Faculty of Resource Science and Technology

Computational Analysis of Epstein-Barr Virus *Bam*H1 A Rightward Transcript (BART) MicroRNA (miRNA) Regulation on Messenger and Long Non-Coding RNAs in Nasopharyngeal Cancer

Daphne Olivia Jawai Anak Sadai

Master of Science
2024

Computational Analysis of Epstein-Barr Virus *Bam*H1 A Rightward Transcript (BART) MicroRNA (miRNA) Regulation on Messenger and Long Non-Coding RNAs in Nasopharyngeal Cancer

Daphne Olivia Jawai Anak Sadai

A thesis submitted

In fulfillment of the requirements for the degree of Master of Science

(Genetics)

Faculty of Resource Science and Technology
UNIVERSITI MALAYSIA SARAWAK
2024

DECLARATION

I declare that the work in this thesis was carried out in accordance with the regulations of Universiti Malaysia Sarawak. Except where due acknowledgements have been made, the work is that of the author alone. The thesis has not been accepted for any degree and is not concurrently submitted in candidature of any other degree.

……

Signature

Name: Daphne Olivia Jawai Anak Sadai

Matric No.: 20020007

Faculty of Resource Science and Technology

Universiti Malaysia Sarawak

Date : 19/3/2024

ACKNOWLEDGEMENT

First and foremost, I would like to express my gratitude to my supervisor, Prof. Dr. Edmund Sim Ui Hang, for his invaluable supervision and guidance in completing this study. Besides, I appreciate the knowledge he shared with me and his assistance in writing my thesis throughout this study. Without his help, this study would not have been a success and completed.

Next, special thanks to my fellow labmates and friends, Ms. Adrianne Vong, Ms. Alyaa Syafiqah, and Mr. Alvin Wee, for sharing the knowledge and have consistently and patiently helped and guided me throughout my study. It was soothing to have them to confide in when I felt like giving up and I was also grateful for the emotional support they gave me during the ups and downs in completing this study. I am thankful for the times we have spent together and the safety they offered throughout this journey.

Furthermore, I would like to thank the Faculty of Resource Science and Technology of UNIMAS for providing access to the facilities, particularly the Immunology Laboratory and Animal and Tissue Culture Laboratory, which enabled me to learn and gain new knowledge by giving me the opportunity to conduct laboratory-based experiments.

Lastly, I would like to thank my parents, Mrs. Pepiet and Mr. Sadai, for their endless support and monetary. My mother, especially, for her endless emotional support, understanding, and words of encouragement, which have been fuelling me to complete this study when I feel hopeless. Thank you so much for allowing this journey to happen.

ABSTRACT

The interaction and regulation amongst messenger RNA (mRNA), microRNA (miRNA) and long non-coding RNA (lncRNA) have been identified as causative of nasopharyngeal cancer (NPC) when their expression levels are dysregulated by each other and affecting the biological and molecular functions of affected nasopharynx cells. This interaction is known as competing endogenous RNA (ceRNA) regulatory activity. However, the study of regulation mechanisms of *Bam*HI A rightward transcript (BART) miRNAs encoded by Epstein-Barr virus (EBV) on ceRNA in the pathogenesis of NPC is unavailable. According to previous studies, the Epstein-Barr virus was deemed one of NPC causative factors through its regulation of ceRNA activity using BART miRNAs. Therefore, in this study, candidate BART miRNAs, mRNAs, miRNAs and lncRNAs were obtained by integrating multi-level RNAs expressions data through bioinformatics analysis by constructing the cross-regulatory network. The cross-regulatory network was the combination of mRNA-miRNA-lncRNA ceRNA network with EBV miRNAs-mRNAs network. The results revealed six EBV miRNAs (ebv-miR-BART21-3p, ebv-miR-BART19-3p, ebv-miR-BART15, ebv-miR-BART2-5p, ebv-miR-BART20-3p and ebv-miR-BART11-5p) were interacting with four mRNAs (EYA4, EYA1, EBF1 and MACROD2). In addition, these mRNAs were interacting with six miRNAs (hsa-miR-1246, hsa-miR-93-5p, hsa-miR-16-5p, hsa-miR-135b-5p, hsa-miR-211-5p and hsa-miR-1305) where these miRNAs were interacting with three lncRNAs which are CASC2, TPTE2P1 and ARHGEF26-AS1. The data attained showed that these BART miRNAs may deregulate these mRNAs and affect the expressions of both miRNAs and lncRNAs. This gives insight into understanding the regulatory mechanisms of BART miRNAs on ceRNA activity. This study also revealed the potential functions of BART miRNAs and lncRNAs acting similarly as either tumour

suppressors or oncogenes. Furthermore, the data obtained suggested that BART miRNAs affect DNA repair regulation and apoptosis in NPC by down-regulating the mRNAs, which pertains to poor overall survival (OS) in NPC patients. Overall, these findings exhibited the potential role of these BART miRNAs, mRNAs and lncRNAs as diagnostic and prognostic biomarkers in EBV-induced NPC carcinogenesis based on their ability to target and regulate the expressions of each other.

Keywords: Nasopharyngeal Cancer, EBV, Bioinformatics, BART miRNAs, Long Non-Coding RNA

Analisis Komputasi ke atas Pengawalan MikroRNA (miRNA) Transkrip BamHI A Rightward (BART) Diekrespikan Oleh Virus Epstein-Barr Ke atas RNA Bukan Pengekodan yang Panjang Dalam Kanser Nasofarink

ABSTRAK

Interaksi dan pengawalan di antara pengutus RNA (mRNA), mikroRNA (miRNA) dan RNA bukan pengekodan yang panjang (lncRNA) telah dibuktikan sebagai penyebab dalam pembentukan kanser nasofarink (NPC) apabila tahap ekpresi mereka disregulasi oleh sama sendiri dan hal ini didapati memberi kesan terhadap fungsi-fungsi biologikal dan molekular di dalam sel-sel nasofarink yang terkesan. Interaksi ini dikenali sebagai aktiviti pengawalan persaingan asid ribonukleik yang berasal dari dalam (ceRNA). Walau bagaimanapun, kajian terhadap mekanisma pengawalan oleh miRNA transkrip BamHI A rightward (BART) yang diekspreskan oleh virus Epstein-Barr (EBV) keatas ceRNA didalam pembentukkan NPC belum pernah dikaji. Menurut kajian-kajian terdahulu, Epstein-Barr virus merupakan salah satu faktor penyebab kepada NPC melalui pengawalan keatas aktiviti ceRNA dengan menggunakan miRNA BART. Oleh itu, kajian ini calon miRNA BART, mRNA, miRNA dan lncRNA mampu didedahkan dengan mengabungkan data ekpresi berbilang peringkat asid ribonukleik (RNA) melalui analisa bioinformatik dengan membina rangkaian silang kawal selia. Rangkaian silang kawal selia ini merupakan pengabungan antara rangkaian mRNA-miRNA-lncRNA ceRNA dengan rangkaian EBV miRNA-mRNA. Hasil kajian ini menunjukkan bahawa enam miRNA EBV (ebv-miR-BART21-3p, ebv-miR-BART19-3p, ebv-miR-BART15, ebv-miR-BART2-5p, ebv-miR-BART20-3p dan ebv-miR-BART11-5p) berinteraksi dengan empat mRNA (EYA4, EYA1, EBF1 dan MACROD2). Sehubungan dengan itu, mRNA tersebut juga berinteraksi dengan enam miRNA (hsa-miR-1246, hsa-miR-93-5p, hsa-miR-16-5p, hsa-miR-135b-5p, hsa-miR-211-5p dan hsa-miR-1305) dimana miRNA tersebut juga

berinteraksi dengan tiga lncRNA iaitu CASC2, TPTE2P1 dan ARHGEF26-AS1. Data yang diperolehi ini menunjukkan bahawa miRNA BART mendisregulasi mRNA tersebut dan memberi kesan terhadap ekspresi-ekpresi miRNA dan lncRNA. Ini memberi pengertian terhadap kefahaman tentang mekanisma pengawalan miRNA BART ke atas aktiviti pengawalan ceRNA. Kajian ini juga telah mendedahkan potensi miRNA BART dan lncRNA berfungsi menyerupai pembantut tumor dan onkogene. Tambahan pula, data yang diperolehi mencadangkan bahawa miRNA BART memberi kesan terhadap pengawalan pemulihan asid deoksiribonukleik (DNA) dan apoptosis di dalam NPC dengan menurunkan ekspresi mRNA dan hal ini dikaitkan dengan kelangsungan hidup secara keseluruhan yang buruk di dalam pesakit-pesakit NPC. Secara keseluruhannya, penemuan-penemuan ini mempamerkan potensi miRNA BART, mRNA, miRNA dan lncRNA sebagai biomarker diagnostik dan prognostik dalam NPC karsinogensis yang disebabkan oleh EBV berdasarkan keupayaan mereka dalam mensasarkan dan mengawal ekspresi satu sama lain.

Kata kunci: *Kanser nasofarinks, virus Epstein-Barr, mikroRNA BART, RNA bukan pengekodan yang panjang*

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

LIST OF ABBREVIATIONS

2D	Two-dimensional
3D	Three-dimensional
αvβ	Alpha v beta
ADP-ribosylation	Adenine dinucleotide phosphate-ribosylation
AGO2	Human Argonaute 2
AKR1B1	Aldo-keto reductase family 1 member B1
ARHGEF26-AS1	Rho guanine Nucleotide exchange factor 26 antisense RNA 1
ASOs	Antisense oligonucleotides
ATM	Ataxia-telangiectasia-mutated
BART	*Bam*HI fragment A rightward
BC	Breast cancer
BHRF1	*Bam*HI fragment H rightward open reading frame 1
BL	Burkitt lymphoma
BP	Biological Processes
CASC2	Cancer susceptibility candidate 2
CC	Cellular Component
ceRNA	Competing endogenous RNA
ceRNET	ceRNA network
CCDC113	Coiled-Coil Domain Containing 113
CCRT	Chemoradiotherapy
circRNA	circular RNA
CLASH-seq	Cross-linking, ligation, and sequencing of hybrids

CLASH-seq	Cross-linking, ligation, and sequencing of hybrids
CLEAR-CLIP	Covalent ligation and endogenous Argonaute-bound RNA
CLIP-seq	Cross-linking and immunoprecipitation and high-throughput sequencing
CMTM3	CKLF-like MARVEL transmembrane domain containing 3
CT	Computed tomography
CR2	Complement receptor type 2
CRC	Colorectal cancer
CTL	Cytotoxic T-lymphocytes
cvhRNAs	Competitive viral and host RNAs
DAVID	Database of Annotation, Visualization and Discovery software
DBD	DNA-binding domain
DEG	Differentially expressed gene
DEL	Differentially expressed lncRNA
DEmiRNA	Differentially expressed miRNA
DDR	DNA damage response
DICE1	Deleted in cancer 1
DLBCL	Diffuse large B-cell lymphoma
DNA	Deoxyribonucleic acid
DNAH	Dynein axonemal heavy chain
DNAI	Dynein Axonemal Intermediate Chain
DNALI1	Dynein light intermediate polypeptide 1
DSB	Double-stranded break
DYNLRB2	Dynein light chain roadblock-type 2
DYNC2H1	Cytoplasmic Dynein 2 Heavy Chain

E2F3	E2F transcription factor 3
EA	Early antigen
EAd/BMRF1	Early antigen diffuse
EBNA1	Epstein-Barr nuclear antigen-1
EBV	Epstein-Barr virus
EBVaGC	EBV-associated gastric cancer
EGA	European Genome-phenome Archive
ENKUR	Enkurin
EBERs	Epstein-Barr small encoded RNAs
EBF1	Early B-cell factor transcription factor 1
EGFR	Epidermal growth factor receptor
ESCC	Esophageal squamous cell carcinoma
EYA	Coactivator and phosphatase
FANK1	Fibronectin type 3 and ankyrin repeat domains 1
GC	Gastric cancer
GENCODE	Genome ENCyclopedia Of DNA Elements
GEO	Gene Expression Omnibus
GEO2R	Gene Expression Omnibus to R language
gB	Glycoprotein B
gH	Glycoprotein H
gL	Glycoprotein L
gp42	Glycoprotein 42
GC	Gastric cancer
GO	Gene Ontology
GWAS	Genome-wide association studies