A NOVEL POPULATION GENETIC VARIATION ASSESSMENT ALGORITHM WITH GRAPHICAL PERSPECTIVE

By

Sagy Elsayed Mohamed Elzalabany

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
2019

A NOVEL POPULATION GENETIC VARIATION ASSESSMENT ALGORITHM WITH GRAPHICAL PERSPECTIVE

By
Sagy Elsayed Mohamed Elzalabany

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

Associate Prof. Dr. Sahar Fawzi
Associate Professor
Department of Biomedical Engineering
and Systems
Faculty of Engineering, Some University

Prof. Dr. Olfat Shaker
Head of Department of Medical
Biochemistry and Molecular Biology
Faculty of Medicine, Cairo University

FACULTY OF ENGINEERING, CAIRO UNIVERSITY
GIZA, EGYPT
2019

A NOVEL POPULATION GENETIC VARIATION ASSESSMENT ALGORITHM WITH GRAPHICAL PERSPECTIVE

By

Sagy Elsayed Mohamed Elzalabany

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

Associate Prof. Dr. Sahar Ali Fawzi	Thesis Main Advisor
Prof. Dr. Olfat Gamil Shaker	Advisor
Prof. Dr. Ayman Mohammed Eldeib	Internal Examiner
Prof. Dr. Samia Abdelrazik Mashaly	External Examiner

FACULTY OF ENGINEERING, CAIRO UNIVERSITY
GIZA, EGYPT
2019

Engineer's Name: Sagy Elsayed Mohamed Elzalabany
Date of Birth: 14/4/1969
Nationality: Egyptian
E-mail: sagy.elzalabany69@eng1.cu.edu.eg
Phone: +201001062826
Address: 172 Jasmin 6 New Cairo Egypt

Registration Date: 1/10/2014
Awarding Date: 11th March 2019
Degree: Doctor of Philosophy
Department: Biomedical Engineering and Systems

Supervisors: Associate Prof. Dr. Sahar Fawzi
Prof. Dr. Olfat Shaker

Examiners:

Associate Prof. Dr. Sahar Ali Fawzi	(Thesis main advisor)
Prof. Dr. Olfat Gamil Shaker	(Advisor)
Prof. Dr. Ayman Mohammed Eldeib	(Internal examiner)
Prof. Dr. Samia Abdelrazik Mashaly	(External examiner)

Title of Thesis:

A Novel Population Genetic Variation Assessment Algorithm With Graphical Perspective

Key Words:

Bioinformatics; Population Genetics; Allele Frequency; F_{ST} Genetic Distance; STR Autosomal Loci;

Summary:

DNA fingerprinting is the main step to assist identification of individuals based on their respective DNA profiles. This research aims to establish a complete DNA fingerprinting map for Egyptian population using 9 Short Tandem Repeats (SRTs) loci D3S1358, VWA, FGA, TH01, TPOX, CSF1PO, D5S818, D13S317 and D7S820. Then, the genetic characteristics of allele frequencies data were studied for 15 autosomal STRs loci (D19S433, FGA, vWA, CSF1PO, TPOX, TH01, D5S818, D7S820, D13S317, D21S11, D8S1179, D18S51, D2S1338, D16S539, and D3S1358) for more than 60 populations and ethnic groups. Moreover, for all the populations under study, a visual and statistical comparative analysis were made to obtain the global gene profile, which is average of populations' alleles frequencies, and populations discriminative genetic features such as, the highest frequencies alleles, the lowest frequencies alleles, and rare alleles. Besides, building an Allele Frequency Quality Control System (AFQCS) which automatically compares, evaluates, and then updates its database with the validated tested input data.

Acknowledgments

"First and foremost praises and thanks to Allah, the Lord, the Most Gracious, and the Most Merciful. You give me the power to be believe in my passion and pursue my dreams. You provided me the opportunity to step in the excellent world of science. To be able to step strong and smooth in this way, I could never have done this without faith I have in you, the Almighty."

I would like to express my sincere gratitude to my advisors **Associate Prof. Dr. Sahar Fawzi** and **Prof. Dr. Olfat Shaker** for their support, patience, motivation, enthusiasm, and immense knowledge. Their guidance helped me in all the time of research and writing of thesis. I could not have imagined having a better advisors and mentors for my PhD study.

I would like to thank **Dr. Tarek Taha**, for being a part of this research and his insightful comments, kind cooperation, and encouragement.

My sincere thanks also goes to **Prof. Dr. Ahmed Elbialy**, **Prof. Dr. Ahmed Hesham** Kandil, and **Prof. Dr. Ayman Eldeib** for their help, support and valuable advice.

Besides my advisors, I would like to thank my family and friends: my parents **Mr. Elsayed Elzalabany**, **Dr. Thoria Mady**, my brother **Eng. Samer Elzalabany**, and my friend **Dr. Hanafy Meleis** for their support, help and continuous encouragements and kindness. They gave me the strength and the reason to carry on, and I can never thank them enough.

I also would like to thank my colleges' **Dr. Ahmed zakaria**, **Eng. Mostafa yosry**, **Eng. Moh. Mohsen**, **Eng. Ayman Anwar**, **Eng. Sherif Sayed**, **Eng. Moh. Elmarakby**, **Eng. Amr Kamal Eldin**, and **Eng. Kamal Mohamed** for their help and support.

I owe my deepest gratitude and endless thanks to my wife **Eng. Soha Waly** and my daughters who have sacrificed beyond measure to reach this day.

Lastly, I offer my regards and blessings to all those who supported me in any aspect during the completion of this work.

Dedication

I dedicate this work to whom my heart felt thanks;

My great beloved Dad, my kind beloved mom,

my lovely wife Soha Waly and my daughters Samar, Sama, Sohaila,

and my dear brother

Samer Elzalabany

*I thank **Allah** for choosing you to be my family.*

Thank you for supporting me with kindness, patience and love

Yours,

Sagy Elzalabany

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AFQC	Allele Frequency Quality Control
AL	Allelic Ladder
bp	base pair
CCMs	Correlation Coefficient Matrices
CMP	Combined Match Probability
CODIS	Combined DNA Index System
DNA	Deoxyribonucleic acid
FS	Full Sibling
GGP	Global Gene Profile
HWE	Hardy-Weinberg equilibrium
HWP	Hardy-Weinberg Principle
IQR	Inter Quartile Range
LD	Linkage Disequilibrium
MDS	Multidimensional Scaling
NMDS	Non-metric Multidimensional Scaling
PCR	Polymerase Chain Reaction
PIC	Polymorphism Information Content
SNP	Single-Nucleotide Polymorphisms
SSRs	Simple Sequence Repeats
STR	Short Tandem Repeats
UI	Unrelated Individuals

Publications

1. Tarek Taha, Sagy Elzalabany, Olfat Shaker, Sahar Fawzi, "Genetic Variation for Nine Autosomal Short Tandem Repeat (STRs) in Egyptian Population" Journal of Forensic Medicine, 2018.

2. Sagy Elzalabany, Tarek Taha, Sahar Fawzi, Olfat Shaker, "A novel allele frequency trajectories template to discriminate genetic similarity among populations" Meta Gene, 2019.