UNIVERSITI TEKNOLOGI MARA

IDENTIFICATION OF NOVEL GENETIC VARIANTS AND *IN-SILICO* PATHWAY ANALYSIS IN PATIENTS WITH DILATED CARDIOMYOPATHY

RAFEZAH BINTI RAZALI

PhD

May 2019

AUTHOR'S DECLARATION

I declare that the work in this thesis was carried out in accordance with the regulations of Universiti Teknologi MARA. It is original and is the results of my own work, unless otherwise indicated or acknowledged as referenced work. This thesis has not been submitted to any other academic institution or non-academic institution for any degree or qualification.

I, hereby, acknowledge that I have been supplied with the Academic Rules and Regulations for Post Graduate, Universiti Teknologi MARA, regulating the conduct of my study and research.

Name of Student	:	Rafezah Binti Razali
Student I.D. No.	:	2011426208
Programme	:	Doctor of Philosophy – MD990
Faculty	:	Medicine
Thesis Title	:	*Identification of Novel Genetic Variants and In-Silico Pathway Analysis in Patients With Dilated Cardiomyopathy*
Signature of Student	:	……..
Date	:	May 2019

ABSTRACT

Dilated cardiomyopathy (DCM), a genetically heterogeneous heart muscle disease is the third leading cause of heart failure. DCM usually leads to severe congestive heart failure and an increased risk of sudden cardiac death. The limitations of heart donor and transplant complications necessitate the development of new treatment strategies targeting the molecular mechanisms involving the cardiomyocytes. Identification of genetic variants and its molecular pathways may be useful in providing mechanistic insights in DCM and its progression to heart failure. More need be known about genetics of DCM especially among patients in South East Asia including Malaysia. This study attempted to identify the underlying genetic variants that may lead to DCM by utilizing the whole exome sequencing. The study also explores the presence of the identified genetic variants among the family members of DCM patients and to investigate the role of the identified genes involved in the pathogenesis of DCM among Malaysian patients using in-silico analyses. 1982 subjects were screened from various community health screening programs (1978 males; 33 females); and 29 from hospital admissions (26 males; 3 females) were recruited. One subject from the community screening was diagnosed as DCM, thus giving the estimated prevalence of DCM in the current study as 0.5%. This prevalence finding serves as a first report of DCM prevalence in Malaysia. 111 first degree family members of the DCM patients were also recruited for Sanger sequencing to determine if the genetic variants found in DCM patients were also present among the family members. The genetic variants from exome sequencing data were compared with publicly available reference databases. The variants were subsequently annotated using the annotation tools in order to predict whether an amino acid substitution effects were damaging (the variants affect protein function and structure). 223 novel damaging variants and 114 non-synonymous *TTN* variants were identified in the DCM patients. 62 non-synonymous *TTN* variants (54.4%) identified in 83% of DCM patients have minor allele frequency of less than 1% and thus were classified as *TTN* rare variants. The novel damaging variants were not found among the family members of DCM patients suggesting the lack of familial involvement in these DCM patients. The genes underlying the novel damaging variants when performed David Pathway analysis were significantly enriched in the Krueppel-associated box (KRAB) domain binding (Benjamini corrected p=0.00004), myosin complex (Benjamini corrected p=0.004), calcium-mediated signaling (Benjamini corrected p=0.005) and cytoskeleton (Benjamini corrected p=0.008) pathways. In the present study, novel candidate genes for DCM have been identified, including Myosin light chain 2 (*MYL2*), Myosin light chain 7 (*MYL7*), Kruppel-associated box domain (KRAB) Zinc-fingers gene family, zinc finger MYM-type containing 4 (*ZMYM4*), Interleukin 23 receptor (*IL23R*), F-box protein 22 (*FBX022*), phospolipase C gamma 1 (*PLCG1*), Ryanodine receptor 1 (*RYR1*), Integrin beta 7 (*ITGB7*) and GNAS complex locus (*GNAS*) genes that involve in the sarcomere assembly in providing structural integrity of cardiomyocytes, autoimmune and inflammatory response, calcium signalling and regulation of actin cytoskeletal pathways. In addition, there are 21 recurrent variants that were observed to have unknown effect in DCM pathogenesis, hence classified as variants of uncertain significance (VUS). These imply that DCM is genetically heterogeneous disease that involves the activation of multiple signalling pathways. This study serves as first report on the genetic of DCM in Malaysia and South East Asia countries.

ACKNOWLEDGEMENT

In the name of Allah, the Most Gracious, the Most Merciful.

Alhamdulillah, great thanks to Allah and His blessing for giving me the strength and health to complete this study.

This thesis would not have been possible without the guidance and assistance of several individuals who in one way or another contributed and extended their valuable assistance in preparation and completion of this study.

I would like to express my sincere gratitude to my supervisors Emeritus Professor Dato' Dr Khalid Yusoff, Professor Dr Hoh Boon Peng and Associate Professor Dr Siti Hamimah Syeikh Abdul Kadir for their guidance, support and encouragement throughout this study and thesis writing.

I am thankful to my beloved parents (Al-Fatihah), husband and daughter for their prayers and endless moral support.

My appreciation goes to Professor Dr Teo Yik Ying and his team members at Saw Swee Hock School of Public Health, National University of Singapore, Dr Chan Kok Gan and his team members at High Impact Research Central Facilities, University of Malaya, Puan Norlaila Danuri and Cik Fashieha Basir at Non Invasive Cardiac Lab (NICL) Faculty of Medicine UiTM, Puan Fadhlina Abd Majid and Cik Najmin Abu Bakar at Centre for Translational Research and Epidemiology (CenTRE), Faculty of Medicine, UiTM, En Abu Talhah Abdul Aziz at Institute of Medical Molecular Biothechnology (IMMB), Cik Sarina Ali and Puan Wan Juliana in Chemical Pathology Unit of Department of Pathology and all the staff and postgraduate students in Institute Molecular Medicine and Biotechnology (IMMB), Chemical Pathology Unit of Pathology Department and Centre for Translational Research and Epidemiology (CenTRE) with whom I have privilege of working with.

Finally, I would like to convey my regards and gratitude to the patients and family members recruited in this study.

TABLE OF CONTENTS