Cairo University

CONSTRUCTING GENE REGULATORY NETWORK USING TIME SERIES GENE EXPRESSION

By

Shaimaa Mohammed Elembaby Hassan Sharaf Eldien

A Thesis Submitted to the
Faculty of Engineering at Cairo University
in Partial Fulfillment of the
Requirements for the Degree of
DOCTOR OF PHILOSOPHY
in
Biomedical Engineering and Systems

FACULTY OF ENGINEERING, CAIRO UNIVERSITY
GIZA, EGYPT
2020

CONSTRUCTING GENE REGULATORY NETWORK USING TIME SERIES GENE EXPRESSION

By
Shaimaa Mohammed Elembaby Hassan Sharaf Eldien

A Thesis Submitted to the
Faculty of Engineering at Cairo University
in Partial Fulfillment of the
Requirements for the Degree of
DOCTOR OF PHILOSOPHY
in
Biomedical Engineering and Systems

Under the Supervision of

Prof. Dr. Manal Abdel Wahed	**Dr. Vidan Fathy Ghoniem**
……………………………….	……………………………….
Professor of Biomedical Engineering and Systems Faculty of Engineering, Cairo University	Assistant Professor Biomedical Engineering Faculty of Engineering, Helwan University

FACULTY OF ENGINEERING, CAIRO UNIVERSITY
GIZA, EGYPT
2020

CONSTRUCTING GENE REGULATORY NETWORK USING TIME SERIES GENE EXPRESSION

By

Shaimaa Mohammed Elembaby Hassan Sharaf Eldien

A Thesis Submitted to the
Faculty of Engineering at Cairo University
in Partial Fulfillment of the
Requirements for the Degree of
DOCTOR OF PHILOSOPHY
in
Biomedical Engineering and Systems

Approved by the
Examining Committee

Prof. Dr. Manal Abdel Wahed Abdel Fatah, Thesis Main Advisor

____

Prof. Dr. Sahar Ali Fawzi, Internal Examiner

____

Prof. Dr. Mai Said Mabrouk, External Examiner
- Head of Biomedical Department, Engineering Faculty, MUST.

FACULTY OF ENGINEERING, CAIRO UNIVERSITY
GIZA, EGYPT
2020

Engineer's Name: Shaimaa Mohammed Elembaby Hassan Sharaf Eldien
Date of Birth: 22/2/1980
Nationality: Egyptian
E-mail: ShaimaaMElembaby@gmail.com
Phone: 01060203704
Address: 12 Rawdat Elmokattam Compound

Registration Date: 1/10 /2011.
Awarding Date: …./…./2020.
Degree: Doctor of Philosophy
Department: Biomedical Engineering and Systems.

Supervisors:

Prof. Dr. Manal Abdel Wahed Abdel Fatah
Dr. Vidan Fathy Ghoneim

Examiners:

Prof. Dr. Mai Said Mabrouk (External examiner)
- Head of Biomedical Department, Engineering Faculty, MUST.

Prof. Dr. Sahar A. Fawzy (Internal examiner)
Prof. Dr. Manal Abdel Wahed (Thesis main advisor)

Title of Thesis:

Constructing Gene Regulatory Network Using Time Series Gene Expression Microarrays

Key Words:
Gene Regulatory Network; Gene Expression;

Summary:

Inferring Gene Regulatory Network (GRN) is important to understand genetic changes occurring in cells. GRNs help in designing drugs and vaccines. As changes occurred in cell are dynamic, time series gene expression is very important to understand how genes regulates each other. There are many competitions to infer GRN, here we used DREAM4 in-silico 10 genes, DREAM 4 insilico-100 genes and DREAM 5. As H1N1 is the most endemic disease recently and still does not have a known medication. To comprehend alterations in genes due to H1N1 infection, we attempted to build GRN based on gene expressions of H1N1 infected cells. We improved two algorithms to infer GRN GENIE 3 and PLSNET by one-way ANOVA and produce two new hybrids Algorithms ANOVAPLSN and ANOVAG3 which record best results to infer large scale GRN from time series gene Expression. we used other methods to infer GRN as sparse linear Equation, Distance matrices and Artificial neural network to infer GRN.

Disclaimer

I hereby declare that this thesis is my own original work and that no part of it has been submitted for a degree qualification at any other university or institute.

I further declare that I have appropriately acknowledged all sources used and have cited them in the references section.

Name: Shaimaa Mohamed elembaby Hassan Sharaf Eldien.

Date: 26/4/2020

Signature: Shaimaa M. Elembaby

Acknowledgments

In this section, I want to provide acknowledgements to my supervisions Prof. Dr. Manal Abdel Wahed and Dr. Vidan Fathi Ghonim who gave me support and encouragement to complete my thesis. I thank also Prof. Dr. Khaled Wahba who taught me Biomedical statistics by an easily way. I want to thank All my teachers of Biomedical engineering and systems department because they do all efforts to teach us and encourage us to search. I Thanks my parents, my husband, my sisters, my brother, my daughters and my son who encourage me to complete my studies.

Table of Contents

list of Tables

List of Figures

Nomenclature

Gene a unit of inheritance is transferred from parents to its Childs and consist of long strand of DNA sequence.

DNA Deoxyribonucleic acid chains of nucleic acids contains genetic instructions used in the development and functioning of all known living organisms.

RNA Ribonucleic Acid formed by DNA and found in the cytoplasm.

Gene Expression a gene inheritable information is gained from synthesis of a functional gene products (functional RNA) as proteins, small nuclear RNA (snRNA) genes or transfer RNA (tRNA) in non-protein coding genes .

Transcription synthesis of RNA under direction of DNA this process is before translation.

Translation mature mRNA is translated into polypeptide chain of amino acids then to protein according genetic codes.

Transcription factor (TF) DNA-binding protein (promoter) to initiate the biological process in the cell.

A gene regulatory network (GRN) directed graph represents a collection of molecular regulators .The regulator can be DNA, RNA, protein and complexes of these that interact with each other and with other substances in the cell. The interaction can be direct or indirect (through transcribed RNA or translated protein).

An adjacency matrix is a square zero diagonal matrix used to represent GRN or any finite graph between some nodes **.**

ACO	**Ant Colony Optimization**
ALO	**Ant Lion Optimizer**
AUPR	**Area Under Precision Recall Curve**
AUROC	**Area Under Receive Operating Characteristic Curve.**
DBN	**Dynamic Bayesian Network**
DE	**Differential Evolution**
EA	**Evolutionary Algorithm**
GA	**Genetic Algorithm**

GWO	**Grey Wolf Optimizer**
GRN	**Gene Regulatory Network**
mRNA	**Messenger RNA**
MI	**Mutual Information**
PDE	**Partial Differential Equation**
PLDE	**Piecewise-Linear Differential Equation**
PSO	**Particle Swarm Optimization**
QDE	**Qualitative Differential Equation**
RNN	**Recurrent Neural Network**
SI	**Swarm Intelligence**
SsDNA	**Single-Stranded DNA**