DNA SEQUENCE ALLIGNMENT ACCELERATOR

USING MATLAB

NORAMINAH BINTI ISMAIL

FACULTY OF ELECTRICAL ENGINEERING

UNIVERSITI TEKNOLOGI MARA

MALAYSIA

ACKNOWLEDGEMENT

In the name of Allah the Most Gracious and Most Merciful

First of all I would like to forward my thankfulness to **Allah S.W.T** for giving me the strength, courage and opportunity to pull it through thus granting me the capability to proceed successfully. I will not be able to finish my thesis without the assistance and guidance of several people. I would therefore like to offer my sincere thanks to all of them who has contribute in different way

To my supervisor, En.Syed Mutalib Al - Junid; your valuable assistance in guidance and encouragement is much appreciated.

To my husband; I am most grateful for all your love and support and for the endless prayers - this helped me get through some difficult times in the course of this study. Thank you for being so patient with me.

To my family, thank you for the love, encouragement and support through out the years of my studies.

To my friends and others who indirectly involved in this project, thank you so much for your help and support in many ways.

ABSTRACT

This paper presents the development of DNA sequence alignment accelerator using MATLAB. The Smith Waterman algorithm has been developed and simulated for nucleotide sequence alignment in optimizing accelerating speed. The development of computational method become the key factor in biomedical engineering and application as the rate of the biological sequence databases has exponentially growth thus increasing the complexity of the DNA sequences. The software development has been run in Intel Core i5 Processor 2.40GHz with 64 bit operating system. Comparison between random and generated sequences alignment has been made. The accelerated run time taken from random sequence alignment over generated sequence average of 50.12% faster than generated sequences.

TABLE OF CONTENT

DECLARATION	i
ACKNOWLEDGEMENT	iii
ABSTRACT	iv
TABLE OF CONTENT	v
LIST OF FIGURES	vii
LIST OF TABLES	vii
LIST OF ABBREVIATION	ix

CHAPTER 1 : INTRODUCTION	1
1.1 Background of Study	1
1.2 Problem Statement	3
1.3 Objectives	5
1.4 Scope of Work	5
1.5 Thesis Organization	6

CHAPTER 2 : LITERATURE REVIEW	7
2.1 Deoxyribonucleic Acid	7
2.2 DNA Sequence Alignment	9
2.3 Smith Waterman Algorithm	10
2.4 History of Smith Waterman Algorithm	14
2.5 MATLAB	15
2.6 Bioinformatics Application	17

CHAPTER 1

INTRODUCTION

1.1 BACKGROUND OF STUDY

As the size of the biomedical database doubled approximately every 18 months [1], the exponentially growth rate led to the increment of the demand for most recent and advance computational sequencing method for biological applications in tend to find the similarities of the genes of unknown function. The encoding of genetic DNA sequences which consist of four fundamental molecular unit which called nucleotides implement to the various vital functions where the whole of the DNA sequence of one living organism is called a genome of that organism and known as the four nitride bases that can be treated as strings: adenine (A), cytosine (C), guanine (G) and thymine (T) [2].

In order to cater current situation where the DNA sequence become important especially in medical applications, by which the useful information can be extracted from