BIOMOLECULAR ASPECTS OF SECOND ORDER LIMIT LANGUAGE

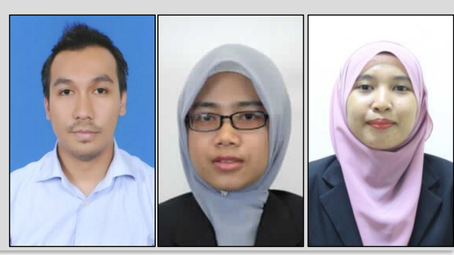
Muhammad Azrin Ahmad[a]*, Nor Haniza Sarmin[b], Fong Wan Heng[b], Mohd Firdaus Abdul-Wahab[c], Yuhani Yusof[a] & Noraziah Adzhar[a]

[a] Applied & Industrial Mathematics (AIMs), Faculty of Industrial Sciences & Technology, Universiti Malaysia Pahang.

[b]Department of Mathematical Sciences, Faculty of Science, Universiti Teknologi Malaysia, Malaysia.

[c]Department of Bioscience & Health Sciences, Faculty of Biosciences & Medical Engineering, Universiti Teknologi Malaysia, Malaysia.

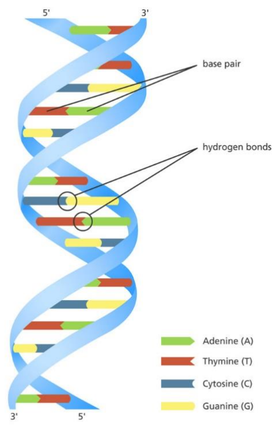
Email: azrin@ump.edu.my*

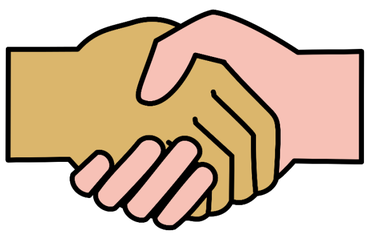


MOTIVATION

DNA plays its significant role in hereditary processes

The presentation of DNA molecules as a series of alphabets, namely A, C, G and T inspired the mathematical modelling of the slicing system





Mathematical modelling of the splicing system is an interdisciplinary study between formal language theory and informational macromolecules.

Second order limit language is an extension of first order limit language (limit language), $L(S)$ which is a distinct set of strings from $L(S)$.

MATHEMATICAL MODELLING

Let $S=(A,I,R)$ be a Y-G splicing system consisting of a set of alphabets, $A=\{a,c,g,t\}$, a set of initial strings, $I=\{\alpha gatc\beta gatc\gamma\}$, and a set of rules, $R=\{r\}$ such that $r=(1;gatc,1:1;gatc,1)$ where $\alpha,\beta,\gamma\in A^*$.

Based on the definition of the second order limit language, $L_2(S)\cap L(S)=\varnothing$ since $L_1(S)\subseteq L(S)$ thus $L_1(S)\not\subset L_2(S)$. Hence, by further splicing the set of strings of $L(S)$, the second order limit language is presented in the following general form:

$$L_2(S)\xrightarrow{R} I\cup\begin{Bmatrix}\alpha(gatc\beta\cup gatc\beta')^* gatc\alpha',\\ \gamma'(gatc\beta'\cup gatc\beta)^* gatc\gamma,\\ \alpha(gatc\beta'\cup gatc\beta)^* gatc\gamma\end{Bmatrix}.$$

NOVELTY: Validating second order limit language through biological process

ACHIEVEMENT: Obtain COPYRIGHT from ICC UTM

PUBLICATION: Biomolecular Aspects of Second Order Limit Language, Malaysian Journal of Fundamental and Applied Sciences (ISI). In Press

MARKET: MOHE, NEB

EXPERIMENTAL FRAMEWORK

Select suitable DNA

Fragment 1 of this lambda phage DNA is the fragment of interest for the initial string *I* since it contains two times the site for the restriction enzyme *Dpn*II. The genome location for the strand is between 5396 and 5549 which gives 154bp long.

Fragment 1*: A-DpnII* site-*B-DpnII* site-*D*

$|A|=68\text{bp}$

$|DpnII|=4\text{bp}$

$|B|=38\text{bp}$

$|DpnII|=4\text{bp}$

$|D|=40\text{bp}.$

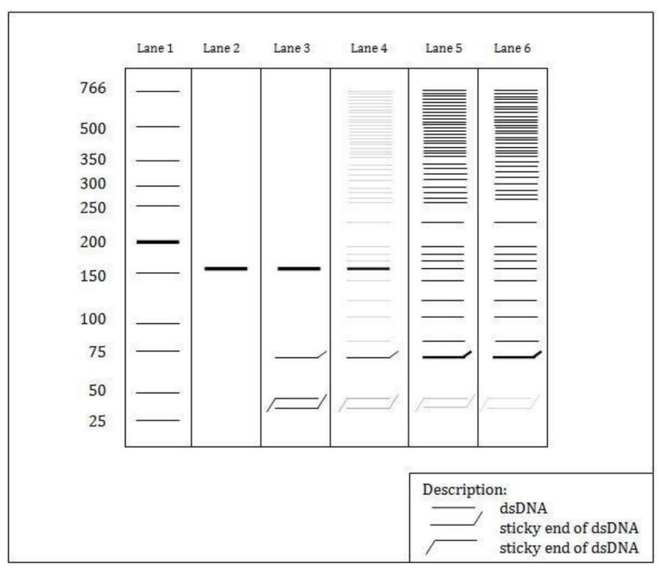
Visualising the gel under UV rays using UV transilluminator

Performing polyacrylamide gel electrophorosis

CONCLUSION

Table 1 The size (bp) of predicted molecules

No.	Molecule	Size (bp)
1.	A	68
2.	B	42
3.	D	44
4.	$A-B-D$	154
5.	$A-D$	112
6.	$A-A'$	140
7.	$D'-D$	84
8.	$A-B-A'$	182
9.	$D'-B-D$	126
10.	$A-B'-D$	154
11.	$A-B-B'-D$	196
12.	$A-B-B'-A'$	224
13.	$A-B'-B-D$	196
14.	$A-B'-B-A'$	224
15.	$D'-B-B'-D$	168
16.	$D'-B'-B-D$	168



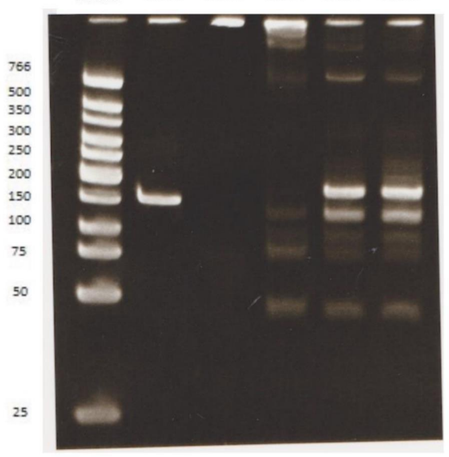


Fig. 2 Predicted gel of the second order limit language

Fig. 3 Gel photo with the splicing pattern of enzyme *Dpn*II.

In conclusion, it was shown that the second order limit language is proven to exist through an experiment. It can be concluded that the mathematical model of the second order limit language has been verified experimentally since the dsDNA molecules produced in the experiment are as predicted in the model.

REFERENCES

1. Head, T. Formal Language Theory and DNA: An Analysis of the Generative Capacity of Specific Recombinant Behaviors. Bull. Math. Biol. 1987. 49 (6):737-759
2. Goode, E. and Pixton, D. Wet Splicing System. DIMAC Series in Discrete Mathematics and Theoretical Computer Science. 73-83; 1999.
3. Watson, J . D. and Crick, F. H. C. Molecular Structure of Nucleic Acids. Nature. 1953. 171(4356): 737-738.

ACKNOWLEDGEMENT

The authors gratefully acknowledge the financial supports received from Universiti Malaysia Pahang, UMP (Vot RDU1703278)