1: [11] Yusof-Goode Splicing System
If $r=(a,x,b:c,x,d)\in R$ 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 monoids generated by A with the concatenation operation and 1 as the identity element.

Next, the original definition of persistent, which was proposed by Head based on a biological problem in order to present the recombination process mathematically, is viewed.

Definition 2.2: [4] Persistent
Let $S=(A,I,B,C)$ 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 $ucxfq$.

The definition of persistent according to Y-G notation can be written as below.
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 $(c,x,d:e,x,f)$ 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 $ucxfq$.□

The concept of permanent was introduced by Gatterdam[7] as a proper case of persistent based on the ideas from DNA recombination. However, in the case of considering two initial strings and two rules to form DNA hybrid, persistent and permanent are equal as investigated in this research. Thus, the definition of permanent is given next.

Definition 2.3: [7] Permanent
Let $S=(A,I,B,C)$ 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 $ucxfq$.

In the following section, the results and discussions of this research are presented.

3.0 RESULTS AND DISCUSSION

As In this section, the main results from this investigation are presented. In the world of biology, particularly in terms of DNA recombination, predicting the persistency and permanency of recombinant DNA fragments by mathematical theorems are the most interesting issues. The process of second stage recombinant DNA molecules can be achieved if the splicing system has a persistent property. Naturally, the sequences of all restriction enzymes, which are chosen from New England Biolabs (NEB) are not palindromic. Thus, after acting RE on dsDNA, the generated DNA templates are not always persistent as well as permanent. In other words, the sequence of RE does not match with DNA strands and in this situation; RE cannot recognize any specific nucleotides on DNA strands. Therefore, the persistency and permanency of two stages splicing languages are investigated by mathematical lemmas, theorem and corollary. Additionally, an example is provided to support the theorem in the biological point of view. First of all, the two stages splicing languages is defined below.

Definition 3.1: Two Stages Splicing Languages
Let $S=(A,I,R)$ be a splicing system. Furthermore, let $L(S)$ be the set of stage one splicing languages produced by splicing system S and $L'(S)$ be the set of stage two splicing languages produced by S that consists of $L(S)$ and all splicing languages that can be resulted by splicing $L(S)$. Then, the union of stage one and stage two splicing languages are called two stages splicing languages.

In the following lemma, the persistency of two stages splicing languages with respect to two initial strings and two rules, where the first rule cuts the first initial string and the second rule cuts the second string, is investigated.

Lemma 3.1: Let $S=(A,I,R)$ be a Y-G splicing system consisting of two initial strings (with two cutting sites), and two rules, where the crossing site of first rule is palindromic and the crossing site of the second rule is non-palindromic. If the first rule cuts the first initial string and the second rule cuts the second string at two specific places, then the set of two stages splicing languages that is produced by Y-G splicing system is persistent.□

Proof: Suppose $S=(A,I,R)$ is a Y-G splicing system consisting two initial strings $s_1,s_2\in I$ and two rules $r_1,r_2\in R$. Thus, the rules $r_1,r_2\in R$ are presented as $r_1=(a_1,a_1a_2,a_2:a_1,a_1a_2,a_2)$ and $r_2=(b_1,b_1b_2,b_2:b_1,b_1b_2,b_2)$ such that a_1 is complementary with a_2, b_1 is not complementary with b_2 and vice-versa and $a_1,a_2,b_1,b_2\in A^*$. Let s_1 and s_2 be initial strings in A^*, which have the forms $s_1=\alpha a_1a_1a_2a_2a_1a_1a_2a_2\beta$ and $s_2=\gamma b_1b_1b_2b_2b_1b_1b_2b_2\delta$, where the strings α and α', β and β' are complementary, $[\alpha/\alpha']$ and $[\beta/\beta']$, and s_1 can be cut by r_1 and s_2 can be cut by r_2 at two specific places, Using Y-G splicing system and applying rule r_1 on s_1 and rule r_2 on s_2 the following 4 DNA splicing languages will be generated at stage one, namely

$$\alpha a_1(a_1a_2a_2a_1)^k a_1a_2a_2\beta,\ \alpha a_1(a_1a_2a_2a_1)^k a_1a_2a_2\alpha'$$
$$\beta' a_1(a_1a_2a_2a_1)^k a_1a_2a_2\beta,\ \gamma b_1(b_1b_2b_2b_1)^k b_1b_2b_2\delta, \text{ where } k\in\mathbb{N}.$$

Since the family of stage one splicing languages, $L(S)$ have the sites to be carved by the rules above. Therefore, when splicing occurs among the family of stage one splicing languages, the following splicing languages will be resulted at stage two,

$$\alpha a_1(a_1a_2a_2a_1)^n a_1a_2a_2\beta,\ \alpha a_1(a_1a_2a_2a_1)^n a_1a_2a_2\alpha'$$
$$\beta' a_1(a_1a_2a_2a_1)^n a_1a_2a_2\beta,\ \gamma b_1(b_1b_2b_2b_1)^n b_1b_2b_2\delta, \text{ where } n=k+i,\ i=0,1,2,\ldots,k.$$

Now the persistency of two stages splicing languages should be proven. To prove that the families of two stages splicing languages are persistent, the pattern with the same crossing should be considered. According to the definition of persistent, if a_1a_2 is a sub segment of $\alpha a_1a_1a_2$, that is the crossing of $\alpha a_1a_1a_2a_2a_1a_1a_2a_2\beta$. This sub segment a_1a_2 contains an occurrence of the crossing of a site in the resulted string $\alpha a_1(a_1a_2a_2a_1)^k a_1a_2a_2\alpha'$ and all splicing languages of two stages that have a pattern of $(a_1,a_1a_2,a_2:a_1,a_1a_2,a_2)$. Also, by using the same approach the persistency of splicing languages, which have a pattern of $(b_1,b_1b_2,b_2:b_1,b_1b_2,b_2)$ can be proven. Consequently, the families of two stages splicing languages are persistent.■

In the next lemma, the persistency of two stages splicing languages with respect to two initial strings and two rules, where the first rule cuts the first initial string and both splicing rules cut the second string, is discussed.

Lemma 3.2: Let $S=(A,I,R)$ be a Y-G splicing system consisting of two initial strings (with two cutting sites), and two rules, where the crossing site of first rule is palindromic and the crossing site of the second rule is non-palindromic. If the first rule cuts the first initial, and both first rule and second rule cut the second string at two specific places, then the set of two stages splicing languages that is produced by Y-G splicing system is persistent.□

Proof: Suppose $S=(A,I,R)$ is a Y-G splicing system consisting two initial strings $s_1,s_2\in I$ and two rules

$r_1, r_2 \in R$. Thus, the rules r_1 and r_2 are represented as $r_1 = (a_1, a_1a_2, a_2 : a_1, a_1a_2, a_2)$ and $r_2 = (b_1, b_1b_2, b_2 : b_1, b_1b_2, b_2)$, respectively, such that a_1 is complementary with a_2, b_1 is not complementary with b_2 and vice-versa and $a_1, a_2, b_1, b_2 \in A^*$. Let s_1 and s_2 be initial strings in A^*, which have the forms $s_1 = \alpha a_1a_1a_2a_2a_1a_1a_2a_2\beta$ and $s_2 = \gamma a_1a_1a_2a_2b_1b_1b_2b_2\delta$, where s_1 can be cut by r_1 and s_2 by both rules $r_1, r_2 \in R$ at two specific places. Let the complementary relations of the strings be in the form of $[\alpha/\alpha'], [\beta/\beta'], [\gamma/\gamma']$ and $[\delta/\delta']$. Splicing the initial strings by applying rule r_1 on s_1 and $r_1, r_2 \in R$ on s_2, the following 10 DNA splicing languages will be generated at stage one as listed below,

$$\alpha a_1(a_1a_2a_2a_1)^k a_1a_2a_2\beta,\ \alpha a_1(a_1a_2a_2a_1)^k a_1a_2a_2\alpha',$$
$$\beta' a_1(a_1a_2a_2a_1)^k a_1a_2a_2\beta,\ \alpha a_1(a_1a_2a_2a_1)^k a_1a_2a_2b_1b_1b_2b_2\delta,$$
$$\alpha a_1(a_1a_2a_2a_1)^k a_1a_2a_2\gamma',\ \beta' a_1(a_1a_2a_2a_1)^k a_1a_2a_2b_1b_1b_2b_2\delta,$$
$$,\ \gamma a_1(a_1a_2a_2a_1)^k a_1a_2a_2\beta,\ \gamma a_1(a_1a_2a_2a_1)^k a_1a_2a_2\gamma',$$
$$\gamma a_1(a_1a_2a_2a_1)^k a_1a_2a_2b_1b_1b_2b_2\delta,$$
$$\delta' b_2'b_2'b_1'b' a_1(a_1a_2a_2a_1)^k a_1a_2a_2b_1b_1b_2b_2\delta, \text{ where } k \in \mathbb{N}.$$

To generate the stage two splicing languages, the rules r_1 and r_2 are added to the resulted DNA splicing languages of stage one. By splicing operation the following DNA splicing languages will be produced at stage two namely,

$$\alpha a_1(a_1a_2a_2a_1)^n a_1a_2a_2\beta,\ \alpha a_1(a_1a_2a_2a_1)^n a_1a_2a_2\alpha',$$
$$\beta' a_1(a_1a_2a_2a_1)^n a_1a_2a_2\beta,\ \alpha a_1(a_1a_2a_2a_1)^n a_1a_2a_2b_1b_1b_2b_2\delta,$$
$$\alpha a_1(a_1a_2a_2a_1)^n a_1a_2a_2\gamma',\ \beta' a_1(a_1a_2a_2a_1)^n a_1a_2a_2b_1b_1b_2b_2\delta,$$
$$,\ \gamma a_1(a_1a_2a_2a_1)^n a_1a_2a_2\beta,\ \gamma a_1(a_1a_2a_2a_1)^n a_1a_2a_2\gamma',$$
$$\gamma a_1(a_1a_2a_2a_1)^n a_1a_2a_2b_1b_1b_2b_2\delta,$$
$$\delta' b_2'b_2'b_1'b' a_1(a_1a_2a_2a_1)^n a_1a_2a_2b_1b_1b_2b_2\delta, \text{ where}$$
$$n = k + i,\ i = 0, 1, 2, \ldots, k.$$

To prove the families of two stages splicing languages are persistent, the pattern with same crossing should be considered. According to the definition of persistent if a_1a_2 is a sub segment of $\alpha a_1a_1a_2$, that is crossing of $\alpha a_1a_1a_2a_2a_1a_1a_2a_2\beta$. This sub segment a_1a_2 also contains an occurrence of the crossing of a site in the obtained splicing language $\alpha a_1(a_1a_2a_2a_1)^k a_1a_2a_2\gamma'$ and all splicing languages of two stages which have a pattern in the form of $(a_1, a_1a_2, a_2 : a_1, a_1a_2, a_2)$. Similarly, the persistency of those splicing languages of two stages, which contain the sequence of rule $(b_1, b_1b_2, b_2 : b_1, b_1b_2, b_2)$, can be proven. Consequently, the families of two stages splicing languages are persistent.■

In Lemma 3.3, the persistency of two stages splicing languages with respect to two initial strings and two rules, where both rules are applied on the first initial string and the second rule is applied on the second string, is investigated.

Lemma 3.3: Let $S = (A, I, R)$ be a Y-G splicing system consisting of two initial strings (with two cutting sites), and two rules, where the crossing site of first rule is palindromic and the crossing site of the second rule is non-palindromic. If both first and the second rules cut the first initial string, and the second rule cuts the second string at two specific places, then the set of two stages splicing languages that is produced by Y-G splicing system is persistent.□

Proof: Suppose $S = (A, I, R)$ is a Y-G splicing system consisting two initial strings $s_1, s_2 \in I$ and two rules $r_1, r_2 \in R$. Thus, the rules r_1 and r_2 are represented as $r_1 = (a_1, a_1a_2, a_2 : a_1, a_1a_2, a_2)$ and $r_2 = (b_1, b_1b_2, b_2 : b_1, b_1b_2, b_2)$, respectively, such that a_1 is complementary with a_2, b_1 is not complementary with b_2 and vice-versa, and $a_1, a_2, b_1, b_2 \in A^*$. Let $s_1 = \alpha a_1a_1a_2a_2b_1b_1b_2b_2\beta$ and $s_2 = \gamma b_1b_1b_2b_2b_1b_1b_2b_2\delta$ be two initial strings in A^*, where s_2 can be cut by r_2 and s_1 by both rules $r_1, r_2 \in R$ from two specific places. Suppose the complementary of α, β, γ and δ are α', β', γ' and δ', respectively. Splicing s_1 and s_2 using Y-G approach and applying rule r_2 on s_2 and $r_1, r_2 \in R$ on s_1, the following 6 DNA splicing languages will be generated at stage one namely,

$$\alpha a_1a_1a_2a_2b_1(b_1b_2b_2b_1)^k b_1b_2b_2\beta,\ \gamma b_1(b_1b_2b_2b_1)^k b_1b_2b_2\delta,$$
$$\gamma b_1(b_1b_2b_2b_1)^k b_1b_2b_2\beta,\ \alpha a_1a_1a_2a_2b_1(b_1b_2b_2b_1)^k b_1b_2b_2\delta,$$
$$\beta' b_2'b_2'b_1'b_1' a_1a_1a_2a_2b_1(b_1b_2b_2b_1)^k b_1b_2b_2\beta,$$
$$\beta' b_2'b_2'b_1'b_1' a_1a_1a_2a_2b_1(b_1b_2b_2b_1)^k b_1b_2b_2\delta, \text{ where } k \in \mathbb{N}.$$

To obtain the set of stage two splicing languages, the rules r_1 and r_2 are applied on the resulted DNA splicing languages of stage one. When splicing takes place among them, the following DNA splicing languages will be resulted at stage two,

$$\alpha a_1a_1a_2a_2b_1(b_1b_2b_2b_1)^n b_1b_2b_2\beta,\ \gamma b_1(b_1b_2b_2b_1)^n b_1b_2b_2\delta,$$
$$\gamma b_1(b_1b_2b_2b_1)^n b_1b_2b_2\beta,\ \alpha a_1a_1a_2a_2b_1(b_1b_2b_2b_1)^n b_1b_2b_2\delta,$$
$$\beta' b_2'b_2'b_1'b_1' a_1a_1a_2a_2b_1(b_1b_2b_2b_1)^n b_1b_2b_2\beta,$$
$$\beta' b_2'b_2'b_1'b_1' a_1a_1a_2a_2b_1(b_1b_2b_2b_1)^n b_1b_2b_2\delta,$$
$$\text{where } n = k + i,\ i = 0, 1, 2, \ldots, k$$

Now the persistency of two stages splicing languages should be proven. To prove the families of two stages splicing languages are persistent, the pattern with same crossing should be considered. According to the definition of persistent, if the string a_1a_2 is a sub segment of string $\alpha a_1a_1a_2a_2b_1b_1b_2$, that is crossing of a site in $\alpha a_1a_1a_2a_2b_1b_1b_2b_2\beta$. This sub segment a_1a_2 also contains an occurrence of the crossing of a site in the yielding string $\alpha a_1a_1a_2a_2b_1(b_1b_2b_2b_1)^k b_1b_2b_2\delta$ and also the rest of splicing languages with the pattern of $(a_1, a_1a_2, a_2 : a_1, a_1a_2, a_2)$. By using the same approach, the persistency of those two stages splicing languages, which have a pattern in the form of $(b_1, b_1b_2, b_2 : b_1, b_1b_2, b_2)$, can be proven. Consequently, the families of two stages splicing languages are persistent.■

In the last lemma, the persistency of two stages splicing languages with respect to two initial strings and two rules, where both splicing rules are applied on each of the initial strings, is considered and discussed.

Lemma 3.4: Let $S = (A, I, R)$ be a Y-G splicing system consisting of two initial strings (with two cutting sites) and two rules, where the crossing site of first rule is palindromic and the crossing site of the second rule is non-palindromic. If both first and second rules cut each of the initial strings at two specific crossing sites, then the set of two stages splicing languages that is produced by Y-G splicing system is persistent.□

Proof: Suppose $S=(A,I,R)$ is a Y-G splicing system consisting two initial strings $s_1,s_2 \in I$ and two rules $r_1,r_2 \in R$. Thus, the rules r_1 and r_2 are represented as $r_1=(a_1,a_1a_2,a_2:a_1,a_1a_2,a_2)$ and $r_2=(b_1,b_1b_2,b_2:b_1,b_1b_2,b_2)$, respectively, such that a_1 is complementary with a_2, b_1 is not complementary with b_2 and vice-versa and $a_1,a_2,b_1,b_2 \in A^*$. Let s_1 and s_2 be initial strings in A^* which have the forms $s_1=\alpha a_1a_1a_2a_2b_1b_1b_2b_2\beta$ and $s_2=\gamma a_1a_1a_2a_2b_1b_1b_2b_2\delta$, where $[\alpha/\alpha'],[\beta/\beta'],[\gamma/\gamma']$ and $[\delta/\delta']$ are complementary, and both strings s_1 and s_2 can be cut be by $r_1,r_2 \in R$ at two specific places. Therefore, applying rules $r_1,r_2 \in R$ on strings s_1 and s_2 the following 8 splicing languages will be generated besides the initial string I itself, at stage one namely,

$$\alpha a_1a_1a_2a_2\alpha' ,\ \delta'b_2'b_2'b_1'b_1'a_1a_1a_2a_2b_1b_1b_2b_2\delta ,\ \alpha a_1a_1a_2a_2\gamma' ,$$
$$\gamma a_1a_1a_2a_2\gamma' ,\ \gamma a_1a_1a_2a_2b_1b_1b_2b_2\beta ,\ \beta'b_1'b_1'b_2'b_2'a_1a_1a_2a_2b_1b_1b_2b_2\beta ,$$
$$\alpha a_1a_1a_2a_2b_1b_1b_2b_2\delta ,\ \beta'b_2'b_2'b_1'b_1'a_1a_1a_2a_2b_1b_1b_2b_2\delta .$$

However, if the splicing occurs among the resulted splicing languages of stage one; no distinct splicing languages will be resulted at stage two. To show the families of two stages splicing languages are persistent, the pattern with the same crossing should be considered. Therefore, if $y=a_1a_2$ is a sub segment of $\alpha a_1a_1a_2$, that is crossing of $\alpha a_1a_1a_2a_2b_1b_1b_2b_2\beta$. This sub segment $y=a_1a_2$ also contains an occurrence of the crossing of a site in $\alpha a_1a_1a_2a_2b_1b_1b_2b_2\delta$ and all splicing languages that have a pattern of $(a_1,a_1a_2,a_2:a_1,a_1a_2,a_2)$. Similarly, the persistency of those splicing languages, which contain the sequence of rule $(b_1,b_1b_2,b_2:b_1,b_1b_2,b_2)$ can be proven. Consequently, the families of two stages splicing languages are persistent.■

In the following theorem, the persistency and permanency of two stages DNA splicing languages, which are produced by Y-G splicing system consisting two initial strings (with the two cutting sites) and two rules where the crossing site of the first rule is palindromic and the crossing site of the second rule is non-palindromic, is investigated.

Theorem 3.1: Let $S=(A,I,R)$ be a Y-G splicing system consisting of two initial strings (with two cutting sites) and two rules, where the crossing site of first rule is palindromic and the crossing site of the second rule is non-palindromic. Then, the set of two stages splicing languages, which is produced by S, is persistent.□

Proof: Suppose $S=(A,I,R)$ is a Y-G splicing system consisting two initial strings $s_1,s_2 \in I$ and two rules $r_1,r_2 \in R$. Thus, the rules r_1 and r_2 are represented as $r_1=(a_1,a_1a_2,a_2:a_1,a_1a_2,a_2)$ and $r_2=(b_1,b_1b_2,b_2:b_1,b_1b_2,b_2)$, respectively, such that a_1 is complementary with a_2, b_1 is not complementary with b_2 and vice-versa and $a_1,a_2,b_1,b_2 \in A^*$. To prove the persistency of two stages splicing languages, four cases need to be considered according to the cutting sites of initial strings:

Case 1: The first rule is applied on the first initial string, and the second rule on the second initial string,

Case 2: The first rule is only applied on the first initial string, and both first and second rules are applied on the second initial string,

Case 3: Both splicing rules are applied on the first initial string, and second rule is only applied on the second initial string,

Case 4: Both of the splicing rules are applied on each of the initial strings.

Case 1, Case 2, Case 3 and Case 4 can be proven by Lemma 1, Lemma 2, Lemma 3 and Lemma 4, respectively. Thus, Theorem 3.1 is proven.■

The following corollary can be obtained from Theorem 3.1, which shows the above splicing languages are also permanent.

Corollary 3.1: Let $S=(A,I,R)$ be a Y-G splicing system consisting of two initial strings (with two cutting sites) and two rules, where the crossing site of first rule is palindromic and the crossing site of the second rule is non-palindromic. Then, the set of two stages splicing languages, which is produced by S, is permanent.□

To apply the theorem in biology, a test tube consisting of two initial DNA strands named s_1 and s_2 is considered. Suppose that $Hpa\mathrm{II}$ and $Ava\mathrm{II}$ are two restriction enzymes added in the test tube, which are supplied with NEB CutSmart™ Buffer to make the process of DNA recombination convenient. When an appropriate ligase is added, the recombinant DNA strands that will arise during this process are persistent and also permanent. This process of DNA recombination is presented mathematically in the following example.

Example 3.1: Let $S=(A,I,R)$ be a Y-G splicing system such that $A=\{a,c,g,t\}$ and the set of splicing rules R consists of two rules namely, $(c;cg,g{:}c;cg,g)$ and $(g;gwc,c:g;gwc,c)$, where $w=a$ or t. Suppose $s_1=pccggccggq$ and $s_2=uggwccggwccv$ where $p,q,u,v \in A^*$, are two arbitrary initial strings in I. Since the crossing site of the first rule is palindromic and the crossing site of the second rule is non-palindromic, thus according to Case 1 of Theorem 3.1 (Lemma 3.1**)**, the two stages DNA splicing languages that are produced by S are persistent and permanent. ■

In the last section, the conclusion to this paper is stated.

4.0 CONCLUSION

The authors would like to express their appreciations to Ministry of Education Malaysia and Research and Innovation Department of Universiti Malaysia Pahang for funding this research through RDU Vote: 130354 and RAGS Grant Vote: 131404. The first author also dedicates his hearty acknowledgement to the Ministry of Higher Education (MOHE) of Afghanistan for his master scholarship and Kabul Education University (KEU) for his study leave.

Acknowledgement

The authors would like to express their appreciations to Ministry of Education Malaysia and Research and Innovation Department of Universiti Malaysia Pahang for funding this research through RDU Vote: 130354 and RAGS Grant Vote: 131404. The first author also dedicates his hearty acknowledgement to the Ministry of Higher Education (MOHE) of Afghanistan for his master scholarship and Kabul Education University (KEU) for his study leave.

References

[1] Paolella, P. 1998. *Introduction to Molecular Biology*. McGraw-Hill.

[2] Walker, J. M. and Rapley, R. 2009. *Molecular Biology and Biotechnology*. Royal Society of Chemistry.

[3] Walker, J. M. and Rapley, R. 2008. *Molecular Biomethods*. Humana Press.

[4] 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.

[5] Fong, W. H. 2008. *Modelling Splicing System Using Formal Language Theory*. Ph.D Thesis. Universiti Teknologi Malaysia.

[6] Karimi, F., Sarmin, N. H. and Fong, W. H. 2010. An Investigation on Characteristics of Splicing System and Splicing Languages. *Proceeding of SKASM*. 39-43

[7] Gatterdam, R. 1992. Algorithm for Splicing System. *SIAM Journal of Computing*. 21: 507-520.

[8] Karimi, F., Sarmin, N. H. and Fong, W. H. 2012. The Characterization of Different Splicing Systems. *International Journal of Modern Physics: Conference Series*. 9: 89-94.

[9] Karimi, F., Sarmin, N. H. and Fong, W. H. 2011. Some Sufficient Conditions for Persistent Splicing Systems. *Australian Journal of Basic and Applied Sciences*. 5(1): 20-24.

[10] Karimi, F., Sarmin, N. H. and Fong, W. H. 2010. Some Sufficient Conditions for Permanent Splicing Systems. *Proceeding of Faculty of Science Postgraduate Conference (FSPGC 2010)*. 80-84.

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

[12] Yusof, Y., Sarmin, N. H. and Fong, W. H. 2014. The Concept of Persistent and Permanent in Non-Semi Simple Splicing System. *AIP Conference Proceeding*. 586-590.

[13] Mudaber, M. H., Yusof, Y. and Mohamad, S. M. 2014. Some Sufficient Conditions for Persistency and Permanency of Two Stages DNA Splicing Languages via Yusof-Goode Approach. *AIP Conference Proceeding*. 1605: 591-595.

[14] Mudaber, M. H., Yusof, Y. and Mohamad, S.M. 2014. Some Relations between Two Stages DNA Splicing Languages. *AIP Conf. Proc*. 1602: 254-259.