

Generalisations of Splicing Languages in DNA Splicing Systems Involving Two Palindromic Restriction Enzymes

Nurul Izzaty Ismail, Wan Heng Fong*, Nor Haniza Sarmin

Department of Mathematical Sciences, Faculty of Science, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia

Abstract DNA splicing system is initiated by Head to mathematically model a relation between formal language theory and DNA molecules. In DNA splicing systems, DNA molecules are cut and recombined in specific ways with the existence of enzymes, which are also known as endonucleases, to produce further molecules. The resulting molecules are depicted as splicing languages by using concepts in formal languages theory. A sequence of restriction enzyme that reads the same forward and backward is called as a palindromic rule. Previously, researches on different types of splicing languages have been done. In this research, generalisations of splicing languages resulting from DNA splicing systems with non-overlapping cutting sites of two palindromic restriction enzymes are presented as theorems using the induction method. The results from this research are beneficial for researchers in the field of DNA computing since it contributes to the development of splicing languages generated from DNA splicing systems with different palindromic restriction enzymes by using these generalisations.

Keywords: splicing language, palindromic, restriction enzyme, splicing system

Introduction

In 1987, research on splicing systems was initiated by Head to relate between formal language theory and molecular biology [1]. Splicing systems are mathematically modelled from biological process of recombinant deoxyribonucleic acid (DNA) where DNA molecules are cut and recombined when reacting with ligase and restriction enzymes [1]. Restriction enzymes are also known as endodeoxyribonucleases that function as catalyst for chemical reaction in biological systems [2]. In the modelling of a splicing system, nitrogenous bases, nucleotides and restriction enzymes are denoted as symbols, strings and rules respectively using formal language theory.

The field of formal language theory is used in mathematics, computer science, and linguistics which include some topics in discrete mathematics that involve set theory, functions, logic and mathematical reasoning, knowledge of programming languages and syntactical aspects of natural languages [3]. The formal language L consists of a set of strings (words) of symbols derived from an alphabet Σ [3]. Some operations of regular expressions in formal languages which are used in this research such as concatenation, union and star closure [3], are shown respectively in the following:

$$\begin{aligned} L(a \cdot b) &= L(a) L(b) = \{ab\}, \\ L(a + b) &= L(a) \cup L(b) = \{a,b\}, \text{ and} \\ L(a^*) &= (L(a))^* = \{a^n : n \geq 0, n \in \mathbb{Z}^+\} \end{aligned}$$

where Σ is a given alphabet with regular expressions $a, b \in \Sigma$.

*For correspondence:
fwh@utm.my

Received: 20 Dec 2018

Accepted: 26 Feb 2019

© Copyright Ismail. This article is distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use and redistribution provided that the original author and source are credited.

The molecules resulting from DNA splicing systems are called the splicing language [1], which comprises of three sets using formal language theory. The first set consists of Watson-Crick complementarity in the formation of double-stranded DNA (dsDNA) where adenine (A) pairs with thymine (T), while cytosine (C) pairs with guanine (G) [3]. The second set is the set of initial strings representing the DNA molecules. The third set contains the rules for restriction enzymes, where the rule for the cleavage pattern of restriction enzymes is formed as a triple: left context, crossing and right context [4]. The symbols \downarrow and \uparrow indicate the cutting sites of the restriction enzymes.

Over the years, several models had been extended and developed such as Paun [5], Pixton [6], and Yusof-Goode (Y-G) [7] splicing systems. The comparison between Head's and these three splicing models were given in [7]. Besides that, Laun also developed a new model of splicing system known as Goode-Pixton splicing system in [8]. Moreover, a new extension of splicing systems known as a simple splicing system had also been introduced [9].

The modelling of splicing languages resulting from DNA splicing systems involving palindromic sequences is used in this research. In DNA splicing systems, DNA molecule can read in two ways: forward and backward. Palindrome is a sequence that reads the same forward and backward [10]. A sequence of DNA molecule which is exactly the same both ways is called a palindromic string. In the first paper introducing splicing system, Head mentioned the concept of palindromic string in DNA splicing systems [1]. In 2012, the definition of palindromic string had been proposed by Yusof [11].

Besides, palindromic sequences can also be recognised in the sequence of restriction enzymes [12]. Previously, the modelling of one and two stages splicing languages from Y-G splicing systems with palindromic crossings was presented in [13, 14]. Furthermore, the study of second order limit languages from Y-G splicing system with palindromic crossing and recognition site was discussed in [15]. In order to generalise splicing languages from different models of splicing system, the sequences of restriction enzymes have been taken into consideration. Research on the generalisations of splicing languages resulting from DNA splicing systems involving one palindromic restriction enzyme had been done [16].

In this paper, the splicing languages from DNA splicing systems involving two palindromic restriction enzymes for the same and different crossings are generalised and presented as theorems. In nature, the cutting sites may overlap which affect the enzymatic activities in the splicing system [17]. Thus, this research focuses on non-overlapping cutting site of two palindromic restriction enzymes to obtain the generalised splicing language.

PRELIMINARIES

In this research, Head's splicing system is used to generalise the resulting splicing languages from DNA splicing systems with two palindromic restriction enzymes. The definitions of Head's splicing system and a splicing language are stated in the following.

Definition 1 [1] Splicing System and Splicing Language

A splicing system $S = (A, I, B, C)$ consists of a finite alphabet A , a finite set I of initial strings in A^* , and finite sets B and C of triples (c, x, d) with c, x and d in A^* . Each such triple in B or C is called a pattern. For each such triple the string $cx d$ is called a site and the string x is called a crossing. Patterns in B are called left patterns and patterns in C are called right patterns. The language $L = L(S)$ generated by S consists of the strings in I and all strings that can be obtained by adjoining to $ucxfq$ and $pexdv$ whenever $ucxdv$ and $pexfq$ are in L and (c, x, d) and (e, x, f) are patterns of the same hand. A language, L is a splicing language if there exists a splicing system S for which $L = L(S)$.

The splicing language is the set of DNA strings resulting from DNA splicing systems which comprise of the set of DNA symbols, initial strings and cleavage pattern of restriction enzymes.

Next, generalisations of splicing languages in DNA splicing system are done involving palindromic sequences for restriction enzymes. The definition of a palindromic string is stated next.

Definition 2 [11] Palindromic String

A string I of a dsDNA is said to be palindromic if the sequence from the left to the right side of the upper single strand is equal to the sequence from the right to the left side of the lower single strand.

For instance, the enzyme $EcoRI$ $\begin{matrix} 5' - GAATTC - 3' \\ 3' - CTTAAG - 5' \end{matrix}$ is a palindromic restriction enzyme as the upper single strand of enzyme $EcoRI$ matches with the lower single strand when read from backward.

RESULTS

In this reseach, the generalisation of splicing languages from DNA splicing systems with two palindromic restriction enzymes are determined and proved. The splicing language resulting from DNA splicing systems with two palindromic restriction enzymes for the same crossing is generalised and presented as Theorem 1.

Theorem 1

Given $S = (A, I, B, C)$ is a splicing system in which $A = \left\{ \begin{matrix} A, C, G, T \\ T, G, C, A \end{matrix} \right\}$ is the set of dsDNA symbols, $I = \left\{ N_1 N_1 \dots N_1 X_1 Y X_2 M M \dots M W_1 Y W_2 N_2 N_2 \dots N_2 \right\}$ is the set consisting of an initial string with non-overlapping cutting sites of two palindromic rules $\begin{matrix} X_1 Y X_2 \\ X_1' Y' X_2' \end{matrix}$ and $\begin{matrix} W_1 Y W_2 \\ W_1' Y' W_2' \end{matrix}$ where $\begin{matrix} N_1 & X_1 & Y & X_2 & M & W_1 & W_2 \\ N_1' & X_1' & Y' & X_2' & M' & W_1' & W_2' \end{matrix}$ and $\begin{matrix} N_2 \\ N_2' \end{matrix}$ are variables used to denote any arbitrary dsDNA and $N_1', X_1', Y', X_2', M', W_1', W_2'$ and N_2' are complementaries for $N_1, X_1, Y, X_2, M, W_1, W_2,$ and N_2 respectively, set $B = \left\{ \left(\begin{matrix} X_1, Y, X_2 \\ X_1' Y' X_2' \end{matrix} \right), \left(\begin{matrix} W_1, Y, W_2 \\ W_1' Y' W_2' \end{matrix} \right) \right\}$ is the set of cleavage pattern for the rules where $\begin{matrix} Y \\ Y' \end{matrix}$ is the crossing and set C is the empty set, the resulting splicing language consists of strings of the form

$$\begin{aligned} & \left(\begin{matrix} N_1 N_1 \dots N_1 X_1 & + & N_2' N_2' \dots N_2' W_1 \\ N_1' N_1' \dots N_1' X_1' & + & N_2 N_2 \dots N_2 W_1' \end{matrix} \right) \\ & \left\{ Y \left(\begin{matrix} X_2 M M \dots M W_1 & + & W_2 M' M' \dots M' X_1 \\ X_2' M' M' \dots M' W_1' & + & W_2' M M \dots M X_1' \end{matrix} \right) \right\}^{n-1} Y \\ & \left(\begin{matrix} W_2 N_2 N_2 \dots N_2 & + & X_2 N_1' N_1' \dots N_1' \\ W_2' N_2' N_2' \dots N_2' & + & X_2' N_1 N_1 \dots N_1 \end{matrix} \right) \end{aligned} \tag{1}$$

where $n \in \mathbb{Z}^+$ represents multiple copies of strings and $\left\{ \begin{matrix} X_1 Y X_2 & W_1 Y W_2 \\ X_1' Y' X_2' & W_1' Y' W_2' \end{matrix} \right\} \notin \left\{ \begin{matrix} N_1 N_1 \dots N_1 & M M \dots M & N_2 N_2 \dots N_2 \\ N_1' N_1' \dots N_1' & M' M' \dots M' & N_2' N_2' \dots N_2' \end{matrix} \right\}$ indicates no other cutting site is present in strings $\begin{matrix} N_1 N_1 \dots N_1 \\ N_1' N_1' \dots N_1' \end{matrix}$, $\begin{matrix} M M \dots M \\ M' M' \dots M' \end{matrix}$ and $\begin{matrix} N_2 N_2 \dots N_2 \\ N_2' N_2' \dots N_2' \end{matrix}$.

Proof.

Suppose the rules $\begin{matrix} X_1 Y X_2 \\ X_1' Y' X_2' \end{matrix}$ and $\begin{matrix} W_1 Y W_2 \\ W_1' Y' W_2' \end{matrix}$ are palindromic, so the base sequence of the rules can read the same backward and forward:

$$\begin{matrix} X_1 Y X_2 & = & X_2' Y' X_1' \\ X_1' Y' X_2' & = & X_2 Y X_1 \end{matrix}$$

and

$$\begin{matrix} W_1 Y W_2 & = & W_2' Y' W_1' \\ W_1' Y' W_2' & = & W_2 Y W_1 \end{matrix}$$

Then $\begin{matrix} X_1 & = & X_2' & Y & = & Y' & X_2 & = & X_1' & W_1 & = & W_2' \\ X_1' & = & X_2 & Y' & = & Y & X_2' & = & X_1 & W_1' & = & W_2 \end{matrix}$ and $\begin{matrix} W_2 & = & W_1' \\ W_2' & = & W_1 \end{matrix}$.

The initial string $\begin{matrix} N_1 N_1 \dots N_1 X_1 Y X_2 M M \dots M W_1 Y W_2 N_2 N_2 \dots N_2 \\ N_1' N_1' \dots N_1' X_1' Y' X_2' M' M' \dots M' W_1' Y' W_2' N_2' N_2' \dots N_2' \end{matrix}$ with the cutting sites $\begin{matrix} X_1 Y X_2 \\ X_1' Y' X_2' \end{matrix}$ and $\begin{matrix} W_1 Y W_2 \\ W_1' Y' W_2' \end{matrix}$ is shown respectively in the following:

$$\begin{matrix} N_1 N_1 \dots N_1 X_1 \downarrow Y & X_2 M M \dots M W_1 Y W_2 N_2 N_2 \dots N_2 \\ N_1' N_1' \dots N_1' X_1' & Y' \uparrow X_2' M' M' \dots M' W_1' Y' W_2' N_2' N_2' \dots N_2' \end{matrix} \tag{2}$$

for the first cutting site and

$$\begin{matrix} N_1 N_1 \dots N_1 X_1 Y X_2 M M \dots M W_1 \downarrow Y & W_2 N_2 N_2 \dots N_2 \\ N_1' N_1' \dots N_1' X_1' Y' X_2' M' M' \dots M' W_1' & Y' \uparrow W_2' N_2' N_2' \dots N_2' \end{matrix} \tag{3}$$

for the second cutting site.

Due to the complementary nature of double helical structure of DNA, the initial string can be written 180 degree wise [18].

$$\begin{matrix} N_2' N_2' \dots N_2' W_2' Y' W_1' M' M' \dots M' X_2' Y' X_1' N_1' N_1' \dots N_1' \\ N_2 N_2 \dots N_2 W_2 Y W_1 M M \dots M X_2 Y X_1 N_1 N_1 \dots N_1 \end{matrix} \tag{4}$$

Since $\frac{X_1}{X_1'} = \frac{X_2'}{X_2} Y = Y' \frac{X_2}{X_2'} = \frac{X_1' W_1}{X_1' W_1'} = \frac{W_2'}{W_2}$ and $\frac{W_2}{W_2'} = \frac{W_1'}{W_1}$, string (4) becomes

$$\begin{aligned} & N_2' N_2' \dots N_2' W_1 Y W_2 M' M' \dots M' X_1 Y X_2 N_1' N_1' \dots N_1' \\ & N_2 N_2 \dots N_2 W_1' Y' W_2' M M \dots M X_1' Y' X_2' N_1 N_1 \dots N_1 . \end{aligned} \tag{5}$$

The string (5) with the cutting sites $\frac{X_1 Y X_2}{X_1' Y' X_2'}$ and $\frac{W_1 Y W_2}{W_1' Y' W_2'}$ is shown respectively in the following:

$$\begin{aligned} & N_2' N_2' \dots N_2' W_1 \downarrow Y W_2 M' M' \dots M' X_1 Y X_2 N_1' N_1' \dots N_1' \\ & N_2 N_2 \dots N_2 W_1' Y' \uparrow W_2' M M \dots M X_1' Y' X_2' N_1 N_1 \dots N_1 \end{aligned} \tag{6}$$

and

$$\begin{aligned} & N_2' N_2' \dots N_2' W_1 Y W_2 M' M' \dots M' X_1 \downarrow Y X_2 N_1' N_1' \dots N_1' \\ & N_2 N_2 \dots N_2 W_1' Y' W_2' M M \dots M X_1' Y' \uparrow X_2' N_1 N_1 \dots N_1 . \end{aligned} \tag{7}$$

Using the rules $\frac{X_1 Y X_2}{X_1' Y' X_2'}$ and $\frac{W_1 Y W_2}{W_1' Y' W_2'}$, the string (2) combines with (3), (6) and (7) which gives:

$$\begin{aligned} & N_1 N_1 \dots N_1 X_1 Y W_2 N_2 N_2 \dots N_2 \\ & N_1' N_1' \dots N_1' X_1' Y' W_2' N_2' N_2' \dots N_2' , \end{aligned} \tag{8}$$

$$\begin{aligned} & N_1 N_1 \dots N_1 X_1 Y X_2 M M \dots M W_1 Y X_2 M M \dots M W_1 Y W_2 N_2 N_2 \dots N_2 \\ & N_1' N_1' \dots N_1' X_1' Y' X_2' M' M' \dots M' W_1' Y' X_2' M' M' \dots M' W_1' Y' W_2' N_2' N_2' \dots N_2' , \\ & N_1 N_1 \dots N_1 X_1 Y W_2 M' M' \dots M' X_1 \downarrow Y X_2 N_1' N_1' \dots N_1' \\ & N_1' N_1' \dots N_1' X_1' Y' W_2' M M \dots M X_1' Y' \uparrow X_2' N_1 N_1 \dots N_1 , \end{aligned} \tag{9}$$

$$\begin{aligned} & N_2' N_2' \dots N_2' W_1 Y X_2 M M \dots M W_1 \downarrow Y W_2 N_2 N_2 \dots N_2 \\ & N_2 N_2 \dots N_2 W_1' Y' X_2' M' M' \dots M' W_1' Y' \uparrow W_2' N_2' N_2' \dots N_2' , \end{aligned} \tag{10}$$

$$\begin{aligned} & N_1 N_1 \dots N_1 X_1 Y X_2 N_1' N_1' \dots N_1' \\ & N_1' N_1' \dots N_1' X_1' Y' X_2' N_1 N_1 \dots N_1 , \end{aligned} \text{ and} \tag{11}$$

$$\begin{aligned} & N_2' N_2' \dots N_2' W_1 Y W_2 M' M' \dots M' X_1 Y X_2 M M \dots M W_1 Y W_2 N_2 N_2 \dots N_2 \\ & N_2 N_2 \dots N_2 W_1' Y' W_2' M M \dots M X_1' Y' X_2' M' M' \dots M' W_1' Y' W_2' N_2' N_2' \dots N_2' . \end{aligned}$$

The results of the combination of string (3) with (6) and (7) are

$$\begin{aligned} & N_1 N_1 \dots N_1 X_1 Y X_2 M M \dots M W_1 Y W_2 M' M' \dots M' X_1 Y X_2 N_1' N_1' \dots N_1' \\ & N_1' N_1' \dots N_1' X_1' Y' X_2' M' M' \dots M' W_1' Y' W_2' M M \dots M X_1' Y' X_2' N_1 N_1 \dots N_1 , \\ & N_2' N_2' \dots N_2' W_1 Y W_2 N_2 N_2 \dots N_2 \\ & N_2 N_2 \dots N_2 W_1' Y' W_2' N_2' N_2' \dots N_2' , \end{aligned} \tag{12}$$

$$\begin{aligned} & N_1 N_1 \dots N_1 X_1 Y X_2 M M \dots M W_1 Y X_2 N_1' N_1' \dots N_1' \\ & N_1' N_1' \dots N_1' X_1' Y' X_2' M' M' \dots M' W_1' Y' X_2' N_1 N_1 \dots N_1 , \end{aligned} \text{ and}$$

$$\begin{aligned} & N_2' N_2' \dots N_2' W_1 Y W_2 M' M' \dots M' X_1 Y W_2 N_2 N_2 \dots N_2 \\ & N_2 N_2 \dots N_2 W_1' Y' W_2' M M \dots M X_1' Y' W_2' N_2' N_2' \dots N_2' . \end{aligned}$$

Next, when string (6) combines with string (7), the other two new strings arise:

$$\begin{aligned} & N_2' N_2' \dots N_2' W_1 Y X_2 N_1' N_1' \dots N_1' \\ & N_2 N_2 \dots N_2 W_1' Y' X_2' N_1 N_1 \dots N_1 , \end{aligned} \text{ and} \tag{13}$$

$$\begin{aligned} & N_2' N_2' \dots N_2' W_1 Y W_2 M' M' \dots M' X_1 Y W_2 M' M' \dots M' X_1 Y X_2 N_1' N_1' \dots N_1' \\ & N_2 N_2 \dots N_2 W_1' Y' W_2' M M \dots M X_1' Y' W_2' M M \dots M X_1' Y' X_2' N_1 N_1 \dots N_1 . \end{aligned}$$

By using induction, this theorem can be proved. For $n = 1$, the string (1) is true since it is stated as strings (4), (6), (12) and (13). Next, let $n = k \in \mathbb{Z}^+$, string (1) becomes

$$\begin{aligned} & \left(\begin{matrix} N_1 N_1 \dots N_1 X_1 + N_2' N_2' \dots N_2' W_1 \\ N_1' N_1' \dots N_1' X_1' + N_2 N_2 \dots N_2 W_1' \end{matrix} \right) \\ & \left\{ Y \left(X_2 M M \dots M W_1 + W_2 M' M' \dots M' X_1 \right) \right\}^{k-1} Y \\ & \left(\begin{matrix} W_2 N_2 N_2 \dots N_2 + X_2 N_1' N_1' \dots N_1' \\ W_2' N_2' N_2' \dots N_2' + X_2' N_1 N_1 \dots N_1 \end{matrix} \right) . \end{aligned} \tag{14}$$

The following strings are among the strings in (14):

$$\begin{matrix} N_1 N_1 \dots N_1 X_1 (Y X_2 M M \dots M W_1)^{k-1} Y W_2 N_2 N_2 \dots N_2 \\ N_1' N_1' \dots N_1' X_1' (Y' X_2' M' M' \dots M' W_1') \end{matrix} \begin{matrix} Y W_2 N_2 N_2 \dots N_2 \\ Y' W_2' N_2' N_2' \dots N_2' \end{matrix}, \tag{15}$$

$$\begin{matrix} N_1 N_1 \dots N_1 X_1 (Y X_2 M M \dots M W_1)^{k-1} Y X_2 N_1' N_1' \dots N_1' \\ N_1' N_1' \dots N_1' X_1' (Y' X_2' M' M' \dots M' W_1') \end{matrix} \begin{matrix} Y X_2 N_1' N_1' \dots N_1' \\ Y' X_2' N_1 N_1 \dots N_1 \end{matrix}, \tag{16}$$

$$\begin{matrix} N_1 N_1 \dots N_1 X_1 (Y W_2 M' M' \dots M' X_1)^{k-1} Y W_2 N_2 N_2 \dots N_2 \\ N_1' N_1' \dots N_1' X_1' (Y' W_2' M M \dots M X_1') \end{matrix} \begin{matrix} Y W_2 N_2 N_2 \dots N_2 \\ Y' W_2' N_2' N_2' \dots N_2' \end{matrix},$$

$$\begin{matrix} N_1 N_1 \dots N_1 X_1 (Y W_2 M' M' \dots M' X_1)^{k-1} Y X_2 N_1' N_1' \dots N_1' \\ N_1' N_1' \dots N_1' X_1' (Y' W_2' M M \dots M X_1') \end{matrix} \begin{matrix} Y X_2 N_1' N_1' \dots N_1' \\ Y' X_2' N_1 N_1 \dots N_1 \end{matrix}, \tag{17}$$

$$\begin{matrix} N_2' N_2' \dots N_2' W_1 (Y X_2 M M \dots M W_1)^{k-1} Y W_2 N_2 N_2 \dots N_2 \\ N_2 N_2 \dots N_2 W_1' (Y' X_2' M' M' \dots M' W_1') \end{matrix} \begin{matrix} Y W_2 N_2 N_2 \dots N_2 \\ Y' W_2' N_2' N_2' \dots N_2' \end{matrix}, \tag{18}$$

$$\begin{matrix} N_2' N_2' \dots N_2' W_1 (Y X_2 M M \dots M W_1)^{k-1} Y X_2 N_1' N_1' \dots N_1' \\ N_2 N_2 \dots N_2 W_1' (Y' X_2' M' M' \dots M' W_1') \end{matrix} \begin{matrix} Y X_2 N_1' N_1' \dots N_1' \\ Y' X_2' N_1 N_1 \dots N_1 \end{matrix},$$

$$\begin{matrix} N_2' N_2' \dots N_2' W_1 (Y W_2 M' M' \dots M' X_1)^{k-1} Y W_2 N_2 N_2 \dots N_2 \\ N_2 N_2 \dots N_2 W_1' (Y' W_2' M M \dots M X_1') \end{matrix} \begin{matrix} Y W_2 N_2 N_2 \dots N_2 \\ Y' W_2' N_2' N_2' \dots N_2' \end{matrix}, \text{ and} \tag{19}$$

$$\begin{matrix} N_2' N_2' \dots N_2' W_1 (Y W_2 M' M' \dots M' X_1)^{k-1} Y X_2 N_1' N_1' \dots N_1' \\ N_2 N_2 \dots N_2 W_1' (Y' W_2' M M \dots M X_1') \end{matrix} \begin{matrix} Y X_2 N_1' N_1' \dots N_1' \\ Y' X_2' N_1 N_1 \dots N_1 \end{matrix}. \tag{20}$$

By expanding strings (15), (16), (17), (18), (19) and (20), the strings can be written respectively as:

$$\begin{matrix} N_1 N_1 \dots N_1 X_1 \downarrow Y X_2 M M \dots M W_1 (Y X_2 M M \dots M W_1)^{k-2} \\ N_1' N_1' \dots N_1' X_1' Y' \uparrow X_2' M' M' \dots M' W_1' (Y' X_2' M' M' \dots M' W_1') \end{matrix} \begin{matrix} Y W_2 N_2 N_2 \dots N_2 \\ Y' W_2' N_2' N_2' \dots N_2' \end{matrix}, \tag{21}$$

$$\begin{matrix} N_1 N_1 \dots N_1 X_1 \downarrow Y X_2 M M \dots M W_1 (Y X_2 M M \dots M W_1)^{k-2} \\ N_1' N_1' \dots N_1' X_1' Y' \uparrow X_2' M' M' \dots M' W_1' (Y' X_2' M' M' \dots M' W_1') \end{matrix} \begin{matrix} Y X_2 N_1' N_1' \dots N_1' \\ Y' X_2' N_1 N_1 \dots N_1 \end{matrix}, \tag{22}$$

$$\begin{matrix} N_1 N_1 \dots N_1 X_1 (Y W_2 M' M' \dots M' X_1)^{k-2} Y W_2 M' M' \dots M' X_1 \downarrow Y \\ N_1' N_1' \dots N_1' X_1' (Y' W_2' M M \dots M X_1') Y' \uparrow \\ X_2 N_1' N_1' \dots N_1' \\ X_2' N_1 N_1 \dots N_1 \end{matrix}, \tag{23}$$

$$\begin{matrix} N_2' N_2' \dots N_2' W_1 (Y X_2 M M \dots M W_1)^{k-2} Y X_2 M M \dots M W_1 \downarrow Y \\ N_2 N_2 \dots N_2 W_1' (Y' X_2' M' M' \dots M' W_1') Y' \uparrow \\ W_2 N_2 N_2 \dots N_2 \\ W_2' N_2' N_2' \dots N_2' \end{matrix}, \tag{24}$$

$$\begin{matrix} N_2' N_2' \dots N_2' W_1 \downarrow Y W_2 M' M' \dots M' X_1 (Y W_2 M' M' \dots M' X_1)^{k-2} \\ N_2 N_2 \dots N_2 W_1' Y' \uparrow W_2' M M \dots M X_1' (Y' W_2' M M \dots M X_1') \end{matrix} \begin{matrix} Y W_2 N_2 N_2 \dots N_2 \\ Y' W_2' N_2' N_2' \dots N_2' \end{matrix}, \text{ and} \tag{25}$$

$$\begin{matrix} N_2' N_2' \dots N_2' W_1 \downarrow Y W_2 M' M' \dots M' X_1 (Y W_2 M' M' \dots M' X_1)^{k-2} \\ N_2 N_2 \dots N_2 W_1' Y' \uparrow W_2' M M \dots M X_1' (Y' W_2' M M \dots M X_1') \end{matrix} \begin{matrix} Y X_2 N_1' N_1' \dots N_1' \\ Y' X_2' N_1 N_1 \dots N_1 \end{matrix}. \tag{26}$$

Using the rules $\frac{X_1 Y X_2}{X_1' Y' X_2'}$ and $\frac{W_1 Y W_2}{W_1' Y' W_2'}$, the above resulting strings combine with (2), (3), (6), (7), (9) and (10). Hence, the other recombinations can be shown through six cases.

Case 1: The string (3) combines with (21) and (22) which produces new strings

$$\begin{matrix} N_1 N_1 \dots N_1 X_1 Y X_2 M M \dots M W_1 Y X_2 M M \dots M W_1 \\ N_1' N_1' \dots N_1' X_1' Y' X_2' M' M' \dots M' W_1' Y' X_2' M' M' \dots M' W_1' \end{matrix} \begin{matrix} (Y X_2 M M \dots M W_1)^{k-2} Y W_2 N_2 N_2 \dots N_2 \\ (Y' X_2' M' M' \dots M' W_1') Y' W_2' N_2' N_2' \dots N_2' \end{matrix} \tag{27}$$

and

$$\begin{aligned}
 &N_1 N_1 \dots N_1 X_1 Y X_2 M M \dots M W_1 Y X_2 M M \dots M W_1 \\
 &N'_1 N'_1 \dots N'_1 X'_1 Y' X'_2 M' M' \dots M' W'_1 Y' X'_2 M' M' \dots M' W'_1 \\
 &\qquad\qquad\qquad \left(Y X_2 M M \dots M W_1 \right)^{k-2} Y X_2 N'_1 N'_1 \dots N'_1 \\
 &\qquad\qquad\qquad \left(Y' X'_2 M' M' \dots M' W'_1 \right) Y' X'_2 N_1 N_1 \dots N_1 .
 \end{aligned} \tag{28}$$

Case 2: The string (9) combines with (25) which produces new strings

$$\begin{aligned}
 &N_1 N_1 \dots N_1 X_1 Y W_2 M' M' \dots M' X_1 Y W_2 M' M' \dots M' X_1 \\
 &N'_1 N'_1 \dots N'_1 X'_1 Y' W'_2 M M \dots M X'_1 Y' W'_2 M M \dots M X'_1 \\
 &\qquad\qquad\qquad \left(Y W_2 M' M' \dots M' X_1 \right)^{k-2} Y W_2 N_2 N_2 \dots N_2 \\
 &\qquad\qquad\qquad \left(Y' W'_2 M M \dots M X'_1 \right) Y' W'_2 N'_2 N'_2 \dots N'_2 .
 \end{aligned} \tag{29}$$

Case 3: The string (6) combines with (23) which produces new strings

$$\begin{aligned}
 &N_1 N_1 \dots N_1 X_1 \left(Y W_2 M' M' \dots M' X_1 \right)^{k-2} Y W_2 M' M' \dots M' X_1 Y W_2 \\
 &N'_1 N'_1 \dots N'_1 X'_1 \left(Y' W'_2 M M \dots M X'_1 \right) Y' W'_2 M M \dots M X'_1 Y' W'_2 \\
 &\qquad\qquad\qquad M' M' \dots M' X_1 Y X_2 N'_1 N'_1 \dots N'_1 \\
 &\qquad\qquad\qquad M M \dots M X'_1 Y' X'_2 N_1 N_1 \dots N_1 .
 \end{aligned} \tag{30}$$

Case 4: The string (2) combines with (24) which produces new strings

$$\begin{aligned}
 &N'_2 N'_2 \dots N'_2 W_1 \left(Y X_2 M M \dots M W_1 \right)^{k-2} Y X_2 M M \dots M W_1 Y X_2 \\
 &N_2 N_2 \dots N_2 W'_1 \left(Y' X'_2 M' M' \dots M' W'_1 \right) Y' X'_2 M' M' \dots M' W'_1 Y' X'_2 \\
 &\qquad\qquad\qquad M M \dots M W_1 Y W_2 N_2 N_2 \dots N_2 \\
 &\qquad\qquad\qquad M' M' \dots M' W'_1 Y' W'_2 N'_2 N'_2 \dots N'_2 .
 \end{aligned} \tag{31}$$

Case 5: The string (10) combines with (22) which produces new strings

$$\begin{aligned}
 &N'_2 N'_2 \dots N'_2 W_1 Y X_2 M M \dots M W_1 Y X_2 M M \dots M W_1 \\
 &N_2 N_2 \dots N_2 W'_1 Y' X'_2 M' M' \dots M' W'_1 Y' X'_2 M' M' \dots M' W'_1 \\
 &\qquad\qquad\qquad \left(Y X_2 M M \dots M W_1 \right)^{k-2} Y X_2 N'_1 N'_1 \dots N'_1 \\
 &\qquad\qquad\qquad \left(Y' X'_2 M' M' \dots M' W'_1 \right) Y' X'_2 N_1 N_1 \dots N_1 .
 \end{aligned} \tag{32}$$

Case 6: The string (7) combines with (25) and (26) which produces new strings

$$\begin{aligned}
 &N'_2 N'_2 \dots N'_2 W_1 Y W_2 M' M' \dots M' X_1 Y W_2 M' M' \dots M' X_1 \\
 &N_2 N_2 \dots N_2 W'_1 Y' W'_2 M M \dots M X'_1 Y' W'_2 M M \dots M X'_1 \\
 &\left(Y W_2 M' M' \dots M' X_1 \right)^{k-2} Y W_2 N_2 N_2 \dots N_2 \\
 &\left(Y' W'_2 M M \dots M X'_1 \right) Y' W'_2 N'_2 N'_2 \dots N'_2
 \end{aligned} \tag{33}$$

and

$$\begin{aligned}
 &N'_2 N'_2 \dots N'_2 W_1 Y W_2 M' M' \dots M' X_1 Y W_2 M' M' \dots M' X_1 \\
 &N_2 N_2 \dots N_2 W'_1 Y' W'_2 M M \dots M X'_1 Y' W'_2 M M \dots M X'_1 \\
 &\left(Y W_2 M' M' \dots M' X_1 \right)^{k-2} Y X_2 N'_1 N'_1 \dots N'_1 \\
 &\left(Y' W'_2 M M \dots M X'_1 \right) Y' X'_2 N_1 N_1 \dots N_1 .
 \end{aligned} \tag{34}$$

By simplifying strings (27), (28), (29), (30), (31), (32), (33) and (34), the resulting strings are:

$$\begin{aligned}
 &N_1 N_1 \dots N_1 X_1 \left(Y X_2 M M \dots M W_1 \right)^{(k+1)-1} Y W_2 N_2 N_2 \dots N_2 \\
 &N'_1 N'_1 \dots N'_1 X'_1 \left(Y' X'_2 M' M' \dots M' W'_1 \right) Y' W'_2 N'_2 N'_2 \dots N'_2 , \\
 &N_1 N_1 \dots N_1 X_1 \left(Y X_2 M M \dots M W_1 \right)^{(k+1)-1} Y X_2 N'_1 N'_1 \dots N'_1 \\
 &N'_1 N'_1 \dots N'_1 X'_1 \left(Y' X'_2 M' M' \dots M' W'_1 \right) Y' X'_2 N_1 N_1 \dots N_1 , \\
 &N_1 N_1 \dots N_1 X_1 \left(Y W_2 M' M' \dots M' X_1 \right)^{(k+1)-1} Y W_2 N_2 N_2 \dots N_2 \\
 &N'_1 N'_1 \dots N'_1 X'_1 \left(Y' W'_2 M M \dots M X'_1 \right) Y' W'_2 N'_2 N'_2 \dots N'_2 , \\
 &N_1 N_1 \dots N_1 X_1 \left(Y W_2 M' M' \dots M' X_1 \right)^{(k+1)-1} Y X_2 N'_1 N'_1 \dots N'_1 \\
 &N'_1 N'_1 \dots N'_1 X'_1 \left(Y' W'_2 M M \dots M X'_1 \right) Y' X'_2 N_1 N_1 \dots N_1 , \\
 &N'_2 N'_2 \dots N'_2 W_1 \left(Y X_2 M M \dots M W_1 \right)^{(k+1)-1} Y W_2 N_2 N_2 \dots N_2 \\
 &N_2 N_2 \dots N_2 W'_1 \left(Y' X'_2 M' M' \dots M' W'_1 \right) Y' W'_2 N'_2 N'_2 \dots N'_2 , \\
 &N'_2 N'_2 \dots N'_2 W_1 \left(Y X_2 M M \dots M W_1 \right)^{(k+1)-1} Y X_2 N'_1 N'_1 \dots N'_1 \\
 &N_2 N_2 \dots N_2 W'_1 \left(Y' X'_2 M' M' \dots M' W'_1 \right) Y' X'_2 N_1 N_1 \dots N_1 , \\
 &N'_2 N'_2 \dots N'_2 W_1 \left(Y W_2 M' M' \dots M' X_1 \right)^{(k+1)-1} Y W_2 N_2 N_2 \dots N_2 \\
 &N_2 N_2 \dots N_2 W'_1 \left(Y' W'_2 M M \dots M X'_1 \right) Y' W'_2 N'_2 N'_2 \dots N'_2 , \text{ and}
 \end{aligned}$$

$$\begin{matrix} N'_2 N'_2 \dots N'_2 W_1 & (Y W_2 M' M' \dots M' X_1)^{(k+1)-1} & Y X_2 N'_1 N'_1 \dots N'_1 \\ N_2 N_2 \dots N_2 W'_1 & (Y' W'_2 M M \dots M X'_1) & Y' X'_2 N_1 N_1 \dots N_1 \end{matrix}$$

Therefore, the resulting splicing language can be summarised

$$\begin{aligned} & \left(\begin{matrix} N_1 N_1 \dots N_1 X_1 + N'_2 N'_2 \dots N'_2 W_1 \\ N'_1 N'_1 \dots N'_1 X'_1 + N_2 N_2 \dots N_2 W'_1 \end{matrix} \right) \\ & \left\{ Y \left(X_2 M M \dots M W_1 + W_2 M' M' \dots M' X_1 \right) \right\}^{(k+1)-1} Y \\ & \left\{ Y' \left(X'_2 M' M' \dots M' W'_1 + W'_2 M M \dots M X'_1 \right) \right\} Y' \\ & \left(\begin{matrix} W_2 N_2 N_2 \dots N_2 + X_2 N'_1 N'_1 \dots N'_1 \\ W'_2 N'_2 N'_2 \dots N'_2 + X'_2 N_1 N_1 \dots N_1 \end{matrix} \right). \end{aligned} \tag{35}$$

So, string (35) is true since it depicts string (1) when $n = k + 1$. Hence, Theorem 1 is proved. \square

Next, Theorem 2 shows the generalisation of splicing language from DNA splicing systems with two palindromic rules for different crossings.

Theorem 2

Given $S = (A, I, B, C)$ is a splicing system in which $A = \begin{Bmatrix} A, C, G, T \\ T, G, C, A \end{Bmatrix}$ is the set of dsDNA symbols, $I = \{ N_1 N_1 \dots N_1 X_1 Y X_2 M M \dots M W_1 Z W_2 N_2 N_2 \dots N_2 \}$ is the set consisting of an initial string with non-overlapping cutting sites of two palindromic rules $\begin{matrix} X_1 Y X_2 \\ X'_1 Y' X'_2 \end{matrix}$ and $\begin{matrix} W_1 Z W_2 \\ W'_1 Z' W'_2 \end{matrix}$ where $\begin{matrix} N_1 & X_1 & Y & X_2 & M & W_1 & Z \\ N'_1 & X'_1 & Y' & X'_2 & M' & W'_1 & Z' \end{matrix}$ W_2 and N_2 are variables used to denote any arbitrary dsDNA and $N'_1, X'_1, Y', X'_2, M', W'_1, Z', W'_2$ and N'_2 are complementaries for $N_1, X_1, Y, X_2, M, W_1, Z, W_2$, and N_2 respectively, set $B = \left\{ \begin{pmatrix} X_1, Y, X_2 \\ X'_1, Y', X'_2 \end{pmatrix}, \begin{pmatrix} W_1, Z, W_2 \\ W'_1, Z', W'_2 \end{pmatrix} \right\}$ is the set of cleavage pattern for rules where $\begin{matrix} Y \\ Y' \end{matrix}$ and $\begin{matrix} Z \\ Z' \end{matrix}$ are the different crossings and set C is the empty set, the resulting splicing language consists of strings of the form

$$\begin{aligned} & \left\{ \begin{matrix} N_1 N_1 \dots N_1 X_1 Y X_2 (M M \dots M W_1 Z W_2 M' M' \dots M' X_1 Y X_2)^{n-1} \\ N'_1 N'_1 \dots N'_1 X'_1 Y' X'_2 (M' M' \dots M' W'_1 Z' W'_2 M M \dots M X_1 Y' X'_2) \end{matrix} \right\} + \\ & \left\{ \begin{matrix} (N'_1 N'_1 \dots N'_1 + M M \dots M W_1 Z W_2 N_2 N_2 \dots N_2) \\ (N_1 N_1 \dots N_1 + M' M' \dots M' W'_1 Z' W'_2 N'_2 N'_2 \dots N'_2) \end{matrix} \right\} \\ & \left\{ \begin{matrix} N'_2 N'_2 \dots N'_2 W_1 Z W_2 (M' M' \dots M' X_1 Y X_2 M M \dots M W_1 Z W_2)^{n-1} \\ N_2 N_2 \dots N_2 W'_1 Z' W'_2 (M M \dots M X'_1 Y' X'_2 M' M' \dots M' W'_1 Z' W'_2) \end{matrix} \right\} \\ & \left\{ \begin{matrix} (N_2 N_2 \dots N_2 + M' M' \dots M' X_1 Y X_2 N'_1 N'_1 \dots N'_1) \\ (N'_2 N'_2 \dots N'_2 + M M \dots M X'_1 Y' X'_2 N_1 N_1 \dots N_1) \end{matrix} \right\} \end{aligned} \tag{36}$$

where $n \in \mathbb{Z}^+$ represents multiple copies of strings and $\begin{pmatrix} X_1 Y X_2 & W_1 Z W_2 \\ X'_1 Y' X'_2 & W'_1 Z' W'_2 \end{pmatrix} \notin \{ N_1 N_1 \dots N_1 M M \dots M N_2 N_2 \dots N_2 \}$ indicates no other cutting site is present in strings $\begin{matrix} N_1 N_1 \dots N_1 \\ N'_1 N'_1 \dots N'_1 \\ M M \dots M \\ M' M' \dots M' \end{matrix}$ and $\begin{matrix} N_2 N_2 \dots N_2 \\ N'_2 N'_2 \dots N'_2 \end{matrix}$.

Proof.

Suppose the rules $\begin{matrix} X_1 Y X_2 \\ X'_1 Y' X'_2 \end{matrix}$ and $\begin{matrix} W_1 Z W_2 \\ W'_1 Z' W'_2 \end{matrix}$ are palindromic, so the base sequence of rules can read the same backward and forward:

$$\begin{matrix} X_1 Y X_2 & = & X'_2 Y' X'_1 \\ X'_1 Y' X'_2 & = & X_2 Y X_1 \end{matrix}$$

and

$$\begin{matrix} W_1 Z W_2 & = & W'_2 Z' W'_1 \\ W'_1 Z' W'_2 & = & W_2 Z W_1 \end{matrix}$$

Then $\begin{matrix} X_1 & = & X'_2 & Y & = & Y' & X_2 & = & X'_1 & W_1 & = & W'_2 & Z & = & Z' & \text{and} & W_2 & = & W'_1 \\ X'_1 & = & X_2 & Y' & = & Y & X'_2 & = & X_1 & W'_1 & = & W_2 & Z' & = & Z & \text{and} & W'_2 & = & W_1 \end{matrix}$.

The initial string $N_1 N_1 \dots N_1 X_1 Y X_2 M M \dots M W_1 Z W_2 N_2 N_2 \dots N_2$ with the cutting site $\begin{matrix} X_1 Y X_2 \\ X'_1 Y' X'_2 \end{matrix}$ and $\begin{matrix} W_1 Z W_2 \\ W'_1 Z' W'_2 \end{matrix}$

is shown respectively in the following:

$$\begin{matrix} N_1 N_1 \dots N_1 X_1 \downarrow Y & X_2 M M \dots M W_1 Z W_2 N_2 N_2 \dots N_2 \\ N'_1 N'_1 \dots N'_1 X'_1 & Y' \uparrow X'_2 M' M' \dots M' W'_1 Z' W'_2 N'_2 N'_2 \dots N'_2 \end{matrix} \quad (37)$$

for the first cutting site and

$$\begin{matrix} N_1 N_1 \dots N_1 X_1 Y X_2 M M \dots M W_1 \downarrow Z & W_2 N_2 N_2 \dots N_2 \\ N'_1 N'_1 \dots N'_1 X'_1 Y' X'_2 M' M' \dots M' W'_1 & Z' \uparrow W'_2 N'_2 N'_2 \dots N'_2 \end{matrix} \quad (38)$$

for the second cutting site.

The initial string can be written 180 degree wise as

$$\begin{matrix} N'_2 N'_2 \dots N'_2 W'_2 Z' W'_1 M' M' \dots M' X'_2 Y' X'_1 N'_1 N'_1 \dots N'_1 \\ N_2 N_2 \dots N_2 W_2 Z W_1 M M \dots M X_2 Y X_1 N_1 N_1 \dots N_1 \end{matrix} \quad (39)$$

Since $\begin{matrix} X_1 = X'_2 & Y = Y' & X_2 = X'_1 & W_1 = W'_2 & Z = Z' & \text{and} & W_2 = W'_1 \\ X'_1 = X_2 & Y' = Y & X'_2 = X_1 & W'_1 = W_2 & Z' = Z & \text{and} & W'_2 = W_1 \end{matrix}$, string (39) becomes

$$\begin{matrix} N'_2 N'_2 \dots N'_2 W_1 Z W_2 M' M' \dots M' X_1 Y X_2 N'_1 N'_1 \dots N'_1 \\ N_2 N_2 \dots N_2 W'_1 Z' W'_2 M M \dots M X'_1 Y' X'_2 N_1 N_1 \dots N_1 \end{matrix} \quad (40)$$

The string (40) with the cutting site $\begin{matrix} X_1 Y X_2 \\ X'_1 Y' X'_2 \end{matrix}$ and $\begin{matrix} W_1 Z W_2 \\ W'_1 Z' W'_2 \end{matrix}$ is shown respectively in the following:

$$\begin{matrix} N'_2 N'_2 \dots N'_2 W_1 \downarrow Z & W_2 M' M' \dots M' X_1 Y X_2 N'_1 N'_1 \dots N'_1 \\ N_2 N_2 \dots N_2 W'_1 & Z' \uparrow W'_2 M M \dots M X'_1 Y' X'_2 N_1 N_1 \dots N_1 \end{matrix} \quad (41)$$

and

$$\begin{matrix} N'_2 N'_2 \dots N'_2 W_1 Z W_2 M' M' \dots M' X_1 \downarrow Y & X_2 N'_1 N'_1 \dots N'_1 \\ N_2 N_2 \dots N_2 W'_1 Z' W'_2 M M \dots M X'_1 & Y' \uparrow X'_2 N_1 N_1 \dots N_1 \end{matrix} \quad (42)$$

Using the rule $\begin{matrix} X_1 Y X_2 \\ X'_1 Y' X'_2 \end{matrix}$, the string (37) combines with (42) which gives:

$$\begin{matrix} N_1 N_1 \dots N_1 X_1 Y X_2 N'_1 N'_1 \dots N'_1 \\ N'_1 N'_1 \dots N'_1 X'_1 Y' X'_2 N_1 N_1 \dots N_1 \end{matrix} \quad (43)$$

and

$$\begin{matrix} N'_2 N'_2 \dots N'_2 W_1 Z W_2 M' M' \dots M' X_1 Y X_2 M M \dots M W_1 \downarrow Z & W_2 \\ N_2 N_2 \dots N_2 W'_1 Z' W'_2 M M \dots M X'_1 Y' X'_2 M' M' \dots M' W'_1 & Z' \uparrow W'_2 \\ & W_2 N_2 N_2 \dots N_2 \\ & W'_2 N'_2 N'_2 \dots N'_2 \end{matrix} \quad (44)$$

Using the rule $\begin{matrix} W_1 Z W_2 \\ W'_1 Z' W'_2 \end{matrix}$, the results of the combination of string (38) with (41) are:

$$\begin{matrix} N_1 N_1 \dots N_1 X_1 Y X_2 M M \dots M W_1 Z W_2 M' M' \dots M' X_1 \downarrow Y & X_2 \\ N'_1 N'_1 \dots N'_1 X'_1 Y' X'_2 M' M' \dots M' W'_1 Z' W'_2 M M \dots M X'_1 & Y' \uparrow X'_2 \\ & X_2 N'_1 N'_1 \dots N'_1 \\ & X'_2 N_1 N_1 \dots N_1 \end{matrix} \quad (45)$$

and

$$\begin{matrix} N'_2 N'_2 \dots N'_2 W_1 Z W_2 N_2 N_2 \dots N_2 \\ N_2 N_2 \dots N_2 W'_1 Z' W'_2 N'_2 N'_2 \dots N'_2 \end{matrix} \quad (46)$$

By using induction, this theorem can be proved. For $n = 1$, the string (36) is true since it is stated as strings (37), (40), (43) and (46). Next, let $n = k \in \mathbb{Z}^+$:

$$\left\{ \begin{matrix} N_1 N_1 \dots N_1 X_1 Y X_2 (M M \dots M W_1 Z W_2 M' M' \dots M' X_1 Y X_2)^{k-1} \\ N'_1 N'_1 \dots N'_1 X'_1 Y' X'_2 (M' M' \dots M' W'_1 Z' W'_2 M M \dots M X'_1 Y' X'_2) \\ (N'_1 N'_1 \dots N'_1 + M M \dots M W_1 Z W_2 N_2 N_2 \dots N_2) \\ (N_1 N_1 \dots N_1 + M' M' \dots M' W'_1 Z' W'_2 N'_2 N'_2 \dots N'_2) \end{matrix} \right\} + \left\{ \begin{matrix} N'_2 N'_2 \dots N'_2 W_1 Z W_2 (M' M' \dots M' X_1 Y X_2 M M \dots M W_1 Z W_2)^{k-1} \\ N_2 N_2 \dots N_2 W'_1 Z' W'_2 (M M \dots M X'_1 Y' X'_2 M' M' \dots M' W'_1 Z' W'_2) \\ (N_2 N_2 \dots N_2 + M' M' \dots M' X_1 Y X_2 N'_1 N'_1 \dots N'_1) \\ (N'_2 N'_2 \dots N'_2 + M M \dots M X'_1 Y' X'_2 N_1 N_1 \dots N_1) \end{matrix} \right\} \quad (47)$$

The following strings are among the strings in (47):

$$\begin{aligned}
 &N_1 N_1 \dots N_1 X_1 \downarrow Y \quad X_2 \left(M M \dots M W_1 Z W_2 M' M' \dots M' X_1 Y X_2 \right)^{k-1} \\
 &N'_1 N'_1 \dots N'_1 X'_1 \quad Y' \uparrow X'_2 \left(M' M' \dots M' W'_1 Z' W'_2 M M \dots M X'_1 Y' X'_2 \right) \\
 &N'_1 N'_1 \dots N'_1 \\
 &N_1 N_1 \dots N_1 ,
 \end{aligned} \tag{48}$$

$$\begin{aligned}
 &N_1 N_1 \dots N_1 X_1 \downarrow Y \quad X_2 \left(M M \dots M W_1 Z W_2 M' M' \dots M' X_1 Y X_2 \right)^{k-1} \\
 &N'_1 N'_1 \dots N'_1 X'_1 \quad Y' \uparrow X'_2 \left(M' M' \dots M' W'_1 Z' W'_2 M M \dots M X'_1 Y' X'_2 \right) \\
 &M M \dots M W_1 Z W_2 N_2 N_2 \dots N_2 \\
 &M' M' \dots M' W'_1 Z' W'_2 N'_2 N'_2 \dots N'_2 ,
 \end{aligned} \tag{49}$$

$$\begin{aligned}
 &N'_2 N'_2 \dots N'_2 W_1 \downarrow Z \quad W_2 \left(M' M' \dots M' X_1 Y X_2 M M \dots M W_1 Z W_2 \right)^{k-1} \\
 &N_2 N_2 \dots N_2 W'_1 \quad Z' \uparrow W'_2 \left(M M \dots M X'_1 Y' X'_2 M' M' \dots M' W'_1 Z' W'_2 \right) \\
 &N_2 N_2 \dots N_2 \quad \text{and} \\
 &N'_2 N'_2 \dots N'_2 ,
 \end{aligned} \tag{50}$$

$$\begin{aligned}
 &N'_2 N'_2 \dots N'_2 W_1 \downarrow Z \quad W_2 \left(M' M' \dots M' X_1 Y X_2 M M \dots M W_1 Z W_2 \right)^{k-1} \\
 &N_2 N_2 \dots N_2 W'_1 \quad Z' \uparrow W'_2 \left(M M \dots M X'_1 Y' X'_2 M' M' \dots M' W'_1 Z' W'_2 \right) \\
 &M' M' \dots M' X_1 Y X_2 N'_1 N'_1 \dots N'_1 \\
 &M M \dots M X'_1 Y' X'_2 N_1 N_1 \dots N_1 .
 \end{aligned} \tag{51}$$

Using the rules $\frac{X_1 Y X_2}{X'_1 Y' X'_2}$ and $\frac{W_1 Y W_2}{W'_1 Y' W'_2}$, the above resulting strings combine with (44) and (45). Hence, the other recombinations can be shown through two cases.

Case 1: The string (45) combines with (48) and (49) which produces new strings

$$\begin{aligned}
 &N_1 N_1 \dots N_1 X_1 Y X_2 M M \dots M W_1 Z W_2 M' M' \dots M' X_1 Y X_2 \\
 &N'_1 N'_1 \dots N'_1 X'_1 Y' X'_2 M' M' \dots M' W'_1 Z' W'_2 M M \dots M X'_1 Y' X'_2 \\
 &\left(M M \dots M W_1 Z W_2 M' M' \dots M' X_1 Y X_2 \right)^{k-1} N'_1 N'_1 \dots N'_1 \\
 &N_1 N_1 \dots N_1
 \end{aligned} \tag{52}$$

and

$$\begin{aligned}
 &N_1 N_1 \dots N_1 X_1 Y X_2 M M \dots M W_1 Z W_2 M' M' \dots M' X_1 Y X_2 \\
 &N'_1 N'_1 \dots N'_1 X'_1 Y' X'_2 M' M' \dots M' W'_1 Z' W'_2 M M \dots M X'_1 Y' X'_2 \\
 &\left(M M \dots M W_1 Z W_2 M' M' \dots M' X_1 Y X_2 \right)^{k-1} M M \dots M W_1 Z W_2 \\
 &\left(M' M' \dots M' W'_1 Z' W'_2 M M \dots M X'_1 Y' X'_2 \right) M' M' \dots M' W'_1 Z' W'_2 \\
 &N_2 N_2 \dots N_2 \\
 &N'_2 N'_2 \dots N'_2 .
 \end{aligned} \tag{53}$$

Case 2: The string (44) combines with (50) and (51) which produces new strings

$$\begin{aligned}
 &N'_2 N'_2 \dots N'_2 W_1 Z W_2 M' M' \dots M' X_1 Y X_2 M M \dots M W_1 Z W_2 \\
 &N_2 N_2 \dots N_2 W'_1 Z' W'_2 M M \dots M X'_1 Y' X'_2 M' M' \dots M' W'_1 Z' W'_2 \\
 &\left(M' M' \dots M' X_1 Y X_2 M M \dots M W_1 Z W_2 \right)^{k-1} N_2 N_2 \dots N_2 \\
 &\left(M M \dots M X'_1 Y' X'_2 M' M' \dots M' W'_1 Z' W'_2 \right) N'_2 N'_2 \dots N'_2
 \end{aligned} \tag{54}$$

and

$$\begin{aligned}
 &N'_2 N'_2 \dots N'_2 W_1 Z W_2 M' M' \dots M' X_1 Y X_2 M M \dots M W_1 Z W_2 \\
 &N_2 N_2 \dots N_2 W'_1 Z' W'_2 M M \dots M X'_1 Y' X'_2 M' M' \dots M' W'_1 Z' W'_2 \\
 &\left(M' M' \dots M' X_1 Y X_2 M M \dots M W_1 Z W_2 \right)^{k-1} M' M' \dots M' X_1 Y X_2 \\
 &\left(M M \dots M X'_1 Y' X'_2 M' M' \dots M' W'_1 Z' W'_2 \right) M M \dots M X'_1 Y' X'_2 \\
 &N'_1 N'_1 \dots N'_1 \\
 &N_1 N_1 \dots N_1 .
 \end{aligned} \tag{55}$$

By simplifying strings (52), (53), (53) and (55), the resulting strings are:

$$\begin{aligned}
 &N_1 N_1 \dots N_1 X_1 Y X_2 \left(M M \dots M W_1 Z W_2 M' M' \dots M' X_1 Y X_2 \right)^{(k+1)-1} \\
 &N'_1 N'_1 \dots N'_1 X'_1 Y' X'_2 \left(M' M' \dots M' W'_1 Z' W'_2 M M \dots M X'_1 Y' X'_2 \right) \\
 &N'_1 N'_1 \dots N'_1 \\
 &N_1 N_1 \dots N_1 ,
 \end{aligned}$$

$$\begin{aligned}
 &N_1 N_1 \dots N_1 X_1 Y X_2 \left(M M \dots M W_1 Z W_2 M' M' \dots M' X_1 Y X_2 \right)^{(k+1)-1} \\
 &N'_1 N'_1 \dots N'_1 X'_1 Y' X'_2 \left(M' M' \dots M' W'_1 Z' W'_2 M M \dots M X'_1 Y' X'_2 \right) \\
 &M M \dots M W_1 Z W_2 N_2 N_2 \dots N_2 \\
 &M' M' \dots M' W'_1 Z' W'_2 N'_2 N'_2 \dots N'_2 ,
 \end{aligned}$$

$$\begin{aligned}
 &N'_2 N'_2 \dots N'_2 W_1 Z W_2 \left(M' M' \dots M' X_1 Y X_2 M M \dots M W_1 Z W_2 \right)^{(k+1)-1} \\
 &N_2 N_2 \dots N_2 W'_1 Z' W'_2 \left(M M \dots M X'_1 Y' X'_2 M' M' \dots M' W'_1 Z' W'_2 \right)
 \end{aligned}$$

$$N_2 N_2 \dots N_2 \quad \text{and} \\ N_2' N_2' \dots N_2'$$

$$N_2' N_2' \dots N_2' W_1 Z W_2 \left(M' M' \dots M' X_1 Y X_2 M M \dots M W_1 Z W_2 \right)^{(k+1)-1} \\ N_2 N_2 \dots N_2 W_1' Z' W_2' \left(M M \dots M X_1' Y' X_2' M' M' \dots M' W_1' Z' W_2' \right) \\ M' M' \dots M' X_1 Y X_2 N_1' N_1' \dots N_1' \\ M M \dots M X_1' Y' X_2' N_1 N_1 \dots N_1 .$$

Therefore, the resulting splicing language can be summarised:

$$\left\{ \begin{array}{l} N_1 N_1 \dots N_1 X_1 Y X_2 \left(M M \dots M W_1 Z W_2 M' M' \dots M' X_1 Y X_2 \right)^{(k+1)-1} \\ N_1' N_1' \dots N_1' X_1' Y' X_2' \left(M' M' \dots M' W_1' Z' W_2' M M \dots M X_1' Y' X_2' \right) \\ \left(N_1' N_1' \dots N_1' + M M \dots M W_1 Z W_2 N_2 N_2 \dots N_2 \right) \\ \left(N_1 N_1 \dots N_1 + M' M' \dots M' W_1' Z' W_2' N_2' N_2' \dots N_2' \right) \end{array} \right\} + \\ \left\{ \begin{array}{l} N_2' N_2' \dots N_2' W_1 Z W_2 \left(M' M' \dots M' X_1 Y X_2 M M \dots M W_1 Z W_2 \right)^{(k+1)-1} \\ N_2 N_2 \dots N_2 W_1' Z' W_2' \left(M M \dots M X_1' Y' X_2' M' M' \dots M' W_1' Z' W_2' \right) \\ \left(N_2 N_2 \dots N_2 + M' M' \dots M' X_1 Y X_2 N_1' N_1' \dots N_1' \right) \\ \left(N_2' N_2' \dots N_2' + M M \dots M X_1' Y' X_2' N_1 N_1 \dots N_1 \right) \end{array} \right\} . \tag{56}$$

So, string (56) is true since it depicts string (36) when $n = k + 1$. Thus, Theorem 2 is proved. □

Conclusion

In this research, the generalisations of splicing languages in DNA splicing systems involving two (non-overlapping) palindromic rules for the same and different crossings are presented as Theorem 1 and 2 respectively. Both theorems are proved using the induction method. The generalised splicing languages consist of the combination of a set of double-stranded DNA strings resulting from DNA splicing systems using operations in formal languages; concatenation, union and star-closure. The results from this research can be used to obtain splicing languages from splicing systems with two rules which are taken from any palindromic restriction enzymes. For further research, modelling of splicing systems with some cases of overlapped and non-palindromic restriction enzymes can be discussed.

Acknowledgement

The first author would like to thank UTM for supporting her study through Zamalah Scholarship. The second and third authors would also like to thank the Ministry of Higher Education (MOHE) and Research Management Centre (RMC), Universiti Teknologi Malaysia (UTM) for the financial funding through Fundamental Research Grant Scheme (FRGS/1/2018/STG06/UTM/02/10).

References

- [1] T. Head, *B. Math. Biol* 49 (1987) 737-759.
- [2] T. Palmer, P.L. Bonner, *Enzymes: Biochemistry, Biotechnology, Clinical Chemistry*, 2nd ed, Woodhead Publishing Limited, UK, 2007.
- [3] P. Linz, *An Introduction to Formal Languages and Automata*, 4th ed, Jones and Bartlett Publisher, USA, 2006.
- [4] T. Head, *Discrete. Appl. Math.* 87 (1998) 139-147.
- [5] G. Păun, *Discrete. Appl. Math.* 70 (1996) 57-79.
- [6] D. Pixton, *Discrete. Appl. Math.* 69 (1996) 101-124.
- [7] Y. Yusof, N.H. Sarmin, W.H. Fong, T.E. Goode, M.A. Ahmad, *An Analysis of Four Variants of Splicing System, Proceeding of the 20th National Symposium on Mathematical Sciences - Research in Mathematical Sciences: A Catalyst for Creativity and Innovation (SKSM 2012), December 18-20, 2012, AIP Conference Proceedings, Melville, NY, 2013, p. 888-895.*
- [8] E. Goode, D. Pixton, in: Jonoska, N., Păun, G., Rozenberg, G. (Eds.), *Aspects of Molecular Computing, Lecture Notes in Computer Science*, Springer-Verlag, Germany, 2004, p. 189-201.
- [9] T.E.G. Laun, *Constants and Splicing Systems*, Ph.D. Thesis, State University of New York, 1999.
- [10] I. Tomohiro, S. Inenaga, M. Takeda, *Theor. Comput. Sci.* 483 (2013) 162-170.
- [11] Y. Yusof, *DNA Splicing System Inspired by Bio Molecular Operations*, Ph.D. Thesis, Universiti Teknologi Malaysia, 2012.
- [12] H.M. Eun, *Enzymology Primer for Recombinant DNA Technology*, Academic Press, USA, 1996.
- [13] W.L. Lim, Y. Yusof, M.H. Mudaber, *Modeling of DNA single stage splicing language via Yusof-Goode approach: One string with two rules*, *Proceeding of the 2nd ISM International Statistical Conference 2014 (ISM-II)*, Kuantan, Malaysia, August 12-14, 2014, AIP Conference Proceedings, Melville, NY, 2015, p. 695-699.

- [14] M.H. Mudaber, Y. Yusof, M.S. Mohamad, Some Sufficient Conditions for Persistency and Permanency of Two Stages DNA Splicing Languages via Yusof-Goode Approach, Proceeding of the 21th National Symposium on Mathematical Sciences (SKSM21), Penang, Malaysia, November 6-8, 2013, AIP Conference Proceedings, Melville, NY, 2014, p. 591-595.
- [15] M.A. Ahmad, N.H. Sarmin, Y. Yusof, W.H. Fong, Some Restrictions on the Existence of Second Order Limit Language, Proceeding of the 22th National Symposium on Mathematical Sciences (SKSM22), Selangor, Malaysia, November 24-26, 2014, AIP Conference Proceedings, Melville, NY, 2015, p. 020048.
- [16] W.H. Fong, N.I. Ismail, MATEMATIKA 34 (2018) 59-71.
- [17] R.M. Horton, S.N. Ho, J.K. Pullen, H.D. Hunt, Z. Cai, L.R. Pease, in: Wu, R. (Eds.), Methods in enzymology, Academic Press, 1995, p. 363-372.
- [18] J.D. Watson, T.A. Baker, S.P. Bell, A. Gann, M. Levine, R. Losick, Molecular Biology of the Gene, 7th ed, Cold Spring Harbor Laboratory Press, New York, 2014.