IDENTIFICATION OF RESTRICTION ENDONUCLEASE SITES IN DNA SEQUENCES

Project report submitted by

ROUVIM OLGA

In partial satisfaction of the Award of Diploma in the Computer Studies Course of the Higher Technical Institute, Nicosia, CYPRUS.

Project Supervisor: Mr. C. Makarounas,
BSc in Computer Science,
Lecturer,
Computer Studies Course, H.T.I.

External Assessor: Ms. K. Christodoulou,
BSc in Genetics,
Msc in Applied Molecular Biology and Biotechnology,
Geneticist,
Cyprus Institute of Neurology and Genetics.

Type of Project: INDIVIDUAL : X

GROUP : _____

June - 1991.

INSTITUTE 1757

HIGHER TECHNICAL INSTITUTE

Diploma Project In Computer Studies

1990 - 1991

IDENTIFICATION OF RESTRICTION ENDONUCLEASE SITES IN DNA SEQUENCES

BY

ROUVIM OLGA

SUMMARY

The purpose of this system is to computerize the process of identifying restriction endonuclease sites in DNA sequences. The idea was derived from the need, of the geneticists in the Cyprus Institute of Neurology and Genetics, to keep all the enzyme records, to search for their sites in any DNA sequence and to produce reports, concerning the searching.
The basic objective of this project, was to develop a system that would be user-friendly and efficient, to fulfill the need of having results without any delays.

The first phase of the project, was the INVESTIGATION phase. Processing, carried out manually, was studied, and the feasibility of the project, was decided upon.

The second phase, was the ANALYSIS AND GENERAL DESIGN phase. System Specifications were extracted, along with the specified needs of the Institute. Also, a preliminary outline of the computerized system was designed.

The third phase, was the DETAILED DESIGN AND IMPLEMENTATION phase, during which, the new system was designed with every detailed aspect. Also, the system was fully tested in order

to meet the user performance criteria.

For the last, INSTALLATION phase, the developed system, was installed on the existing hardware.

CONTENTS

CHAPTERS		PAGE
CHAPTER 1:	INTRODUCTION	1.
2.1	INTRODUCTION, PURPOSE, CONTEXT	2.
2.2	INTRODUCTION TO GENETICS ENGINEERING . . .	3.
CHAPTER 2:	INITIAL INVESTICATION	5.
3.1	INTRODUCTION	6.
3.2	PURPOSE OF THE SYSTEM	7.
3.3	GENERAL DESCRIPTION OF THE SYSTEM . . .	8.
CHAPTER 3:	FEASIBILITY STUDY	10.
3.1	INTRODUCTION	11.
3.2	FINANCIAL FEASIBILITY	12.
3.3	OPERATIONAL FEASIBILITY	13.
3.4	TECHNICAL FEASIBILITY	14.
3.5	HUMAN FACTOR FEASIBILITY	15.
CHAPTER 4:	ANALYSIS ANG GENERAL DESIGN . . .	16.
4.1	INTRODUCTION	17.
4.2	INPUTS	18.
4.3	OUTPUTS	24.
4.4	DATA STORES	26.
4.5	PROCESSES	32.
4.6	DATA STRUCTURES	62.
4.7	DATA ELEMENTS	81.
CHAPTER 5:	DETAILED DESIGN, IMPLEMENTATION AND INSTALLATION . .	90.
5.1	DETAILED DESIGN AND CODING	91.
5.2	NEW SYSTEM DESIGN SPECIFICATION	92.
5.3	PROCESSING DESCRIPTION	93.
5.4	DATA FILES	94.
5.5	SECURITY AND CONTROL	94.
5.6	SYSTEM FUNCTION	94.

CHAPTER PAGE

5.7 PRELIMINARY DETAILED DESIGN AND IMPLEMENTATION PLAN . . . . . 95.
5.8 PRELIMINARY SYSTEM PLAN . . . . . 95.
5.9 USER TRAINING OUTLINE . . . . . 96.
5.10 PRELIMINARY INSTALLATION PLAN . . . . . 97.

CHAPTER 6: DETAILED SYSTEM DESCRIPTION . . . 98.
6.1 PROGRAMS DESCRIPTION . . . . . 99.

CHAPTER 7: TESTING . . . . . 106.
7.1 INTRODUCTION . . . . . 107.
7.2 TESTING CONDITIONS . . . . . 108.

CHAPTER 8: SCREENS AND LAYOUTS . . . . . 110.

APPENDICES . . . . . 126.
APPENDIX A : DATA FLOWCHARTS . . . . . 127.
APPENDIX B : DATA DICTIONARY . . . . . 136.
APPENDIX C : SYSTEM FLOWCHARTS . . . . . 140.
APPENDIX D : NORMALIZATION . . . . . 145.
APPENDIX E : DATA STRUCTURES . . . . . 146.
APPENDIX F : DICTIONARY . . . . . 147.