Faculty of Medicine & Health Sciences

The Role of Epstein-Barr virus *LMP1* Gene Expression in Chromosome Rearrangements in Nasopharyngeal Carcinoma

Mohd Aminudin Mustapha

Doctor of Philosophy
2025

The Role of Epstein-Barr virus LMP1 Gene Expression in Chromosome Rearrangements in Nasopharyngeal Carcinoma

Mohd Aminudin Mustapha

A thesis submitted

In fulfillment of the requirements for the degree of Doctor of Philosophy

(Medical Sciences)

Faculty of Medicine & Health Sciences
UNIVERSITI MALAYSIA SARAWAK
2025

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: Mohd Aminudin Bin Mustapha

Matric No.: 16010173

Faculty of Medicine & Health Sciences

Universiti Malaysia Sarawak

Date :

ACKNOWLEDGEMENT

This thesis arose in part, out of 8 years of research that has been done since I joined Para-clinical Sciences, Faculty of Medicine and Health Sciences, Universiti Malaysia Sarawak (UNIMAS). During this period, I have worked with a great number of people who contributed in assorted ways to the research and the making of this thesis, deserve special mention. It is a pleasure to convey my thanks and gratitude to them all in my humble acknowledgment.

Firstly, praise is to Allah S.W.T for giving me strength, good health, motivation, and patience in completing my PhD's research.

I would like to record my gratitude to my beloved supervisor, Prof. Dr. Sim Sai Peng @ Sim Sai Ping @ Samirah Abdullah, Para-clinical Sciences, Faculty of Medicine and Health Sciences, Universiti Malaysia Sarawak (UNIMAS), for her supervision, advice, and guidance from the very early stage of this research as well as giving me extraordinary experiences throughout the work. Above all and the most needed, she provided me unflinching encouragement and support in various ways. Her truly scientific intuition has made her as a constant oasis of ideas and passions in science, which exceptionally inspired and enriched my growth as a student, a researcher, and a scientist I want to be. I am indebted to her more than she knows.

We would like to thank Universiti Malaysia Sarawak (UNIMAS) for the infrastructure provided. We would like to express our gratitude to Dr. Eng-Lai Tan and Prof. Dr. Choon-Kook Sam for the EBV *LMP1* recombinant plasmid. We are deeply grateful to to Prof. Tsao Sai Wah (The University of Hong Kong, Hong Kong, China) and Prof. Lo Kwok Wai (The

Chinese University of Hong Kong, Hong Kong, China) for providing the normal nasopharyngeal epithelial cell line (NP69). We sincerely thank the Ministry of Higher Education, Malaysia for providing the Fundamental Research Grant Scheme (FRGS) [grant number: FRGS/1/2016/SKK08/UNIMAS/02/1] which supported this work.

Also, I would like to thank all the staffs and students in the Faculty of Medicine and Health Sciences, Universiti Malaysia Sarawak (UNIMAS) especially lecturers; Assoc. Prof. Dr. William Lim Kiong Seng and Dr. Isabel Fong Lim and science officers; Mr Safarudin and Mr Dunstan. Many thanks to all the students; Tan Siew Huang, Dr Grace, Nor Ajua and other staffs for their constant support.

Words fail me to express my love and appreciation to my father, Mr. Mustapha Bin Ab Kadir, my mother, Ms. Che Jah Binti Che Wil, my wife, Ms. Ernie Suhaiza, my daughters Aulia Madeena, Ayra Marissa, Amina Eliya and my son, Ahmad Elyas for being supportive and encouraging me throughout my study period.

Finally, I would like to thank everybody who has contributed and has been important to the successful realisation of this thesis, as well as express my apology that I could not mention personally one by one.

ABSTRACT

Nasopharyngeal carcinoma (NPC) is strongly associated with Epstein-Barr virus (EBV) infection, particularly the expression of the viral oncogene Latent Membrane Protein 1 (*LMP1*). *LMP1* is known to induce oxidative stress, apoptosis, and genomic instability, which may contribute to chromosome rearrangements—a hallmark of cancer progression. However, the precise mechanisms by which *LMP1* expression leads to chromosomal breaks and rearrangements, particularly within the *AF9* gene, remain unclear. This study aimed to establish a transient *LMP1* gene expression system and investigate its effects on oxidative stress, apoptosis, and chromosome breaks, particularly within the *AF9* gene, while also evaluating the role of MAR/SAR sequences in mediating these breaks. *LMP1* gene constructs in pTracer and pcDNA forms were transfected into a normal nasopharyngeal cell line to assess ROS levels and *LMP1*-induced apoptosis through Caspase activity measurement. Inverse PCR (IPCR) was utilized to detect chromosomal breakage in scaffold-attachment region (SAR) and non-SAR regions of the *AF9* gene in *LMP1*-expressing cells. The results showed no significant differences in ROS production between pTracer-*LMP1*, pcDNA-*LMP1*, and their combination compared to the vector control. However, pcDNA-LMP1 exhibited significantly higher Caspase activity than the pcDNA vector control. Furthermore, *LMP1* expression showed significant differences in chromosome breaks within the SAR compared to the control. Statistical analyses revealed a significant association between LMP1 transfections and increased chromosome breaks in the SAR, suggesting a higher risk of chromosomal instability. Conversely, *LMP1* transfections did not significantly affect chromosome breaks in the non-SAR, indicating a region-specific effect. These findings suggest that LMP1 expression may trigger oxidative stress-mediated apoptosis,

which contributes to chromosome rearrangements and plays a role in the pathogenesis of nasopharyngeal carcinoma (NPC).

Keywords: Nasopharyngeal carcinoma, Epstein–Barr virus, latent membrane protein 1, reactive oxygen species, caspase activity scaffold-attachment region,

Peranan Ekspresi Gen LMP1 virus Epstein-Barr dalam Penyusunan Semula Kromosom dalam Karsinoma Nasofaring

ABSTRAK

Karsinoma nasofaring (NPC) mempunyai kaitan rapat dengan jangkitan virus Epstein-Barr (EBV), terutamanya melalui ekspresi onkogen virus Latent Membrane Protein 1 (LMP1). LMP1 diketahui boleh mencetuskan tekanan oksidatif, apoptosis, dan ketidakstabilan genom, yang berkemungkinan menyumbang kepada penyusunan semula kromosom—ciri utama dalam perkembangan kanser. Namun begitu, mekanisme tepat bagaimana ekspresi LMP1 menyebabkan kerosakan dan penyusunan semula kromosom, khususnya dalam gen AF9, masih belum difahami sepenuhnya. Kajian ini bertujuan untuk membangunkan sistem ekspresi gen LMP1 secara sementara dan menyelidik kesannya terhadap tekanan oksidatif, apoptosis, dan kerosakan kromosom, khususnya dalam gen AF9, serta menilai peranan jujukan MAR/SAR dalam pengantaraan kerosakan ini. Konstruk gen LMP1 dalam bentuk pTracer dan pcDNA telah ditransfeksi ke dalam garis sel nasofaring normal untuk menilai tahap ROS dan apoptosis yang disebabkan oleh LMP1 melalui pengukuran aktiviti Caspase. PCR songsang (IPCR) telah digunakan untuk mengesan kerosakan kromosom di kawasan scaffold-attachment region (SAR) dan kawasan bukan-SAR dalam gen AF9 pada sel yang mengekspresikan LMP1. Keputusan menunjukkan tiada perbezaan ketara dalam penghasilan ROS antara pTracer-LMP1, pcDNA-LMP1 dan gabungannya berbanding kawalan vektor. Walau bagaimanapun, pcDNA-LMP1 menunjukkan aktiviti Caspase yang jauh lebih tinggi berbanding kawalan vektor pcDNA. Tambahan pula, ekspresi LMP1 menunjukkan perbezaan ketara dalam kerosakan kromosom di kawasan SAR berbanding kawalan. Analisis statistik menunjukkan terdapat hubungan signifikan antara transfeksi LMP1 dan peningkatan kerosakan kromosom dalam kawasan SAR, mencadangkan risiko

yang lebih tinggi terhadap ketidakstabilan kromosom. Sebaliknya, transfeksi LMP1 tidak menunjukkan kesan yang signifikan terhadap kerosakan kromosom dalam kawasan bukan-SAR, menunjukkan kesan yang khusus kepada kawasan tertentu.

Kata kunci: *Kanser nasofaring, virus Epstein-Barr, Gen protein membran laten 1, spesies oksigen reaktif, aktiviti caspase*

TABLE OF CONTENTS

	Page
DECLARATION	i
ACKNOWLEDGEMENT	ii
ABSTRACT	iv
ABSTRAK	vi
TABLE OF CONTENTS	viii
LIST OF TABLES	xiv
LIST OF FIGURES	xvi
LIST OF ABBREVIATIONS	xvii
CHAPTER 1: INTRODUCTION & LITERATURE REVIEW	1
1.1 Nasopharyngeal Carcinoma (NPC)	1
1.1.1 Incidence of Nasopharyngeal Carcinoma (NPC)	2
1.1.2 Risk Factors of Nasopharyngeal Carcinoma (NPC)	5
1.1.3 Genetic Alteration in Nasopharyngeal Carcinoma (NPC)	8
1.2 Epstein-Barr Virus (EBV) in Nasopharyngeal Carcinoma (NPC)	10
1.2.1 Epstein-Barr Virus (EBV) Reactivation in Nasopharyngeal Carcinoma (NPC)	13
1.3 *LMP1* Gene in Nasopharyngeal Carcinoma (NPC)	15
1.3.1 Implication of *LMP1* in Nasopharyngeal Carcinoma (NPC) Development	17
1.4 Oxidative Stress in Nasopharyngeal Carcinoma (NPC)	18

1.5 Association of Apoptosis with Chromosomal Rearrangement in Nasopharyngeal Carcinoma (NPC) 20

1.5.1 Apoptosis and Its Hallmark 21

1.6 The Implication of Apoptosis and Chromosomal Rearrangement in Nasopharyngeal Carcinoma (NPC) 23

1.7 Association of Matrix Attachment Region/ Scaffold Attachment Region (MAR/SAR) Sequences with Chromosomal Rearrangements 23

1.8 The *AF9* Gene 27

1.9 Rationale of Study 27

1.10 Hypothesis 28

1.11 Objectives 28

CHAPTER 2: MATERIALS AND METHODS 29

2.1 Plasmid and Bacteria Culture 29

2.2 Extraction of Plasmid 30

2.3 Plasmid DNA 31

2.4 Cell Line 33

2.5 Transient Transfection 34

2.6 Protein Extraction for Western Blotting 36

2.6.1 Cell Lysis 36

2.6.2 Protein Quantification by Bradford Assay 36

2.7 Western Blot 37

2.7.1 Preparation of Sodium Dodecyl-Sulfate Polyacrylamide Gel Electrophoresis (SDS-PAGE) 37

2.7.2 Sodium Dodecyl-Sulfate Polyacrylamide Gel Electrophoresis (SDS-PAGE) 38

2.7.3 Transferring of Protein onto Polyvinylidene Fluoride (PVDF) Membrane 39

2.8 Transfection for Oxidative Stress and Caspase Activity Measurement 41

2.8.1 Measurement of Oxidative Stress in *LMP1*-Expression Cells 41

2.8.2 Measurement of Caspases Activity in *LMP1*-Expression Cells 43

2.9 Identification of Chromosomal Breaks within the *AF9* Gene using Inverse PCR (IPCR). 44

2.9.1 Genomic DNA Extraction 44

2.9.2 First Restriction Enzyme (RE) Digestion 46

2.9.3 Klenow Fill-In and Cyclisation 46

2.9.4 Ethanol Precipitation 47

2.9.5 Second and Third Restriction Enzyme (RE) Digestion for SAR 47

2.9.6 Removal of Excessive Unincorporated Nucleotide 48

2.9.7 Nested IPCR of the *AF9* SAR 49

2.9.8 Nested IPCR of the *AF9* non-SAR 51

2.9.9 Identify Chromosomal Breaks 53

2.9.10 Agarose Gel Electrophoresis and Gel Extraction 53

2.9.11 DNA Sequencing 54

2.10 Statistical Analysis 55

CHAPTER 3: **RESULTS** 56

3.1 Transient Transfection of NP69 Cells: Observation by Fluorescence Microscopy 56

3.2 Detection of *LMP1* Gene Expression by Western Blotting 58

3.3 Measurement of Oxidative Stress in *LMP1*-Expression Cells. 60

3.3.1 Association of Oxidative Stress with Expression of *LMP1*. 61

3.3.2 Association of Oxidative Stress with Green Fluorescence Protein (GFP) Expression. 63

3.4 Measurement of Caspases Activity in *LMP1*-expressing Cells 64

3.4.1 Association of Caspase Activity with *LMP1* Expression 64

3.4.2 Association of Caspase Activity with Green Fluorescence Protein (GFP) Expression 67

3.5 Association of Combination of Oxidative Stress and Caspase Activity with *LMP1* Expression 68

3.6 Risk Association of Oxidative Stress and Caspase Activity with *LMP1* Expression 70

3.6.1 Risk Association of pTracer, pTracer-*LMP1*, pcDNA and pcDNA-*LMP1* Transfections with Oxidative Stress. 71

3.6.2 Risk Association of pTracer, pTracer-*LMP1*, pcDNA and pcDNA-*LMP1* Transfection with Caspase Activity 72

3.6.3 Risk Association of Combination of Oxidative Stress and Caspase Activity with *LMP1* Expression 73

3.7 Identification of *LMP1* Expression-Induced Chromosomal Breaks in Scaffold/Matrix Attachment Region (SAR) and non-SAR of *AF9* Gene. 74

3.7.1 Association of Chromosome Breaks with *LMP1* Expression in the SAR and Non-SAR of *AF9* Gene. 76

3.7.2 Risk association of Number of Chromosomal Breaks in the SAR and the Non-SAR of the *AF9* Gene with *LMP1* Expression. 80

3.8 Sequencing of IPCR Products 83

3.9 Chapter Summary 86

CHAPTER 4: **DISCUSSION** 89

4.1 Model of Mechanism of EBV-Induced Carcinogenesis via *LMP1* Expression and DNA Damage. 89

4.2 Association of *LMP1* Expression with Nasopharyngeal Carcinoma (NPC) 93

4.3 Association of Oxidative Stress with *LMP1* Expression 94

4.4 Association of Caspases Activity with *LMP1* Expression 97

4.5 Association of Combination of Oxidative Stress and Caspases Activity with *LMP1* Expression 100

4.6 Association of Chromosomal Breaks within the *AF9* Gene with *LMP1* Expression. 101

4.6.1 *LMP1* Expression-Induced Chromosomal Breaks Within the *AF9* Gene Scaffold Attachment Region (SAR) and Non-SAR. 102

4.7 Limitations 106

CHAPTER 5: CONCLUSION 107

REFERENCES 108

APPENDICES 125

LIST OF TABLES

		Page
Table 2.1:	Keratinocyte-SFM (KSFM) Medium used in this study	34
Table 2.2:	Seeding amounts of NP69 cells in different culture vessel to achieve a confluency of 70-85%.	35
Table 2.3:	Preparation of 10% SDS-PAGE Mixture.	38
Table 2.4:	Nucleotide Sequences and Nucleotide Positions of the Primers Used in the First Round and Second Round of IPCR for the SAR of the *AF9* Gene.	50
Table 2.5:	PCR Programme Employed in the First Round of Nested IPCR for the SAR of the *AF9* Gene.	50
Table 2.6:	PCR Programme Used in the Second Round of Nested IPCR for the SAR of the *AF9* Gene.	51
Table 2.7:	Nucleotide Sequence and Nucleotide Position of the Primers Used in the Nested PCR for Non-SAR of the *AF9* Gene	52
Table 2.8:	PCR programme used in the first round of nested IPCR for the non-SAR of the *AF9* gene.	52
Table 2.9:	PCR Programme used in the Second Round of Nested IPCR for the Non-SAR of the *AF9* Gene.	53
Table 3.1:	Mean and Standard Deviation of Oxidative Stress in Transfected NP69.	62
Table 3.2:	Association of Oxidative Stress with *LMP1* Expression.	63
Table 3.3:	Mean and Standard Deviation of Oxidative Stress in Vector Control Transfected NP69	64
Table 3.4:	Association of Oxidative stress with Vector Control Transfected Cells	64
Table 3.5:	Mean and Standard Deviation of Caspace Activity in Transfected NP69	65
Table 3.6:	Association of Caspase Activity with *LMP1* Expression	66
Table 3.7:	Mean and Standard Deviation of Caspace Activity in Vector Control Transfected NP69	67

Table 3.8: Association of Caspase Activity with Vector Control Transfected Cells 67

Table 3.9: Mean and Standard Deviation of Combination of Oxidative Stress and Caspace Activity in Transfected NP69 69

Table 3.10: Association of Oxidative Stress and Caspase Activity with *LMP1* Expression 70

Table 3.11: Risk Association of pTracer, pTracer-*LMP1*, pcDNA and pcDNA-*LMP1* Transfections with Oxidative Stress. 71

Table 3.12: Risk Association of pTracer, pTracer-*LMP1*, pcDNA and pcDNA-*LMP1* Transfections with Caspase Activity. 72

Table 3.13: Risk Association of Combination of Oxidative Stress and Caspase Activity with pTracer, pTracer-*LMP1*, pcDNA and pcDNA-*LMP1* Transfections. 74

Table 3.14: Mean and Standard Deviation of the Number of Chromosomal Breaks in the SAR of the *AF9* Gene in Transfected NP69 Cells. 77

Table 3.15: Association of Number of Chromosome Breaks in the SAR with *LMP1* Expression 78

Table 3.16: Mean and Standard Deviation of Number of Chromosomal Breaks in the Non-SAR of the *AF9* Gene in Transfected NP69 Cells. 79

Table 3.17: Association of Chromosome Break in Non-SAR of the *AF9* Gene with *LMP1* Expression. 80

Table 3.18: Risk Association of pTracer, pTracer-*LMP1*, pcDNA and pcDNA-*LMP1* Transfection with Number of Chromosomal Breaks in the SAR of the *AF9* Gene. 82

Table 3.19: Risk Association of pTracer, pTracer-*LMP1*, pcDNA and pcDNA-*LMP1* Transfection with Chromosomal Breaks in the Non-SAR of the *AF9* Gene. 83

Table 3.20: Breakpoints detected within the *AF9* SAR in *LMP1* Transfected cells. 84

Table 3.21: Breakpoints detected within the *AF9* non-SAR in *LMP1* Transfected cells. 86

LIST OF FIGURES

		Page
Figure 1.1:	The global distribution of NPC incidence in 2022 is illustrated through multiple visual representations.	3
Figure 1.2:	Linear representation of the wildtype EBV latent genome.	13
Figure 1.3:	Nuclear Matrix Model: MARs play various roles, including controlling gene expression, silencing, and coordinating DNA replication.	25
Figure 2.1:	Diagram showing a plasmid map of the pTracer-*LMP1* (approximately 7,265 bp).	32
Figure 2.2:	Diagram showing a plasmid map of the pcDNA-*LMP1* (approximately 6835 bp).	32
Figure 2.3:	Distribution of potential MAR/SAR sites predicted in the AF9 gene.	44
Figure 2.4:	Manipulation of gDNA in preparation for nested IPCR.	48
Figure 3.1:	Bright-field microscopy of NP69 cells transfected with the vector control pTracer.	57
Figure 3.2:	Dark-field microscopy of NP69 cells transfected with the vector control pTracer.	57
Figure 3.3:	Bright-field microscopy of pTracer-*LMP1* transfected NP69 cells.	58
Figure 3.4:	Dark-field microscopy of the pTracer- *LMP1* transfected NP69 cells with the GFP expression.	58
Figure 3.5:	Immunodetection of LMP1 by the S12 antibody on Western blot. Lane 1 represents the positive control B95 cell line.	60
Figure 3.6:	A representative gel picture shows the cleavages of the *AF9* gene identified by IPCR within the SAR.	75
Figure 3.7:	A representative gel image illustrates the cleavages in the AF9 gene detected by IPCR within the non-SAR region.	76
Figure 4.8:	Schematic representation of the molecular mechanism by which Epstein-Barr Virus (EBV) infection leads to carcinogenesis.	89

LIST OF ABBREVIATIONS

NPC	Nasopharyngeal Carcinoma
EBV	Epstein-Barr virus
DNA	Deoxyribose Nucleic Acid
MNNG	N-methyl-N'-nitro-N-nitrosoguanidine
LMP1	latent membrane protein
iNOS	inducible nitric oxide synthase
EBNA1	Epstein-Barr nuclear antigen 1
EBER1/2	Epstein-Barr virus (EBV)-encoded RNA 1
BARTs	Binder of ArlTwo
miRNAs	MicroRNAs
CTAR	C-terminal activating regions
TNF	tumour necrosis factor
TRAF	receptor-associated factor
ROS	reactive oxygen species
BZLF1	*Bam*HI Z fragment leftward open reading frame 1
TNFR	tumour necrosis factor receptor
TRAF	tumour necrosis factor receptor associated factor
STAT	signal transducers and activators of transcription
JNK	Janus kinase
RNS	reactive nitrate species
$O^{2\cdot -}$	superoxide anion
H_2O_2	hydrogen peroxide

NADPH	nicotinamide adenine dinucleotide phosphate
HOCl	hydroxyl radicals and hydrochloric acid
MLL	Mixed Lineage Leukaemia
HMW	high molecular- weight
MAR/SAR	matrix-attachment region or scaffold-associated region
CAD	caspase activated DNAase
BER	Base Excision Repair pathway
ATP	Adenosine triphosphate
ALL	acute lymphocytic leukaemia
MDS	myelodysplastic syndrome
t-AML	therapy related leukaemia
LB	Luria broth
rEGF	recombinant Epidermal Growth Factor
BPE	Bovine Pituitary Extract
PMSF	Phenylmethylsulphonyl fluoride
NaCl	sodium chloride
DCFH-DA	Dichlorofluorescin diacetate
IPCR	inverse polymerase chain reaction
UV	ultraviolet

CHAPTER 1

INTRODUCTION & LITERATURE REVIEW

1.1 Nasopharyngeal Carcinoma (NPC)

Nasopharyngeal carcinoma (NPC) is an epithelial cancer that begins in the mucosal lining of the nasopharynx, often occurring in the pharyngeal recess (fossa of Rosenmüller). While NPC and other epithelial head and neck tumors originate from similar cells or tissues, they are distinctly different. The World Health Organization (WHO) classifies nasopharyngeal carcinoma (NPC) into three subtypes based on the degree of epithelial differentiation: keratinizing squamous cell carcinoma (Type I), non-keratinizing squamous cell carcinoma (Type II), and undifferentiated or poorly differentiated carcinoma (Type III) (Shanmugaratnam, 1978). The non-keratinizing subtype is the most common in endemic regions and is strongly associated with Epstein–Barr virus (EBV) infection, whereas keratinizing subtypes are primarily observed in non-endemic areas of nasopharyngeal carcinoma (NPC) (Chen et al., 2019).

The TNM Classification for nasopharyngeal carcinoma (NPC), established by the International Union Against Cancer (UICC) and the American Joint Committee on Cancer (AJCC), serves as an essential prognostic tool for predicting patient outcomes and guiding treatment decisions. Advances in imaging and therapy have refined the system, enhancing its predictive accuracy (Guo et al., 2019). Integrating clinical features and molecular markers could further improve the existing prognostic model, enhancing its effectiveness in guiding personalized treatment strategies for nasopharyngeal carcinoma (NPC) (Lai et al., 2021).

With recent advancements in oncology, early-stage nasopharyngeal carcinoma (NPC) patients (stages I and II) achieve a 5-year overall survival rate of 94.0% (Jen et al., 2020). However, due to its anatomical location, vague clinical symptoms, and limited awareness, most NPC cases are diagnosed at advanced stages, leading to a poor prognosis. Symptoms such as headache, epistaxis, nasal obstruction, ocular issues, and neck masses are often misinterpreted due to their nonspecific nature, contributing to a misdiagnosis rate of 43.4% (Wu et al., 2016). This delay in diagnosis results in a median interval of 23.5 weeks from symptom onset to treatment initiation (K. H. Wang et al., 2017). For patients diagnosed at advanced stages, the 5-year overall survival rate drops to 89.2% for stage III and 73.7% for stage IV (Jen et al., 2020). Enhancing early diagnosis rates is crucial for reducing NPC-related mortality (Jiromaru et al., 2022).

Nasopharyngeal carcinoma (NPC) exhibits a distinct geographical and ethnic distribution, suggesting a significant role of genetic variations and environmental factors such as dietary habits, alcohol consumption, and smoking in its development. Over the past two decades, NPC incidence has shown an increasing trend, while mortality rates have declined (Su et al., 2024). Although research has identified several risk factors, the biological mechanisms underlying NPC remain inadequately understood. Therefore, further studies are needed to explore the complex interplay of these factors, enhance screening and early detection strategies, and develop more effective treatments for advanced NPC.

1.1.1 Incidence of Nasopharyngeal Carcinoma (NPC)

Nasopharyngeal carcinoma (NPC) is rare in most regions worldwide, with an age-standardized incidence rate (ASIR) of 1.3 per 100,000 people (Ferlay et al., 2024). In 2022, global cancer statistics reported 120,416 new NPC cases and 73,476 deaths, ranking it 23rd

among newly diagnosed cancers and 21st in cancer-related mortality (Ferlay et al., 2024). NPC incidence is approximately three times higher in males (86,257 cases) than in females (34,159 cases), possibly due to the protective effects of sex hormones, inherited susceptibility loci on the X chromosome, or lifestyle differences (Zuo et al., 2019). Socioeconomic status has also been linked to NPC risk, with lower-income populations facing greater exposure to risk factors (Chen et al., 2021). Additionally, individuals with lower socioeconomic status are more likely to be diagnosed at later stages, leading to poorer survival outcomes (London et al., 2022). NPC demonstrates a strong geographic and ethnic preference (Figure 1.1)(Sung et al., 2021).

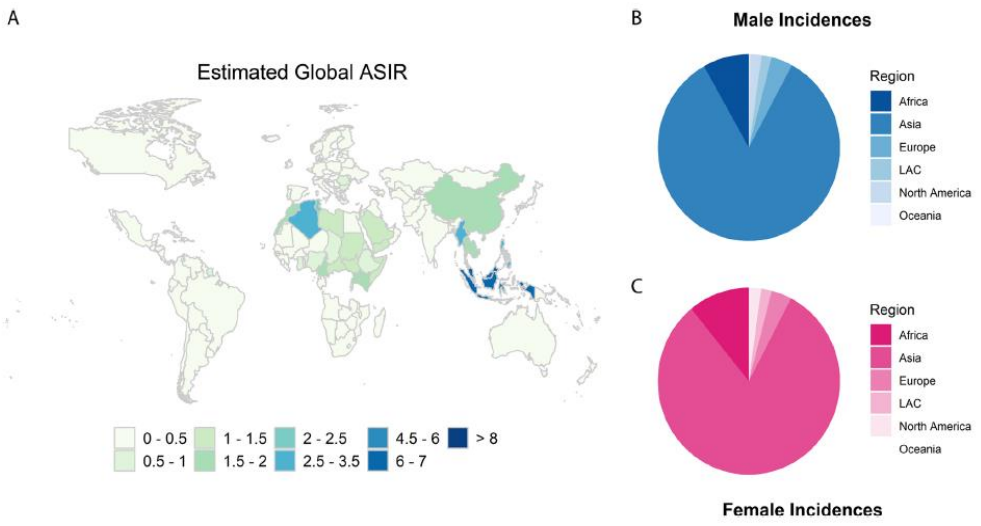

Figure 1.1: The global distribution of NPC incidence in 2022 is illustrated through multiple visual representations. A: features a heat map depicting the estimated global age-standardized incidence rate (ASIR). B: presents a pie chart showing the distribution of NPC cases among males, while C: displays a pie chart representing the distribution among females. The data was obtained from the Global Cancer Observatory (Ferlay et al., 2024). ASIR refers to the age-standardized incidence rate, and LAC stands for Latin America and the Caribbean.

Asia accounted for 83.3% of global nasopharyngeal carcinoma (NPC) cases (100,298 incidences) and 83.6% of NPC-related deaths (61,442 fatalities) (Ferlay et al., 2024). Among these, East Asia contributed 52.5% (52,610 cases), followed by Southeast Asia with 35.8%

(35,889 cases) and South-Central Asia with 9.2% (9,248 cases) (Ferlay et al., 2024). China reported the highest number of cases, totaling 51,010, with an age-standardized incidence rate (ASIR) of 2.4 per 100,000 people (Ferlay et al., 2024). Cancer registry data from 2013 to 2017 revealed a north-to-south increasing gradient in NPC incidence within China, with ASIRs varying nearly 50-fold, from 0.5 per 100,000 in North China to 25.0 per 100,000 in South China (Bray et al., 2024). High incidence rates were observed among the Cantonese-speaking population in southeastern China, particularly in Foshan City (ASIR: 24.0 per 100,000 males, 8.3 per 100,000 females), Luoding City (20.0 per 100,000 males, 7.3 per 100,000 females), Zhongshan City (19.4 per 100,000 males, 6.7 per 100,000 females), Zhuhai (17.2 per 100,000 males, 7.3 per 100,000 females), and Guangzhou (12.9 per 100,000 males, 4.5 per 100,000 females) (Bray et al., 2024).

According to the International Agency for Research on Cancer, there were about 129,000 new cases of NPC in 2018, making up only 0.7% of all cancer diagnoses that year (Bray et al., 2018). It is rare in most populations around the world, with an incidence of below 1 per 100,000 persons per year in Europe and the United States; however, in southern China and Southeast Asia, NPC is endemic, with an incidence rate of 20–30 per 100,000 persons per year (W. H. Jia & H. D. Qin, 2012). In 2008, approximately 84,400 new cases of NPC were reported, resulting in around 51,600 deaths, accounting for roughly 0.7% of the global cancer burden (Jemal, 2011).

The highest incidence was observed in South-Eastern Asia compared to the Americas, Europe, Africa, and Central and Eastern Asia (Jemal et al., 2011). The incidence of NPC is higher in males than in females, with a male-to-female ratio of approximately 2.5 in China in 2015 (W. Chen et al., 2016). NPC is more prevalent in males than females across

both developing and developed countries, with incidence rates typically 2 to 3 times higher in males in high-resource nations. In developing regions, the male-to-female rate ratio is often significantly greater (Jemal, 2011). The remarkable geographical variations in NPC prevalence results from the complex development of this carcinoma (Chang & Adami, 2006). The persistence of this malignancy in certain geographic regions suggests that genetic and/or stable environmental risk factors significantly contribute to its development. Additionally, the involvement of the Epstein-Barr virus (EBV) in the pathogenesis of NPC adds another layer of complexity to its aetiology.

The incidence and mortality rates of nasopharyngeal carcinoma (NPC) vary significantly across different regions and populations. Endemic areas, such as Southeast Asia, Southern China, and Northeast India, consistently report higher NPC rates, whereas North America and Europe are classified as non-endemic regions with lower prevalence (Su et al., 2024). These differences are influenced by various social factors beyond genetic or ancestral background. Socioeconomic conditions play a key role, as lower education levels can reduce awareness of NPC prevention and early detection. Additionally, limited access to healthcare services, including screening and treatment facilities, often leads to delayed diagnosis and medical intervention. Cultural beliefs and social norms further influence healthcare-seeking behaviors related to NPC.

1.1.2 Risk Factors of Nasopharyngeal Carcinoma (NPC)

Over the years, research has provided deeper insights into the complex, multifactorial pathogenesis of nasopharyngeal carcinoma (NPC), which arises from the interplay of environmental, genetic, and viral factors (Hung et al., 2023). This interplay likely contributes to the distinct geographic and ethnic patterns observed in NPC incidence. Among low-risk

populations, NPC cases typically follow a bimodal age distribution, with an initial peak between ages 15–24, likely influenced by susceptibility genes and Epstein–Barr virus (EBV) infection, and a second peak between ages 65–79, which may be driven by environmental factors such as heavy smoking. In contrast, high-risk populations exhibit a unimodal age distribution, with a single peak occurring between ages 45–59, where early epigenetic events are thought to obscure the first peak (Bray et al., 2008).

Environmental exposure to various risk factors, such as tobacco smoking, dietary elements, and inhalants, has been linked to the development of nasopharyngeal carcinoma (NPC) (Lin et al., 2021). Tobacco smoking is a well-established cause of NPC, particularly the differentiated subtype, which is more prevalent in low-risk populations. The likelihood of developing NPC increases with smoking intensity, duration, cumulative exposure, and early smoking initiation (Lin et al., 2021). Smokers face a 32% higher risk of NPC compared to non-smokers, with this risk rising to 67% for those consuming more than 16 cigarettes per day. Although quitting smoking leads to a slight reduction in NPC risk (Xue et al., 2013), passive smoke exposure during childhood (OR = 1.24) and adulthood (OR = 1.3) has been linked to an increased NPC risk in non-smokers (Chang et al., 2017).

The heightened risk associated with smoking may be due to repeated Epstein-Barr virus (EBV) reactivation, as well as DNA damage or mutations (Hu et al., 2019). While emerging evidence suggests a connection between thirdhand smoke exposure and various cancers, conclusive research directly linking it to NPC remains lacking (Arfaeinia et al., 2023). Additionally, e-cigarette use has been associated with an increased risk of head and neck cancers. The aerosol or liquid in e-cigarettes contains carcinogenic substances, such as formaldehyde, metals, and polycyclic hydrocarbons, which have been linked to NPC risk

(Szukalska et al., 2020). In vitro studies indicate that exposure to e-cigarette liquid induces oxidative stress, DNA damage, and cellular changes (Szukalska et al., 2020). However, research specifically examining the relationship between e-cigarette use and NPC is still limited. Further studies are essential to better understand the complex interplay between traditional tobacco use, thirdhand smoke exposure, and e-cigarette consumption in NPC development, which could help shape effective public health policies and interventions.

Several risk factors have been identified for NPC, including geographic location and ethnicity. NPC is more common in regions such as Southeast Asia, North Africa, and the Arctic. People of Chinese or Southeast Asian descent are at a higher risk (Simons, 2011). Additionally, having a family history of NPC increases the risk, and certain genetic factors, such as specific human leukocyte antigen (HLA) types, are associated with a higher risk (Su et al., 2013). Bile acid (BA), a component of gastric duodenal contents, has been recognised as a carcinogen. BA-induced apoptosis has been suggested to play a role in human malignancies. BA-induced apoptosis might contribute to mediating chromosome rearrangements in nasopharyngeal carcinoma (NPC) (Tan et al., 2018). Furthermore, consuming salt-cured fish and other preserved foods containing high levels of nitrosamines is linked to a higher risk of NPC (Ward et al., 2000). A diet low in fresh fruits and vegetables may also contribute to the risk. Exposure to certain chemicals and dust, such as wood dust, can also increase the risk. Smoking and heavy alcohol consumption are potential risk factors, although their association with NPC is less strong compared to other factors. NPC can occur at any age but is most commonly diagnosed in people between 30 and 50 years old, and males are more likely to develop NPC than females (Okekpa et al., 2019). In addition, EBV has long been recognized as being associated with nasopharyngeal carcinoma (NPC)

(Bouvard et al., 2009; Niedobitek, 2000). Its direct role in NPC pathogenesis remains unclear investigation.

1.1.3 Genetic Alteration in Nasopharyngeal Carcinoma (NPC)

Genetic factors play a significant role in NPC, as evidenced by its distinct geographical and racial distribution, the higher susceptibility of migrants from endemic to non-endemic regions compared to local populations, and the presence of familial history in over 10% of patients from endemic areas (Cao et al., 2014). Various genetic studies utilizing linkage and association analyses have identified risk genes linked to NPC susceptibility (Siak et al., 2021). These genes are involved in crucial biological processes, including immune response, metabolism, DNA repair, and cell cycle regulation (Ferlay et al., 2024).

Research into the genetic landscape of NPC has highlighted the importance of immune-related genes in disease susceptibility. In particular, the human leukocyte antigen (HLA) loci on chromosome 6p21 have been consistently associated with an increased risk of NPC. Additionally, non-HLA innate immune genes within the major histocompatibility complex (MHC) region, such as Tripartite motif-containing (TRIM) proteins, as well as genes encoding cytokines and chemokines such as interleukins (ILs) play a crucial role in regulating host immune response and controlling viral infections, thereby influencing NPC risk (Ning et al., 2020). Overall, genetic variations in immune-related genes contribute to NPC susceptibility, underscoring the need for further research to unravel the underlying mechanisms and develop targeted strategies for prevention and treatment.

Multiple genetic abnormalities have been linked to nasopharyngeal carcinoma, suggesting that its tumorigenesis involves various genetic alterations. These alterations include chromosomal gains or losses (Chen et al., 1999; Hui et al., 1999; Fang et al., 2001;

Yan et al., 2001; Li et al., 2006), loss of heterozygosity (LOH) (Hui et al., 1996; Mutirangura et al., 1998; Lo et al., 2000; Shao et al., 2000), homozygous deletions (Huang et al., 1994; Lo et al., 1995; Hu et al., 1996; Hui et al., 2005), hypermethylation of tumour suppressor gene promoters (Lo et al., 1996; Kwong et al., 2002; Tan et al., 2006; Yanatatsaneejit et al., 2008), and telomere shortening (Xiao et al., 1999; Wen et al., 2002).

Early karyotype studies consistently identified numerical and structural abnormalities of chromosome 3, such as deletions, duplications, and translocations, in established cell lines and xenografts (Bernheim et al., 1993). Additionally, various chromosomal rearrangements, including gains and losses, have been observed in NPC patients (Chien et al., 2001). Studies on the loss of heterozygosity in the short arm of chromosome 3 in NPC have revealed frequent deletions in a small panel of primary tumours (Huang et al., 1991). Lo et al. (2012) proposed that the loss of regions on chromosomes 3p and 9p is an early event in the transformation of normal nasopharyngeal epithelium. Their study also showed high frequencies of genomic deletions in these regions not only in dysplastic lesions but also in histologically normal epithelia from endemic regions (Lo et al., 2012).

Deletions of chromosome 9p were identified in 66.7% of dysplastic nasopharyngeal lesions (Chan et al., 2002). These findings imply that the loss of heterozygosity (LOH) on chromosome 9p is an early event in the tumourigenesis of nasopharyngeal carcinoma (NPC). The heightened risk of NPC among southern Chinese populations may be associated with the early loss of genetic material, as indicated by 9p LOH in nasopharyngeal tissue from both high- and low-risk regions (Chan et al., 2002). The minimal common deletion region could be pinpointed to 9p21.1, indicating that the inactivation of one or more tumour

suppressor genes in this area may play a crucial role in the development of NPC (Yang et al., 2000). The studies also identified frequent rearrangements of chromosomes 1 and 12, along with deletions in chromosomes 11 and 14. These pre-existing genetic alterations in precursor cells may play a crucial role in cell immortalization, susceptibility to EBV infection, and the maintenance of permanent viral latency (Tsang et al., 2014). Additionally, several studies have pinpointed potential tumour suppressor genes and oncogenes (Hu et al., 2012). However, the mechanisms driving the non-random chromosome rearrangements in NPC remain unclear. Various mechanisms have been proposed for chromosome rearrangements in other malignancies.

1.2 Epstein-Barr Virus (EBV) in Nasopharyngeal Carcinoma (NPC)

EBV is a human herpesvirus that consists of double-stranded deoxyribonucleic acid (DNA), similar to other herpesviruses, with a genome of approximately 180 kb containing about 80 open reading frames (Baer et al., 1984). EBV has infected more than 90% of the adult human population and is closely associated with malignancies such as lymphoma, NPC and gastric cancer (Chien & Chen, 2003; Morales-Sanchez & Fuentes-Panana, 2014; Pagano, 1999). EBV infection is found to be involved in the development of about 20% of human cancers with approximately 2 million cases per year (de Martel et al., 2012). Study by Tao et al (2006) suggested that, infection with EBV is very common, affecting all human populations and is largely asymptomatic and life-long (Tao et al., 2006). One of the characteristics of herpesviruses is their ability to produce lifelong infections in the host. Once individuals are infected, they remain infected until death, and this usually occurs in healthy carriers (Gourzones et al., 2013). Epstein-Barr Virus infects B cells of the immune system and epithelial cells. It latently persists in the host B cells for life (Xu et al., 2015). Many researchers have reported a connection between EBV infection and the development of head

and neck cancer (Al-Thawadi et al., 2020; Jalouli et al., 2010; Jalouli et al., 2012; Prabhu & Wilson, 2016; Shao et al., 2023). About 98% of NPC cases are from healthy carriers of the Epstein-Barr virus (EBV), which is found in the majority of ethnic groups worldwide (Kutok & Wang, 2006). The link between EBV infection and NPC was initially identified when patients exhibited high levels of serum antibodies against EBV antigens, such as viral capsid antigen (VCA) and early antigen diffuse (EAd/BMRF1) (GunvÉN et al., 1970). The presence of the EBV genome in NPC cells was subsequently demonstrated through *in situ* hybridisation (Wolf et al., 1975).

Stable EBV infection was established in multiple lines of immortalised nasopharyngeal epithelial cells and NPC cells (Tsang et al., 2012). Lo et al (2003) reported that stress–resistance phenotype could also be observed in nasopharyngeal epithelial cells expressing the EBV-encoded LMP1 (Lo et al., 2003). Stable EBV infection and the expression of latent EBV genes in genetically abnormal nasopharyngeal epithelial cells are believed to drive the clonal expansion of these EBV-infected cells (Tsao et al., 2017). Therefore, the establishment of latent EBV infection in precancerous nasopharyngeal epithelium is considered an early and crucial step in the development of NPC (Tsao et al., 2017).

EBV infection primarily occurs in childhood and typically remains asymptomatic due to immune responses. Today, about 90% of the global adult population is infected with EBV, with the most common transmission route being close contact with the saliva of an infected individual (Kgatle et al., 2017). In NPC, EBV persists in a latent state, with restricted expression of viral genes such as *LMP1/2*, Epstein-Barr nuclear antigen 1 (*EBNA1*), Epstein-Barr virus (EBV)-encoded RNA 1/2 (*EBER1/*2), and BARTs (Binder of

Arl2). *LMP1* can activate cellular DNA methyltransferase via c-Jun NH2-terminal kinase signalling (Tsai et al., 2006) and upregulate the PcG protein BMI-1 (Dutton et al., 2007), which is associated with epigenetic alterations in NPC. On the other hand, latent episomal EBV genomes are subjected to host cell-dependent epigenetic modifications including DNA methylation and histone modifications, which are important for maintenance of EBV latent infection (Minarovits, 2006). In cells stably infected with EBV, various latent gene proteins like LMP1 and LMP2A promote rapid clonal expansion and transformation (Rancan et al., 2015). Discovering variability in the *LMP1* gene and its functions could be pivotal in understanding EBV's role in NPC carcinogenesis.

The expression of these EBV latent proteins consistently activates multiple signalling pathways, increases genetic instability, induces epigenetic changes, modulates the microenvironment, and suppresses the host immune response during the early stages of cancer development (Lo et al., 2012). The epigenetic impact is evident through the promoter methylation of tumour suppressor genes, which is believed to contribute to NPC cell growth and invasion (Zhang et al., 2013). Other clonal genetic and epigenetic changes accumulate under continuous selection pressure. *LMP1* and *LMP2A* expression is suppressed by various mechanisms such as EBV microRNAs (miRNAs), DNA methylation, and transcription factors to avoid cytotoxic effects and immune response in most primary tumours (Takacs et al., 2010). Their oncogenic function is preserved by various permanent genetic alterations that develop during tumour progression. EBV was the first human virus identified with oncogenic potential (Epstein et al., 1964).

Epigenetic mechanisms related to EBV involve a complex regulatory network that induces heritable, non-mutative genomic changes, impacting both viral and host DNA. From

the viral perspective, the ability to reprogram gene expression across different latency types and between latent and lytic stages highlights the crucial role of epigenetic modulation in facilitating successful EBV infection. The EBV genome encodes various nuclear antigens (EBNAs), membrane proteins (LMPs), microRNAs (ebv-miRs), and non-coding RNAs (ncRNAs; EBERs) (Rowe et al., 1992) (Figure 1), many of which are selectively expressed depending on the latency type.

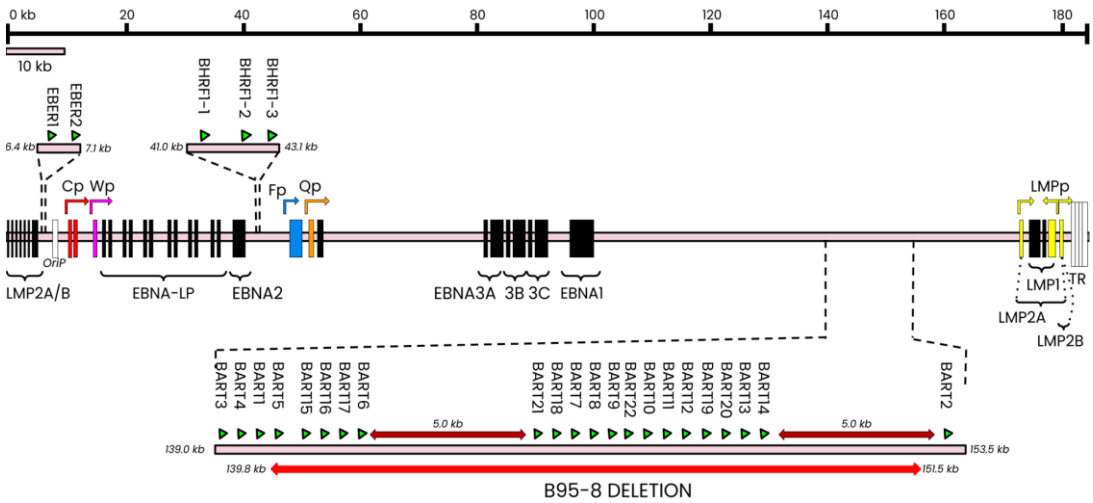

Figure 1.2: Linear representation of the wildtype EBV latent genome. The genome contains four latent promoters (Cp, Wp, Qp, LMPp) that generate mRNA transcripts for eight key latent antigens. The major latent antigens are shown as black bars, representing their coding exon positions. The LMP coding region crosses the terminal repeats and loops back to the genome's starting point.Some spontaneous deletions occur in the BART region (red underline), as seen in the B95-8 laboratory strain (Price & Luftig, 2014).

1.2.1 Epstein-Barr Virus (EBV) Reactivation in Nasopharyngeal Carcinoma (NPC)

In an infected host, virus tends to establish a lifelong latent infection, during which the virus alternates between the latent state and lytic replication. Likewise, EBV establishes a latent infection that can periodically reactivate into the productive lytic cycle (Kenney & Mertz, 2014). The switch between latency and lytic phases is tightly regulated by both viral and cellular factors. Upon reactivation, EBV's lytic genes are expressed in a temporally controlled sequence. The shift from latent to lytic phase is marked by the expression of two

viral immediate early genes, Zebra (also known as BZLF1, Zta, or Z) and Rta (also known as BRLF1 or R). In latently infected cells, the expression of BZLF1 and BRLF1 is suppressed by multiple cellular transcriptional repressors. Evidence suggests that lytic induction of EBV can occur in EBV-associated diseases (Arvey et al., 2012; Tsai et al., 2013).

Research has shown that EBV lytic reactivation is a significant risk factor for the development of EBV-associated diseases. It primarily contributes to genome instability, evasion of host immune responses, resistance to cell death, and the promotion of tumour development, progression, and invasiveness (Young et al., 2016). Repeated chemical reactivation of EBV has been demonstrated to increase genome instability, thereby accelerating tumour progression in NPC cells (Fang et al., 2009). Additionally, the N-nitroso compound N-methyl-N'-nitro-N-nitrosoguanidine (MNNG) was found to work in conjunction with 12-O-tetradecanoylphorbol-1,3-acetate (TPA)/sodium butyrate to enhance EBV reactivation (Huang et al., 2010). This treatment also raised reactive oxygen species (ROS) levels and increased genome instability in NPC cells. Moreover, the expression of the EBV latency gene, latent membrane protein 1 (*LMP1*), has been shown to promote genome instability, observed as nonclonal chromosomal aberrations in a Burkitt's lymphoma cell line (B. Gruhne et al., 2009).

Pre-clinical and clinical evidence suggests that oxidative stress plays a role in EBV lytic reactivation, with some proposing that EBV-positive tumours are driven by reactive oxygen species (Bonner & Arbiser, 2014). This indicates that oxidative stress could be a crucial factor in the carcinogenesis of EBV-induced cancers and may serve as a potential pharmacological target for cancer prevention and treatment. Notably, EBV lytic replication

has been shown to significantly impact genetic stability, revealing a new mechanism in EBV-induced cancer (Shumilov et al., 2017).

Hong et al. demonstrated that lytically infected cells in EBV-positive tumours might contribute to tumour formation by releasing paracrine growth factors and angiogenesis factors, such as vascular endothelial growth factor (VEGF), potentially aiding viral pathogenesis (Hong et al., 2005). The establishment of a latent transforming infection in epithelial cells, along with genetic alterations that either facilitate latent infection or work synergistically with EBV transforming proteins, are likely key steps in the development of NPC. Despite the high prevalence of EBV in the human population, certain genome-specific factors contributing to NPC pathogenesis remain unidentified (Shannon-Lowe & Rickinson, 2019).

1.3 *LMP1* Gene in Nasopharyngeal Carcinoma (NPC)

The *LMP1* gene of EBV is a key oncoprotein in EBV-associated malignancies, playing a significant role in the initiation and progression of NPC. *LMP1* is one of seven critical genes that can accurately predict the survival outcomes of NPC patients (Wang et al., 2011). However, the mechanisms by which *LMP1* contributes to the radioresistance of NPC remain largely unclear. *LMP1*, a crucial EBV oncogene, has been shown to transform rodent fibroblasts in vitro and induce tumours in nude mice (Moorthy & Thorley-Lawson, 1993). The oncogenic potential of *LMP1*, leading to B cell transformation, is thought to be due to its strong functional resemblance to members of the tumour necrosis factor receptor (TNFR) family, such as CD40 and TNFR1(Kulwichit et al., 1998).

LMP1 plays a crucial role in EBV-induced B-cell transformation and is also highly expressed during the lytic phase of viral replication. Several studies suggest that *LMP1* may

be a key inducer of reactive oxygen species (ROS). Cerimele et al. used a tetracycline-inducible epithelial cell line to investigate whether *LMP1* can trigger ROS generation. They found that adding tetracycline significantly reduced *LMP1* expression and ROS production, implicating *LMP1* as a major ROS inducer (Cerimele et al., 2005). The C-terminal activating regions (CTAR) of the LMP1 protein interact with members of the tumour necrosis factor (TNF) receptor-associated factor (TRAF) families. Although ROS activation has not previously been linked to *LMP1* signaling, TRAF activation by other ligands has been shown to induce ROS signaling, suggesting a potential mechanism for *LMP1*-mediated ROS induction. Additionally, Chen et al. demonstrated that EBV infection of B lymphocytes could lead to ROS accumulation, and reducing ROS levels was associated with a selective down-regulation of viral *LMP1*(X. Chen et al., 2016).

Spontaneous and bleomycin-induced micronucleus formation was found to be more in *LMP1*-expressing cells (Liu et al., 2004). *LMP1*'s role in promoting genome instability was suggested to be attributed to its ability in inhibiting DNA repair, whereby the involvement of oxidative stress was moderate (B. Gruhne et al., 2009). *LMP1*'s involvement in oxidative and nitrative DNA damage is evident. This was demonstrated by observation that its expression triggers the expression of inducible nitric oxide synthase (iNOS) (Yu et al., 2002). Oxidative and nitrative DNA lesions has long been associated with EBV-positive patients (Ma et al., 2008). The oncogenic potential of *LMP1* is also associated with its modulation on stress-induced apoptosis. Both etoposide- and cisplatin-triggered apoptosis were enhanced by *LMP1* expression. This potentiation is thought to be at early stage of apoptosis (Lo et al., 2012).

1.3.1 Implication of *LMP1* in Nasopharyngeal Carcinoma (NPC) Development

The role of *LMP1* in the pathogenesis of NPC remains speculative. The expression of *LMP1* in NPC biopsies varies considerably, with studies generally agreeing that about 20%–40% of tumors express *LMP1* at the protein level (Tao et al., 2006). One study found that all six early, preinvasive NPC (NPC in situ) lesions analyzed expressed *LMP1*, suggesting a critical role for *LMP1* in the early stages of NPC pathogenesis (Pathmanathan et al., 1995). Additionally, *LMP1* expression has been shown to inhibit the AMPK/LKB1 signaling pathways, thereby promoting cellular growth and survival (Lo et al., 2013).

LMP1 influences EBV lytic reactivation through multiple mechanisms. Recent studies have highlighted various roles of *LMP1* in this process. The expression of *LMP1* contributes to EBV lytic reactivation, and its inducible expression during the viral lytic cycle is crucial for virus production. The absence of *LMP1* significantly impairs virus release into culture supernatant fractions, leading to reduced infection efficiency(Ahsan et al., 2005). EBV utilizes *LMP1* for dual purposes during its life cycle (Hu et al., 2017). First, in latently infected B lymphocytes, the steady-state expression of *LMP1* is essential for maintaining the transformation status. Second, at the onset of the virus production cycle, the strong induction of *LMP1* facilitates efficient virus release from cells. Thus, *LMP1* is one of the most critical viral proteins required for EBV's life cycle, survival, and transmission from host to host.

The carboxyl terminus of *LMP1* contains consensus tumour necrosis factor receptor-associated factor (TRAF) binding domains, which can constitutively activate signal transducers and activators of transcription (STAT), Janus kinase (JNK), and nuclear factor (NF) κB pathways for cell survival and growth (Dawson et al., 2012). *LMP1* is also considered as a classical viral oncogene that activates multiple cell signalling pathways,

including NFκB, MAPK, and PI3K, to drive tumourigenesis (Dawson et al., 2012). As a key effector in EBV-driven B cell transformation and an established “transforming” gene, *LMP1* displays oncogenic properties in rodent fibroblasts and induces profound morphological and phenotypic effects in epithelial cells (Dawson et al., 2012).

An analysis of how elevated ROS levels support *LMP1* expression revealed that they selectively inhibit viral microRNAs that target the *LMP1* transcript (X. Chen et al., 2016). *LMP1* has been shown to induce ROS generation, and the oxidative stress microenvironment can also influence *LMP1* expression by targeting viral microRNAs. The researchers identified a potential mechanism involving the disruption of signaling pathways initiated by *LMP1* (Cheerathodi & Meckes Jr, 2018).

1.4 Oxidative Stress in Nasopharyngeal Carcinoma (NPC)

In the human body, oxygen is essential for metabolic processes that produce the energy necessary for life. During cellular respiration, oxygen acts as the final electron acceptor in the four-stage electron transport chain. However, sometimes electrons "leak" from this chain, leading to an incomplete reduction of the oxygen molecule, which results in the formation of reactive oxygen species (ROS) (Karpińska & Gromadzka, 2013). ROS are free radicals, molecules that contain at least one unpaired electron in their outer electron shell, making them highly reactive as they seek to pair their electrons by either donating to or taking from other molecules. The formation of these free radicals is strongly influenced by physical factors such as ionizing radiation, ultraviolet radiation, ultrasound, and elevated temperatures. Although ROS are by-products of oxygen metabolism and play important roles in cell signaling and homeostasis, oxidative stress occurs when the body's antioxidants

cannot neutralize the harmful effects of these free radicals. Normal oxygen is relatively unreactive, but certain metabolic processes can convert it into a highly reactive oxidant.

Epstein-Barr Virus (EBV) reactivation and the expression of latency genes could play an important role in inducing chromosome rearrangements, perhaps via the production of reactive oxygen species (ROS) and the involvement of apoptosis It is known that ROS trigger apoptosis, leading to chromosome breaks in regions frequently rearranged in nasopharyngeal carcinoma (NPC) (Tan et al., 2016). During repeated EBV reactivation and latency establishment, ROS may be produced, causing repeated oxidative damage to the cell's genome. This can lead to apoptosis, and in the process of DNA repair, errors might occur, allowing cells with rearranged chromosomes to survive and continue replicating, thus contributing to NPC carcinogenesis. Reactive oxygen species (e.g., H_2O_2, hydroxyl radical, and superoxide) cause injuries to various cellular macromolecules. These damages have been proposed to contribute to the development of cancer (Ames, 1989; Wiseman & Halliwell, 1996). The Base Excision Repair (BER) pathway, which involves the hOGG1 and XRCC1 genes, is responsible for repairing oxidatively damaged DNA (Hoeijmakers, 2001); (Marsin et al., 2003). Polymorphisms in these genes have been associated with an increased risk of NPC (Cho et al., 2003), further supporting the role of oxidative stress in NPC development.

Oxidative stress has been implicated as a key factor in the development of cancer (Klaunig & Kamendulis, 2004). Given the strong association between oxidative stress and the risk factors for nasopharyngeal carcinoma (NPC), exploring its role in the molecular mechanisms driving chromosome rearrangements in NPC is particularly compelling. Oxidative stress can trigger apoptosis (Sun et al., 2018a), a process characterized by distinct

morphological changes such as chromatin condensation, nuclear fragmentation, cytoplasmic organelle compaction, cell shrinkage, and membrane blebbing (Shimizu et al., 2007). Chromosomal breakages, which are crucial in the early stages of chromosome rearrangement and apoptotic DNA fragmentation, often occur in specific chromatin regions like matrix attachment regions/scaffold attachment regions (MAR/SAR) (Shimizu et al., 2007; Strick et al., 2006b; Strissel et al., 2000). Additionally, the primary apoptotic nuclease, CAD, has been observed to associate with the nuclear matrix in apoptotic cells (Lechardeur et al., 2004). Research indicates that in both normal nasopharyngeal epithelial cells and NPC cells, oxidative stress-induced apoptosis leads to chromosome breaks in the *AF9* gene located on chromosome 9p22 (Tan et al., 2018a). Caspase-activated DNase (CAD) is likely a key mediator of these oxidative stress-induced chromosomal breaks. Multiple chromosome breaks have been identified within the *AF9* region, previously associated with translocations in a patient with acute lymphoblastic leukaemia (ALL). These observations suggest that oxidative stress-induced apoptosis may significantly contribute to chromosome rearrangements in NPC (Tan et al., 2016).

1.5 Association of Apoptosis with Chromosomal Rearrangement in Nasopharyngeal Carcinoma (NPC)

Although chromosomal abnormalities are frequently observed in NPC, the precise molecular mechanisms behind these abnormalities remain unclear. However, growing evidence suggests that apoptotic nucleases are responsible for the initial chromosomal breakage that leads to translocations in leukemic cells(Betti et al., 2003; Sim & Liu, 2001). A similar mechanism has been proposed for chromosome rearrangements in NPC. Recent studies have indicated that apoptotic nucleases may also be involved in mediating chromosome breaks in NPC cells during oxidative stress, potentially serving as the initial

event leading to chromosome rearrangement (Boon & Sim, 2015a, 2015b). According to research by Tan et al. (2016), oxidative stress-induced apoptosis is suggested as one of the mechanisms contributing to chromosomal rearrangements in NPC (Tan et al., 2016). Additionally, bile acid-induced apoptosis has also been shown to play a role in chromosome rearrangement in NPC (Tan & Sim, 2018).

Oxidative stress is significant as it can promote the expansion of mutant cell clones by temporarily altering the expression of genes that regulate cell proliferation or cell death (Trueba et al., 2004). Tumour cells can develop various strategies to evade apoptosis and ensure their survival (Kim et al., 2015). During EBV infection, the virus produces several anti-apoptotic factors, including BHRF1, BARF1, EBNA1, ERERs, and miR-BARTs, which are crucial in helping the infected cells resist apoptosis.

1.5.1 Apoptosis and Its Hallmark

Apoptosis is a natural process of cell death that plays a crucial role in various biological systems (Ellis et al., 1991). It is characterized by distinctive morphological features, including cell shrinkage, fragmentation into membrane-bound apoptotic bodies, and rapid engulfment by neighbouring cells. These changes are largely driven by a class of cysteine proteases known as caspases (Hengartner, 2000). Several caspases have been identified, with approximately two-thirds involved in apoptosis (Earnshaw et al., 1999). Caspases typically exist as inactive pro-enzymes and are activated when apoptosis is initiated, enabling them to cleave specific proteins, thereby either inactivating or activating these target proteins (Alenzi et al., 2010); (Hengartner, 2000).

Various proteins are involved in the apoptotic cascade, and their presence can be identified using standard protein detection techniques. The progression and rate of apoptosis

are often assessed by measuring the activity of caspases protease enzymes that selectively cleave peptide bonds in specific substrates. So far, 11 caspase isoenzymes have been identified in humans, with seven of them (caspase 2, 3, 6, 7, 8, 9, and 10) participating in apoptosis (McIlwain et al., 2013). Apoptosis plays a crucial role in both physiological and pathological processes. Dysregulation, either through excessive apoptosis (Mattson, 2000) or insufficient apoptosis (Wong, 2011), can contribute to disease development. Therefore, detecting apoptosis is essential for improving strategies to slow or even prevent disease progression. Moreover, apoptosis detection is a key factor in evaluating the effectiveness of potential new medications and assessing the cytotoxicity of various chemicals.

During apoptosis, the activation of caspases leads to the fragmentation of genomic DNA into high molecular weight (HMW) DNA and smaller fragments known as the internucleosomal DNA ladder (Oberhammer et al., 1993). These HMW DNA fragments, ranging from 50 to 300 kb, correspond to DNA-loop structures(Filipski et al., 1990), that interact with the nuclear matrix via matrix-attachment region (MAR) or scaffold-associated region (SAR) sequences (Cockerill & Garrard, 1986). The formation of HMW DNA during apoptosis is believed to result from DNA loop excision at the MAR/SAR sequence at the base of the DNA loop (Lagarkova et al., 1995). Caspase-activated DNase (CAD) is one of the key enzymes involved in this process Under normal conditions, CAD exists in an inactive complex with its inhibitor, the inhibitor of CAD (ICAD). Upon induction of apoptosis, caspase-3 cleaves ICAD, releasing the active CAD, which then cleaves the genomic DNA into HMW DNA and the internucleosomal DNA ladder (Sakahira et al., 1998).

Caspase-activated DNase (CAD) appears to have multiple functions. It is known primarily as the apoptotic nuclease, but it has also been implicated in chromosome

rearrangements commonly observed in leukaemia (Nicholas & Sim, 2012). Additionally, CAD has been shown to promote cell differentiation by inducing DNA strand breaks (Larsen et al., 2010). As an enzyme involved in apoptosis, CAD plays a crucial role in chromosome rearrangements, particularly in leukaemia (Sim & Liu, 2001). Research suggests that CAD may also be instrumental in chromosomal cleavages driven by oxidative stress-induced apoptosis. This has led to the proposal of a model in which oxidative stress-induced apoptosis mediates chromosomal rearrangements in NPC (Tan et al., 2016).

1.6 The Implication of Apoptosis and Chromosomal Rearrangement in Nasopharyngeal Carcinoma (NPC)

Apoptosis was found to be triggered by oxidative stress resulting from the excessive production of reactive oxygen species (ROS) such as H_2O_2, hydroxyl radical, and superoxide (Zhao & Wang, 2012). Oxidative stress has been shown to induce apoptosis in endothelial cells in an ATP-rich environment (Lelli et al., 1998). It also triggers the formation of High Molecular Weight (HMW) DNA fragments, an early step in apoptosis, in leukemic cells (Lelli et al., 1998). Evidence suggests that NPC cells under stress experience apoptosis-induced chromosome breaks (Lelli et al., 1998). Although apoptosis is primarily a process of cell death, it has also been linked to chromosome rearrangements. For instance, apoptosis induced by chemotherapeutic drugs has been associated with chromosome breaks in the Mixed Lineage Leukaemia (*MLL*) gene, a gene commonly involved in chromosome translocations (Sim and Liu, 2001).

1.7 Association of Matrix Attachment Region/ Scaffold Attachment Region (MAR/SAR) Sequences with Chromosomal Rearrangements

Eukaryotic cells are organized into multiple organelles, with a well-defined nucleus that houses genetic material. Human DNA, approximately 3 meters in length, is tightly

compacted to fit within the small nucleus. However, this compaction does not render the DNA inactive; instead, it remains dynamically accessible for controlled gene expression. The nuclear matrix, a three-dimensional RNA–protein network, provides structural support for the organized compaction of DNA (Heng et al., 2004). Chromatin is arranged into loops, anchored to the nuclear matrix by specific DNA sequences known as Matrix Attachment Region/ Scaffold Attachment Region (MAR/SAR) (Capco et al., 1982). Various proteins, called MAR/SAR binding proteins (MAR/SARBPs), interact with MAR/SAR s to facilitate chromatin looping, which is essential for key cellular processes such as DNA replication, transcription, chromatin-to-chromosome transition, and DNA repair (Stein et al., 2003). Interestingly, MAR/SAR sequences lack strict sequence conservation (Laemmli et al., 1992), but their structural features appear to be preserved and functionally significant (Tikhonov et al., 2000). These sequences possess characteristics such as origins of replication (OriC), AT-rich regions, kinked and curved DNA, TG-rich segments, MAR signatures, and Topoisomerase-II binding sites (Tikhonov et al., 2000).

In 1999, Croft et al. highlighted the role of the nuclear matrix in regulating gene expression on chromosomes 18 and 19. Their study found that genes on chromosome 19, which is positioned internally within the nucleus and closely associated with the nuclear matrix, exhibit high transcriptional activity. In contrast, chromosome 18, which tends to occupy a peripheral location in the nucleus, shows lower gene expression levels (Croft et al., 1999). Similarly, MAR/SARs have been shown to enhance transgene expression and stability across various organisms (Zhao et al., 2017), underscoring their critical role in the organization and function of genetic material.

The interaction between MAR/SARs and the nuclear matrix has also been extensively studied in different biological conditions, including diseases (Barboro et al., 2012). Given their influence on genome topology and function, both MAR/SARs and the nuclear matrix present promising targets for therapeutic interventions. However, despite significant advancements in chromatin biology, a comprehensive genome-wide map of MAR/SARs in the human genome is still lacking. MAR/SARs seem to be involved in invasion-related biological processes and may also serve as a structural framework for organizing these nuclear activities (Fig. 4)(Drennan et al., 2010).

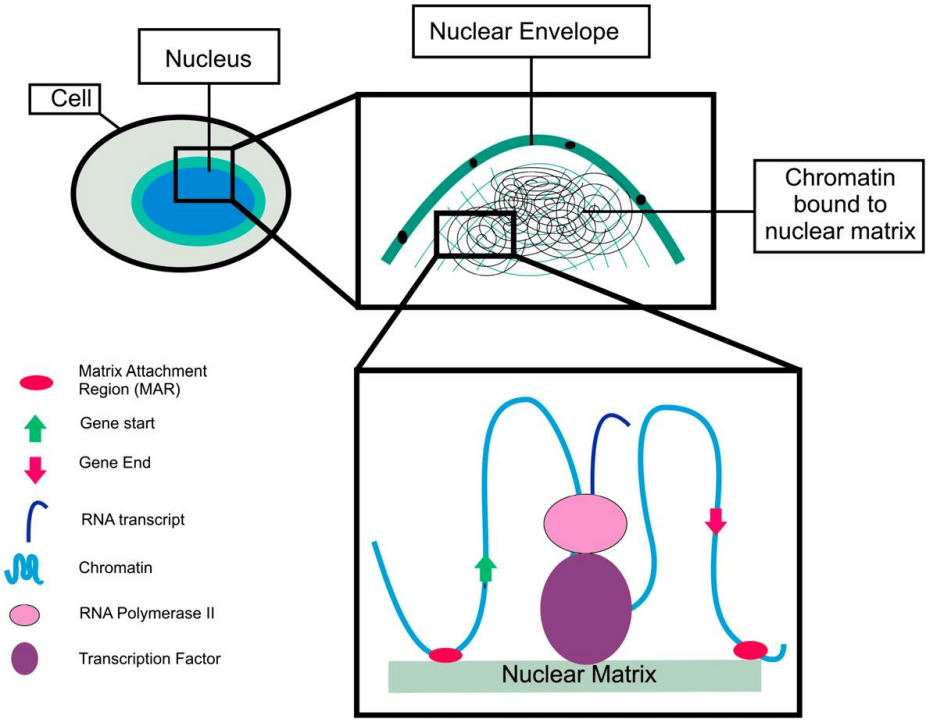

Figure 1.3: Nuclear Matrix Model: MARs play various roles, including controlling gene expression, silencing, and coordinating DNA replication. They are often found near the beginning and end of genes, helping position them for activation, regulatory binding, or post-transcriptional control. When MARs are located within genes, they are linked to gene silencing. These MARs may be part of a broader group of elements that influence cell behavior in different ways.

Matrix association region/scaffold attachment region (MAR/SAR) sequences function as binding sites for DNA loop structures to nuclear scaffold/matrix proteins

(Broeker et al., 1996). These sequences have DNA unwinding properties, enabling them to play a key role in facilitating the access of protein factors involved in transcription, replication, chromosome condensation, and apoptosis (Bode et al., 2000; Göhring et al., 1997). Additionally, their unwinding properties make MAR/SAR sequences fragile and particularly susceptible to DNA breakage (Bode et al., 2000; Legault et al., 1997).

Chromosomal breakage is an early event in both apoptotic DNA fragmentation and chromosome rearrangement. Previous studies have shown that these breaks often occur within specific regions containing particular chromatin structural elements, such as MAR/SAR sequences (Strick et al., 2006b; Strissel et al., 2000). During the early stages of apoptosis, the DNA loop structure is cleaved at its base, where it is anchored to the nuclear matrix or scaffold via the MAR/SAR sequence (Berezney et al., 1996; Oberhammer et al., 1993). The key apoptotic nuclease, CAD, has been found to associate with the nuclear matrix in cells undergoing apoptosis (Lechardeur et al., 2004). MAR/SAR sequences are not only involved in apoptosis but are also associated with chromosomal rearrangements and have been implicated in illegitimate recombination events (Broeker et al., 1996). Notably, the breakpoint cluster region of the Mixed Lineage Leukaemia (*MLL*) gene, frequently involved in translocations in leukaemia patients, contains two MAR/SAR sequences (Broeker et al., 1996). Similarly, MAR/SAR sequences have been identified in the *ABL* gene (Tan et al., 2018a) and in *AF9* gene (Strissel et al., 2000), both of which are involved in chromosomal translocations (Tan et al., 2018a). MAR/SAR sequences are also believed to play a crucial role in determining the location of chromosome breaks mediated by oxidative stress-induced apoptosis, with CAD being the key nuclease involved. These chromosomal breaks may ultimately lead to chromosome aberrations in nasopharyngeal epithelial cells(Tan et al., 2018b).

1.8 The *AF9* Gene

The *AF9* gene was chosen for this study because it is located in the 9p21-22 region, which is frequently deleted in NPC, and is commonly involved in translocations in leukaemia (Alonso et al., 2008; Shao et al., 2000). *AF9* is one of the most frequent fusion partners of the *MLL* gene at 11q23. The fusion of these genes, resulting from the t(9;11)(p22;q23) translocation, is primarily observed in de novo acute myelogenous leukaemia (AML). The *MLL-AF9* fusion gene is also found, though less commonly, in acute lymphocytic leukaemia (ALL), myelodysplastic syndrome (MDS), and therapy-related leukaemia (t-AML) (Swansbury et al., 1998; Strissel et al., 2000). Two MAR/SARs have been experimentally identified: the first (SAR1) is located in intron 4, and the second (SAR2) spans parts of exons 5-7 (Strissel et al., 2000).

Multiple studies have demonstrated that the *MLL-AF9* fusion gene is critical in stem cell development and leukemogenesis. To investigate its role, a homologous recombination knock-in approach was used to generate the *MLL-AF9* fusion gene in mouse embryonic stem cells. Introducing the *MLL-AF9* fusion gene into the mouse model caused an initial, non-malignant expansion of myeloid precursors, followed by an accumulation of malignant cells in bone marrow and other tissues. These findings indicate that the *MLL-AF9* oncogene disrupts normal haematopoietic cell proliferation and the survival of myeloid progenitors (Corral et al., 1996; Dobson et al., 1999; Pession et al., 2003).

1.9 Rationale of Study

Chromosome rearrangements are commonly observed in non-random locations in the genomes of nasopharyngeal carcinoma (NPC) patients. However, the mechanism leading to these rearrangements is largely unknown. Epstein-Barr virus (EBV) has long been known

to be associated with NPC, but its direct role in the pathogenesis of NPC is still under investigation. It is known that repeated EBV activation and the expression of some latency genes result in genome instability, possibly via ROS production. Additionally, ROS-induced apoptosis results in chromosome breaks within regions of the chromosome that are commonly rearranged in NPC. Therefore, the current study aims to investigate the link between latency gene expression, ROS production, and chromosome breaks within regions that are commonly rearranged in NPC. This study aims to shed light on the pathogenesis of NPC in relation to EBV.

1.10 Hypothesis

We hypothesise that during repeated EBV activation and expression of latency genes, ROS is produced, damaging the genome and leading to apoptosis. During attempted repair, errors occur, resulting in cells bearing rearranged chromosomes surviving apoptosis and contributing to the pathogenesis of NPC.

1.11 Objectives

I. To evaluate the effect of *LMP1* gene expression on oxidative stress generation.

II. To assess the induction of apoptosis in *LMP1* expressing cells.

III. To analyse *LMP1* expression-induced chromosome breaks within the *AF9* gene.

IV. To compare Scaffold Attached Regions (SAR) and non-SAR in order to investigate the role of MAR/SAR sequence in mediating *LMP1*-expression induced chromosome breaks.

CHAPTER 2

MATERIALS AND METHODS

2.1 Plasmid and Bacteria Culture

To prepare selective agar, four grams of Luria broth (LB) agar was mixed with 100 ml of distilled water to prepare agar plates. The agar mixture was autoclaved for 1 hour and 30 minutes. After autoclave, agar was cooled down until 55˚C before added with 100 µl of 50 mg/ml of ampicillin to a final concentration of 50 µg/ml. An agar mixture was poured into the plates and waited for 30 minutes to ensure the agar has hardened. Plates were sealed to avoid contamination.

Luria broth (LB) is a nutrient-rich media used to culture vector bacteria in the current study. The addition of agar to LB results in the formation of a gel that bacteria can grow on, as they are unable to digest the agar but can gather nutrition from the LB within. The addition of an antibiotic, ampicillin allows selection of only bacteria containing the plasmid which carries the ampicillin resistance gene.

Bacteria strains harbouring plasmids pTracer, pTracer-*LMP1*, pcDNA or pcDNA-*LMP1* were streaked on agar plate and incubated overnight at 37˚C. After incubation, selected plasmid bacteria colony was transferred into 100 ml of LB broth with 50 µg/ml ampicillin for culture at 37˚C overnight. Luria broth (LB) also referred to as Lysogeny Broth, this medium is used for general bacterial cultivation and growth. It supports the growth of a wide variety of bacteria and is commonly used to grow *Escherichia coli* for molecular biology applications.

2.2 Extraction of Plasmid

Extraction of plasmid was performed using Purelink ™ Hipure Plasmid Midiprep Kit (Invitrogen, Thermo Fisher Scientific) following the manufacturer's protocol. Briefly, 10 ml of Equilibration Buffer (EQI) was added into the HiPure Midi Column that was placed inside a 50 ml tube. The solution was drained by gravity.

From the 100 ml of overnight culture, 50 ml was centrifuged at 700 x g for 10 minutes. The supernatant was removed, 10 ml of Resuspension Buffer (R3) containing RNase A was added to resuspend the pellet until homogenous. Then, 4 ml of Lysis Buffer (L7) was added and mixed gently by inverting the tube for 5 times. The mixture was incubated at room temperature for 5 minutes.

Subsequently, 4 ml of Precipitation Buffer (N3) was added and mixed immediately by inverting the tube. The mixture was centrifuged at 12,000 x g for 10 minutes at room temperature. After centrifugation, the supernatant was loaded into HiPure Midi Column that has been equilibrated with EQ1 previously and the solution was allowed to drain by gravity flow. The column was then washed with 10 ml of Wash Buffer (W8) twice.

After all the solution has drained, new sterile 50 ml tube was placed under the column and 5 ml of Elution Buffer was added into the column. After the solution has drained, the solution was transferred into five 2 ml centrifuge tubes and 0.7 ml isopropanol was added to each tube. Solution was well mixed and centrifuged at 12,000 x g for 30 minutes at 4˚C. After centrifugation, supernatant was discarded and 0.6 ml 70% ethanol was added to the pellet. The mixture was then centrifuged at 12,000 x g for 5 minutes at 4˚C. The supernatant was removed after centrifugation. The pellet was air dried for 10 minutes and subsequently resuspend with 40 µl of TE buffer. For long term storage the extracted plasmid was kept at

-20˚C. Plasmid DNA for both pTracer-*LMP1* and pcDNA-*LMP1* were confirmed by DNA sequencing before proceeding to transfection.

2.3 Plasmid DNA

The plasmid used in this study were: pTracerTM-EF/V5-His B (pTracer) vector, ptracer-*LMP1* (*LMP1* expression plasmid), pcDNA 3.1/V5-His-TOPO (pcDNA) vector and pcDNA-*LMP1* (*LMP1* expression plasmid). pTracer and pTracer-*LMP1* were used because they contained the green fluorescence protein (GFP) gene. Expression of GFP allows us to estimate transfection efficiency during optimisation. pcDNA-B95 is a *LMP1* expression plasmid obtained from Professor Sam Choon Kook from Universiti Malaya. pcDNA -B95 contains the *LMP1* gene isolated from the EBV genome positive marmoset cell line B95. This pcDNA-B95 was renamed pcDNA-*LMP1* in this study. The *LMP1* gene was previously subcloned into pTracer and named pTracer-*LMP1* (Yee, 2010). A map of pTracer-*LMP1* is as shown in Figure 2.1, while the map of pcDNA-*LMP1* is as shown in Figure 2.2.

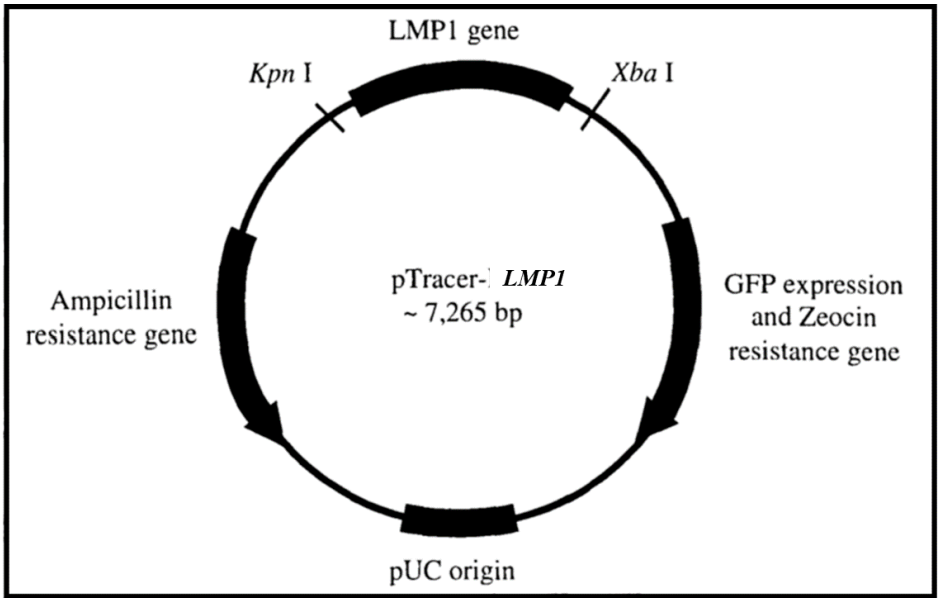

Figure 2.1: Diagram showing a plasmid map of the pTracer-*LMP1* (approximately 7,265 bp). pTracer vector (approximately 5867 bp) was digested with *Kpn* I and *Xba* I and ligated with the *LMP1* gene (approximately 1398 bp) excised from pcDNA-*LMP1*.

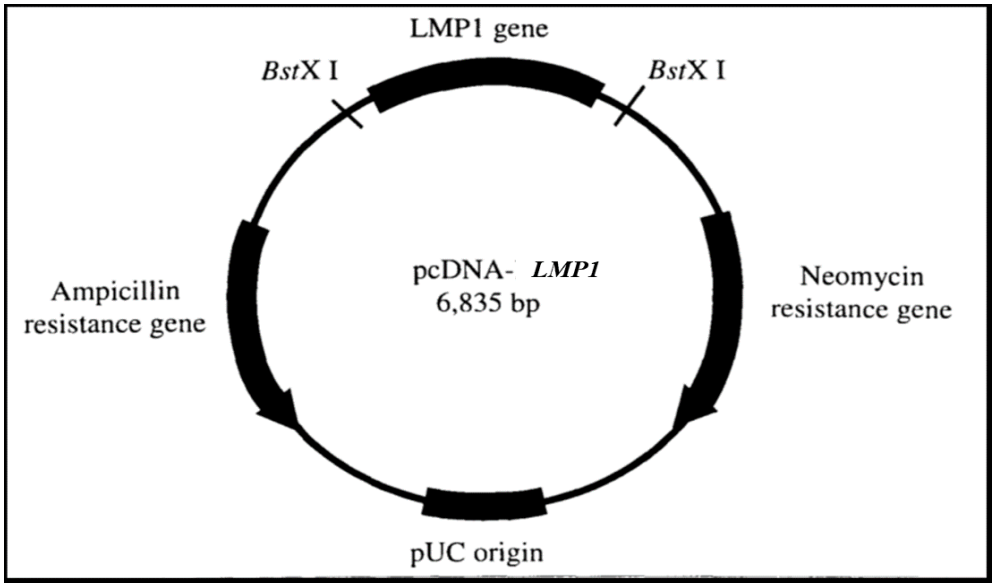

Figure 2.2: Diagram showing a plasmid map of the pcDNA-*LMP1* (approximately 6835 bp). pcDNA-*LMP1* was digested with *Bst*X I in order to excise out the *LMP1* gene (approximately 1354 bp) and self-ligated to obtain the pcDNA vector (approximately 5481 bp).

2.4 Cell Line

Cell line NP69 is an immortalized nasopharyngeal epithelial cell line established by transfecting cells with the SV40 large T oncogene (Tsao et al., 2002). This cell line, provided by the late Prof. Tsao Sai Wah and Prof. Lo Kwok Wai from the University of Hong Kong and The Chinese University of Hong Kong respectively, retains the normal characteristics of nasopharyngeal epithelial cells and is non-tumorigenic. In a study by Huang et al. (2015), NP69 was used as a model to investigate the mechanisms involved in the tumorigenesis of nasopharyngeal carcinoma (NPC).(Huang et al., 2015).

NP69 cells were maintained in Keratinocyte-SFM (KSFM) medium, which was supplemented with recombinant Epidermal Growth Factor (rEGF) (4-5 ng/ml), Bovine Pituitary Extract (BPE) (40-50 µg/ml), 2% (v/v) heat-inactivated fetal bovine serum, and 1% penicillin-streptomycin (GIBCO, Invitrogen, USA). The cells were incubated in a humidified 5% CO2 incubator at 37˚C (Forma Direct Heat CO2 Incubator, Model 311, Thermo Scientific, USA). Table 2.1 details the KSFM medium used in this study. Cells were cultured in T-25 flasks (Corning, USA) and passaged every two days upon reaching 70% confluency. For passaging, the culture medium was aspirated, and the cells were trypsinised with 0.7 ml of TrypLE Express (GIBCO, Invitrogen, USA). After a 4-minute incubation at 37˚C in the CO2 incubator, the flask was tapped to detach the cells. The cells were then resuspended in 4 ml of culture medium, and 2.5 x 10^4 cells were transferred to a new T-25 flask containing 5 ml of fresh culture medium.

Table 2.1: Keratinocyte-SFM (KSFM) Medium used in this study

Medium	Contents
Culture	Keratinocyte-SFM (KSFM) medium supplemented with recombinant Epidermal Growth Factor (rEGF) (4-5 ng/ml), Bovine Pituitary Extract (BPE) (40-50 ug/ml), 2% (v/v) heat-inactivated foetal bovine serum and 1% penicillin-streptomycin
Seeding	Keratinocyte-SFM (KSFM) medium supplemented with recombinant Epidermal Growth Factor (rEGF) (4-5 ng/ml), Bovine Pituitary Extract (BPE) (40-50 ug/ml) and 20% (v/v) heat-inactivated foetal bovine serum
Transfection	Keratinocyte-SFM (KSFM) medium supplemented with recombinant Epidermal Growth Factor (rEGF) (4-5 ng/ml) and Bovine Pituitary Extract (BPE) (40-50 ug/ml).

2.5 Transient Transfection

NP69 cells were cultured in Keratinocyte-SFM medium (GIBCO, Invitrogen, USA) supplemented with human recombinant Epidermal Growth Factor 1-53 (EGF 1-53), Bovine Pituitary Extract (BPE), 2% FBS (GIBCO, Invitrogen, USA), and 1% Penicillin/Streptomycin (100 U/mL Penicillin and 100 µg/mL Streptomycin) (Fisher Bioreagents, USA). The cells were maintained at 70% confluency in T-25 flasks (Corning, USA). To detach the cells, 700 µL of TrypLE Express (GIBCO, Invitrogen, USA) was used for each T-25 flask. After detachment, cells were resuspended in 5 mL of culture medium and thoroughly mixed. The cells were then seeded into multi-well plates at approximately 70-85% confluency, with seeding details provided in Table 2.2. The cells were incubated overnight before transfection, and the seeding medium was replaced with transfection medium.

For transfection, Lipofectamine 3000 reagent (Thermo Fisher Scientific, USA) was utilized. The optimal amount of Lipofectamine 3000 reagent was determined and diluted in Opti-MEM medium. A DNA master mix was prepared by diluting DNA in Opti-MEM medium, followed by the addition of P3000 reagent and thorough mixing. The diluted DNA was then combined with the diluted Lipofectamine 3000 reagent in a 1:1 ratio and incubated at room temperature for 10 minutes to form the DNA-lipid complex. This complex was added to the cells in transfection medium, and after 4 hours of transfection, the transfection medium was replaced with fresh culture medium.

Following transient transfection with pTracer-LMP1, cells were observed under a fluorescence microscope to confirm successful transfection. Transient transfection allows for temporary gene expression without integrating the genes into the host genome, with expression lasting for a limited period (Khan, 2013). Based on GFP fluorescence intensity measured at intervals from 0 to 72 hours, optimal time points for LMP1 Western blotting were identified. Cells were collected at 24-, 48-, and 72-hours post-transfection, and Western blot analysis was performed to determine the time point with the highest LMP1 protein expression.

Table 2.2: Seeding amounts of NP69 cells in different culture vessel to achieve a confluency of 70-85%.

Type of culture vessel	Amount of cell-seeded at confluency of 70-85%	Final Volume of Seeding medium (ml) present in each well/dish
96 well plate	2.0×10^3 - 3.0×10^3 cells	0.1
24 well plate	4.0×10^4 - 5.0×10^4 cells	0.5
6 well plate	8.2×10^4 - 9.5×10^4 cells	2.0
60 mm dish	1.8×10^5 - 2.5×10^5 cells	4.0

2.6 Protein Extraction for Western Blotting

2.6.1 Cell Lysis

NP69 cells were cultured in a 60 mm dish and subjected to transfection. After the transfection process, the culture medium was aspirated, and the cells were rinsed with cold 1X PBS. The cells were then lysed using 0.5 ml of a triple detergent buffer containing 50 mM Tris.Cl (pH 8), 150 mM sodium chloride, 0.02% sodium azide, 0.1% sodium dodecyl sulfate (SDS), 1% Nonidet P-40 (NP-40), and 0.5% sodium deoxycholate. Protease Inhibitor Cocktail (5 µl) and phenylmethylsulfonyl fluoride (PMSF) (5 µg) were freshly added to the lysis buffer. The cells were lysed on ice for 5 minutes, and the lysate was collected into a 1.5 ml microcentrifuge tube by scraping with an inverted sterile 1 ml pipette tip.

2.6.2 Protein Quantification by Bradford Assay

The Bradford Assay is widely used to quantify proteins in the range of 1 to 10 µg by comparing the binding of an unknown protein to Coomassie Brilliant Blue dye with that of a standard protein. To prepare the dye, 10 mg of Coomassie Brilliant Blue G-250 (Sigma) was dissolved in 5 ml of 95% ethanol. Then, 10 ml of 85% phosphoric acid was added, and the mixture was diluted with distilled water to a final volume of 100 ml. Before use, this dye solution was diluted fivefold

2.6.2.1 Establishment of Protein Standard Curve

A standard curve was generated using 2.5 µg, 5 µg, 7.5 µg, and 10 µg of γ-globulin, which corresponded to 5, 10, 15, and 20 µl of a 0.5 mg/ml γ-globulin solution, diluted with 0.15 M NaCl to a final volume of 100 µl. Each concentration was prepared in duplicate. A blank was prepared using 100 µl of 0.15 M NaCl. To each tube containing 100 µl of the standard protein solution, 1 ml of Coomassie Brilliant Blue solution was added, and the

mixture was vortexed. The solutions were then incubated at room temperature for 5 minutes. Absorbance was measured using a microcuvette with a 1 cm path length at 595 nm (A595). The standard curve was plotted by graphing the absorbance values (A595) against the concentrations of γ-globulin.

2.6.2.2 Measuring Concentration of sample protein

Three microliters of crude lysate were added to 97 µl of 0.15 M NaCl as a diluent. One milliliter of Coomassie Brilliant Blue was then added to each diluted lysate, followed by a 5-minute incubation at room temperature. The absorbance at 595 nm (A595) was measured and recorded. The concentration of each lysate was determined using the standard curve.

2.7 Western Blot

2.7.1 Preparation of Sodium Dodecyl-Sulfate Polyacrylamide Gel Electrophoresis (SDS-PAGE)

A denaturing discontinuous gel was prepared using the Laemmli method. Spacer glass plates with 1.55 mm integrated spacers and short plates were assembled in a casting frame, then secured in a casting stand to create a tight seal with the rubber gasket. The resolving gel mixture was prepared, consisting of 10% acrylamide/bis-acrylamide (Sigma), 0.38 M Tris (pH 8.8), 0.1% SDS, and 1 mM ethylenediaminetetraacetic acid (EDTA) in a final volume of 5 ml. Seven microliters of TEMED (Amresco, USA) and ammonium persulfate (APS) to a final concentration of 0.1% were added last. Once mixed, the gel solution was immediately pipetted into the gap between the glass plates, filling it to a level 1 cm below the edge of the 10-well comb. To prevent oxidation, which could inhibit polymerization, and to create a flat interface, isopropanol was gently layered on top of the

resolving gel. The gel was allowed to solidify for 20 minutes, after which the isopropanol was removed using filter paper.

Next, the stacking gel was prepared, consisting of 4% acrylamide/bis-acrylamide (Sigma), 0.38 M Tris (pH 6.8), 0.1% SDS (Amresco, USA), and 1 mM EDTA in a final volume of 2.5 ml. To initiate polymerization, 3.5 µl of TEMED and APS to a final concentration of 0.1% were added last. The stacking gel mixture was then mixed and immediately poured onto the solidified resolving gel between the glass plates. The comb was carefully inserted to avoid bubble formation. The preparation details for the SDS-PAGE mixture are provided in Table 2.3.

Table 2.3: Preparation of 10% SDS-PAGE Mixture.

Reagent	Resolving gel (10%)	Stacking gel (4%)
30% acrylamide-bis-acrylamide (Sigma)	1.67 ml	0.335 ml
1.5 M Tris pH 8.8	1.25 ml	-
0.5 M Tris pH 6.8	-	0.625 ml
20% SDS (Amresco)	25 ul	25 ul
dH_2O	2.025 ml	1.5 ml
100 mM EDTA	50 ul	25 ul
10% APS	50 ul	25 ul
TEMED	3.5 ul	1.75 ul
Total	~5 ml	~2.5 ml

2.7.2 Sodium Dodecyl-Sulfate Polyacrylamide Gel Electrophoresis (SDS-PAGE)

After determining the protein concentration using the Bradford Assay, 30 µg of each sample protein was loaded into the gel lanes. The required sample volume was calculated from the standard curve and transferred to a clean 1.5 ml microcentrifuge tube. An equal volume of 2X SDS Sample Buffer, containing 1X Tris-Cl SDS (pH 6.8), 20% glycerol, 4%

SDS, 0.2% 2-mercaptoethanol (ME), 0.001% bromophenol blue, and dH2O, was added to each sample and mixed thoroughly. The samples were then boiled for 10 minutes and centrifuged at 20,000x g for 5 minutes at room temperature. The supernatant was then loaded into the designated gel lanes. Prestained PageRuler (Fermentas) was used as a protein marker and loaded into the first and last lanes of each gel. The gel cassettes were assembled and placed into the clamping frame, and electrophoresis was conducted at 120 V for approximately one and a half hours using 1X Running Buffer, which contained 25 mM Tris (base) (J.T. Baker, UK), 192 mM Glycine (Amresco, USA), and 0.01% SDS.

2.7.3 Transferring of Protein onto Polyvinylidene Fluoride (PVDF) Membrane

2.7.3.1 Equilibration of Gel and PVDF Membrane

After electrophoresis, the gel was removed from the glass plates and briefly rinsed with an adequate volume of dH_2O before proceeding with blotting. The gel was equilibrated in blotting buffer [25 mM Tris (base) (J.T. Baker), 192 mM Glycine (Amresco), 20% Methanol, and dH_2O] for 20 minutes. During this time, the PVDF membrane, initially opaque, was activated by gently mixing it in methanol with a rocking motion for about 10 seconds until it turned translucent. The membrane was then washed with dH_2O for 2 minutes and immersed in blotting buffer for 5 minutes. Ice blocks made from blotting buffer were prepared in advance and stored in a -80°C freezer.

2.7.3.2 Protein Transfer

Protein transfer was carried out under cold conditions. The blotting buffer was pre-chilled, and the blotting tank was placed in a container surrounded by ice. A magnetic stir bar was added to the blotting tank to ensure even circulation of the cold, pre-chilled blotting buffer. The blotting sandwich was assembled by placing a fibre pad on the black side of the gel holder cassette, followed by two pieces of 3 mm Whatman filter papers, the protein gel,

the PVDF membrane, another two pieces of 3 mm Whatman filter papers, and a final fibre pad. A clean glass rod was used to roll over the sandwich, removing any trapped air bubbles. The transparent side of the gel holder cassette was then closed and tightened. The blotting sandwich was placed in the lower buffer chamber with the gel facing the cathode and the PVDF membrane closer to the anode. This setup ensured that the negatively charged, denatured proteins would migrate toward the positive pole (anode) and be transferred onto the PVDF membrane. The blotting process was conducted at 100 V for 1 hour.

2.7.3.3 Protein Gel Staining and Blocking

After blotting, the sandwich was disassembled. The gel was briefly rinsed with dH2O, while the PVDF membrane was washed with dH2O for 2 minutes. The gel was then stained using Gel Code Blue Stain (Pierce, USA). Meanwhile, the PVDF membrane, or blot, was blocked overnight at 4°C with a blocking solution containing 5% skimmed milk in 1X Phosphate Buffer Saline-Tween 20 (0.05%) (PBST).

2.7.3.4 Immunodetection and Image Development

Blot was incubated in blocking buffer containing monoclonal primary antibody: anti-*LMP1* antibody (clone S12, BD PharMingen), diluted 1: 3,000 in blocking buffer for 1 hour at room temperature with mild rocking. Following that, three times of washing with PBS, 0.05% Tween 20 was carried out at room temperature with mild rocking (10 minutes for the first washing, 5 minutes each for the second and third washing). The blot was then incubated with anti-mouse horseradish peroxidase-conjugated secondary antibody (Pierce), diluted 1: 67,000 in blocking buffer, for 1 hour at room temperature. Subsequently, washing was carried out at room temperature with mild rocking (10 min for the first washing, 5 min each for the second and third washing). The blot was then placed onto cling wrap with the protein side facing up. Thereafter, 2 ml of the chemiluminescent substrate (a freshly mixed

solution of 1 ml each of SuperSignal™ West Pico Luminol/Enhancer and SuperSignal™ West Pico Stable Peroxide) was added to the blot and incubated for 5 minutes. The excess substrate was removed by gently draining it off. The blot was then covered with cling wrap and imaged using the Syngene ChemiGenius system (USA).

2.8 Transfection for Oxidative Stress and Caspase Activity Measurement

After the third passage, $2x10^3$ NP69 cells were seeded into a 96-well plate containing 100 µl of seeding medium and incubated overnight at 37°C. The next day, the seeding medium was replaced with 100 µl of transfection medium. In a 0.2 ml centrifuge tube, 0.2 µg of plasmid DNA (10 µl), 0.2 µl of p3000 reagent, and 0.3 µl of 3000 reagent were combined and incubated at room temperature for 10 minutes. After incubation, the mixture was added to the 96-well plate containing NP69 cells in transfection medium and incubated for 4 hours. After the incubation period, the transfection medium was replaced with 100 µl of culture medium.

2.8.1 Measurement of Oxidative Stress in *LMP1*-Expression Cells

In this study, two types of plasmid DNA comprising pTracer and pcDNA were used. The pTracer set comprised of pTracer-*LMP1* and pTracer vector control as negative control. While the pcDNA set consisted of pcDNA-*LMP1* and pcDNA vector control as negative control. The cells at 70% confluency were trypsinised in the T-25 flask with TrypLE express. Then cells were incubated at 37˚C for 4 minutes and 4 ml of culture medium was subsequently added. Cells were seeded at a density of $2x10^3$ cells per well in 100 µl of seeding medium in a black 96-well plate with a clear bottom (Corning,USA) and allowed to adhere overnight.

For transfection with pTracer-*LMP1* and the vector control pTracer, CellROX was used to examine the effect of *LMP1* expression on reactive oxygen species (ROS) production. pTracer plasmids containing a fluorescent reporter gene, enabling the visualisation and confirmation of successful transfection. When combined with CellROX, this dual approach offers comprehensive insights into both transfection efficiency and oxidative stress levels within the same cell population. CellROX provides a sensitive and specific measure of ROS levels, allowing for the assessment of oxidative stress in the transfected cells.

Transfection with pcDNA and pcDNA-*LMP1* was carried out for 4 hours, after which the transfection medium was replaced with culture medium. Forty-eight hours post-transfection, ROS production was evaluated using 2,7'-Dichlorofluorescin diacetate (DCFH-DA) (Sigma). DCFH-DA is nonfluorescent until hydrolyzed by esterases and oxidized by ROS within the cells. For the oxidative stress assay, the culture medium was removed, and the cells were incubated with 100 µl of transfection medium containing 10 µM DCFH-DA at 37°C for 30 minutes. The cells were then washed three times with transfection medium, and fluorescence intensity (excitation at 485 nm, emission at 530 nm) was measured using a multimode microplate reader (SpectraMax® iD3). Each transfection condition was replicated 20 times.

Due to the overlap in excitation and emission wavelengths between GFP and the DCFH-DA method, this method could not be used to detect ROS generation in pTracer transfectants. Instead, the CellROX Deep Red assay (ThermoFisher Scientific, USA) was used to measure ROS production in NP69 cells transfected with pTracer and pTracer-LMP1. Forty-eight hours post-transfection, the culture medium was removed, and the cells were

incubated with transfection medium containing CellROX reagent at a final concentration of 5 µM at 37°C for 30 minutes. The medium was then removed, and the cells were washed three times with 1X PBS. Fluorescence intensity was measured using a multimode microplate reader (Molecular Devices, SpectraMax® iD3) with excitation at 640 nm and emission at 665 nm. Each transfection condition was replicated 20 times.

2.8.2 Measurement of Caspases Activity in *LMP1*-Expression Cells

To examine if *LMP1* expression induces apoptosis, the activity of caspases was determined in cells transfected with pTracer-*LMP1*, pcDNA-*LMP1*, as well as the vector control. Transfection of NP69 cells with pTracer-*LMP1*, pTracer vector control, pcDNA-*LMP1*, and pcDNA vector control was performed in a white-walled, clear-bottom 96-well plate (Corning). Four hours post-transfection, the transfection medium was replaced with culture medium. Forty-eight hours later, the culture medium was removed, and the cells were washed once with 1X PBS buffer at room temperature. The Caspase-Glo® 3/7 Assay kit was prepared according to the manufacturer's instructions. The Caspase-Glo® 3/7 Buffer and lyophilized Caspase-Glo® 3/7 substrate were first brought to room temperature. The buffer was then added to the amber bottle containing the Caspase-Glo® 3/7 substrate, and the solution was gently inverted until the substrate was fully dissolved, forming the Caspase-Glo® 3/7 reagent. A mixture of 100 µl of 1X PBS buffer and 100 µl of the Caspase-Glo® 3/7 reagent was added to each well. The plate was incubated at room temperature for 30 minutes, with gentle shaking on a horizontal shaker for 30 seconds before and after incubation. Luminescence was measured using a multimode microplate reader (SpectraMax® iD3).

2.9 Identification of Chromosomal Breaks within the *AF9* Gene using Inverse PCR (IPCR).

Chromosomal breakage serves as an initial event in both apoptotic DNA fragmentation and chromosome rearrangement. Earlier studies have indicated that these breaks commonly occur in specific regions containing distinct chromatin structural elements, such as the matrix attachment region/scaffold attachment region (MAR/SAR) (Strick et al., 2006b; Strissel et al., 2000). Strissel et al. identified two MAR/SAR regions within the AF9 gene, which were designated as SAR1 and SAR2 (Strissel et al., 2000). Figure 1 illustrates the distribution of the predicted MAR/SAR sites within the AF9 gene.

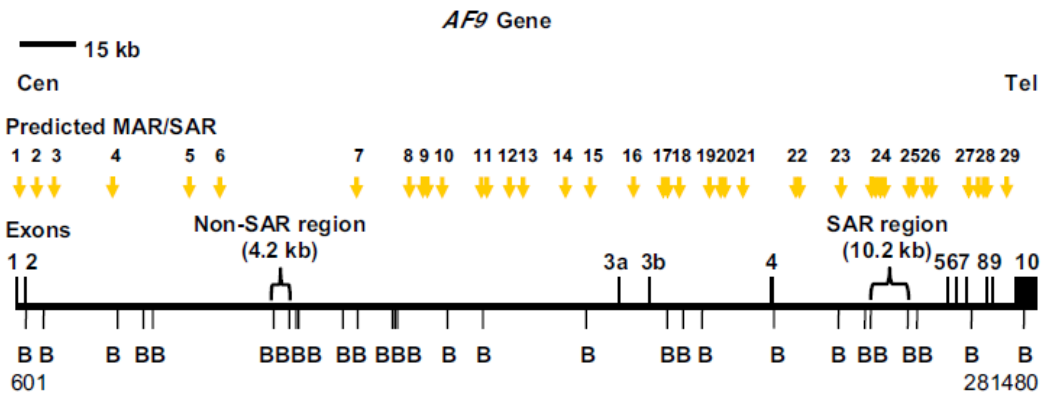

Figure 2.3: Distribution of potential MAR/SAR sites predicted in the AF9 gene. The AF9 genomic map from nucleotide positions 601–281,480 is depicted above [EMBL:ENSG00000171843] (Tan & Sim, 2019) illustrated above [EMBL:ENSG00000171843].

2.9.1 Genomic DNA Extraction

Genomic DNA (gDNA) extraction was performed using Wizard® Genomic DNA Purification Kit by (Promega). After 4 hours of transfection and 48 hours of post-transfection, the cells were trypsinised. The cells were then harvested and transferred into a 1.5 ml microcentrifuge tube. The cells were centrifuged at 13,000–16,000 × g for 10 seconds. After that, supernatant was removed leaving behind the cell pellet. The pellet was washed

using 200 μl of PBS and vortexed vigorously to resuspend the cells before centrifuging the tube at 13,000–16,000 × g for 10 seconds. After that, 600 μl of Nuclei Lysis Solution was added to the suspension and pipetted up and down to lyse the cells. 17.5 μl of 20 mg/ml Proteinase K and 3 μl of RNase Solution were added to the nuclear lysate and the sample was mixed by inverting the tube 2–5 times. The solution was incubated for 15–30 minutes at 37°C. Then the solution was incubated at room temperature for 5 minutes.

Two hundred microlitre of Protein Precipitation Solution was added into the solution and vortexed vigorously at high speed for 20 seconds. Then, the solution was chilled on ice for 5 minutes before centrifuged for 4 minutes at 13,000–16,000 × g. The precipitated protein formed a tight white pellet, the supernatant containing the DNA (leaving the protein pellet behind) was transferred to a clean 1.5 ml microcentrifuge tube containing 600 μl of room temperature isopropanol. Then, solution was gently mixed by inversion until the white thread-like strands of DNA formed a visible mass and the solution was centrifuged for 1 minute at 13,000–16,000 × g at room temperature. The DNA was visible as a small white pellet. After aspirating the supernatant, 600 μl of room temperature 70% ethanol was added into the tube, and the tube was gently inverted several times to wash the DNA pellet, followed by centrifuging the tube for 1 minute at 13,000–16,000 × g at room temperature. The supernatant was removed, and the tube was air dried for 10–15 minutes. One hundred microliter of DNA Rehydration Solution was added to the tube and the DNA was rehydrated by incubating at 65°C for 1 hour. The solution was periodically mixed by gently tapping the tube. Lastly, the DNA was stored at 2–8°C before use.

2.9.2 First Restriction Enzyme (RE) Digestion

The genomic DNA (gDNA) was extracted and prepared for nested Inverse Polymerase Chain Reaction (IPCR) based on the Bowtell Lab Manual (1998) from the Peter MacCallum Cancer Institute, Australia, with slight modifications. The gDNA was digested using *Bam*H I to produce shorter, specific DNA fragments. The reaction included 100 units of *Bam*H I, 1X NEBuffer 3, and 1X bovine serum albumin (BSA) (NEB, USA) in a total volume of 60 µl, and was incubated overnight (16 hours) at 37°C. *Bam*H I digestion creates 4-base pair staggered 5' overhangs (GATC), whereas apoptotic nucleases like CAD typically generate blunt ends (Widlak & Garrard, 2005). As a result, the intact targeted DNA fragment has *Bam*H I sites with staggered overhangs on both ends. In contrast, the cleaved targeted DNA fragment has one end with a *Bam*H I staggered overhang and the other with a blunt end from the apoptotic nuclease. To prepare for cyclization, the digested DNA underwent a Klenow fill-in to create blunt-ended fragments.

2.9.3 Klenow Fill-In and Cyclisation

The Klenow fill-in reaction was performed using 2 µg of DNA, 2 units of DNA Polymerase I Large (Klenow) Fragment, 33 µM dNTP mix, and 1X NEBuffer 2 in a total volume of 50 µl at 25°C for 15 minutes. The enzyme was subsequently heat-inactivated at 75°C for 20 minutes. Cyclisation was then carried out with 2000 units of T4 DNA ligase and 1X T4 DNA ligase buffer in a final volume of 500 µl, incubated at 16°C for 16 hours. The large reaction volume was chosen to promote intramolecular ligation while reducing intermolecular ligation. The ligase was heat-inactivated at 65°C for 10 minutes.

2.9.4 Ethanol Precipitation

Ethanol precipitation was performed following the method outlined by Moore and Dowhan (2003), with slight modifications (Moore & Dowhan, 2002). All centrifugation steps were conducted at room temperature. To 250 µl of the DNA solution from section 2.9.3, 250 µl of sodium acetate (pH 5.2) and 625 µl of ice-cold absolute ethanol were added, and the tube was inverted several times to mix. The mixture was then incubated at -80°C for 15 minutes, followed by centrifugation at 20,800 x g for 15 minutes at room temperature. After discarding the supernatant, the pellet was washed with 1 ml of 70% ethanol and centrifuged again at 20,800 x g for 10 minutes at room temperature. The supernatant was discarded, the pellet was air-dried for 5 minutes, and then dissolved in 60 µl of TE (pH 8).

2.9.5 Second and Third Restriction Enzyme (RE) Digestion for SAR

As illustrated in Figure 2.3, the second restriction enzyme (RE2), *Kpn* I, was used to linearize the cyclized DNA. Shorter PCR products indicate cleaved fragments, while longer PCR products indicate intact fragments. The third restriction enzyme (RE3), *Nde* I, was employed to linearize the cyclized DNA and prevent amplification of the intact fragment. PCR could be performed for the cleaved fragment, but not for the intact fragment as shown in Figure 2.3. The second RE reaction was conducted with 20 units of *Kpn* I, 1X NEBuffer 1, and 1X BSA in a total volume of 40 µl. The third RE reaction involved 20 units of *Nde* I and 1X NEBuffer 4 in a total volume of 40 µl. Both reactions were incubated at 37°C for 16 hours.

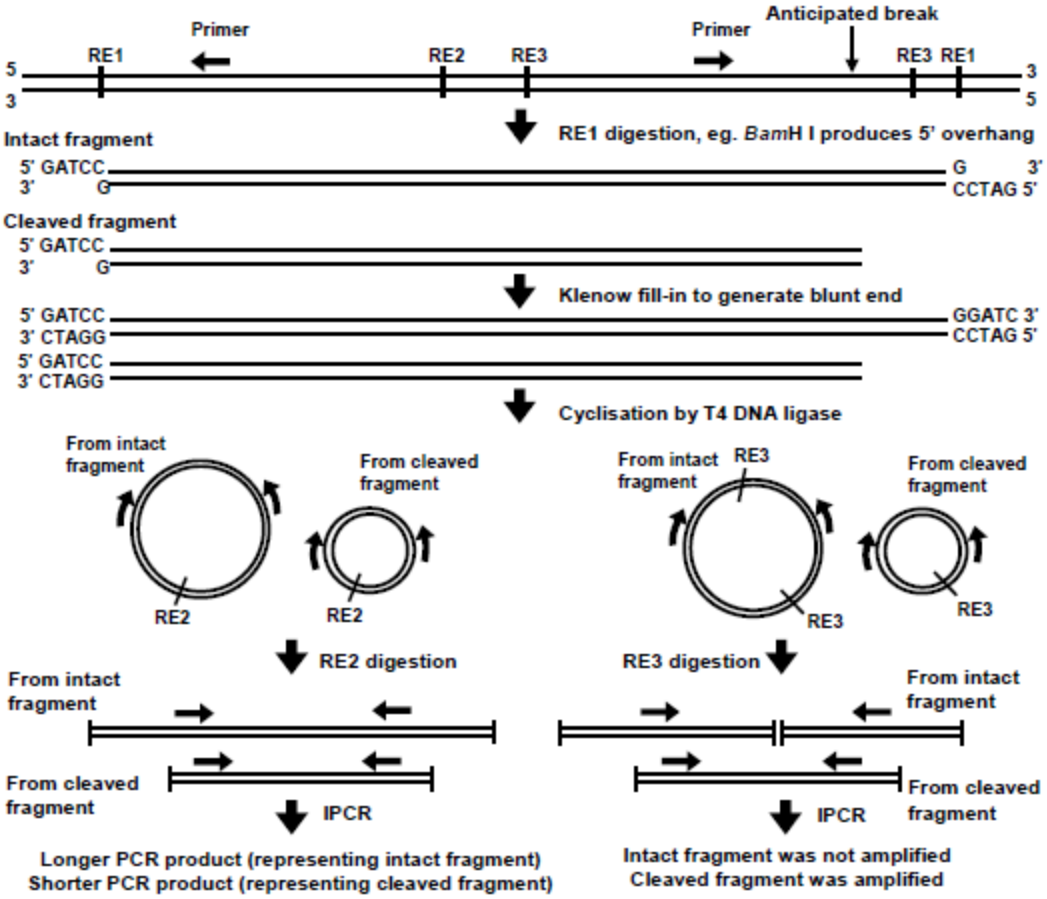

Figure 2.4: Manipulation of gDNA in preparation for nested IPCR. Initially, the extracted gDNA was digested with RE1, generating 5' overhangs, while apoptotic nucleases like CAD predominantly produced blunt ends. To enable cyclisation, Klenow fill-in was performed to convert these overhangs into blunt ends. After cyclisation with T4 DNA ligase, the DNA was digested with either RE2 or RE3. RE2 was used to linearize the cyclized fragments, whereas RE3 served to linearize the cyclized fragments and simultaneously block the amplification of intact fragments, ensuring that only cleaved fragments were amplified.

2.9.6 Removal of Excessive Unincorporated Nucleotide

Process of removal of excessive unincorporated nucleotide was conducted by following QIAquick Nucleotide Removal Kit protocol. First, 400 µl of Buffer PN was added to the 40 µl of second or third RE reaction and mixed. Then, the mixture was added into the QIAquick spin column which was placed into a 2 ml collection tube. The mixture was centrifuged at 8,500 x g for 1 minute at room temperature. The flow-through was discarded and 750 µl of buffer PE was added. The mixture was centrifuged at 8,500 x g for 1 minute

at room temperature. The flow-through was discarded and centrifugation was continued at 17,900 x g for 1 minute at room temperature. Collection tube was replaced with 1.5 ml microcentrifuge tube and 50 µl of Buffer EB was added into the spin column. Then, the column was incubated at room temperature for 5 minutes before it was centrifuged at 17,900 x g for 1 minute at room temperature.

2.9.7 Nested IPCR of the *AF9* SAR

The optical density (O.D.) of the purified DNA sample was measured at 260 nm to determine its concentration prior to performing nested IPCR. The PCR reaction was set up in a total volume of 20 µl using a 0.2 ml PCR tube (Eppendorf). The first round of IPCR was carried out with 200 ng of DNA template, 1X HF buffer (containing 1.5 mM magnesium chloride), 200 µM dNTP mix, 0.5 µM each of the reverse primer (AF9 236451 R) and forward primer (AF9 245385 F), and 0.4 unit of Phusion High-Fidelity DNA Polymerase. For the negative control, ultra-pure water was used instead of DNA. The DNA sequence for the region of interest in the SAR region of the AF9 gene is detailed in Appendix E. The nucleotide sequences and positions of the primers used in IPCR are listed in Table 2.4, and the PCR conditions are provided in Table 2.5. Nested IPCR was performed according to the method described by Tan et al. (2016) (Tan et al., 2016).

The first round IPCR product was then diluted five times by adding 20 µl of the first round PCR product to 80 µl of sterile ultra-pure water. Second round of IPCR was carried out with 2 µl of diluted DNA and primers *AF9* 236211 R and *AF9* 245507 F. The primers sequences and PCR condition are as detailed in Table 2.4 and Table 2.6.

Table 2.4: Nucleotide Sequences and Nucleotide Positions of the Primers Used in the First Round and Second Round of IPCR for the SAR of the *AF9* Gene.

Primer	Sequence (5'-3')	Nucleotide Position (Ensembl: ENSG00000171843)
AF9 236451 R (First round)	ATTCTAGACCCCAAAAAATTCTCAG	236427-236451
AF9 245385 F (First round)	CTCTTAATGCCACTGCCATGA	245385-245405
AF9 236211 R (Second round)	CATATCCTTTTCATACCTGG	236192-236211
AF9 245507 F (Second round)	ATTGGTGTCAATCAAATGC	245507-245525

Table 2.5: PCR Programme Employed in the First Round of Nested IPCR for the SAR of the *AF9* Gene.

Step	Temperature (°C)	Time
Hot start	80	Forever
Initial denaturation	98	30 seconds
Denaturation	98	10 seconds
Annealing	69	30 seconds
Extension	72	15 seconds
Go to Denaturation	29 cycles	
Final extension	72	10 minutes
Cooling	20	Forever

Table 2.6: PCR Programme Used in the Second Round of Nested IPCR for the SAR of the *AF9* Gene.

Step	Temperature (°C)	Time
Hot start	80	Forever
Initial denaturation	98	30 seconds
Denaturation	98	10 seconds
Annealing	57	30 seconds
Extension	72	15 seconds
Go to denaturation	29 cycles	
Final extension	72	10 minutes
Cooling	20	Forever

2.9.8 Nested IPCR of the *AF9* non-SAR

For the non-SAR region, *Bam*H I was used as the first restriction enzyme (RE1), as detailed in section 2.9.2 and illustrated in Figure 2.3. This was followed by Klenow fill-in, cyclisation (section 2.9.3), and ethanol precipitation (section 2.9.4). The second restriction enzyme digestion (RE2) utilized *Hind* III to linearize the cyclized DNA. The third restriction enzyme digestion (RE3) involved *Xba* I, which linearized the cyclized DNA and prevented the amplification of intact fragments during IPCR, as shown in Figure 2.3. Finally, excess unincorporated nucleotides were removed using the QIAquick Nucleotide Removal Kit, following the protocol described in section 2.9.6.

Nested IPCR was performed following the method outlined by Tan et al. (2018)(Tan et al., 2018b). The composition of the nested IPCR for both rounds was similar to that used for the SAR region, as described in section 2.9.7. In the first round of IPCR, the primers AF9 71653 R and AF9 74399 F were used. For the second round, 2 µl of the first-round PCR product, diluted fivefold, was used with the primers *AF9* 71282 R and AF9 74494 F. The

DNA sequence of the non-SAR region of the *AF9* gene is provided in Appendix G. The nucleotide sequences and positions of the primers used in the IPCR are listed in Table 2.7, and the PCR conditions for both rounds are provided in Tables 2.8 and 2.9.

Table 2.7: Nucleotide Sequence and Nucleotide Position of the Primers Used in the Nested PCR for Non-SAR of the *AF9* Gene

Primer	Sequence (5'-3')	Position (Ensembl: ENSG00000171843)
AF9 71653 R (First round)	TACCAAACATTTTGAGTCCTACAG	71630-71653
AF9 74399 F (First round)	GGCATTCAGGTGAGTAGTTTATTC	74399-74422
AF9 71282 R (Second round)	AGCAGTAGACTTTTGTAACCTCAC	71259-71282
AF9 74494 F (Second round)	AGGGGATGACTTTTCTTCAATC	74494-74515

Table 2.8: PCR programme used in the first round of nested IPCR for the non-SAR of the *AF9* gene.

Step	Temperature (°C)	Time
Hot start	80	Forever
Initial denaturation	98	30 seconds
Denaturation	98	10 seconds
Annealing	64	30 seconds
Extension	72	15 seconds
Go to Denaturation	29 cycles	
Final extension	72	10 minutes
Cooling	20	Forever

Table 2.9: PCR Programme used in the Second Round of Nested IPCR for the Non-SAR of the *AF9* Gene.

Step	Temperature (°C)	Time
Hot start	80	Forever
Initial denaturation	98	30 seconds
Denaturation	98	10 seconds
Annealing	63	30 seconds
Extension	72	15 seconds
Go to denaturation	29 cycles	
Final extension	72	10 minutes
Cooling	20	Forever

2.9.9 Identify Chromosomal Breaks

In gel electrophoresis, an intact chromosome appears as a single, high-molecular-weight band because the DNA remains unbroken and migrates as a single large fragment. However, when chromosomal breaks occur, the DNA is fragmented into smaller pieces. These smaller fragments result in multiple bands on the gel or, if the breaks are extensive and random, a smeared appearance instead of distinct bands. The number of chromosomal breaks can be estimated using the formula B = F - 1, where B is the number of breaks and F is the total number of DNA fragments (bands) observed on the gel; for example, if four bands are present, the calculation B = 4 - 1 = 3 indicates that three chromosomal breaks have occurred.

2.9.10 Agarose Gel Electrophoresis and Gel Extraction

IPCR products were analysed by separating them on a 1% agarose gel. Electrophoresis was carried out in 0.5X TBE (45 mM Tris Base, 45 mM boric acid, 1 mM EDTA, pH 8) at 90 volts for 80 minutes. The gel was stained with ethidium bromide (10

µg/ml) and visualized using a UV transilluminator (Vilber Lourmat). Images of the gel were captured with a gel documentation system (Syngene). Chromosome breaks were quantified by counting the number of breaks in each IPCR reaction. Bands corresponding to cleaved DNA fragments of the *AF9* gene were extracted and purified using the QIAquick Gel Extraction Kit. The purified fragments were then sent for sequencing to verify the amplification region and identify the breakpoints.

For gel extraction, the QIAquick Gel Extraction Kit (QIAGEN) protocol was followed. Briefly, the IPCR bands were excised from the agarose gel and weighed. Three volumes of Buffer QG were added to one volume of gel (approximately 100 mg of gel equals 100 µl). The mixture was incubated at 50°C for 10 minutes or until the gel slice completely dissolved, with inversion every 3 minutes during incubation. The solution was then transferred to a QIAquick column in a 2 ml collection tube and centrifuged at 17,900 x g for 1 minute. After discarding the flow-through, 750 µl of Buffer PE was added. The tube was incubated at room temperature for 5 minutes before being centrifuged again at 17,900 x g for 1 minute. The flow-through was discarded, and the centrifugation was repeated at 17,900 x g for 1 minute. The column was then placed in a new 1.5 ml microcentrifuge tube, and 30 µl of ultra-pure water was added. After incubating at room temperature for 5 minutes, the tube was centrifuged at 17,900 x g for 1 minute. The extracted DNA was then stored at -20°C.

2.9.11 DNA Sequencing

Gel purified IPCR product was outsourced for sequencing at Apical Scientific Sdn Bhd to confirm that amplification was at the *AF9* gene and to map the breakpoints.

2.10 Statistical Analysis

Association of *LMP1* gene expression with oxidative stress, caspase activity and chromosomal break were analysed using Statistical Package for Social Sciences (SPSS) version 23 by performing Independent Sample T test. The association analysis to obtain the Odds Ratio (OR) for risk association was done using "2-way Contingency Table Analysis".

For the analysis of the risk association of oxidative stress and caspase activity with *LMP1* expression, the original data need to be categorised. Categorisation was done based on the mean value from Table 3.1 for oxidative stress. For caspase activity, the data was categorised based on the mean value from Table 3.3. The data were categorised such that values equal to or below the mean were considered as 0 (lower), while values above the mean were considered as 1 (higher).

CHAPTER 3

RESULTS

3.1 Transient Transfection of NP69 Cells: Observation by Fluorescence Microscopy

In order to study the effect of *LMP1* expression, NP69 cells were transiently transfected with either the vector control plasmids (pcDNA and pTracer) or the *LMP1* recombinant plasmids (pcDNA-*LMP1* and pTracer- *LMP1*). As shown in Figure 3.1, the pTracer-transfected NP69 cells have normal morphology. In addition, the same dark field microscopy of the pTracer-transfected NP69 cells shows defined morphology of the attached epithelial cells with different GFP expression level and fluorescence intensity (Figure 3.2). On the other hand, as shown in Figure 3.3, several of the pTracer-*LMP1* transfected cells were rounded and floating with some blebbing cells. The floating and blebbing cells were also expressing the GFP as shown in Figure 3.4. In addition, some attached fluorescent cells were also observed in the same slide. pTracer vectors use a CMV promoter to strongly express GFP in mammalian cells. Transfected cells glow green under a fluorescence microscope, while non-transfected cells remain dark. Transfection efficiency is estimated as the percentage of GFP-positive cells in the total cell population.

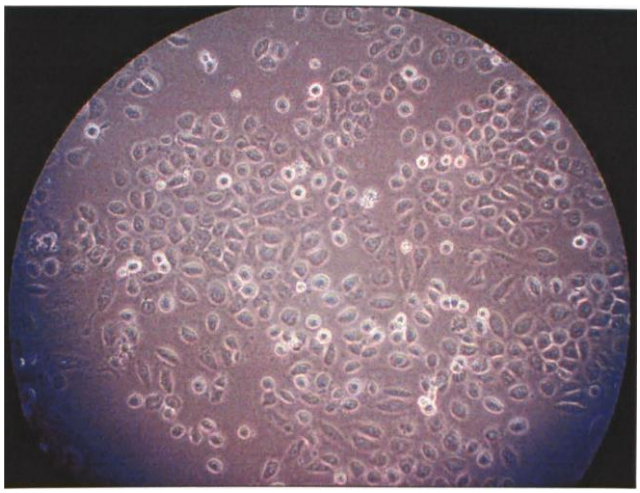

Figure 3.1: Bright-field microscopy of NP69 cells transfected with the vector control pTracer. Bright-field microscopy shows well-growing cells with defined morphology of an attached monolayer of epithelial cells in a 60 mm dish at 200x magnification.

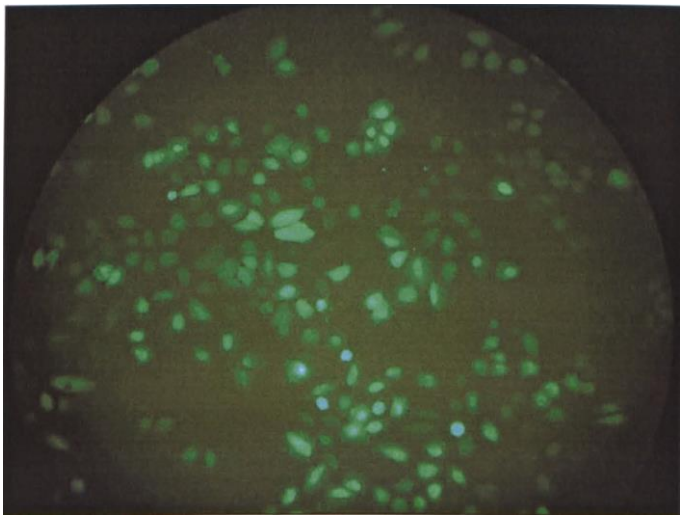

Figure 3.2: Dark-field microscopy of NP69 cells transfected with the vector control pTracer. Dark-field microscopy shows a well-defined, growing morphology of attached epithelial cells, each displaying varying levels of GFP expression as indicated by fluorescence intensity, at 200x magnification.

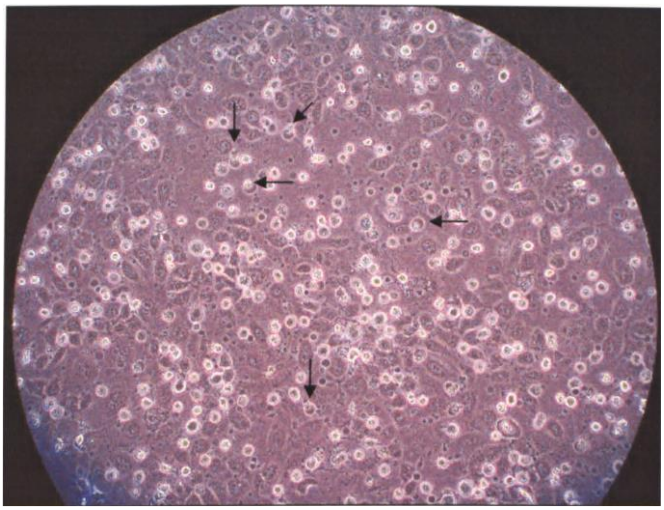

Figure 3.3: Bright-field microscopy of pTracer-*LMP1* transfected NP69 cells. The bright-field microscopy of pTracer-*LMP1*-transfected NP69 cells revealed several rounded floating cells and some blebbing cells, as indicated by the dark arrows, in the 60 mm dish at 200x magnification.

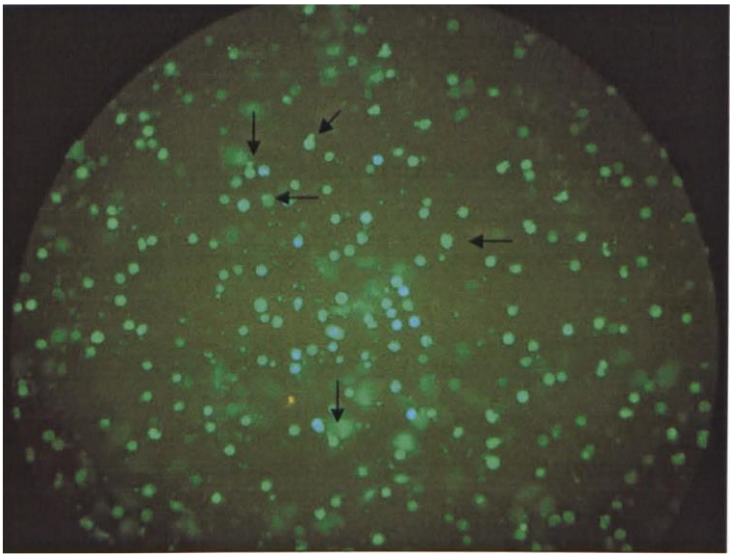

Figure 3.4: Dark-field microscopy of the pTracer- *LMP1* transfected NP69 cells with the GFP expression. The dark-field microscopy (x200) revealed several fluorescent cells appearing rounded and floating, with some exhibiting blebbing (as indicated by the dark arrows). Additionally, some of the attached cells also showed GFP expression.

3.2 Detection of *LMP1* Gene Expression by Western Blotting

To confirm the expression of the *LMP1* protein in the transfected cells, Western Blotting was carried out for the cells transfected with either pTracer-*LMP1* and pcDNA-*LMP1* or the vector control plasmids (pcDNA and pTracer). Anti-*LMP1* antibody (S12) was

used for the detection of *LMP1* gene expression. As shown in Figure 3.5, LMP1 protien was expressed from both plasmids (pTracer-*LMP1* and pcDNA-*LMP1*), as a protein of 60 kDa. Lane 2 shows no expression of *LMP1* in pTracer vector control transfected cells and lane 3 shows expression of *LMP1* in pTracer-*LMP1* transfected cells with size approximately 60 kDa. Lane 4 shows no expression of *LMP1* in pcDNA vector control transfected cells and lane 5 shows expression of *LMP1* in pcDNA-*LMP1* transfected cells with size approximately 60 kDa. Lane 1 shows positive control of B95 cell line. The B95 cell line, an *in vitro*-immortalised lymphoblastoid cell line, is an EBV-positive cell line. The blurry band in Figure 3.5 may be resulted from poor gel polymerization, causing incomplete acrylamide gel formation and leading to uneven protein migration. Meanwhile, the slanted bands in Figure 3.5 could be due to an uneven gel surface during polymerization, resulting in irregular protein migration.

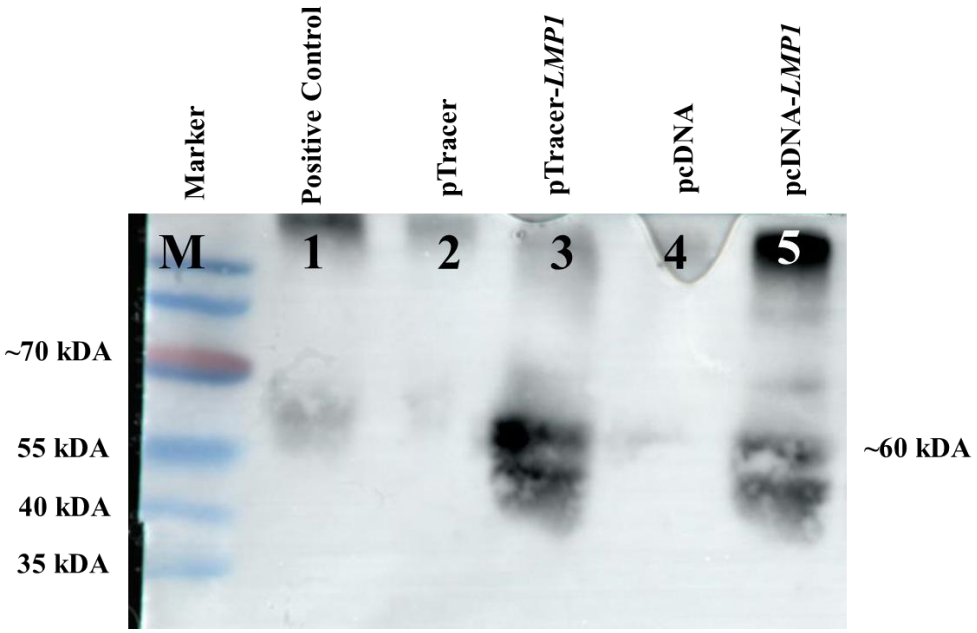

Figure 3.5: Immunodetection of LMP1 by the S12 antibody on Western blot. Lane 1 represents the positive control B95 cell line. NP69 cells were either transfected with vector control pTracer (lane 2) and pcDNA (lane 4), or *LMP1*-expression plasmids pTracer-*LMP1* (lane 3) and pcDNA-*LMP1* (lane 5). Lane 4 likely contains an unwanted artefact, possibly caused by non-specific antibody binding or protein degradation. Cell lysate was analysed on a 10% SDS-PAGE. M represents the PageRuler™ Prestained Protein Ladder (Fermentas).

3.3 Measurement of Oxidative Stress in *LMP1*-Expression Cells.

To examine the effect of *LMP1* expression on reactive oxygen species (ROS) production, NP69 cells were transfected with either *LMP1*-expression plasmid pcDNA-*LMP1* or the vector control pcDNA. Subsequently, ROS levels were measured, and fluorescence microscopy was performed using the fluorescent probe 2',7'-dichlorofluorescein diacetate (DCFH-DA). This probe is nonpolar and cell-permeable, remaining nonfluorescent until it is hydrolyzed by intracellular esterases and oxidized by ROS. Upon cleavage of its acetate groups by intracellular esterases, DCFH-DA is converted to polar dichlorofluorescein (DCFH), which becomes trapped in viable cells. In the presence of ROS, the nonfluorescent DCFH is quickly oxidized into the highly fluorescent

dichlorofluorescein (DCF) (Bass et al., 1983). Hydrogen peroxide (H2O2), a well-established inducer of oxidative stress, was used as a positive control.

3.3.1 Association of Oxidative Stress with Expression of *LMP1*.

Oxidative stress levels were measured across 100 replications (N) of cells transfected with either the pTracer or pcDNA vectors, with and without the *LMP1* expression. As shown in Table 3.1, cells transfected with pTracer-*LMP1* and the pTracer vector control displayed similar mean oxidative stress levels (~17,981 and ~18,043, respectively), both with relatively high standard deviations.

In contrast, both pcDNA-*LMP1* and its corresponding vector control exhibited much higher mean oxidative stress levels (~49,427 and ~49,169, respectively), with slightly lower variability than the pTracer groups. However, the oxidative stress levels between pcDNA-*LMP1* and its vector control remained closely matched, indicating that the expression of *LMP1* in the pcDNA system also did not result in a notable increase in oxidative stress relative to its control.

Due to the use of different oxidative stress detection kits for the pTracer and pcDNA systems, the data were binarized based on each group's mean value—classified as 0 (lower) or 1 (higher). When these binarized data sets were combined, both *LMP1*-expressing and vector-only cells showed an identical mean score of 0.41 (±0.492). This further supports the observation that, across both vector systems, *LMP1* expression did not consistently shift oxidative stress levels above or below the average, suggesting no clear association between *LMP1* and oxidative stress in this experimental context.

Table 3.1: Mean and Standard Deviation of Oxidative Stress in Transfected NP69.

Transfection Plasmid DNA	N	Mean	Std. Deviation	Std. Error Mean
pTracer	100	18043.47	11527.655	1152.765
pTracer-*LMP1*	100	17981.12	11996.745	1199.674
pcDNA	100	49168.78	9034.687	903.469
pcDNA-*LMP1*	100	49426.75	7489.795	748.979
pTracer & pcDNA	200	0.41	0.492	0.035
pTracer-*LMP1* & pcDNA-*LMP1*	200	0.41	0.492	.035

Table 3.2 illustrates the relationship between oxidative stress and *LMP1* expression using pTracer-*LMP1* and pcDNA-*LMP1*. The comparison between pTracer-*LMP1* transfection and the production of oxidative stress, relative to pTracer vector control cells, does not show a statistically significant association, as indicated by a t-value of 0.37 and a p-value of 0.970. For pcDNA-*LMP1* transfected cells, the association between oxidative stress and *LMP1* expression also does not show statistical significance, with a t-value of -0.220 and a p-value of 0.826, when compared to pcDNA vector control cells. When the analysis of association was done using a combination of both pTracer-*LMP1* and pcDNA-*LMP1* transfection data, the association of *LMP1* expression with oxidative stress value was also found to be not statistically significant, with a t-value of less than 0.001 and a p-value of 1.000.

Table 3.2: Association of Oxidative Stress with *LMP1* Expression.

Transfection Plasmid DNA		t-value	df	Sig. (2-tailed)	95% Confidence Interval of the Difference	
					Lower	Upper
pTracer & pTracer-*LMP1*	Equal variances assumed	.037	198	.970	-3218.608	3343.308
	Equal variances not assumed	.037	197.686	.970	-3218.640	3343.340
pcDNA & pcDNA-*LMP1*	Equal variances assumed	-.220	198	.826	-2572.236	2056.296
	Equal variances not assumed	-.220	191.423	.826	-2572.725	2056.785
pTracer & pTracer-*LMP1*	Equal variances assumed	.000	398	1.000	-.097	.097
pcDNA & pcDNA-*LMP1*	Equal variances not assumed	.000	398.000	1.000	-.097	.097

3.3.2 Association of Oxidative Stress with Green Fluorescence Protein (GFP) Expression.

pTracer contains the green fluorescence protein (GFP) gene. This GFP gene expression was reported to affect oxidative stress (Kalyanaraman & Zielonka, 2017), thus we tried to investigate the effect of GFP expression in oxidative stress induction in NP69. The data in Table 3.3 compares oxidative stress levels in cells transfected with vector-only controls (pTracer and pcDNA), using a binary classification system based on each group's respective mean oxidative stress level. Despite differences in detection kits between the two vector systems, the classification allowed for a consistent comparison of "higher" and "lower" oxidative stress across both groups.

The results show that both pcDNA and pTracer vector control cells had nearly identical proportions of samples classified as having higher oxidative stress, with mean values of 0.41 and 0.40, respectively. The standard deviations for both groups (~0.49) indicate similar variability in oxidative stress classification.

Table 3.3: Mean and Standard Deviation of Oxidative Stress in Vector Control Transfected NP69

Transfection Plasmid DNA	N	Mean	Std. Deviation	Std. Error Mean
pcDNA	100	.41	.494	.049
pTracer	100	.40	.492	.049

Table 3.4 shows the association of oxidative stress with GFP expression. The association between oxidative stress and GFP expression in pTracer is not statistically significant, with a t-value of 0.143and a p-value of 0.886, when compared to pcDNA.

Table 3.4: Association of Oxidative stress with Vector Control Transfected Cells

Plasmid DNA		t-value	Df	Sig. (2-tailed)	95% Confidence Interval of the Difference	
					Lower	Upper
pcDNA & pTracer	Equal variances assumed	.143	198	.886	-.128	.148
	Equal variances not assumed	.143	197.997	.886	-.128	.148

3.4 Measurement of Caspases Activity in *LMP1*-expressing Cells

3.4.1 Association of Caspase Activity with *LMP1* Expression

The caspase activity data presented in Table 3.5 suggests that LMP1 expression does not consistently lead to increased apoptosis in transfected cells. In the pTracer system, cells

expressing *LMP1* showed slightly lower mean caspase activity (513.18) compared to their vector control (529.99), although the large standard deviations in both groups indicate high variability and no clear difference.

In the pcDNA system, *LMP1*-transfected cells also showed a modest increase in caspase activity (304.30) compared to their vector control (271.52), but again, the difference is small, and variability remains relatively high.

When data from both vector systems were combined, the mean caspase activity in LMP1-expressing cells (408.78) was only marginally higher than in vector control cells (400.76), with overlapping standard deviations. These results indicate that *LMP1* expression does not strongly or consistently activate caspase 3/7 activity, suggesting it may not significantly induce apoptosis under the tested conditions.

Table 3.5: Mean and Standard Deviation of Caspace Activity in Transfected NP69

Transfection Plasmid DNA	N	Mean	Std. Deviation	Std. Error Mean
pTracer	100	529.99	322.257	32.226
pTracer -*LMP1*	100	513.18	413.327	41.333
pcDNA	100	271.52	95.827	9.583
pcDNA-*LMP1*	100	304.30	119.322	11.932
pTracer & pcDNA	200	400.76	270.218	19.107
pTracer -*LMP1* & pcDNA-*LMP1*	200	408.74	320.992	22.698

Table 3.6 shows the association of caspase activity with *LMP1* transfection in both pTracer-*LMP1* and pcDNA-*LMP1* transfected cells. The observed association between

pTracer-*LMP1* transfection and caspase activity was found not to be statistically significant, with a t-value of 0.321 and a p-value of 0.749, compared to pTracer vector control transfected cells. For pcDNA-*LMP1* transfection, the association with caspase activity was found to be statistically significant, with a t-value of -2.142 and a p-value of 0.033, compared to pcDNA vector control transfected cells. When the association analysis was performed using the combined data from both pTracer-LMP1 and pcDNA-*LMP1* transfections, the relationship between LMP1 expression and caspase activity was found to be not statistically significant, as indicated by a t-value of -0.269 and a p-value of 0.788.

Table 3.6: Association of Caspase Activity with *LMP1* Expression

Transfection Plasmid DNA		**t-value**	**Df**	**Sig. (2-tailed)**	**95% Confidence Interval of the Difference**	
					Lower	**Upper**
pTracer & pTracer-*LMP1*	Equal variances assumed	.321	198	.749	-86.545	120.165
	Equal variances not assumed	.321	186.885	.749	-86.583	120.203
pcDNA & pcDNA-*LMP1*	Equal variances assumed	-2.142	198	.033	-62.959	-2.601
	Equal variances not assumed	-2.142	189.187	.033	-62.968	-2.592
pTracer & pTracer-*LMP1*	Equal variances assumed	-.269	398	.788	-66.313	50.343
pcDNA & pcDNA-*LMP1*	Equal variances not assumed	-.269	386.756	.788	-66.318	50.348

3.4.2 Association of Caspase Activity with Green Fluorescence Protein (GFP) Expression

To investigate the role of GFP in caspase activity when expressed together in pTracer, we compared the caspase activity in pTracer and pcDNA transfected cells. Table 3.7 shows the mean and standard deviation of caspase activity in cells transfected with pTracer and pcDNA. In pcDNA vector control transfected cells, the mean caspase activity is 271.52 with a standard deviation of 95.827. Meanwhile, pTracer vector-transfected cells exhibited a mean of 529.99 with a standard deviation of 322.257.

Table 3.7: Mean and Standard Deviation of Caspace Activity in Vector Control

Transfection Plasmid DNA	N	Mean	Std. Deviation	Std. Error Mean
pcDNA	100	271.52	95.827	9.583
pTracer	100	529.99	322.257	32.226

Transfected NP69

Table 3.8 shows the association of caspase activity with GFP expression. The association between caspase activity and GFP expression is statistically significant, with a t-value of -7.688 and a p-value less than 0.05.

Table 3.8: Association of Caspase Activity with Vector Control Transfected Cells

Plasmid DNA		t-value	Df	Sig. (2-tailed)	95% Confidence Interval of the Difference	
					Lower	Upper
pcDNA & pTracer	Equal variances assumed	-7.688	198	.000	-324.770	-192.170
	Equal variances not assumed	-7.688	116.372	.000	-325.057	-191.883

3.5 Association of Combination of Oxidative Stress and Caspase Activity with *LMP1* Expression

The data in Table 3.9 examines the combined effect of *LMP1* expression on oxidative stress and caspase activity, with values categorized as either low (0) or high (1) based on the mean values from previous tables. Across 400 transfection replicates, the results show minimal differences between LMP1-expressing cells and their corresponding vector controls.

In the pTracer system, *LMP1*-transfected cells had a mean score of 0.40, compared to 0.43 in vector controls. Similarly, in the pcDNA system, *LMP1*-transfected cells showed a mean of 0.44, slightly lower than 0.47 in vector-transfected cells. These small differences, coupled with high standard deviations (~0.49–0.50), suggest that *LMP1* expression does not significantly influence the combined levels of oxidative stress and caspase activity when compared to the baseline levels seen in vector-only controls.

When data from both systems were pooled, the combined mean for *LMP1*-transfected cells was 0.45, compared to 0.41 for vector controls—again indicating only a slight difference. Taken together, the findings suggest that *LMP1* does not have a strong or consistent effect on elevating both oxidative stress and apoptosis, as measured by caspase activity, under these experimental conditions.

Table 3.9: Mean and Standard Deviation of Combination of Oxidative Stress and Caspace Activity in Transfected NP69

Transfection Plasmid DNA	N	Mean	Std. Deviation	Std. Error Mean
pTracer	200	.43	.496	.035
pTracer -*LMP1*	200	.40	.490	.035
pcDNA	200	.47	.500	.035
pcDNA-*LMP1*	200	.44	.498	.035
pTracer & pcDNA	400	.41	.492	.025
pTracer -*LMP1* & pcDNA-*LMP1*	400	.45	.498	.025

Table 3.10 shows the association of the combination of oxidative stress and caspase activity with *LMP1* transfection in both pTracer-*LMP1* and pcDNA-*LMP1* transfected cells. The observed association between pTracer-*LMP1* transfection and the combination of oxidative stress and caspase activity was found to be not statistically significant, with a t-value of 0.609 and a p-value of 0.543, compared to pTracer vector control transfected cells. For pcDNA-*LMP1* transfection, the association with combination of oxidative stress and caspase activity was also found to be not statistically significant, with a t-value of 0.501 and a p-value of 0.617, compared to pcDNA vector control transfected cells. When the analysis of association was done using a combination of both pTracer-*LMP1* and pcDNA-*LMP1* transfection data, the association of *LMP1* expression with the combination of oxidative stress and caspase activity was found to be not statistically significant, with a t-value of -1.213 and a p-value of 0.225.

Table 3.10: Association of Oxidative Stress and Caspase Activity with *LMP1* Expression

Plasmid DNA		t-value	Df	Sig. (2-tailed)	95% Confidence Interval of the Difference	
					Lower	Upper
pTracer & pTracer-*LMP1*	Equal variances assumed	.609	398	.543	-.067	.127
	Equal variances not assumed	.609	397.950	.543	-.067	.127
pcDNA & pcDNA-*LMP1*	Equal variances assumed	.501	398	.617	-.073	.123
	Equal variances not assumed	.501	397.991	.617	-.073	.123
pTracer & pTracer-*LMP1* pcDNA & pcDNA-*LMP1*	Equal variances assumed	-1.213	798	.225	-.111	.026
	Equal variances not assumed	-1.213	797.886	.225	-.111	.026

3.6 Risk Association of Oxidative Stress and Caspase Activity with *LMP1* Expression

Risk association analysis is a tool for identifying factors associated with disease development or progression. This information is crucial for understanding underlying disease mechanisms and implementing preventive measures. To assess the risk association between oxidative stress, caspase activity, and LMP1 expression, the original data must be categorized based on the mean value from Table 3.1 for oxidative stress, and for caspase activity, the data should be categorized according to the mean value from Table 3.5.The data were categorised with values equal to or below the mean value considered as 0 (lower), while values above the mean value were considered as 1 (higher).

3.6.1 Risk Association of pTracer, pTracer-*LMP1*, pcDNA and pcDNA-*LMP1* Transfections with Oxidative Stress.

Table 3.11 presents the analysis of the risk association between oxidative stress and the transfection of pTracer, pTracer-*LMP1*, pcDNA, and pcDNA-*LMP1*. When the risk association of pTracer vector control and pTracer-*LMP1* transfection with oxidative stress was investigated, there was no significant risk for increased oxidative stress, with an odds ratio (OR) of 0.9194 (CI 0.5207-1.6231) and a p-value of 0.7719. Similarly, the analysis of risk association between pcDNA vector control and pcDNA-*LMP1* transfected cells with oxidative stress also did not show significant risk for increased oxidative stress, with an OR of 1.0856 (CI 0.6190-1.9038) and a p-value of 0.7745. Additionally, the combined risk analysis of pTracer, pTracer-LMP1, pcDNA, and pcDNA-LMP1 transfection showed no significant association with increased oxidative stress, with an OR of 1.0000 (CI 0.6708-1.4907) and a p-value of 1.0000.

Table 3.11: Risk Association of pTracer, pTracer-*LMP1*, pcDNA and pcDNA-*LMP1* Transfections with Oxidative Stress.

Transfection Plasmid DNA	Lower	Higher	OR (CI 95%)	p-value
pTracer	60	40	0.9194 (0.5207-1.6231)	0.7719
pTracer-*LMP1*	62	38		
pcDNA	59	41	1.0856 (0.6190-1.9038)	0.7745
pcDNA-*LMP1*	57	43		
pTracer & pcDNA	119	81	1.0000 (0.6708-1.4907)	1.0000
pTracer-LMP1 & pcDNA-*LMP1*	119	81		

3.6.2 Risk Association of pTracer, pTracer-*LMP1*, pcDNA and pcDNA-*LMP1* Transfection with Caspase Activity

Table 3.12 shows the risk association of pTracer, pTracer-*LMP1*, pcDNA, and pcDNA-*LMP1* transfections with caspase activity. When the risk association of pTracer and pTracer-*LMP1* transfections with caspase activity was investigated, pTracer-*LMP1* transfections did not show a significant risk of increased caspase activity, with an odds ratio (OR) of 0.8485 (CI 0.4836-1.4886) and a p-value of 0.5667. For the risk association between pcDNA and pcDNA-*LMP1* transfections with caspase activity, pcDNA-*LMP1* transfections did not show a significant risk of increased caspase activity, with an OR of 0.7552 (CI 0.4331-1.3169) and p-value of 0.3224. Based on the risk association analysis of the combined data from pTracer, pTracer-*LMP1*, pcDNA, and pcDNA-*LMP1* transfections with caspase activity, the combined data from pTracer-*LMP1* and pcDNA-*LMP1* transfections did not show a significant risk of increased caspase activity, with an OR of 0.8777 (CI 0.5941-1.2967) and a p-value of 0.5124.

Table 3.12: Risk Association of pTracer, pTracer-*LMP1*, pcDNA and pcDNA-*LMP1* Transfections with Caspase Activity.

Transfection Plasmid DNA	**Lower**	**Higher**	**OR (CI 95%)**	**P-value**
pTracer	56	44	0.8485 (0.4836-1.4886)	0.5667
pTracer-*LMP1*	60	40		
pcDNA	48	52	0.7552 (0.4331-1.3169)	0.3224
pcDNA-*LMP1*	55	45		
pTracer & pcDNA	114	96	0.8777 (0.5941-1.2967)	0.5124
pTracer-LMP1 & pcDNA-*LMP1*	115	85		

3.6.3 Risk Association of Combination of Oxidative Stress and Caspase Activity with *LMP1* Expression

Table 3.13 presents the analysis of the risk association between oxidative stress combined with caspase activity and the transfection of pTracer, pTracer-*LMP1*, pcDNA, and pcDNA-*LMP1*. When the analysis of risk association of pTracer and pTracer-*LMP1* transfections with the combination of oxidative stress and caspase activity was investigated, pTracer-*LMP1* transfections did not show a significant risk for increased oxidative stress and caspase activity, with an odds ratio (OR) of 0.9012 (CI 0.6042-1.3443) and a p-value of 0.6102. For the analysis of risk association between pcDNA and pcDNA-*LMP1* transfections with the combination of oxidative stress and caspase activity, pcDNA-*LMP1* did not show a significant risk for increased oxidative stress and caspase activity, with an OR of 0.9040 (CI 0.6097-1.3404) and a p-value of 0.6155. Furthermore, the combined analysis of the data from pTracer, pTracer-LMP1, pcDNA, and pcDNA-LMP1 transfections revealed that the combination of pTracer-LMP1 and pcDNA-LMP1 did not significantly increase the risk of oxidative stress and caspase activity, with an OR of 0.9029 (CI 0.6822-1.1949) and a p-value of 0.4749.

Table 3.13: Risk Association of Combination of Oxidative Stress and Caspase Activity with pTracer, pTracer-*LMP1*, pcDNA and pcDNA-*LMP1* Transfections.

Transfection Plasmid DNA	Oxidative Stress & Caspase Activity		OR (CI 95%)	P-value
	Lower	Higher		
pTracer	117	83	0.9012 (0.6042 -1.3443)	0.6102
pTracer-*LMP1*	122	78		
pcDNA	107	93	0.9040 (0.6097-1.3404)	0.6155
pcDNA-*LMP1*	112	88		
pTracer & pcDNA	224	176	0.9029 (0.6822 -1.1949)	0.4749
pTracer-LMP1 & pcDNA-*LMP1*	234	166		

3.7 Identification of *LMP1* Expression-Induced Chromosomal Breaks in Scaffold/Matrix Attachment Region (SAR) and non-SAR of *AF9* Gene.

According to the primer positions, if no breakage occurs within the studied region, the expected IPCR product for the SAR of the AF9 gene will be approximately 944 bp (or ~950 bp). If there is any chromosomal breakage within the region of study, IPCR products of less than 950 bp should be obtained. As shown in Figure 3.6, lanes 1 and 3 show IPCR products of pTracer vector control transfected cells with no chromosomal breakage within the SAR. Lanes 4, 5, and 6 reveal IPCR products of pTracer-*LMP1* transfected cells that have breakage within the SAR. Lanes 10, 11, and 12 present IPCR products of pcDNA vector control cells that have breakage within the SAR. Lanes 13 and 14 represent IPCR products of pcDNA-*LMP1* transfected cells that have breakage within the SAR. 'M' refers to the marker used, GeneRuler 1 kb DNA ladder. A slanted band in both Figure 3.6 and Figure 3.7 occurs due to uneven gel preparation, incorrect sample loading, or poor electrophoresis conditions.

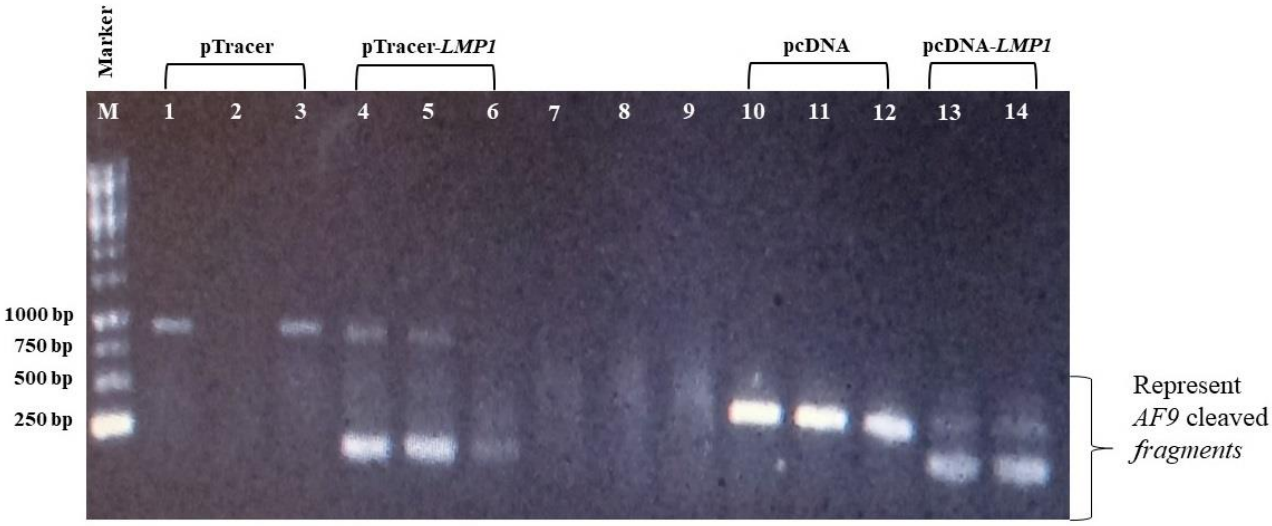

Figure 3.6: A representative gel picture shows the cleavages of the *AF9* gene identified by IPCR within the SAR. NP69 cells were transfected with either pTracer vector control (Lanes 1 and 3), pTracer-*LMP1* (Lanes 4-6), pcDNA vector control (Lanes 10-12), or pcDNA-*LMP1* (Lanes 13 and 14). Genomic DNA extraction and nested IPCR were carried out as detailed in the 'Methods' section. The side bracket indicates the IPCR bands originating from the cleaved AF9 fragments. 'M' represents the 1 kbp DNA ladder.

For the non-SAR region of the *AF9* gene, based on the primer positions, the expected IPCR product, in the absence of any breakage within the studied region, will be approximately 956 bp (or ~950 bp). If there is any breakage within the region of study, IPCR products of less than 950 bp should be obtained. As shown in Figure 3.7, lanes 1, 2, 3, and 4 show IPCR products of pTracer vector control transfected cells with chromosomal breakage within the non-SAR. Lanes 5, 6, 7, and 8 represent IPCR products of pTracer-*LMP1* transfected cells that show chromosomal breakages within the non-SAR. Lanes 9, 10, 11, and 12 reveal IPCR products of pcDNA-*LMP1* transfected cells that show breakages within the non-SAR. Lanes 13, 14, 15, and 16 represent IPCR products of pcDNA vector control transfected cells that have no chromosomal breakage within the non-SAR. 'M' refers to the marker used, GeneRuler 1 kb DNA ladder.

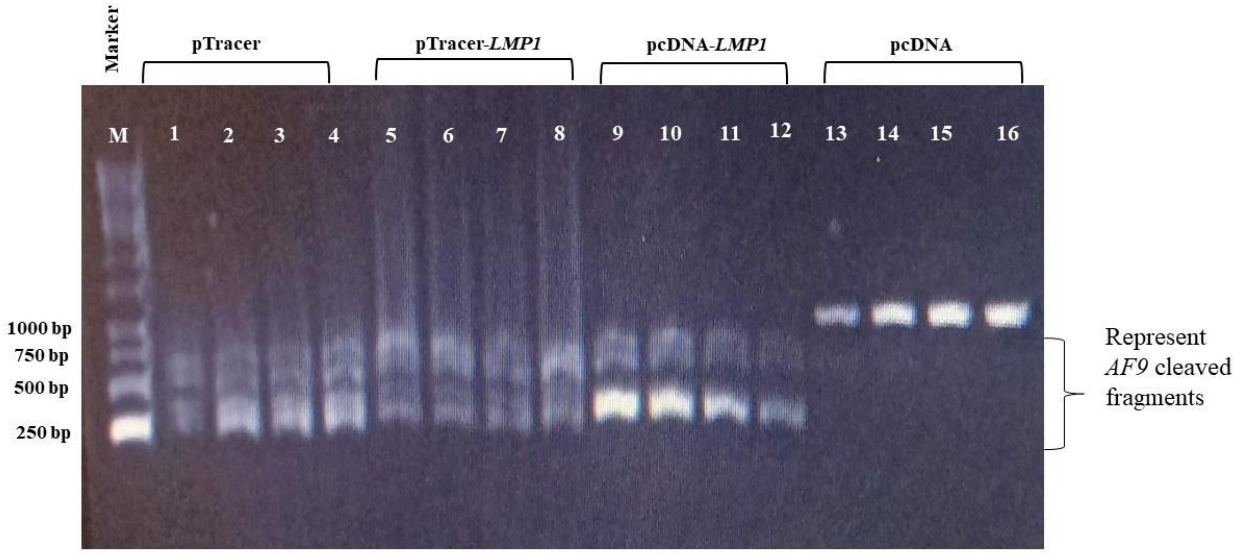

Figure 3.7: A representative gel image illustrates the cleavages in the AF9 gene detected by IPCR within the non-SAR region. NP69 cells were transfected with either the pTracer vector control (Lanes 1-4), pTracer-LMP1 (Lanes 5-8), pcDNA-LMP1 (Lanes 9-12), or pcDNA vector control (Lanes 13-16). Genomic DNA extraction and nested IPCR were conducted as outlined in the 'Methods' section. The side bracket indicates the IPCR bands corresponding to the cleaved fragments of the AF9 gene. 'M' refers to the 1 kbp DNA ladder.

3.7.1 Association of Chromosome Breaks with *LMP1* Expression in the SAR and Non-SAR of *AF9* Gene.

Twenty-five replicates (N) of transfections with pTracer, pTracer-*LMP1*, pcDNA, and pcDNA-*LMP1* were conducted to identify the number of chromosomal breaks within the SAR and non-SAR regions of the *AF9* gene. Table 3.14 shows the mean and standard deviation of the number of chromosomal breaks in the SAR in pTracer, pTracer-*LMP1*, pcDNA, and pcDNA-*LMP1* transfected cells. In pTracer-*LMP1* transfected cells, the mean number of chromosomal breaks was 3.32, with a standard deviation of 1.180. For pTracer vector control transfections, the mean was 1.96, with a standard deviation of 1.241. The mean and standard deviation for pcDNA-*LMP1* transfected cells were 2.12 and 1.013, respectively. For pcDNA vector control transfection, a mean of 1.16 and a standard deviation of 0.987 were observed. When data from both pTracer and pcDNA were combined, the mean number of chromosomal breaks in *LMP1*-expressing cells was 2.72 with a standard deviation

of 1.246, while vector-transfected cells showed a mean of 1.56 and a standard deviation of 1.181.

Table 3.14: Mean and Standard Deviation of the Number of Chromosomal Breaks in the SAR of the *AF9* Gene in Transfected NP69 Cells.

Transfection Plasmid DNA	N	Mean	Std. Deviation	Std. Error Mean
pTracer	25	1.96	1.241	.248
pTracer-*LMP1*	25	3.32	1.180	.236
pcDNA	25	1.16	.987	.197
pcDNA-*LMP1*	25	2.12	1.013	.203
pTracer & pcDNA	50	1.56	1.181	.167
pcDNA-*LMP1* & pcDNA-*LMP1*	50	2.72	1.246	.176

N= number of transfection replication

Table 3.15 shows the association of the number of chromosome breaks in the SAR with *LMP1* transfection in both pTracer-*LMP1* and pcDNA-*LMP1* transfected cells. The observed association between pTracer-*LMP1* transfection and the number of chromosome breaks in the SAR was found to be statistically significant, with a t-value of -3.970 and a p-value of less than 0.001, compared to pTracer vector control transfected cells. Similarly, for pcDNA-*LMP1* transfection, the association with the number of chromosome breaks in the SAR was also found to be statistically significant, with a t-value of -3.394 and a p-value of 0.001, compared to pcDNA vector control transfected cells. When the analysis of association was conducted using a combination of both pTracer-LMP1 and pcDNA-*LMP1* transfection data, the association of *LMP1* expression with the number of chromosome breaks was found to be statistically significant, with a t-value of -4.778 and a p-value of less than 0.001.

Table 3.15: Association of Number of Chromosome Breaks in the SAR with *LMP1* Expression

Transfection Plasmid DNA		t-value	df	Sig. (2-tailed)	95% Confidence Interval of the Difference	
					Lower	Upper
pTracer & pTracer-*LMP1*	Equal variances assumed	-3.970	48	<.001	-2.049	-.671
	Equal variances not assumed	-3.970	47.880	<.001	-2.049	-.671
pcDNA & pcDNA-*LMP1*	Equal variances assumed	-3.394	48	.001	-1.529	-.391
	Equal variances not assumed	-3.394	47.966	.001	-1.529	-.391
pTracer & pTracer-*LMP1*	Equal variances assumed	-4.778	98	<.001	-1.642	-.678
pcDNA & pcDNA-*LMP1*	Equal variances not assumed	-4.778	97.718	<.001	-1.642	-.678

Twenty-five replications (N) each of the transfections were conducted to investigate the association of *LMP1* expression with the number of chromosome breaks in the non-SAR of the *AF9* gene. Table 3.16 shows the mean and standard deviation of chromosome breaks in the non-SAR of the *AF9* gene in cells transfected with pTracer-*LMP1*, pcDNA-*LMP1*, as well as the vectors. In pTracer-*LMP1* transfected cells, the mean number of chromosomal breaks was 2.24, with a standard deviation of 0.523. For cells transfected with the pTracer vector control, the observed mean was 1.92, with a standard deviation of 0.862. In contrast,

pcDNA-*LMP1* transfected cells exhibited a mean of 2.24 and a standard deviation of 1.562, while the pcDNA vector transfection showed a mean of 1.96 with a standard deviation of 1.338. The combined analysis of data from both pTracer-*LMP1* and pcDNA-*LMP1* transfections revealed a mean of 2.24 with a standard deviation of 1.153. For the vector control transfections (pTracer and pcDNA), the combined analysis indicated a mean of 1.94 and a standard deviation of 1.114.

Table 3.16: Mean and Standard Deviation of Number of Chromosomal Breaks in the Non-SAR of the *AF9* Gene in Transfected NP69 Cells.

Transfection Plasmid DNA	**N**	**Mean**	**Std. Deviation**	**Std. Error Mean**
pTracer	25	1.92	.862	.172
pTracer-*LMP1*	25	2.24	.523	.105
pcDNA	25	1.96	1.338	.268
pcDNA-*LMP1*	25	2.24	1.562	.312
pTracer & pcDNA	50	1.94	1.114	.158
pcDNA-*LMP1* & pcDNA-*LMP1*	50	2.24	1.153	.163

Table 3.17 presents the analysis of the association between the number of chromosomal breaks in the non-SAR region of the AF9 gene and *LMP1* expression using pTracer-*LMP1* and pcDNA-*LMP1*. The association between pTracer-*LMP1* transfection and the number of chromosomal breaks in the non-SAR was not statistically significant, with a t-value of -1.587 and a p-value of 0.119, when compared to cells transfected with the pTracer vector control. Similarly, in pcDNA-*LMP1* transfected cells, the association between chromosomal breaks in the non-SAR and *LMP1* expression was also not statistically significant, with a t-value of -0.681 and a p-value of 0.499, compared to cells transfected

with the pcDNA vector control. When combining data from both pTracer-*LMP1* and pcDNA-*LMP1* transfections, the association between *LMP1* expression and the number of chromosomal breaks in the non-SAR remained not statistically significant, with a t-value of -1.323 and a p-value of 0.189.

Table 3.17: Association of Chromosome Break in Non-SAR of the *AF9* Gene with *LMP1* Expression.

Transfection Plasmid DNA		t-value	df	Sig. (2-tailed)	95% Confidence Interval of the Difference	
					Lower	Upper
pTracer & pTracer-*LMP1*	Equal variances assumed	-1.587	48	.119	-.725	.085
	Equal variances not assumed	-1.587	39.548	.121	-.728	.088
pcDNA & pcDNA-*LMP1*	Equal variances assumed	-.681	48	.499	-1.107	.547
	Equal variances not assumed	-.681	46.893	.499	-1.108	.548
pTracer & pTracer-*LMP1*	Equal variances assumed	-1.323	98	.189	-.750	.150
pcDNA & pcDNA-*LMP1*	Equal variances not assumed	-1.323	97.886	.189	-.750	.150

3.7.2 Risk association of Number of Chromosomal Breaks in the SAR and the Non-SAR of the *AF9* Gene with *LMP1* Expression.

For pTracer, pTracer-*LMP1*, pcDNA, and pcDNA-*LMP1*, 25 replications of transfection were performed to investigate the association of *LMP1* expression with

chromosome breaks in the SAR of the *AF9* gene. Table 3.18 shows the risk association of pTracer, pTracer-*LMP1*, pcDNA, and pcDNA-*LMP1* transfection with the number of chromosome breaks in the SAR of the *AF9* gene. When the risk association of pTracer vector control and pTracer-*LMP1* transfected cells with the number of chromosome breaks in the SAR was investigated, the latter showed a significantly higher risk, with an odds ratio (OR) of 21.9293 (CI 1.2092 - 397.7105) and a p-value of 0.0368 for increased number of chromosome breaks in the SAR. Risk association analysis of pcDNA vector control and pcDNA-*LMP1* transfection with the number of chromosome breaks in the SAR reveals that pcDNA-*LMP1* transfection showed a significantly higher risk of chromosome breaks in the SAR, with an OR of 30.8305 (CI 1.7179 - 553.3123) and a p-value of 0.0199. In the analysis of the risk association for the combination of pTracer, pTracer-*LMP1*, pcDNA, and pcDNA-*LMP1* with the number of chromosome breaks in the SAR, pTracer-*LMP1* and pcDNA-*LMP1* showed significantly higher risk of chromosome breaks in the SAR, with an OR of 50.4268 (CI 2.9673 - 856.9706) and a p-value of 0.0067.

Table 3.18: Risk Association of pTracer, pTracer-*LMP1*, pcDNA and pcDNA-*LMP1* Transfection with Number of Chromosomal Breaks in the SAR of the *AF9* Gene.

Transfection Plasmid DNA	**No Chromosomal Breaks**	**Chromosomal Breaks**	**OR (CI 95%)**	**p-value**
pTracer	6	49	21.9293 (1.2092 - 397.7105)	0.0368
pTracer-*LMP1*	0	83		
pcDNA	8	29	30.8305 (1.7179 -553.3123)	0.0199
pcDNA-*LMP1*	0	53		
pTracer & pcDNA	14	78	50.4268 (2.9673 - 856.9706)	0.0067
pTracer-*LMP1* & pcDNA-*LMP1*	0	136		

To investigate the association between *LMP1* expression and the number of chromosomal breaks in the non-SAR region of the *AF9* gene, 25 replicates of transfections were performed. Table 3.19 presents the risk association of pTracer, pTracer-*LMP1*, pcDNA, and pcDNA-*LMP1* transfections with the number of chromosomal breaks in the non-SAR region of the *AF9* gene. When evaluating the risk associated with pTracer and pTracer-LMP1 transfections, the pTracer-*LMP1* transfection did not show a significantly higher risk of increased chromosomal breaks in the non-SAR, with an odds ratio (OR) of 1.1649 (CI 0.0227 - 59.8240) and a p-value of 0.076. Similarly, in the analysis of pcDNA and pcDNA-*LMP1* transfections, pcDNA-*LMP1* transfection also did not exhibit a significantly higher risk for increased chromosomal breaks in the non-SAR, with an OR of 0.0761 (CI 0.0042 - 1.3666) and a p-value of 0.0805. Furthermore, when combining the data from pTracer, pTracer-*LMP1*, pcDNA, and pcDNA-*LMP1* transfections, the combination of pTracer-*LMP1* and pcDNA-*LMP1* did not show a significantly higher risk of chromosomal breaks in the non-SAR, with an OR of 0.0790 (CI 0.0045 - 1.4005) and a p-value of 0.0836.

Table 3.19: Risk Association of pTracer, pTracer-*LMP1*, pcDNA and pcDNA-*LMP1* Transfection with Chromosomal Breaks in the Non-SAR of the *AF9* Gene.

Transfection Plasmid DNA	No Chromosomal Breaks	Chromosomal Breaks	OR (CI 95%)	p-value
pTracer	0	48	1.1649 (0.0227 -59.8240)	0.076
pTracer-*LMP1*	0	56		
pcDNA	0	49	0.0761 (0.0042 - 1.3666)	0.0805
pcDNA-*LMP1*	7	56		
pTracer & pcDNA	0	97	0.0790 (0.0045 -1.4005)	0.0836
pTracer-*LMP1* & pcDNA-*LMP1*	7	115		

3.8 Sequencing of IPCR Products

To verify that the fragments identified through IPCR originated from the cleaved *AF9* gene, several IPCR bands were excised, purified, and sequenced. The sequencing confirmed that these fragments were indeed derived from the cleaved *AF9* gene. The breakpoints identified within the *AF9* SAR in *LMP1*-expressing cells are listed in Table 3.20. Notably, several breakpoints (at coordinates 245527, 245575, and 245596) were mapped within the *AF9* region (coordinates 245252-245612), a region previously reported to be involved in the t (9;11) (p22; q23) translocation, which led to the formation of the *MLL-AF9* fusion gene in an acute lymphoblastic leukaemia (ALL) patient.

The breakpoint at coordinate 245596 is similar to one reported in cultured normal blood cells (at coordinate 245593) and CEM cells (at coordinate 246114) exposed to etoposide (VP16) (Cynthia and Sim, 2012). Additionally, the breakpoint at coordinate 245596 is also similar to those identified in H_2O_2-treated NP69 and HK1 cells (Tan et al., 2016) as well as in TWO4 cells treated with bile acid at pH 7.4 (Tan & Sim, 2018).

Furthermore, some chromosomal breaks (at coordinates 245649 and 245725) are located within the repeat ERE2_EH (coordinates 245627-245728).

Table 3.21 lists the breakpoints identified within the *AF9* non-SAR in *LMP1*-transfected cells. Two chromosomal breaks (at coordinates 75013 and 75081) fall within the second repeat CHARLIE5 (at coordinates 75006-75169).

Table 3.20: Breakpoints detected within the *AF9* SAR in *LMP1* Transfected cells.

Breakpoint	Remarks	Sequencing result
245527	Located at coordinates 245252-245612, this chromosomal break is within the AF9 region, which has been previously noted for translocating with the MLL gene. This leads to the creation of the MLL-AF9 fusion gene in a patient diagnosed with acute lymphoblastic leukemia (ALL) [Ensembl: ENSG00000171843].	Appendix I, 1
245575	Located at coordinates 245252-245612, this chromosomal break occurs within the AF9 region, which has been reported to translocate with the MLL gene, forming the MLL-AF9 fusion gene in an ALL patient [Ensembl: ENSG00000171843].	Appendix I, 2
245596	This breakpoint matches the one observed in TWO4 cells treated with bile acid at pH 7.4 and is located in the AF9 region (coordinates 245252-245612), which has been previously linked to the MLL-AF9 fusion gene in an ALL patient. It differs by three nucleotides from the breakpoint in VP16-treated cultured normal blood cells (coordinate 245593) (Cynthia and Sim, 2012), by five nucleotides from the breakpoint in H2O2-treated NP69 cells (coordinate 245591), and by six nucleotides from the breakpoint in H2O2-treated HK1 cells (coordinate 245590) (Tan et al., 2016).	Appendix I, 3
245649	The chromosomal break is located within the ERE2_EH repeat sequence at coordinates 245627-245728.	Appendix I, 4
245725	The chromosomal break is located within the ERE2_EH repeat sequence at coordinates 245627-245728.	Appendix I, 5
245979		Appendix I, 6

Table 3.21: Breakpoints detected within the *AF9* non-SAR in *LMP1* Transfected cells.

Breakpoint	Remarks	Sequencing result
74653		Appendix J, 1
74752		Appendix J, 2
74786		Appendix J, 3
75013	The chromosomal break is positioned in the CHARLIE5 repeat sequence, covering coordinates 75006 to 75169.	Appendix J, 4
75081	The chromosomal break is positioned in the CHARLIE5 repeat sequence, covering coordinates 75006 to 75169.	Appendix J, 5

3.9 Chapter Summary

The association between oxidative stress and *LMP1* expression was investigated using pTracer-*LMP1*, pcDNA-*LMP1*, and a combination of the data of both transfections. However, the production of oxidative stress was not found to be statistically significant compared to the vector control cells. Additionally, the risk of increased oxidative stress associated with vector control, pTracer-*LMP1*, pcDNA-*LMP1*, and the combination of the data of both transfections was assessed, revealing no significant risks. Furthermore, the role of green fluorescence protein (GFP) in oxidative stress, expressed together with *LMP1* in the pTracer vector, was examined. The association of oxidative stress and GFP expression was found to be not statistically significant when compared to vector control cells.

The analysis of the association of caspase activity with *LMP1* transfection in both pTracer-*LMP1* and pcDNA-*LMP1* transfected cells, as well as the combination of both data, shows that pTracer-*LMP1* transfection with caspase activity was not found to be statistically significant compared to pTracer vector control transfection. For pcDNA-*LMP1* transfection, the association with caspase activity was found to be statistically significant compared to

pcDNA vector control transfected cells. The combined analysis of the pTracer-*LMP1* and pcDNA-*LMP1* transfection data indicated that there was no statistically significant association between *LMP1* expression and caspase activity. The analysis of the risk association of vector control and pTracer-*LMP1*, pcDNA-*LMP1*, and the combination of the data of both transfections with caspase activity was investigated, and there were no significant risks for increased caspase activity. To investigate the role of green fluorescence protein (GFP) in caspase activation when expressed together in pTracer, we compared the vector control pTracer and pcDNA transfected cells. The association between caspase activity and GFP expression was found to have a statistically significant difference compared to vector control cells.

The association of pTracer-*LMP1*, pcDNA-*LMP1* transfection, and their combined data with both oxidative stress and caspase activity was found to be not statistically significant compared to vector control transfected cells. The analysis focused on the risk association of vector control, pTracer-LMP1, pcDNA-LMP1, and their combined transfection data with oxidative stress and caspase activity. The results indicated that neither pTracer-LMP1, pcDNA-LMP1, nor their combined data presented a significant risk of increased oxidative stress and caspase activity.

The observed association between pTracer-*LMP1*, pcDNA-*LMP1*, and the combined data of both transfections with the number of chromosome breaks in the SAR of the *AF9* gene was found to be statistically significant compared to vector control transfected cells. Risk association analysis between vector control and pTracer-*LMP1*, pcDNA-*LMP1*, and the combined data of both pTracer-*LMP1* and pcDNA-*LMP1* transfections with number of

chromosome breaks in the SAR reveals that *LMP1* transfections showed a significantly higher risk of chromosome breaks in the SAR of the *AF9* gene.

The association between pTracer-*LMP1*, pcDNA-*LMP1*, and the combined data from both transfections and chromosome breaks in the non-SAR region of the *AF9* gene is not statistically significant when compared to vector control transfected cells. Additionally, when evaluating the risk association of pTracer-*LMP1*, pcDNA-*LMP1*, and their combined transfections with the number of chromosome breaks in the non-SAR, *LMP1* expression was found to not pose a statistically significant higher risk of chromosome breaks in the non-SAR.

CHAPTER 4

DISCUSSION

4.1 Model of Mechanism of EBV-Induced Carcinogenesis via *LMP1* Expression and DNA Damage.

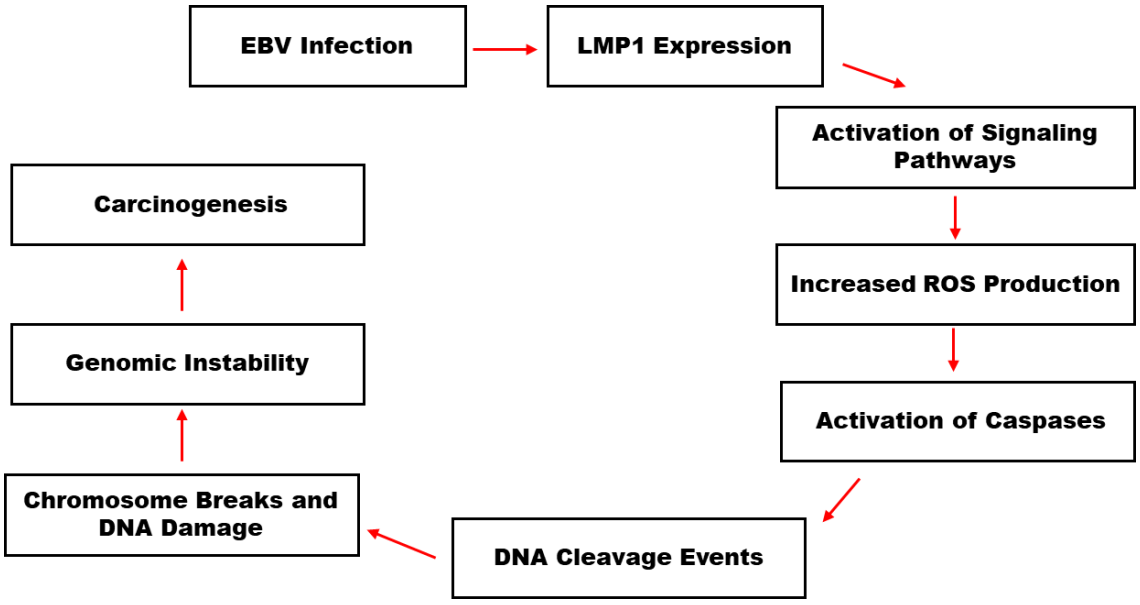

Figure 4.8: Schematic representation of the molecular mechanism by which Epstein-Barr Virus (EBV) infection leads to carcinogenesis. EBV infection induces the expression of Latent Membrane Protein 1 (LMP1), which activates signaling pathways. This results in increased Reactive Oxygen Species (ROS) production, caspase activation, and DNA cleavage events, leading to chromosome breaks, genomic instability, and ultimately carcinogenesis.

The EBV latency protein LMP1 plays a crucial role in driving EBV-associated malignancies in both B cells and epithelial cells (L. W. Wang et al., 2017). As a membrane protein, LMP1 mimics CD40 signaling and activates various downstream pathways (Wang & Ning, 2021). While its role in EBV-infected B cell lymphomas is well studied, the mechanisms by which LMP1 contributes to EBV-associated epithelial cell tumors remain less understood. Additionally, some effects of LMP1 in EBV-infected epithelial cells may differ from those in B cells, as LMP1 is expressed at significantly higher levels in EBV-

transformed lymphoblastoid B-cell lines (LCLs) compared to EBV-infected epithelial cell lines (Singh et al., 2022). Moreover, EBV-infected B cells and epithelial cells exhibit substantial differences in both cellular and viral gene expression. The Epstein-Barr virus (EBV)-encoded LMP1 acts as a key oncogene in nasopharyngeal carcinoma (NPC) (Cao, 2017).

Overexpression of Latent Membrane Protein 1 (*LMP1*) has been associated with inducing oxidative stress in cells (Feng et al., 2021). LMP1 is a viral protein encoded by the Epstein-Barr virus (EBV), a virus implicated in several human cancers, including nasopharyngeal carcinoma (Banko et al., 2021). Studies have demonstrated that *LMP1* expression can lead to increased reactive oxygen species (ROS) production and oxidative stress in cells infected with EBV. This oxidative stress can contribute to various pathological processes, including inflammation, DNA damage, and promotion of cancer progression (Jelic et al., 2021).

LMP1 is an integral membrane protein consisting of a short N-terminal cytoplasmic domain, six transmembrane-spanning loops, and a long C-terminal cytoplasmic tail that contains two key signaling domains, known as C-terminal activating regions (CTAR-1 and CTAR-2) (Dawson et al., 2012). Acting as a constitutively active viral mimic of CD40, a member of the tumor necrosis factor (TNF) receptor superfamily, LMP1 triggers multiple signaling cascades through its CTAR domains. These include NF-κB, PI3K–AKT, ERK–MAPK, JNK, JAK-STAT, and p38/MAPK pathways. Together, these pathways regulate the expression of various downstream cellular targets, ultimately contributing to the oncogenic properties of LMP1 (Lo et al., 2021).

Elevated levels of reactive oxygen species (ROS) are present in nearly all cancers, if not all, as well as during Epstein-Barr virus (EBV) latency (Cerimele et al., 2005). It is well established that LMP1, along with other EBV-derived factors, is a major contributor to ROS production in EBV latency, at least through the JNK-p38/AP1 pathway (Bettina Gruhne et al., 2009). Mitochondria are the primary source of ROS (Stowe & Camara, 2009). In malignant cells, mitochondria exhibit functional and structural deregulation, leading to excessive ROS production and playing a crucial role in ROS-induced DNA damage (Fulda et al., 2010).

Apoptosis is widely believed to act as a tumor suppressor by eliminating cells with DNA damage (Hanahan, 2022). However, recent studies have shown that apoptotic caspases can also contribute to tumor development by inducing DNA damage and promoting genomic instability (Lopez & Bouchier-Hayes, 2022). Classical caspase-dependent apoptosis follows the intrinsic pathway (Sarkar et al., 2022), which is triggered by cellular stresses such as heat shock, DNA damage, and oxidative stress, leading mitochondria to release cytochrome c or Smac/DIABLO into the cytosol. Cytochrome c helps form the apoptosome, a caspase-activating complex made of seven Apaf-1 adaptor molecules, each bound to cytochrome c, along with a caspase-9 dimer (Bhadra, 2022). This activates caspase-9, which then triggers caspases-3 and -7, leading to apoptosis. Both pathways involve chromatin condensation, DNA fragmentation, and membrane changes like phosphatidylserine exposure and blebbing. Stress signals, including DNA damage, can increase p53 levels, which play key roles in growth arrest, DNA repair, and cell death (Bhadra, 2022). Caspases are a family of protease enzymes that play crucial roles in programmed cell death (apoptosis) and inflammation (Kesavardhana et al., 2020). They are often activated in response to various cellular stresses, including oxidative stress. Oxidative stress leads to the accumulation of ROS within cells.

ROS can directly damage cellular components such as DNA, proteins, and lipids, triggering a cascade of cellular responses. One of these responses involves the activation of signalling pathways that ultimately lead to the activation of caspases. Oxidative stress-induced caspase activation represents a critical mechanism by which cells respond to cellular damage and stress, ultimately leading to programmed cell death or apoptosis (Matés et al., 2012).

Caspase activation can indirectly contribute to chromosome breaks, particularly during apoptotic cell death. During apoptosis, caspases are activated through both intrinsic (mitochondrial) and extrinsic (death receptor-mediated) pathways. Activation of caspases results in the cleavage of numerous cellular proteins, leading to characteristic morphological changes such as cell shrinkage, membrane blebbing, and chromatin condensation. While caspase activation itself does not directly cause chromosome breaks, the downstream effects of caspase activation can lead to genomic instability, including chromosome breaks. Caspase-activated DNase (CAD) cleaves chromatin at the scaffold/matrix attachment region (MAR/SAR) sites, releasing high molecular weight (HMW) fragments and eventually cleaves between the nucleosomes, forming the typical nucleosomal ladder seen during apoptosis (Sakahira et al., 1999; Tan et al., 2018a).

Genomic instability is a well-recognized feature of nasopharyngeal carcinoma (NPC). Although various somatic mutations are present, copy-number variations (CNVs) occur more frequently in NPC compared to other head and neck squamous cell carcinomas (HNSCCs), as identified through whole-genome sequencing (WGS) (Zhou et al., 2022). During NPC development, functional genes involved in different signaling pathways undergo alterations through mutations or structural changes(Liu et al., 2022), leading to phenotypic shifts in cellular functions. As NPC progresses, genomic instability worsens due

to accumulating somatic mutations, CNVs, and structural alterations in nasopharyngeal epithelial cells, along with increased adaptive plasticity in response to the tumor microenvironment (TME) (Liu et al., 2022). Eventually, malignant transformation becomes irreversible, and these epithelial cells evolve into invasive NPC cells.

4.2 Association of *LMP1* Expression with Nasopharyngeal Carcinoma (NPC)

EBV infection is frequently associated with nasopharyngeal carcinoma (NPC), and the expression of *LMP1* has been observed in nasopharyngeal carcinoma cells. (Young et al., 1988). Additionally, it has been noted that *LMP1* can exhibit cytotoxic effects on certain cell types when expressed at high levels (Hammerschmidt et al., 1989). Expression of *LMP1* at high levels (27-60 times higher than that in B95-8, an EBV-immortalised B cell line) was reported to mediate apoptosis in human keratinocytes (RHEK-1) (Lu et al., 1996). Furthermore, *LMP1* expression seems to correlate with apoptosis in NPC cells (CNE2) (Liu et al., 2002). This implies that EBV infection and the expression of *LMP1* may possibly induce apoptosis, which subsequently leads to chromosome rearrangement in NPC. Studies have reported that both *LMP1* antigen and *LMP1*-induced tumour-associated antigen (TAA) can induce T cell recognition *in vitro* and *in vivo* (Choi et al., 2021).

LMP1 is also reported to enhance the anti-apoptosis, motility, and invasion of tumour cells, and T cell immune suppression in NPC patients. (Yoshizaki et al., 2013). In the study of early diagnosis and predicting disease progression of nasopharyngeal carcinoma, *LMP1* was found to suppress STAT4 expression in NPC cells (Liu et al., 2023). A study on the oncoprotein 18/stathmin signalling pathway, mediated by cdc2 in nasopharyngeal carcinoma cells, shows that *LMP1* promotes cell proliferation and accelerates cell cycle progression (Lin et al., 2009).

The development of NPC involves multiple factors. EBV infection appears to be a significant factor in NPC carcinogenesis. However, despite EBV infection affecting over 90% of the world population (Chang & Adami, 2006), NPC remains unique in its geographical distribution. This implies that there are other important risk factors contributing to NPC carcinogenesis. Emerging evidence suggests that chronic ear and nose diseases, predominantly chronic rhinitis, sinusitis, and otitis media, increase the risk of developing NPC (Ekburanawat et al., 2010; Huang et al., 2012). Additionally, several lifestyle factors have been reported to induce the transformation of EBV from the latent to the lytic stage and contribute to NPC occurrence, such as salted fish or preserved food (W.-H. Jia & H.-D. Qin, 2012). Moreover, cigarette smoking can induce EBV lytic reactivation and increase the risk of NPC (Xu et al., 2012).

4.3 Association of Oxidative Stress with *LMP1* Expression

Notably, reactive oxygen signalling can distinguish EBV-positive versus EBV-negative Burkitt's lymphoma based on the finding that elevated levels of ROS were observed in EBV-positive tumours but not in EBV-negative tumours (Cerimele et al., 2005). This demonstrates that EBV infection could cause the generation of ROS, leading to high levels of oxidative stress in EBV-positive cells or tumours. Inducible expression of *LMP1* during the viral lytic cycle plays a vitally important role in virus production. This is supported by the observation that, the loss of *LMP1* severely impairs virus release into culture supernatant fractions, resulting in poor infection efficiency (Ahsan et al., 2005). Studies have shown that EBV reactivation and the expression of lytic proteins, such as the BZLF1 and BRLF1, are frequently observed in NPC tissues (Zhang et al., 2018). These lytic proteins can promote cell proliferation, invasion, and immune evasion, contributing to the carcinogenic process. The reactivation of EBV and the subsequent expression of lytic proteins play a significant

role in NPC carcinogenesis (Wu et al., 2018) by promoting cell transformation, altering immune responses, and modulating the tumour microenvironment. Generating oxidative stress is a critical mechanism by which host cells defend against infection by pathogenic microorganisms (Rada & Leto, 2008). *LMP1* was reported to contribute to the radio resistance of NPC by activating several oncogenic signalling axes, including LMP1/JNKs/c-Jun/HIF-1/VEGF (Yang et al., 2015). Reactivation of EBV by *LMP1*-induced high oxidative stress could lead to radio resistance of NPC cells both *in vitro* and *in vivo*, suggesting that EBV reactivation plays an important role in NPC resistance, and *LMP1* is one of the factors that mediate resistance (Hu et al., 2020).

DNA viruses may encode oncogenic genes that are also capable of hijacking host cellular mechanisms to regulate cell survival and propagation. When these oncogenic genes overcome the ability of the host cell machinery to control homeostasis, they trigger the tumour microenvironment associated with an elevated level of mutations that cause malignant transformation and ultimately cancer (Wiseman & Halliwell, 1996). Despite the contrasting findings in the present research, it could still be presumed that differences in signalling and biological properties of the *LMP1* variants contribute to differences in pathogenicity (Edwards et al., 2004). ROS are also involved in the progression – at this stage, the generation of large amounts of ROS may contribute to mutation, increase the amount of matrix metalloproteinases (MMPs), and damage local tissues. Increased levels of DNA modified due to oxidation can contribute to genetic instability and may be crucial in tumour metastases in cancer (Choudhari et al., 2014; Katakwar et al., 2016).

In the current study, *LMP1* expression was carried out using two different plasmid vectors. Transfection was carried out in NP69, a transformed normal nasopharyngeal

epithelial cell line. Our result shows that *LMP1* transfection using both plasmid vectors and the production of oxidative stress do not show statistically significant correlation compared to the vector control cells. The risk association analysis also shows no statistically significant correlation between *LMP1* expression and oxidative stress. This result might be due to the transient transfection performed in the current study which may not provide a consistent level of *LMP1* (Stepanenko & Heng, 2017). A study using NP69 cells transfected with pLNSX-*LMP1* and stably expressing *LMP1* exhibited significantly increased basal ROS level (approximately 8- and 10-fold increase) (Sun et al., 2015). They suggested that *LMP1* expression causes cellular ROS accumulation in nasopharyngeal epithelial cells. *LMP1* induces excessive ROS generation by upregulating the expression of NADPH Oxidase through the JNK/AP-1 signalling pathway in nasopharyngeal epithelial cells without affecting the antioxidant system (Sun et al., 2015). Another study investigating the inner relationship and regulatory mechanism among oxidative stress and EBV reactivation, using established CNE1-*LMP1* and HNE2-*LMP1* cells, also shows that overexpression of *LMP1* significantly increased the level of ROS (Hu et al., 2020). They also determined the role of *LMP1* in ROS generation in transfected cells with *LMP1*. They demonstrated that *LMP1* promotes ROS generation in a stable cell line, while in transient transfection, *LMP1* depletion significantly decreases the level of ROS. The same study also showed that *LMP1* significantly induced EBV reactivation. Their data also illustrated that *LMP1* promotes EBV reactivation through the production of ROS.

GFP expression was shown to be potentially toxic to cells (Liu et al., 1999). GFP was used for protein tracking and enzyme activity detection, enhancing fusion protein folding, and is relatively small (Stepanenko et al., 2013). A study on the formation of reactive oxygen species by enhanced green fluorescent protein (eGFP) shows that eGFP induces

oxidative stress in cells (Ganini et al., 2017). The study showed that in the initiation of the apoptosis cascade, reactive oxygen production induced by GFP has been linked to cellular toxicity and eventually death in GFP-expressing cells (Gambotto et al., 2000). Additionally, neuroblastoma cell lines showed increased sensitivity to cytotoxic agents when transduced with GFP, eGFP, and YFP (Goto et al., 2003). In our current study, pTracer contains the GFP gene which was also expressed during transfection. Considering the possibility of this GFP expression-induced ROS may interfere with *LMP1* expression-induced ROS, a comparison of oxidative stress production was performed between pTracer and pcDNA (does not contain GFP gene). To our surprise, the association between oxidative stress and pTracer transfection was not statistically significant compared to pcDNA transfection. This means that GFP expression did not induce significant ROS. However, since pTracer and pcDNA are two very different plasmids, a more appropriate comparison would be between pTracer and a pTracer with the GFP gene removed.

4.4 Association of Caspases Activity with *LMP1* Expression

Caspase-activated DNase (CAD) seems to be playing multiple roles. CAD is the apoptotic nuclease that is found to play a role in chromosome rearrangement commonly found in leukaemia (Nicholas & Sim, 2012). In addition, CAD was also shown to promote cell differentiation by inducing DNA strand breaks (Larsen et al., 2010). CAD is one of the enzymes involved in apoptosis and plays an important role in chromosome rearrangement mostly found in leukaemia (Hars et al., 2006; Sim & Liu, 2001; Vaughan et al., 2002) as well as in nasopharyngeal carcinoma (Boon & Sim, 2015a). The findings also suggest that CAD may play an important role in chromosomal cleavages mediated by oxidative stress-induced apoptosis as well as bile-acid-induced apoptosis (Boon & Sim, 2015a; Tan & Sim, 2018; Tan et al., 2018a).

CAD is activated by caspase-dependent signal transduction, a proteolytic signalling cascade that alters the activity of numerous substrate proteins (Larsen & Sørensen, 2017). CAD is associated with an inhibitor of CAD (ICAD), and possibly, ICAD serves as a chaperone during the synthesis of CAD (Enari et al., 1998). To promote DNA fragmentation, Caspase 3 activates CAD by proteolytic inactivation of the inhibitor of CAD (Larsen & Sørensen, 2017). Meanwhile, Caspase 3 inactivates ICAD by cleaving at two aspartic acid residues, D117 and D224, destabilising its interaction with CAD and allowing CAD dimerisation and subsequent DNA fragmentation (Sakahira et al., 1998). In a study of mouse ICAD and human CAD, expression of ICAD was reported to enhance the expression of endogenous and exogenous CAD (Boon & Sim, 2015b). The study also reported that ICAD expression induced endogenous CAD expression and extensively reduced H_2O_2-induced *MLL* gene cleavage.

In the current study, we assessed the induction of apoptosis by measuring caspase activity in transformed normal nasopharyngeal cells expressing *LMP1* and compared it with nasopharyngeal cells without *LMP1* gene expression. Our results showed that the association value is statistically significant when comparing the mean of caspase activity in pcDNA-*LMP1* and pcDNA vector control transfected cells. Our results are supported by a study that demonstrated that *LMP1* expression from pcDNA-*LMP1* induces apoptosis, as shown by cell blebbing. They also demonstrated that *LMP1* expression induced cleavage of the *MLL* gene at 11q23, indicating that *LMP1* expression induces apoptosis in NPC cells (Yee & Sim, 2010). Interestingly, in the same study by Yee and Sim (2010), *LMP1* expression from pTracer-*LMP1* also demonstrated cell blebbing, indicating apoptosis. On the contrary, in our current study, *LMP1* expression from pTracer-*LMP1* did not show significant activation of caspase activity. This discrepancy could be due to the fact that the NP69 cell line used in

this study is a transformed normal epithelial cell line while an NPC cell line was used by Yee and Sim (Yee & Sim, 2010). It is known that normal cells are more resistant to apoptosis compared to cancer cells (Clarke et al., 2011). Additionally, it is noteworthy that statistical analysis was not performed in the study by Yee and Sim (2010). Interestingly, caspase activity was significantly higher in cells transfected with pTracer than in those transfected with pcDNA. This suggests that GFP expression from pTracer likely induced caspase activity, leading to a high background level that may have masked the effect of LMP1 expression in triggering caspase activity.

The apoptotic nuclease, caspase-activated DNase (CAD), was suggested to play a direct role in mediating chromosome translocation in leukaemia (Hars et al., 2006; Sim & Liu, 2001). *LMP1* expression was associated with a slightly increased proportion of dead cells in MDCK epithelial cells (Brocqueville et al., 2013). *LMP1* was cytotoxic to MDCK cells when expressed at relatively high levels. However, our data for pTracer-*LMP1* transfection did not show statistically significant induction of apoptosis in NP69 cells. The result might be due to the GFP-induced apoptosis mentioned before. This reasoning is supported by a study showing that initiation of the apoptosis cascade has been postulated as a possible mechanism for the toxicity of GFP and cellular death (Liu et al., 1999). They also found that GFP caused redistribution of phosphatidylserine, an indicator of the apoptosis signaling cascade initiation, to the cell surface (Liu et al., 1999). In another study, liver cells transfected with GFP plasmid have been shown to have increased cellular permeability following the initiation of cellular death (Taghizadeh & Sherley, 2008).

For the risk association of *LMP1* and caspase activity, *LMP1* did not show a significant risk for increased caspase activity, which might be due to a lack of sample size.

Small sample sizes often result in lower statistical power, which refers to the ability of a study to detect a true effect if it exists. When statistical power is low, the study may not have enough data to detect meaningful differences or relationships, even if they truly exist in the study. In contrast, in another study that investigated the relationship between *LMP1* and caspase-3, *LMP1* expression was shown to increase the cleavage of caspase-3 and -8 after *LMP1* transfection into HeLa cells (Togi et al., 2016). Additionally, the expression of EBV *LMP1* induced cleavage of the mixed lineage leukaemia (*MLL*) gene at 11q23, which was reduced by caspase inhibitor(Yee & Sim, 2010).

An association study is the most applicable tool to assess gene susceptibility to complex diseases that involve high interaction between genetic and environmental factors. Many complex diseases have a variety of genetic variants that affect the disease risk, even though with minimal effect (Frazer et al., 2009). The current study shows that *LMP1* gene expression in transformed normal nasopharyngeal cells contributes to caspase activation, leading to apoptosis. Although apoptosis is a cell death process (Kerr et al., 1972), upcoming literature has also supported the role of apoptosis in chromosome rearrangement in leukaemia (Hars et al., 2006; Sim & Liu, 2001; Vaughan et al., 2002) as well as in NPC (Tan & Sim, 2018; Tan et al., 2018a; Tan et al., 2016).

4.5 Association of Combination of Oxidative Stress and Caspases Activity with *LMP1* Expression

A previous report has shown that oxidative stress, DNA damage, and mitochondrial membrane depolarisation may all induce cells to undergo apoptosis (Mani et al., 2020). Oxidative stress-induced apoptosis could be one of the mechanisms underlying the chromosomal rearrangements in NPC (Tan et al., 2016). CAD may play an important role in chromosomal cleavages mediated by oxidative stress-induced apoptosis. However, in the

current study, there is no significant association between the combination of oxidative stress and caspase activity with *LMP1* expression. This suggests that there is no significant relationship between the combination of oxidative stress and caspase activity, and the expression of *LMP1*.This result might be due to the expression of *LMP1* and the oxidative stress data were not statistically significant in the current study. In contrast, a study of oxidative stress in acute promyelocytic leukaemia cells suggested that oxidative stress plays an important role in the induction of apoptosis (Jing et al., 1999). In addition, Cladosporol A induces ROS-mediated cell death via apoptosis caused by oxidative stress (Koul et al., 2017). Oxidative stress also induces apoptosis in glioma cells (Sun et al., 2018b). Moreover, oxidative stress was shown to induce apoptosis in normal nasopharyngeal epithelial cells and NPC cells (Tan et al., 2018a). In a study on mitochondria and mitochondrial ROS in Cancer, ROS were found to be crucial in triggering tumour cell death (Yang et al., 2016).

4.6 Association of Chromosomal Breaks within the *AF9* Gene with *LMP1* Expression.

The human *AF9* gene at 9p22 was targeted in this study because this gene is one of the most common fusion partner genes of the *MLL* gene translocation. The *MLL* gene translocation has been strongly associated with acute myelogenous leukaemia (AML). However, it is uncommon in therapy-related AML (t-AML), acute lymphocytic leukaemia (ALL), and myelodysplastic syndromes (MDS) (Strissel et al., 2000; Swansbury et al., 1998). Interestingly, the 9p22 which contains the *AF9* gene is a common chromosomal deletion site in NPC (Shao et al., 2000). Chromosomal breakage is an early event in both apoptotic DNA fragmentation and chromosome rearrangement (Tan et al., 2018b). Previous studies revealed that chromosome breaks tend to fall within certain regions which contain specific chromatin structural elements, such as the matrix attachment region/scaffold

attachment region (MAR/SAR) (Strick et al., 2006b; Strissel et al., 2000). There were two MAR/SARs isolated experimentally in the *AF9* gene. They were designated as SAR1 and SAR2. SAR1 is found in intron 4, whereas SAR2 spans from exons 5 to 7. Two patient breakpoint cluster regions (BCR) have been identified in the *AF9* gene, namely, BCR1 and BCR2. BCR1 is in intron 4, whereas BCR2 encompasses introns 7 to 8. These two BCRs are bordered by SAR1 and SAR2 (Strick et al., 2006b; Strissel et al., 2000). MAR/SAR are at the base of chromatin loop structure, thus indicating that chromatin structure may contribute to defining the site of chromosome rearrangement (Tan et al., 2018b).

4.6.1 *LMP1* Expression-Induced Chromosomal Breaks Within the *AF9* Gene Scaffold Attachment Region (SAR) and Non-SAR

The *MLL* gene contains MAR/SAR sequences that have been implicated in chromosome rearrangement (Broeker et al., 1996; Strick et al., 2006a). High molecular weight (HMW) DNA fragmentation appears to result from excision of chromosomal DNA loops, whose bases interact with the nuclear matrix via the MAR/SAR sequence (Lagarkova et al., 1995; Mirkovitch et al., 1988). MAR/SARs are DNA sequences responsible for chromosomal loop attachment (Sperry et al., 1989). Matrix association region/scaffold attachment region sequences have been found to possess DNA unwinding properties that enable them to play an important role in facilitating the entry of protein factors involved in transcription, replication, chromosome condensation, and apoptosis (Bode et al., 2000; Göhring et al., 1997). Furthermore, these unwinding properties also make MAR/SAR sequences regions of DNA fragility that are prone to breakages (Bode et al., 2000; Legault et al., 1997). Additionally, MAR/SAR sequences are associated with chromosomal rearrangements and have been implicated in illegitimate recombination (Broeker et al., 1996). The breakpoint cluster region (BCR) of the *MLL* gene, which is frequently involved

in translocations in leukaemia patients, has been reported to contain two MAR/SAR sequences (Broeker et al., 1996).

In the current study, the association between the number of chromosome breaks in the SAR and the transfections with pTracer-*LMP1*, pcDNA-*LMP1*, and the combined data of both was found to be statistically significant compared to the vector control-transfected cells. Our result is supported by a study on the association of bile acid treatments with chromosomal breakage in the SAR of the *AF9* gene (Tan & Sim, 2019). It was found that, for the SAR, the cleavage frequencies in bile acid-treated cells were significantly higher than those in the untreated control (Tan & Sim, 2019). In the study of MAR/SAR sequences' role in mediating chromosome breakages in nasopharyngeal epithelial cells, oxidative stress-induced apoptosis results in cleavages within the *AF9* SAR (Tan et al., 2018b). In the same study, H_2O_2 induces cleavages within the *AF9* SAR in a caspase-3-dependent manner. For the risk association analysis performed, *LMP1*-expressing cells showed a significantly higher risk of increased chromosome breakage in the SAR. It is suggested that *LMP1* expression contributes to chromosomal breaks in the SAR of the *AF9* gene in NP69 cells. This result is in accordance with a study which shows that *LMP1* induces the cleavage of the mixed lineage leukaemia gene at 11q23 in a nasopharyngeal carcinoma cell line (Yee & Sim, 2010). Together, our results suggest that *LMP1* expression induces cleavages within the *AF9* SAR.

The association between the number of chromosome breaks in the non-SAR and the transfections with pTracer-*LMP1*, pcDNA-*LMP1*, and the combined data of both was not statistically significant compared to vector control transfected cells. Similarly, for the non-SAR, there was no significant difference in the cleavage frequencies between the bile acid-

treated cells and untreated control cells in NPC (Tan & Sim, 2019). Our finding is further supported by another study which demonstrated that, there was no significant difference between the untreated cells and H_2O_2-treated cells in the cleavage frequency of the *AF9* non-SAR (Tan et al., 2018b). They also observed that cleavages of the non-SAR were not dependent on caspase 3. Based on a previous study (Tan et al., 2018b), MAR/SAR is a preferential site of chromosome breaks. Therefore, fewer or no chromosome breaks were expected to be detected in this non-SAR. To their surprise, there were more cleavages in the non-SAR as compared to SAR. However, the cleavages in the non-SAR were not dependent on H_2O_2 treatment. This observation indicates that there are other chromatin structures that may also contribute to DNA fragility. Indeed, the non-SAR were found to contain numerous repeat elements (Tan et al., 2018b). In addition to the MAR/SAR sequence, repeat elements have also been implicated in mediating chromosome breaks (Jeffs et al., 1998; Strick et al., 2006b). For risk association analysis, *LMP1*-transfected NPC cells showed no significantly higher risk of increased chromosome breaks in the non-SAR of *AF9* gene. It shows that *LMP1* expression did not contribute to chromosomal breakage in the non-SAR of the *AF9* gene. Again, our results support the hypothesis that *LMP1* expression does not significantly contribute to cleavages in the non-SAR of the *AF9* gene.

The observed differences in chromosomal breakage between pTracer and pcDNA transfections likely result from their distinct vector properties, which influence gene expression, protein behavior, and cellular stress responses. These factors must be considered when analyzing or comparing experimental outcomes. The differential number of chromosomal breaks between SAR and non-SAR regions likely reflects the interplay of chromatin structure, transcriptional activity, sequence motifs, and mechanical factors.

Understanding these differences provides insight into genome stability and the cellular response to DNA damage.

MAR/SAR is known as an epigenetic regulator because it functions to help generate and maintain an open chromatin domain that is favourable to transcription and may also facilitate the integration of several copies of the transgene (Harraghy et al., 2012). *LMP1* has been found to hijack the host epigenetic machinery to alter the expression of host genes by activating the chromatin-modifying enzyme poly(ADP-ribose) polymerase 1 (PARP1) (Martin et al., 2016). The study also reported that the expression of *LMP1* alone is sufficient to activate PARP1 and promote a series of PARP1-dependent alterations in chromatin structure. *LMP1* expression has also been found to change the expression of either DNA methyltransferases or histone methyltransferases (Shi et al., 2019), subsequently contributing to epigenetic changes. *LMP1*-expressing cell lines exhibited significantly higher frequencies of chromatid breaks through defective CHK1 activation, suggesting an association between *LMP1* expression and NPC. (Deng et al., 2012). These findings support the idea that *LMP1* expression induces changes and might cause chromosome breaks in the SAR located at the base of chromatin loop structure.

On sequencing data identified a few breakpoints that are similar to those identified in different cells lines as well as in cells induced to undergo apoptosis by different stimuli (Tan & Sim, 2018, 2019; Tan et al., 2018a; Tan et al., 2018b). This indicates that *LMP1* expression and other stimuli may share a similar mechanism in inducing apoptosis-mediated chromosome breaks in the SAR, defined by the chromatin loop structure.

4.7 Limitations

This study provides valuable insights into the role of *LMP1* expression in oxidative stress, apoptosis, and chromosomal instability in nasopharyngeal cells; however, several limitations should be noted. First, the use of transient transfection may have resulted in variable levels and duration of *LMP1* expression, potentially affecting the consistency of oxidative stress detection. The absence of a stable *LMP1*-expressing cell line limits the ability to assess long-term effects and may underestimate the gene's contribution to oxidative stress and genomic instability.

Additionally, different oxidative stress detection kits were employed for separate vector systems (pTracer and pcDNA), which may introduce variability and complicate direct comparisons. The *in vitro* model used in this study does not fully replicate the tumor microenvironment or the systemic interactions present in nasopharyngeal carcinoma *in vivo*.

Moreover, while chromosomal breakage was assessed in the SAR of the AF9 gene, other potentially affected genomic regions were not examined. Although a high number of transfection replicates were conducted, the biological variability inherent to transient systems may have reduced the statistical power to detect subtle effects. Lastly, while we observed associations between *LMP1* expression, caspase activity, and chromosomal breakage, mechanistic studies to validate the direct pathways involved were beyond the scope of this study.

CHAPTER 5

CONCLUSION

This study reinforces the role of Epstein-Barr virus (EBV), particularly the latency gene LMP1, in the development of nasopharyngeal carcinoma (NPC). While no link was found between transient LMP1 expression and oxidative stress in normal nasopharyngeal cells, our findings suggest that stable *LMP1* expression may contribute to oxidative stress, apoptosis, and chromosomal instability. Specifically, *LMP1* is associated with caspase activity and chromosomal breakage within the SAR region of the AF9 gene, a known deletion site in NPC. These results indicate that *LMP1*-driven oxidative stress and apoptosis may promote chromosome rearrangements, contributing to the pathogenesis of NPC. This provides new insights into the molecular mechanisms linking EBV infection to NPC and highlights potential targets for further investigation or therapeutic intervention.

REFERENCES

Ahsan, N., Kanda, T., Nagashima, K., & Takada, K. (2005). Epstein-Barr virus transforming protein LMP1 plays a critical role in virus production. *Journal of Virology*, *79*(7), 4415-4424.

Al-Thawadi, H., Gupta, I., Jabeen, A., Skenderi, F., Aboulkassim, T., Yasmeen, A., Malki, M. I., Batist, G., Vranic, S., & Al Moustafa, A.-E. (2020). Co-presence of human papillomaviruses and Epstein–Barr virus is linked with advanced tumor stage: a tissue microarray study in head and neck cancer patients. *Cancer Cell International*, *20*, 1-13.

Alenzi, F. Q., Lotfy, M., & Wyse, R. (2010). Swords of cell death: caspase activation and regulation. *Asian Pac J Cancer Prev*, *11*(2), 271-280.

Alonso, C. N., Longo, P. L. R., Gallego, M. S., Medina, A., & Felice, M. S. (2008). A novel AF9 breakpoint in MLL-AF9-positive acute monoblastic leukemia. *Pediatric Blood & Cancer*, *50*(4), 869-871.

Ames, B. N. (1989). Endogenous oxidative DNA damage, aging, and cancer. *Free radical research communications*, *7*(3-6), 121-128.

Arfaeinia, H., Ghaemi, M., Jahantigh, A., Soleimani, F., & Hashemi, H. (2023). Secondhand and thirdhand smoke: a review on chemical contents, exposure routes, and protective strategies. *Environmental Science and Pollution Research*, *30*(32), 78017-78029.

Arvey, A., Tempera, I., Tsai, K., Chen, H.-S., Tikhmyanova, N., Klichinsky, M., Leslie, C., & Lieberman, P. M. (2012). An atlas of the Epstein-Barr virus transcriptome and epigenome reveals host-virus regulatory interactions. *Cell host & microbe*, *12*(2), 233-245.

Baer, R., Bankier, A. T., Biggin, M. D., Deininger, P. L., Farrell, P. J., Gibson, T. J., Hatfull, G., Hudson, G. S., Satchwell, S. C., Seguin, C., & et al. (1984). DNA sequence and expression of the B95-8 Epstein-Barr virus genome. *Nature*, *310*(5974), 207-211. http://www.ncbi.nlm.nih.gov/pubmed/6087149

Banko, A., Miljanovic, D., Lazarevic, I., & Cirkovic, A. (2021). A systematic review of Epstein–Barr virus latent membrane protein 1 (LMP1) gene variants in nasopharyngeal carcinoma. *Pathogens*, *10*(8), 1057.

Barboro, P., Repaci, E., D'Arrigo, C., & Balbi, C. (2012). The role of nuclear matrix proteins binding to matrix attachment regions (Mars) in prostate cancer cell differentiation. *PloS one*, *7*(7), e40617.

Bass, D., Parce, J. W., Dechatelet, L. R., Szejda, P., Seeds, M., & Thomas, M. (1983). Flow cytometric studies of oxidative product formation by neutrophils: a graded response to membrane stimulation. *Journal of Immunology (Baltimore, Md.: 1950)*, *130*(4), 1910-1917.

Berezney, R., Mortillaro, M. J., Ma, H., Wei, X., & Samarabandu, J. (1996). The nuclear matrix: a structural milieu for genomic function. *International review of cytology*, *162*, 1-65.

Bernheim, A., Rousselet, G., Massaad, L., Busson, P., & Tursz, T. (1993). Cytogenetic studies in three xenografted nasopharyngeal carcinomas. *Cancer Genet Cytogenet*, *66*(1), 11-15.

Betti, C. J., Villalobos, M. J., Diaz, M. O., & Vaughan, A. T. (2003). Apoptotic stimuli initiate MLL-AF9 translocations that are transcribed in cells capable of division. *Cancer Res*, *63*(6), 1377-1381.

Bhadra, K. (2022). A mini review on molecules inducing caspase-independent cell death: A new route to cancer therapy. *Molecules*, *27*(19), 6401.

Bode, J., Benham, C., Ernst, E., Knopp, A., Marschalek, R., Strick, R., & Strissel, P. (2000). Fatal connections: when DNA ends meet on the nuclear matrix. *Journal of Cellular Biochemistry*, *79*(S35), 3-22.

Bonner, M. Y., & Arbiser, J. L. (2014). The antioxidant paradox: what are antioxidants and how should they be used in a therapeutic context for cancer. *Future medicinal chemistry*, *6*(12), 1413-1422.

Boon, S. S., & Sim, S.-P. (2015a). Inhibitor of caspase-activated DNase expression enhances caspase-activated DNase expression and inhibits oxidative stress-induced chromosome breaks at the mixed lineage leukaemia gene in nasopharyngeal carcinoma cells. *Cancer Cell International*, *15*(1), 1-9.

Boon, S. S., & Sim, S.-P. (2015b). Inhibitor of caspase-activated DNase expression enhances caspase-activated DNase expression and inhibits oxidative stress-induced chromosome breaks at the mixed lineage leukaemia gene in nasopharyngeal carcinoma cells [journal article]. *Cancer Cell International*, *15*(1), 54. https://doi.org/10.1186/s12935-015-0205-1

Bouvard, V., Baan, R., Straif, K., Grosse, Y., Secretan, B., El Ghissassi, F., Benbrahim-Tallaa, L., Guha, N., Freeman, C., Galichet, L., & Cogliano, V. (2009). A review of human carcinogens--Part B: biological agents. *Lancet Oncol*, *10*(4), 321-322.

Bray, F., Colombet, M., Aitken, J., Bardot, A., Eser, S., Galceran, J., Hagenimana, M., Matsuda, T., Mery, L., & Piñeros, M. (2024). Cancer Incidence in Five Continents, Vol. XII (IARC CancerBase No. 19). Lyon: International Agency for Research on Cancer. In.

Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., & Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*, *68*(6), 394-424.

Bray, F., Haugen, M., Moger, T. A., Tretli, S., Aalen, O. O., & Grotmol, T. (2008). Age-incidence curves of nasopharyngeal carcinoma worldwide: bimodality in low-risk populations and aetiologic implications. *Cancer Epidemiology Biomarkers & Prevention*, *17*(9), 2356-2365.

Brocqueville, G., Ndour, P. A., Ouk, T.-S., Le Goff, A., De Witte, C., Mougel, A., Coll, J., Fafeur, V., Le Bourhis, X., & Adriaenssens, E. (2013). LMP1-induced cell death may contribute to the emergency of its oncogenic property. *PloS one*, *8*(4), e60743.

Broeker, P., Super, H. G., Thirman, M. J., Pomykala, H., Yonebayashi, Y., Tanabe, S., Zeleznik-Le, N., & Rowley, J. D. (1996). Distribution of 11q23 breakpoints within the MLL breakpoint cluster region in de novo acute leukemia and in treatment-related acute myeloid leukemia: correlation with scaffold attachment regions and topoisomerase II consensus binding sites.

Cao, S.-M., Chen, S.-H., Qian, C.-N., Liu, Q., & Xia, Y.-F. (2014). Familial nasopharyngeal carcinomas possess distinguished clinical characteristics in southern China. *Chinese Journal of Cancer Research*, *26*(5), 543.

Cao, Y. (2017). EBV based cancer prevention and therapy in nasopharyngeal carcinoma. *NPJ precision oncology*, *1*(1), 10.

Capco, D. G., Wan, K. M., & Penman, S. (1982). The nuclear matrix: three-dimensional architecture and protein composition. *cell*, *29*(3), 847-858.

Cerimele, F., Battle, T., Lynch, R., Frank, D. A., Murad, E., Cohen, C., Macaron, N., Sixbey, J., Smith, K., & Watnick, R. S. (2005). Reactive oxygen signaling and MAPK activation distinguish Epstein–Barr Virus (EBV)-positive versus EBV-negative Burkitt's lymphoma. *Proceedings of the National Academy of Sciences*, *102*(1), 175-179.

Chan, A. S. C., To, K. F., Lo, K. W., Ding, M., Li, X., Johnson, P., & Huang, D. P. (2002). Frequent chromosome 9p losses in histologically normal nasopharyngeal epithelia from southern Chinese. *Int J Cancer*, *102*(3), 300-303.

Chang, E. T., & Adami, H. O. (2006). The enigmatic epidemiology of nasopharyngeal carcinoma. *Cancer Epidemiol Biomarkers Prev*, *15*(10), 1765-1777. https://doi.org/10.1158/1055-9965.EPI-06-0353

Chang, E. T., Liu, Z., Hildesheim, A., Liu, Q., Cai, Y., Zhang, Z., Chen, G., Xie, S.-H., Cao, S.-M., & Shao, J.-Y. (2017). Active and passive smoking and risk of nasopharyngeal carcinoma: a population-based case-control study in southern China. *American journal of epidemiology*, *185*(12), 1272-1280.

Cheerathodi, M. R., & Meckes Jr, D. G. (2018). The Epstein–Barr virus LMP1 interactome: Biological implications and therapeutic targets. *Future virology*, *13*(12), 863-887.

Chen, W., Zheng, R., Baade, P. D., Zhang, S., Zeng, H., Bray, F., Jemal, A., Yu, X. Q., & He, J. (2016). Cancer statistics in China, 2015. *CA Cancer J Clin*, *66*(2), 115-132.

Chen, X., Kamranvar, S. A., & Masucci, M. (2016). Oxidative stress enables Epstein–Barr virus-induced B-cell transformation by posttranscriptional regulation of viral and cellular growth-promoting factors. *Oncogene*, *35*(29), 3807-3816.

Chen, Y.-P., Chan, A. T., Le, Q.-T., Blanchard, P., Sun, Y., & Ma, J. (2019). Nasopharyngeal carcinoma. *The lancet*, *394*(10192), 64-80.

Chen, Y., Chang, E. T., Liu, Z., Liu, Q., Cai, Y., Zhang, Z., Chen, G., Huang, Q.-H., Xie, S.-H., & Cao, S.-M. (2021). Residence characteristics and risk of nasopharyngeal carcinoma in southern China: A population-based case-control study. *Environment international*, *151*, 106455.

Chien, G., Yuen, P. W., Kwong, D., & Kwong, Y. L. (2001). Comparative genomic hybridization analysis of nasopharygeal carcinoma:: consistent patterns of genetic aberrations and clinicopathological correlations. *Cancer Genet Cytogenet*, *126*(1), 63-67.

Chien, Y.-C., & Chen, C.-J. (2003). Epidemiology and Etiology of Nasopharyngeal Carcinoma: Gene-Environment Interaction. *Cancer Reviews: Asia-Pacific*, *01*(01), 1-19. https://doi.org/10.1142/s0219836303000074

Cho, E.-Y., Hildesheim, A., Chen, C.-J., Hsu, M.-M., Chen, I.-H., Mittl, B. F., Levine, P. H., Liu, M.-Y., Chen, J.-Y., & Brinton, L. A. (2003). Nasopharyngeal carcinoma and genetic polymorphisms of DNA repair enzymes XRCC1 and hOGG1. *Cancer Epidemiology and Prevention Biomarkers*, *12*(10), 1100-1104.

Choi, I.-K., Wang, Z., Ke, Q., Hong, M., Paul Jr, D. W., Fernandes, S. M., Hu, Z., Stevens, J., Guleria, I., & Kim, H.-J. (2021). Mechanism of EBV inducing anti-tumour immunity and its therapeutic use. *Nature*, *590*(7844), 157-162.

Choudhari, S. K., Chaudhary, M., Gadbail, A. R., Sharma, A., & Tekade, S. (2014). Oxidative and antioxidative mechanisms in oral cancer and precancer: a review. *Oral Oncol*, *50*(1), 10-18.

Clarke, J. D., Hsu, A., Yu, Z., Dashwood, R. H., & Ho, E. (2011). Differential effects of sulforaphane on histone deacetylases, cell cycle arrest and apoptosis in normal prostate cells versus hyperplastic and cancerous prostate cells. *Molecular nutrition & food research*, *55*(7), 999-1009.

Cockerill, P. N., & Garrard, W. T. (1986). Chromosomal loop anchorage of the kappa immunoglobulin gene occurs next to the enhancer in a region containing topoisomerase II sites. *cell*, *44*(2), 273-282.

Corral, J., Lavenir, I., Impey, H., Warren, A. J., Forster, A., Larson, T. A., Bell, S., McKenzie, A. N., King, G., & Rabbitts, T. H. (1996). An Mll–AF9 fusion gene made by homologous recombination causes acute leukemia in chimeric mice: a method to create fusion oncogenes. *cell*, *85*(6), 853-861.

Croft, J. A., Bridger, J. M., Boyle, S., Perry, P., Teague, P., & Bickmore, W. A. (1999). Differences in the localization and morphology of chromosomes in the human nucleus. *The Journal of cell biology*, *145*(6), 1119-1131.

Dawson, C. W., Port, R. J., & Young, L. S. (2012). The role of the EBV-encoded latent membrane proteins LMP1 and LMP2 in the pathogenesis of nasopharyngeal carcinoma (NPC). *Semin Cancer Biol*, *22*(2), 144-153. https://doi.org/10.1016/j.semcancer.2012.01.004

de Martel, C., Ferlay, J., Franceschi, S., Vignat, J., Bray, F., Forman, D., & Plummer, M. (2012). Global burden of cancers attributable to infections in 2008: a review and synthetic analysis. *Lancet Oncol*, *13*(6), 607-615. https://doi.org/10.1016/S1470-2045(12)70137-7

Deng, W., Pang, P. S., Tsang, C. M., Hau, P. M., Yip, Y. L., Cheung, A. L., & Tsao, S. W. (2012). Epstein-Barr virus-encoded latent membrane protein 1 impairs G2 checkpoint in human nasopharyngeal epithelial cells through defective Chk1 activation. *PloS one*, *7*(6), e39095.

Dobson, C., Warren, A., Pannell, R., Forster, A., Lavenir, I., Corral, J., Smith, A., & Rabbitts, T. (1999). The Mll–AF9 gene fusion in mice controls myeloproliferation and specifies acute myeloid leukaemogenesis. *The EMBO Journal*, *18*(13), 3564-3574.

Drennan, K. J., Linnemann, A. K., Platts, A. E., Heng, H. H., Armant, D. R., & Krawetz, S. A. (2010). Nuclear matrix association: switching to the invasive cytotrophoblast. *Placenta*, *31*(5), 365-372.

Dutton, A., Woodman, C. B., Chukwuma, M. B., Last, J. I., Wei, W., Vockerodt, M., Baumforth, K. R., Flavell, J. R., Rowe, M., Taylor, A. M., Young, L. S., & Murray, P. G. (2007). Bmi-1 is induced by the Epstein-Barr virus oncogene LMP1 and regulates the expression of viral target genes in Hodgkin lymphoma cells. *Blood*, *109*(6), 2597-2603. https://doi.org/10.1182/blood-2006-05-020545

Earnshaw, W. C., Martins, L. M., & Kaufmann, S. H. (1999). Mammalian caspases: structure, activation, substrates, and functions during apoptosis. *Annual review of biochemistry*, *68*(1), 383-424.

Edwards, R. H., Sitki-Green, D., Moore, D. T., & Raab-Traub, N. (2004). Potential selection of LMP1 variants in nasopharyngeal carcinoma. *Journal of Virology*, *78*(2), 868-881.

Ekburanawat, W., Ekpanyaskul, C., Brennan, P., Kanka, C., Tepsuwan, K., Temiyastith, S., Klinvimol, T., Pongnikorn, S., & Sangrajrang, S. (2010). Evaluation of non-viral risk factors for nasopharyngeal carcinoma in Thailand: results from a case-control study. *Asian Pac J Cancer Prev*, *11*(4), 929-932.

Ellis, R. E., Yuan, J., & Horvitz, H. R. (1991). Mechanisms and functions of cell death. *Annual review of cell biology*, *7*(1), 663-698.

Enari, M., Sakahira, H., Yokoyama, H., Okawa, K., Iwamatsu, A., & Nagata, S. (1998). A caspase-activated DNase that degrades DNA during apoptosis, and its inhibitor ICAD. *Nature*, *391*(6662), 43-50.

Fang, C. Y., Lee, C. H., Wu, C. C., Chang, Y. T., Yu, S. L., Chou, S. P., Huang, P. T., Chen, C. L., Hou, J. W., Chang, Y., Tsai, C. H., Takada, K., & Chen, J. Y. (2009). Recurrent chemical reactivations of EBV promotes genome instability and enhances tumor progression of nasopharyngeal carcinoma cells. *Int J Cancer*, *124*(9), 2016-2025. https://doi.org/10.1002/ijc.24179

Feng, J., Chen, Q., Zhang, P., Huang, X., Xie, W., Zhang, H., & Yao, P. (2021). Latent membrane Protein 1 promotes tumorigenesis through upregulation of PGC1β signaling pathway. *Stem Cell Reviews and Reports*, 1-14.

Ferlay, J., Ervik, M., Lam, F., Laversanne, M., Colombet, M., Mery, L., Piñeros, M., Znaor, A., Soerjomataram, I., & Bray, F. (2024). Global cancer observatory: cancer today (version 1.1). Lyon, France: International Agency for Research on Cancer; 2024. In.

Filipski, J., Leblanc, J., Youdale, T., Sikorska, M., & Walker, P. (1990). Periodicity of DNA folding in higher order chromatin structures. *The EMBO Journal*, *9*(4), 1319-1327.

Frazer, K. A., Murray, S. S., Schork, N. J., & Topol, E. J. (2009). Human genetic variation and its contribution to complex traits. *Nature Reviews Genetics*, *10*(4), 241-251.

Fulda, S., Galluzzi, L., & Kroemer, G. (2010). Targeting mitochondria for cancer therapy. *Nature reviews Drug discovery*, *9*(6), 447-464.

Gambotto, A., Dworacki, G., Cicinnati, V., Kenniston, T., Steitz, J., Tüting, T., Robbins, P., & DeLeo, A. (2000). Immunogenicity of enhanced green fluorescent protein (EGFP) in BALB/c mice: identification of an H2-Kd-restricted CTL epitope. *Gene therapy*, *7*(23), 2036-2040.

Ganini, D., Leinisch, F., Kumar, A., Jiang, J., Tokar, E. J., Malone, C. C., Petrovich, R. M., & Mason, R. P. (2017). Fluorescent proteins such as eGFP lead to catalytic oxidative stress in cells. *Redox biology*, *12*, 462-468.

Göhring, F., Schwab, B. L., Nicotera, P., Leist, M., & Fackelmayer, F. O. (1997). The novel SAR-binding domain of scaffold attachment factor A (SAF-A) is a target in apoptotic nuclear breakdown. *The EMBO Journal*, *16*(24), 7361-7371.

Goto, H., Yang, B., Petersen, D., Pepper, K. A., Alfaro, P. A., Kohn, D. B., & Reynolds, C. P. (2003). Transduction of green fluorescent protein increased oxidative stress and enhanced sensitivity to cytotoxic drugs in neuroblastoma cell lines. *Molecular cancer therapeutics*, *2*(9), 911-917.

Gourzones, C., Busson, P., & Raab-Traub, N. (2013). Epstein-Barr virus and the pathogenesis of nasopharyngeal carcinomas. In *Nasopharyngeal Carcinoma* (pp. 42-60). Springer.

Gruhne, B., Sompallae, R., Marescotti, D., Kamranvar, S. A., Gastaldello, S., & Masucci, M. G. (2009). The Epstein–Barr virus nuclear antigen-1 promotes genomic instability via induction of reactive oxygen species. *Proceedings of the National Academy of Sciences*, *106*(7), 2313-2318.

Gruhne, B., Sompallae, R., & Masucci, M. G. (2009). Three Epstein-Barr virus latency proteins independently promote genomic instability by inducing DNA damage, inhibiting DNA repair and inactivating cell cycle checkpoints. *Oncogene*, *28*(45), 3997-4008. https://doi.org/10.1038/onc.2009.258

GunvÉN, P., Klein, G., Henle, G., Henle, W., & Clifford, P. (1970). Epstein–Barr Virus in Burkitt's Lymphoma and Nasopharyngeal Carcinoma: Antibodies to EBV associated Membrane and Viral Capsid Antigens in Burkitt Lymphoma Patients. *Nature*, *228*(5276), 1053-1056. https://doi.org/10.1038/2281053a0

Guo, R., Mao, Y.-P., Tang, L.-L., Chen, L., Sun, Y., & Ma, J. (2019). The evolution of nasopharyngeal carcinoma staging. *The British journal of radiology*, *92*(1102), 20190244.

Hammerschmidt, W., Sugden, B., & Baichwal, V. R. (1989). The transforming domain alone of the latent membrane protein of Epstein-Barr virus is toxic to cells when expressed at high levels. *Journal of Virology*, *63*(6), 2469-2475.

Hanahan, D. (2022). Hallmarks of cancer: new dimensions. *Cancer discovery*, *12*(1), 31-46.

Harraghy, N., Buceta, M., Regamey, A., Girod, P.-A., & Mermod, N. (2012). Using matrix attachment regions to improve recombinant protein production. *Protein Expression in Mammalian Cells: Methods and Protocols*, 93-110.

Hars, E. S., Lyu, Y. L., Lin, C.-P., & Liu, L. F. (2006). Role of apoptotic nuclease caspase-activated DNase in etoposide-induced treatment-related acute myelogenous leukemia. *Cancer Res*, *66*(18), 8975-8979.

Heng, H. H., Goetze, S., Ye, C. J., Liu, G., Stevens, J. B., Bremer, S. W., Wykes, S. M., Bode, J., & Krawetz, S. A. (2004). Chromatin loops are selectively anchored using scaffold/matrix-attachment regions. *Journal of cell science*, *117*(7), 999-1008.

Hengartner, M. O. (2000). The biochemistry of apoptosis. *Nature*, *407*(6805), 770-776.

Hoeijmakers, J. H. (2001). Genome maintenance mechanisms for preventing cancer. *Nature*, *411*(6835), 366-374.

Hong, G. K., Kumar, P., Wang, L., Damania, B., Gulley, M. L., Delecluse, H.-J., Polverini, P. J., & Kenney, S. C. (2005). Epstein-Barr virus lytic infection is required for efficient production of the angiogenesis factor vascular endothelial growth factor in lymphoblastoid cell lines. *Journal of Virology*, *79*(22), 13984-13992.

Hu, C., Wei, W., Chen, X., Woodman, C. B., Yao, Y., Nicholls, J. M., Joab, I., Sihota, S. K., Shao, J.-Y., & Derkaoui, K. D. (2012). A global view of the oncogenic landscape in nasopharyngeal carcinoma: an integrated analysis at the genetic and expression levels. *PloS one*, *7*(7).

Hu, J., Li, H., Luo, X., Li, Y., Bode, A., & Cao, Y. (2017). The role of oxidative stress in EBV lytic reactivation, radioresistance and the potential preventive and therapeutic implications. *Int J Cancer*, *141*(9), 1722-1729.

Hu, J., Li, Y., Li, H., Shi, F., Xie, L., Zhao, L., Tang, M., Luo, X., Jia, W., & Fan, J. (2020). Targeting Epstein-Barr virus oncoprotein LMP1-mediated high oxidative stress suppresses EBV lytic reactivation and sensitizes tumors to radiation therapy. *Theranostics*, *10*(26), 11921.

Hu, T., Lin, C. Y., Xie, S. H., Chen, G. H., Lu, Y. Q., Ling, W., Huang, Q. H., Liu, Q., & Cao, S. M. (2019). Smoking can increase nasopharyngeal carcinoma risk by repeatedly reactivating Epstein-Barr Virus: An analysis of a prospective study in southern China. *Cancer Medicine*, *8*(5), 2561-2571.

Huang, D. P., Lo, K. W., Choi, P. H., Ng, A. Y., Tsao, S. Y., Yiu, G. K., & Lee, J. C. (1991). Loss of heterozygosity on the short arm of chromosome 3 in nasopharyngeal carcinoma. *Cancer Genet Cytogenet*, *54*(1), 91-99.

Huang, P.-Y., Zeng, T.-T., Li, M.-Q., Ban, X., Zhu, Y.-H., Zhang, B.-Z., Mai, H.-Q., Zhang, L., Guan, X.-Y., & Li, Y. (2015). Proteomic analysis of a nasopharyngeal carcinoma cell line and a nasopharyngeal epithelial cell line. *Tumori Journal*, *101*(6), 676-683.

Huang, S. Y., Fang, C. Y., Tsai, C. H., Chang, Y., Takada, K., Hsu, T. Y., & Chen, J. Y. (2010). N-methyl-N'-nitro-N-nitrosoguanidine induces and cooperates with 12-O-tetradecanoylphorbol-1,3-acetate/sodium butyrate to enhance Epstein-Barr virus reactivation and genome instability in nasopharyngeal carcinoma cells. *Chem Biol Interact*, *188*(3), 623-634. https://doi.org/10.1016/j.cbi.2010.09.020

Huang, W.-Y., Lin, C.-C., Jen, Y.-M., Lin, K.-T., Yang, M.-H., Chen, C.-M., Chang, Y.-N., Sung, F.-C., & Kao, C.-H. (2012). Association between adult otitis media and nasopharyngeal cancer: a nationwide population-based cohort study. *Radiotherapy and Oncology*, *104*(3), 338-342.

Hung, S.-H., Yang, T.-H., Cheng, Y.-F., Chen, C.-S., & Lin, H.-C. (2023). Association of nasopharynx cancer with human papillomavirus infections. *Cancers*, *15*(16), 4082.

Jalouli, J., Ibrahim, S. O., Mehrotra, R., Jalouli, M. M., Sapkota, D., Larsson, P.-A., & Hirsch, J.-M. (2010). Prevalence of viral (HPV, EBV, HSV) infections in oral submucous fibrosis and oral cancer from India. *Acta oto-laryngologica*, *130*(11), 1306-1311.

Jalouli, J., Jalouli, M. M., Sapkota, D., Ibrahim, S. O., Larsson, P.-a., & Sand, L. (2012). Human papilloma virus, herpes simplex virus and epstein barr virus in oral squamous cell carcinoma from eight different countries. *Anticancer research*, *32*(2), 571-580.

Jeffs, A. R., Benjes, S. M., Smith, T. L., Sowerby, S. J., & Morris, C. M. (1998). The BCR gene recombines preferentially with Alu elements in complex BCR-ABL translocations of chronic myeloid leukaemia. *Human molecular genetics*, *7*(5), 767-776.

Jelic, M. D., Mandic, A. D., Maricic, S. M., & Srdjenovic, B. U. (2021). Oxidative stress and its role in cancer. *Journal of cancer research and therapeutics*, *17*(1), 22-28.

Jemal, A., Bray, F., Center, M. M., Ferlay, J., Ward, E., & Forman, D. (2011). Global cancer statistics. *CA Cancer J Clin*, *61*(2), 69-90. https://doi.org/10.3322/caac.20107

Jen, C.-W., Tsai, Y.-C., Wu, J.-S., Chen, P.-L., Yen, J.-H., Chuang, W.-K., & Cheng, S. H.-C. (2020). Prognostic classification for patients with nasopharyngeal carcinoma based on American Joint Committee on Cancer staging system T and N categories. *Therapeutic Radiology and Oncology*, *4*.

Jia, W.-H., & Qin, H.-D. (2012). Non-viral environmental risk factors for nasopharyngeal carcinoma: a systematic review. Semin Cancer Biol,

Jia, W. H., & Qin, H. D. (2012). Non-viral environmental risk factors for nasopharyngeal carcinoma: a systematic review. *Semin Cancer Biol*, *22*(2), 117-126. https://doi.org/10.1016/j.semcancer.2012.01.009

Jing, Y., Dai, J., Chalmers-Redman, R. M., Tatton, W. G., & Waxman, S. (1999). Arsenic trioxide selectively induces acute promyelocytic leukemia cell apoptosis via a hydrogen peroxide-dependent pathway. *Blood, the Journal of the American Society of Hematology*, *94*(6), 2102-2111.

Jiromaru, R., Nakagawa, T., & Yasumatsu, R. (2022). Advanced nasopharyngeal carcinoma: current and emerging treatment options. *Cancer Management and Research*, 2681-2689.

Kalyanaraman, B., & Zielonka, J. (2017). Green fluorescent proteins induce oxidative stress in cells: A worrisome new wrinkle in the application of the GFP reporter system to biological systems? *Redox biology*, *12*, 755.

Karpińska, A., & Gromadzka, G. (2013). Oxidative stress and natural antioxidant mechanisms: the role in neurodegeneration. From molecular mechanisms to

therapeutic strategies. *Postepy higieny i medycyny doswiadczalnej (Online)*, *67*, 43-53.

Katakwar, P., Metgud, R., Naik, S., & Mittal, R. (2016). Oxidative stress marker in oral cancer: a review. *Journal of cancer research and therapeutics*, *12*(2), 438.

Kenney, S. C., & Mertz, J. E. (2014). Regulation of the latent-lytic switch in Epstein–Barr virus. Semin Cancer Biol,

Kerr, J. F., Wyllie, A. H., & Currie, A. R. (1972). Apoptosis: a basic biological phenomenon with wideranging implications in tissue kinetics. *British journal of cancer*, *26*(4), 239-257.

Kesavardhana, S., Malireddi, R. S., & Kanneganti, T.-D. (2020). Caspases in cell death, inflammation, and pyroptosis. *Annual review of immunology*, *38*, 567-595.

Kgatle, M. M., Spearman, C. W., Kalla, A. A., & Hairwadzi, H. N. (2017). DNA oncogenic virus-induced oxidative stress, genomic damage, and aberrant epigenetic alterations. *Oxidative medicine and cellular longevity*, *2017*.

Kim, B. M., Hong, Y., Lee, S., Liu, P., Lim, J. H., Lee, Y. H., Lee, T. H., Chang, K. T., & Hong, Y. (2015). Therapeutic implications for overcoming radiation resistance in cancer therapy. *International journal of molecular sciences*, *16*(11), 26880-26913.

Klaunig, J. E., & Kamendulis, L. M. (2004). The role of oxidative stress in carcinogenesis. *Annu. Rev. Pharmacol. Toxicol.*, *44*(1), 239-267.

Koul, M., Kumar, A., Deshidi, R., Sharma, V., Singh, R. D., Singh, J., Sharma, P. R., Shah, B. A., Jaglan, S., & Singh, S. (2017). Cladosporol A triggers apoptosis sensitivity by ROS-mediated autophagic flux in human breast cancer cells. *BMC cell biology*, *18*(1), 1-15.

Kulwichit, W., Edwards, R. H., Davenport, E. M., Baskar, J. F., Godfrey, V., & Raab-Traub, N. (1998). Expression of the Epstein–Barr virus latent membrane protein 1 induces B cell lymphoma in transgenic mice. *Proceedings of the National Academy of Sciences*, *95*(20), 11963-11968.

Kutok, J. L., & Wang, F. (2006). Spectrum of Epstein-Barr virus-associated diseases. *Annu Rev Pathol*, *1*, 375-404. https://doi.org/10.1146/annurev.pathol.1.110304.100209

Laemmli, U. K., Käs, E., Poljak, L., & Adachi, Y. (1992). Scaffold-associated regions: cis-acting determinants of chromatin structural loops and functional domains. *Current opinion in genetics & development*, *2*(2), 275-285.

Lagarkova, M. A., Iarovaia, O. V., & Razin, S. V. (1995). Large-scale fragmentation of mammalian DNA in the course of apoptosis proceeds via excision of chromosomal DNA loops and their oligomers. *Journal of Biological Chemistry*, *270*(35), 20239-20241.

Lai, C., Zhang, C., Lv, H., Huang, H., Ke, X., Zhou, C., Chen, H., Chen, S., & Zhou, L. (2021). A novel prognostic model predicts overall survival in patients with nasopharyngeal carcinoma based on clinical features and blood biomarkers. *Cancer Medicine*, *10*(11), 3511-3523.

Larsen, B. D., Rampalli, S., Burns, L. E., Brunette, S., Dilworth, F. J., & Megeney, L. A. (2010). Caspase 3/caspase-activated DNase promote cell differentiation by inducing DNA strand breaks. *Proceedings of the National Academy of Sciences*, *107*(9), 4230-4235.

Larsen, B. D., & Sørensen, C. S. (2017). The caspase-activated DN ase: apoptosis and beyond. *The FEBS journal*, *284*(8), 1160-1170.

Lechardeur, D., Xu, M., & Lukacs, G. L. (2004). Contrasting nuclear dynamics of the caspase-activated DNase (CAD) in dividing and apoptotic cells. *The Journal of cell biology*, *167*(5), 851-862.

Legault, J., Tremblay, A., & Mirault, M.-E. (1997). Preferential localization of DNA damage induced by depurination and bleomycin in a plasmid containing a scaffold-associated region. *Biochemistry and cell biology*, *75*(4), 369-375.

Lelli, J. L., Jr., Becks, L. L., Dabrowska, M. I., & Hinshaw, D. B. (1998). ATP converts necrosis to apoptosis in oxidant-injured endothelial cells. *Free Radic Biol Med*, *25*(6), 694-702.

Lin, J. H., Wen, C. P., Jiang, C. Q., Yuan, J.-M., Chen, C. J., Ho, S. Y., Gao, W., Zhang, W., Wang, R., & Chien, Y.-C. (2021). Smoking and nasopharyngeal cancer: individual data meta-analysis of six prospective studies on 334 935 men. *International Journal of Epidemiology*, *50*(3), 975-986.

Lin, X., Liu, S., Luo, X., Ma, X., Guo, L., Li, L., Li, Z., Tao, Y., & Cao, Y. (2009). EBV-encoded LMP1 regulates Op18/stathmin signaling pathway by cdc2 mediation in nasopharyngeal carcinoma cells. *Int J Cancer*, *124*(5), 1020-1027.

Liu, H.-S., Jan, M.-S., Chou, C.-K., Chen, P.-H., & Ke, N.-J. (1999). Is green fluorescent protein toxic to the living cells? *Biochemical and biophysical research communications*, *260*(3), 712-717.

Liu, M. T., Chen, Y. R., Chen, S. C., Hu, C. Y., Lin, C. S., Chang, Y. T., Wang, W. B., & Chen, J. Y. (2004). Epstein-Barr virus latent membrane protein 1 induces micronucleus formation, represses DNA repair and enhances sensitivity to DNA-damaging agents in human epithelial cells. *Oncogene*, *23*(14), 2531-2539. https://doi.org/10.1038/sj.onc.1207375

Liu, S.-C., Wang, C.-I., Liu, T.-T., Tsang, N.-M., Sui, Y.-H., & Juang, J.-L. (2023). A 3-gene signature comprising CDH4, STAT4 and EBV-encoded LMP1 for early diagnosis and predicting disease progression of nasopharyngeal carcinoma. *Discover Oncology*, *14*(1), 119.

Liu, X., Deng, Y., Huang, Y., Ye, J., Xie, S., He, Q., Chen, Y., Lin, Y., Liang, R., & Wei, J. (2022). Nasopharyngeal Carcinoma Progression: Accumulating Genomic Instability and Persistent Epstein–Barr Virus Infection. *Current Oncology*, *29*(9), 6035-6052.

Liu, Y., Wang, X., Lo, A. K., Wong, Y. C., Cheung, A. L., & Tsao, S. W. (2002). Latent membrane protein-1 of Epstein-Barr virus inhibits cell growth and induces sensitivity to cisplatin in nasopharyngeal carcinoma cells. *Journal of medical virology*, *66*(1), 63-69.

Lo, A. K.-F., Dawson, C. W., Lung, H. L., Wong, K.-L., & Young, L. S. (2021). The role of EBV-encoded LMP1 in the NPC tumor microenvironment: from function to therapy. *Frontiers in Oncology*, *11*, 640207.

Lo, A. K., Liu, Y., Wang, X. H., Huang, D. P., Yuen, P. W., Wong, Y. C., & Tsao, G. S. (2003). Alterations of biologic properties and gene expression in nasopharyngeal epithelial cells by the Epstein-Barr virus-encoded latent membrane protein 1. *Lab Invest*, *83*(5), 697-709.

Lo, A. K., Lo, K. W., Ko, C. W., Young, L. S., & Dawson, C. W. (2013). Inhibition of the LKB1-AMPK pathway by the Epstein-Barr virus-encoded LMP1 promotes proliferation and transformation of human nasopharyngeal epithelial cells. *J Pathol*, *230*(3), 336-346. https://doi.org/10.1002/path.4201

Lo, K.-W., Chung, G. T.-Y., & To, K.-F. (2012). Deciphering the molecular genetic basis of NPC through molecular, cytogenetic, and epigenetic approaches. *Semin Cancer Biol*, *22*(2), 79-86. https://doi.org/https://doi.org/10.1016/j.semcancer.2011.12.011

London, A. O., Gallagher, L. W., Sharma, R. K., Spielman, D., Golub, J. S., Overdevest, J. B., Yan, C. H., DeConde, A., & Gudis, D. A. (2022). Impact of race, ethnicity, and socioeconomic status on nasopharyngeal carcinoma disease-specific and conditional survival. *Journal of Neurological Surgery Part B: Skull Base*, *83*(05), 451-460.

Lopez, K. E., & Bouchier-Hayes, L. (2022). Lethal and non-lethal functions of caspases in the DNA damage response. *Cells*, *11*(12), 1887.

Lu, J. J.-Y., Chen, J.-Y., Hsu, T.-Y., Yu, W. C., Su, I.-J., & Yang, C.-S. (1996). Induction of apoptosis in epithelial cells by Epstein—Barr virus latent membrane protein 1. *Journal of General Virology*, *77*(8), 1883-1892.

Ma, N., Kawanishi, M., Hiraku, Y., Murata, M., Huang, G. W., Huang, Y., Luo, D. Z., Mo, W. G., Fukui, Y., & Kawanishi, S. (2008). Reactive nitrogen species-dependent DNA damage in EBV-associated nasopharyngeal carcinoma: the relation to STAT3 activation and EGFR expression. *Int J Cancer*, *122*(11), 2517-2525. https://doi.org/10.1002/ijc.23415

Mani, S., Swargiary, G., & Singh, K. K. (2020). Natural agents targeting mitochondria in cancer. *International journal of molecular sciences*, *21*(19), 6992.

Marsin, S., Vidal, A. E., Sossou, M., Ménissier-de Murcia, J., Le Page, F., Boiteux, S., de Murcia, G., & Radicella, J. P. (2003). Role of XRCC1 in the coordination and stimulation of oxidative DNA damage repair initiated by the DNA glycosylase hOGG1. *Journal of Biological Chemistry*, *278*(45), 44068-44074.

Martin, K. A., Lupey, L. N., & Tempera, I. (2016). Epstein-Barr virus oncoprotein LMP1 mediates epigenetic changes in host gene expression through PARP1. *Journal of Virology*, *90*(19), 8520-8530.

Matés, J. M., Segura, J. A., Alonso, F. J., & Márquez, J. (2012). Oxidative stress in apoptosis and cancer: an update. *Archives of toxicology*, *86*, 1649-1665.

Mattson, M. P. (2000). Apoptosis in neurodegenerative disorders. *Nature reviews Molecular cell biology*, *1*(2), 120-130.

McIlwain, D. R., Berger, T., & Mak, T. W. (2013). Caspase functions in cell death and disease. *Cold Spring Harbor perspectives in biology*, *5*(4), a008656.

Minarovits, J. (2006). Epigenotypes of latent herpesvirus genomes. *Curr Top Microbiol Immunol*, *310*, 61-80.

Mirkovitch, J., Gasser, S. M., & Laemmli, U. K. (1988). Scaffold attachment of DNA loops in metaphase chromosomes. *Journal of molecular biology*, *200*(1), 101-109.

Moore, D., & Dowhan, D. (2002). Purification and concentration of DNA from aqueous solutions. *Current protocols in molecular biology*, *59*(1), 2.1. 1-2.1. 10.

Moorthy, R. K., & Thorley-Lawson, D. A. (1993). All three domains of the Epstein-Barr virus-encoded latent membrane protein LMP-1 are required for transformation of rat-1 fibroblasts. *Journal of Virology*, *67*(3), 1638-1646.

Morales-Sanchez, A., & Fuentes-Panana, E. M. (2014). Human viruses and cancer. *Viruses*, *6*(10), 4047-4079. https://doi.org/10.3390/v6104047

Nicholas, C. P., & Sim, S.-P. (2012). Etoposide-induced apoptosis results in chromosome breaks within the AF9 gene: its implication in chromosome rearrangement in leukaemia. *Advances in Bioscience and Biotechnology*, *3*(06), 686.

Niedobitek, G. (2000). Epstein-Barr virus infection in the pathogenesis of nasopharyngeal carcinoma. *Molecular Pathology*, *53*(5), 248-254. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1186977/

Ning, L., Ko, J. M.-Y., Yu, V. Z., Ng, H. Y., Chan, C. K.-C., Tao, L., Lam, S.-Y., Leong, M. M.-L., Ngan, R. K.-C., & Kwong, D. L.-W. (2020). Nasopharyngeal carcinoma MHC region deep sequencing identifies HLA and novel non-HLA TRIM31 and TRIM 39 loci. *Communications Biology*, *3*(1), 759.

Oberhammer, F., Wilson, J., Dive, C., Morris, I., Hickman, J., Wakeling, A., Walker, P. R., & Sikorska, M. (1993). Apoptotic death in epithelial cells: cleavage of DNA to 300 and/or 50 kb fragments prior to or in the absence of internucleosomal fragmentation. *The EMBO Journal*, *12*(9), 3679-3684.

Okekpa, S. I., Mydin, R. B. S., Mangantig, E., Azmi, N. S. A., Zahari, S. N. S., Kaur, G., & Musa, Y. (2019). Nasopharyngeal carcinoma (NPC) risk factors: a systematic review and meta-analysis of the association with lifestyle, diets, socioeconomic and sociodemographic in Asian region. *Asian Pacific Journal of Cancer Prevention: APJCP*, *20*(11), 3505.

Pagano, J. S. (1999). Epstein-Barr virus: the first human tumor virus and its role in cancer. *Proc Assoc Am Physicians*, *111*(6), 573-580.

Pathmanathan, R., Prasad, U., Sadler, R., Flynn, K., & Raab-Traub, N. (1995). Clonal proliferations of cells infected with Epstein-Barr virus in preinvasive lesions related to nasopharyngeal carcinoma. *N Engl J Med*, *333*(11), 693-698. https://doi.org/10.1056/nejm199509143331103

Pession, A., Martino, V., Tonelli, R., Beltramini, C., Locatelli, F., Biserni, G., Franzoni, M., Freccero, F., Montemurro, L., & Pattacini, L. (2003). MLL-AF9 oncogene expression affects cell growth but not terminal differentiation and is downregulated during monocyte–macrophage maturation in AML-M5 THP-1 cells. *Oncogene*, *22*(54), 8671-8676.

Prabhu, S. R., & Wilson, D. F. (2016). Evidence of Epstein-Barr virus association with head and neck cancers: a review. *Journal of the Canadian Dental Association*, *82*, 1-11.

Price, A. M., & Luftig, M. A. (2014). Dynamic Epstein–Barr virus gene expression on the path to B-cell transformation. *Advances in virus research*, *88*, 279-313.

Rada, B., & Leto, T. L. (2008). Oxidative innate immune defenses by Nox/Duox family NADPH oxidases. In *Trends in Innate Immunity* (Vol. 15, pp. 164-187). Karger Publishers.

Rancan, C., Schirrmann, L., Hüls, C., Zeidler, R., & Moosmann, A. (2015). Latent membrane protein LMP2A impairs recognition of EBV-infected cells by CD8+ T cells. *PLoS pathogens*, *11*(6), e1004906.

Rowe, M., Lear, A., Croom-Carter, D., Davies, A., & Rickinson, A. (1992). Three pathways of Epstein-Barr virus gene activation from EBNA1-positive latency in B lymphocytes. *Journal of Virology*, *66*(1), 122-131.

Sakahira, H., Enari, M., & Nagata, S. (1998). Cleavage of CAD inhibitor in CAD activation and DNA degradation during apoptosis. *Nature*, *391*(6662), 96-99.

Sakahira, H., Enari, M., Ohsawa, Y., Uchiyama, Y., & Nagata, S. (1999). Apoptotic nuclear morphological change without DNA fragmentation. *Current Biology*, *9*(10), 543-546.

Sarkar, S., Tribedi, P., & Bhadra, K. (2022). Structure-activity insights of harmine targeting DNA, ROS inducing cytotoxicity with PARP mediated apoptosis against cervical

cancer, anti-biofilm formation and in vivo therapeutic study. *Journal of Biomolecular Structure and Dynamics*, *40*(13), 5880-5902.

Shanmugaratnam, K. (1978). Histological typing of nasopharyngeal carcinoma. *IARC scientific publications*(20), 3-12.

Shannon-Lowe, C., & Rickinson, A. (2019). The global landscape of EBV-associated tumors. *Frontiers in Oncology*, *9*, 713.

Shao, J.-Y., Wang, H.-Y., Huang, X.-M., Feng, Q.-S., Huang, P., Feng, B.-J., Huang, L.-X., Yu, X.-J., Li, J.-T., & Hu, L.-F. (2000). Genome-wide allelotype analysis of sporadic primary nasopharyngeal carcinoma from southern China. *International journal of oncology*, *17*(6), 1267-1342.

Shao, S., Scholtz, L. U., Gendreizig, S., Martínez-Ruiz, L., Florido, J., Escames, G., Schürmann, M., Hain, C., Hose, L., & Mentz, A. (2023). Primary head and neck cancer cell cultures are susceptible to proliferation of Epstein-Barr virus infected lymphocytes. *BMC cancer*, *23*(1), 47.

Shi, F., Zhou, M., Shang, L., Du, Q., Li, Y., Xie, L., Liu, X., Tang, M., Luo, X., & Fan, J. (2019). EBV (LMP1)-induced metabolic reprogramming inhibits necroptosis through the hypermethylation of the RIP3 promoter. *Theranostics*, *9*(9), 2424.

Shimizu, T., Maeno, E., & Okada, Y. (2007). Prerequisite role of persistent cell shrinkage in apoptosis of human epithelial cells. *ACTA PHYSIOLOGICA SINICA-CHINESE EDITION-*, *59*(4), 512.

Shumilov, A., Tsai, M.-H., Schlosser, Y. T., Kratz, A.-S., Bernhardt, K., Fink, S., Mizani, T., Lin, X., Jauch, A., & Mautner, J. (2017). Epstein–Barr virus particles induce centrosome amplification and chromosomal instability. *Nature communications*, *8*(1), 1-15.

Siak, P. Y., Khoo, A. S.-B., Leong, C. O., Hoh, B.-P., & Cheah, S.-C. (2021). Current status and future perspectives about molecular biomarkers of nasopharyngeal carcinoma. *Cancers*, *13*(14), 3490.

Sim, S. P., & Liu, L. F. (2001). Nucleolytic cleavage of the mixed lineage leukemia breakpoint cluster region during apoptosis. *J Biol Chem*, *276*(34), 31590-31595. https://doi.org/10.1074/jbc.M103962200

Simons, M. J. (2011). Nasopharyngeal carcinoma as a paradigm of cancer genetics. *Chinese journal of cancer*, *30*(2), 79.

Singh, D. R., Nelson, S. E., Pawelski, A. S., Cantres-Velez, J. A., Kansra, A. S., Pauly, N. P., Bristol, J. A., Hayes, M., Ohashi, M., & Casco, A. (2022). Type 1 and Type 2 Epstein-Barr viruses induce proliferation, and inhibit differentiation, in infected telomerase-immortalized normal oral keratinocytes. *PLoS pathogens*, *18*(10), e1010868.

Sperry, A. O., Blasquez, V. C., & Garrard, W. T. (1989). Dysfunction of chromosomal loop attachment sites: illegitimate recombination linked to matrix association regions and topoisomerase II. *Proceedings of the National Academy of Sciences*, *86*(14), 5497-5501.

Stein, G. S., Zaidi, S. K., Braastad, C. D., Montecino, M., Van Wijnen, A. J., Choi, J.-Y., Stein, J. L., Lian, J. B., & Javed, A. (2003). Functional architecture of the nucleus: organizing the regulatory machinery for gene expression, replication and repair. *Trends in cell biology*, *13*(11), 584-592.

Stepanenko, A. A., & Heng, H. H. (2017). Transient and stable vector transfection: Pitfalls, off-target effects, artifacts. *Mutation Research/Reviews in Mutation Research*, *773*, 91-103.

Stepanenko, O. V., Stepanenko, O. V., Kuznetsova, I. M., Verkhusha, V. V., & Turoverov, K. K. (2013). Beta-barrel scaffold of fluorescent proteins: folding, stability and role in chromophore formation. *International review of cell and molecular biology*, *302*, 221-278.

Stowe, D. F., & Camara, A. K. (2009). Mitochondrial reactive oxygen species production in excitable cells: modulators of mitochondrial and cell function. *Antioxidants & redox signaling*, *11*(6), 1373-1414.

Strick, R., Zhang, Y., Emmanuel, N., & Strissel, P. L. (2006a). Common chromatin structures at breakpoint cluster regions may lead to chromosomal translocations found in chronic and acute leukemias. *Human genetics*, *119*(5), 479-495.

Strick, R., Zhang, Y., Emmanuel, N., & Strissel, P. L. (2006b). Common chromatin structures at breakpoint cluster regions may lead to chromosomal translocations found in chronic and acute leukemias. *Human genetics*, *119*, 479-495.

Strissel, P. L., Strick, R., Tomek, R. J., Roe, B. A., Rowley, J. D., & Zeleznik-Le, N. J. (2000). DNA structural properties of AF9 are similar to MLL and could act as recombination hot spots resulting in MLL/AF9 translocations and leukemogenesis. *Human molecular genetics*, *9*(11), 1671-1679.

Su, W.-H., Hildesheim, A., & Chang, Y.-S. (2013). Human leukocyte antigens and Epstein–Barr virus-associated nasopharyngeal carcinoma: old associations offer new clues into the role of immunity in infection-associated cancers. *Frontiers in Oncology*, *3*, 299.

Su, Z. Y., Siak, P. Y., Lwin, Y. Y., & Cheah, S.-C. (2024). Epidemiology of nasopharyngeal carcinoma: current insights and future outlook. *Cancer and Metastasis Reviews*, *43*(3), 919-939.

Sun, J., Hu, C., Zhu, Y., Sun, R., Fang, Y., Fan, Y., & Xu, F. (2015). LMP1 increases expression of NADPH oxidase (NOX) and its regulatory subunit p22 in NP69 nasopharyngeal cells and makes them sensitive to a treatment by a NOX inhibitor. *PloS one*, *10*(8), e0134896.

Sun, Y., Wang, C., Wang, L., Dai, Z., & Yang, K. (2018a). Arsenic trioxide induces apoptosis and the formation of reactive oxygen species in rat glioma cells. *Cellular & Molecular Biology Letters*, *23*, 1-10.

Sun, Y., Wang, C., Wang, L., Dai, Z., & Yang, K. (2018b). Arsenic trioxide induces apoptosis and the formation of reactive oxygen species in rat glioma cells. *Cellular & Molecular Biology Letters*, *23*(1), 1-10.

Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., & Bray, F. (2021). Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*, *71*(3), 209-249.

Swansbury, G., Slater, R., Bain, B., Moorman, A., & Secker-Walker, L. (1998). Hematological malignancies with t (9; 11)(p21–22; q23)–a laboratory and clinical study of 125 cases. *Leukemia*, *12*(5), 792-800.

Szukalska, M., Szyfter, K., Florek, E., Rodrigo, J. P., Rinaldo, A., Mäkitie, A. A., Strojan, P., Takes, R. P., Suárez, C., & Saba, N. F. (2020). Electronic cigarettes and head and neck cancer risk—current state of art. *Cancers*, *12*(11), 3274.

Taghizadeh, R. R., & Sherley, J. L. (2008). CFP and YFP, but not GFP, provide stable fluorescent marking of rat hepatic adult stem cells. *BioMed Research International*, *2008*.

Takacs, M., Banati, F., Koroknai, A., Segesdi, J., Salamon, D., Wolf, H., Niller, H. H., & Minarovits, J. (2010). Epigenetic regulation of latent Epstein-Barr virus promoters. *Biochim Biophys Acta*, *1799*(3-4), 228-235. https://doi.org/10.1016/j.bbagrm.2009.10.005

Tan, S.-N., & Sim, S.-P. (2018). Bile acids at neutral and acidic pH induce apoptosis and gene cleavages in nasopharyngeal epithelial cells: implications in chromosome rearrangement. *BMC cancer*, *18*(1), 409.

Tan, S.-N., & Sim, S.-P. (2019). Matrix association region/scaffold attachment region: the crucial player in defining the positions of chromosome breaks mediated by bile acid-induced apoptosis in nasopharyngeal epithelial cells. *BMC Medical Genomics*, *12*(1), 1-27.

Tan, S.-N., Sim, S.-P., & Khoo, A. S.-B. (2018a). Oxidative stress-induced chromosome breaks within the ABL gene: a model for chromosome rearrangement in nasopharyngeal carcinoma. *Human genomics*, *12*(1), 1-21.

Tan, S.-N., Sim, S.-P., & Khoo, A. S. (2018b). Matrix association region/scaffold attachment region (MAR/SAR) sequence: its vital role in mediating chromosome breakages in nasopharyngeal epithelial cells via oxidative stress-induced apoptosis. *BMC molecular biology*, *19*(1), 1-21.

Tan, S.-N., Sim, S.-P., & Khoo, A. S. B. (2016). Potential role of oxidative stress-induced apoptosis in mediating chromosomal rearrangements in nasopharyngeal carcinoma. *Cell & Bioscience*, *6*, 35. https://doi.org/10.1186/s13578-016-0103-9

Tao, Q., Young, L. S., Woodman, C. B., & Murray, P. G. (2006). Epstein-Barr virus (EBV) and its associated human cancers--genetics, epigenetics, pathobiology and novel therapeutics. *Front Biosci*, *11*, 2672-2713.

Tikhonov, A. P., Bennetzen, J. L., & Avramova, Z. V. (2000). Structural domains and matrix attachment regions along colinear chromosomal segments of maize and sorghum. *The Plant Cell*, *12*(2), 249-264.

Togi, S., Hatano, Y., Muromoto, R., Kawanishi, E., Ikeda, O., Hirashima, K., Kon, S., Kitai, Y., Yasui, T., & Oritani, K. (2016). Caspase-dependent cleavage regulates protein levels of Epstein–Barr virus-derived latent membrane protein 1. *FEBS letters*, *590*(6), 808-818.

Trueba, G. P., Sánchez, G. M., & Giuliani, A. (2004). Oxygen free radical and antioxidant defense mechanism in cancer. *Front Biosci*, *9*(Sep 1), 2029-2044.

Tsai, C. L., Li, H. P., Lu, Y. J., Hsueh, C., Liang, Y., Chen, C. L., Tsao, S. W., Tse, K. P., Yu, J. S., & Chang, Y. S. (2006). Activation of DNA methyltransferase 1 by EBV LMP1 Involves c-Jun NH(2)-terminal kinase signaling. *Cancer Res*, *66*(24), 11668-11676. https://doi.org/10.1158/0008-5472.can-06-2194

Tsai, M.-H., Raykova, A., Klinke, O., Bernhardt, K., Gärtner, K., Leung, C. S., Geletneky, K., Sertel, S., Münz, C., & Feederle, R. (2013). Spontaneous lytic replication and epitheliotropism define an Epstein-Barr virus strain found in carcinomas. *Cell reports*, *5*(2), 458-470.

Tsang, C. M., Deng, W., Yip, Y. L., Zeng, M.-S., Lo, K. W., & Tsao, S. W. (2014). Epstein-Barr virus infection and persistence in nasopharyngeal epithelial cells. *Chinese journal of cancer*, *33*(11), 549.

Tsang, C. M., Yip, Y. L., Lo, K. W., Deng, W., To, K. F., Hau, P. M., Lau, V. M., Takada, K., Lui, V. W., Lung, M. L., Chen, H., Zeng, M., Middeldorp, J. M., Cheung, A. L., & Tsao, S. W. (2012). Cyclin D1 overexpression supports stable EBV infection in

nasopharyngeal epithelial cells. *Proc Natl Acad Sci U S A*, *109*(50), E3473-3482. https://doi.org/10.1073/pnas.1202637109

Tsao, S. W., Tsang, C. M., & Lo, K. W. (2017). Epstein–Barr virus infection and nasopharyngeal carcinoma. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1732), 20160270.

Tsao, S. W., Wang, X., Liu, Y., Cheung, Y. C., Feng, H., Zheng, Z., Wong, N., Yuen, P., Lo, A. K., & Wong, Y. (2002). Establishment of two immortalized nasopharyngeal epithelial cell lines using SV40 large T and HPV16E6/E7 viral oncogenes. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*, *1590*(1-3), 150-158.

Vaughan, A., Betti, C., & Villalobos, M. (2002). Surviving apoptosis. *Apoptosis*, *7*, 173-177.

Wang, H.-Y., Sun, B.-Y., Zhu, Z.-H., Chang, E. T., To, K.-F., Hwang, J. S., Jiang, H., Kam, M. K.-M., Chen, G., & Cheah, S.-L. (2011). Eight-signature classifier for prediction of nasopharyngeal carcinoma survival. *Journal of clinical oncology*, *29*(34), 4516-4525.

Wang, K. H., Austin, S. A., Chen, S. H., Sonne, D. C., & Gurushanthaiah, D. (2017). Nasopharyngeal Carcinoma Diagnostic Challenge in a Nonendemic Setting: Our Experience with 101 Patients. *Permanente Journal*, *21*(3).

Wang, L., & Ning, S. (2021). New look of EBV LMP1 signaling landscape. *Cancers*, *13*(21), 5451.

Wang, L. W., Jiang, S., & Gewurz, B. E. (2017). Epstein-Barr virus LMP1-mediated oncogenicity. *Journal of Virology*, *91*(21), 10.1128/jvi. 01718-01716.

Ward, M. H., Pan, W. H., Cheng, Y. J., Li, F. H., Brinton, L. A., Chen, C. J., Hsu, M. M., Chen, I. H., Levine, P. H., & Yang, C. S. (2000). Dietary exposure to nitrite and nitrosamines and risk of nasopharyngeal carcinoma in Taiwan. *Int J Cancer*, *86*(5), 603-609.

Wiseman, H., & Halliwell, B. (1996). Damage to DNA by reactive oxygen and nitrogen species: role in inflammatory disease and progression to cancer. *Biochemical Journal*, *313*(Pt 1), 17.

Wolf, H., Zur Hausen, H., Klein, G., Becker, V., Henle, G., & Henle, W. (1975). Attempts to detect virus-specific DNA sequences in human tumors. III. Epstein-Barr viral DNA in non-lymphoid nasopharyngeal carcinoma cells. *Med Microbiol Immunol*, *161*(1), 15-21. https://doi.org/10.1007/bf02120766

Wong, R. S. (2011). Apoptosis in cancer: from pathogenesis to treatment. *Journal of experimental & clinical cancer research*, *30*, 1-14.

Wu, C.-C., Fang, C.-Y., Huang, S.-Y., Chiu, S.-H., Lee, C.-H., & Chen, J.-Y. (2018). Perspective: contribution of Epstein–Barr virus (EBV) reactivation to the carcinogenicity of nasopharyngeal cancer cells. *Cancers*, *10*(4), 120.

Wu, Z.-X., Xiang, L., Rong, J.-F., He, H.-L., & Li, D. (2016). Nasopharyngeal carcinoma with headaches as the main symptom: a potential diagnostic pitfall. *Journal of cancer research and therapeutics*, *12*(1), 209-214.

Xu, F.-H., Xiong, D., Xu, Y.-F., Cao, S.-M., Xue, W.-Q., Qin, H.-D., Liu, W.-S., Cao, J.-Y., Zhang, Y., & Feng, Q.-S. (2012). An epidemiological and molecular study of the relationship between smoking, risk of nasopharyngeal carcinoma, and Epstein–Barr virus activation. *Journal of the National Cancer Institute*, *104*(18), 1396-1410.

Xu, X., Yang, Z., Chen, Q., Yu, L., Liang, S., Lu, H., & Qiu, Z. (2015). Comparison of clinical characteristics of chronic cough due to non-acid and acid gastroesophageal reflux. *Clin Respir J*, *9*(2), 196-202. https://doi.org/10.1111/crj.12124

Xue, W.-Q., Qin, H.-D., Ruan, H.-L., Shugart, Y. Y., & Jia, W.-H. (2013). Quantitative association of tobacco smoking with the risk of nasopharyngeal carcinoma: a comprehensive meta-analysis of studies conducted between 1979 and 2011. *American journal of epidemiology*, *178*(3), 325-338.

Yang, J.-b., Zhang, X.-m., Deng, L.-w., Tan, G.-l., Zhou, M., Zeng, Z.-y., Cao, L., Shen, S.-r., & Li, G.-y. (2000). Detailed deletion mapping of chromosome 9p21–22 in nasopharyngeal carcinoma. *Chinese Journal of Cancer Research*, *12*, 161-164.

Yang, L., Liu, L., Xu, Z., Liao, W., Feng, D., Dong, X., Xu, S., Xiao, L., Lu, J., & Luo, X. (2015). EBV-LMP1 targeted DNAzyme enhances radiosensitivity by inhibiting tumor angiogenesis via the JNKs/HIF-1 pathway in nasopharyngeal carcinoma. *Oncotarget*, *6*(8), 5804.

Yang, Y., Karakhanova, S., Hartwig, W., D'Haese, J. G., Philippov, P. P., Werner, J., & Bazhin, A. V. (2016). Mitochondria and mitochondrial ROS in cancer: novel targets for anticancer therapy. *Journal of cellular physiology*, *231*(12), 2570-2581.

Yee, P. H.-C., & Sim, S.-P. (2010). High cell density and latent membrane protein 1 expression induce cleavage of the mixed lineage leukemia gene at 11q23 in nasopharyngeal carcinoma cell line. *Journal of biomedical science*, *17*(1), 77.

Yoshizaki, T., Kondo, S., Wakisaka, N., Murono, S., Endo, K., Sugimoto, H., Nakanishi, S., Tsuji, A., & Ito, M. (2013). Pathogenic role of Epstein–Barr virus latent membrane protein-1 in the development of nasopharyngeal carcinoma. *Cancer letters*, *337*(1), 1-7.

Young, L., Dawson, C., Clark, D., Rupani, H., Busson, P., Tursz, T., Johnson, A., & Rickinson, A. (1988). Epstein—Barr virus gene expression in nasopharyngeal carcinoma. *Journal of General Virology*, *69*(5), 1051-1065.

Young, L. S., Yap, L. F., & Murray, P. G. (2016). Epstein–Barr virus: more than 50 years old and still providing surprises. *Nature Reviews Cancer*, *16*(12), 789.

Yu, J. S., Tsai, H. C., Wu, C. C., Weng, L. P., Li, H. P., Chung, P. J., & Chang, Y. S. (2002). Induction of inducible nitric oxide synthase by Epstein-Barr virus B95-8-derived LMP1 in Balb/3T3 cells promotes stress-induced cell death and impairs LMP1-mediated transformation. *Oncogene*, *21*(52), 8047-8061. https://doi.org/10.1038/sj.onc.1205990

Zhang, J.-B., Huang, S.-Y., Wang, T.-M., Dong, S.-Q., He, Y.-Q., Zheng, X.-H., Li, X.-Z., Wang, F., Jianbing, M., & Jia, W.-H. (2018). Natural variations in BRLF1 promoter contribute to the elevated reactivation level of Epstein-Barr virus in endemic areas of nasopharyngeal carcinoma. *EBioMedicine*, *37*, 101-109.

Zhang, S., Li, S., & Gao, J.-L. (2013). Promoter methylation status of the tumor suppressor gene SOX11 is associated with cell growth and invasion in nasopharyngeal carcinoma. *Cancer Cell International*, *13*(1), 109.

Zhao, C.-P., Guo, X., Chen, S.-J., Li, C.-Z., Yang, Y., Zhang, J.-H., Chen, S.-N., Jia, Y.-L., & Wang, T.-Y. (2017). Matrix attachment region combinations increase transgene expression in transfected Chinese hamster ovary cells. *Scientific reports*, *7*(1), 42805.

Zhao, F., & Wang, Q. (2012). The protective effect of peroxiredoxin II on oxidative stress induced apoptosis in pancreatic beta-cells. *Cell Biosci*, *2*(1), 22. https://doi.org/10.1186/2045-3701-2-22

Zhou, Z., Li, P., Zhang, X., Xu, J., Xu, J., Yu, S., Wang, D., Dong, W., Cao, X., & Yan, H. (2022). Mutational landscape of nasopharyngeal carcinoma based on targeted next-generation sequencing: implications for predicting clinical outcomes. *Molecular Medicine*, *28*(1), 55.

Zuo, X.-Y., Feng, Q.-S., Sun, J., Wei, P.-P., Chin, Y.-M., Guo, Y.-M., Xia, Y.-F., Li, B., Xia, X.-J., & Jia, W.-H. (2019). X-chromosome association study reveals genetic susceptibility loci of nasopharyngeal carcinoma. *Biology of sex Differences*, *10*, 1-11.

APPENDICES

Appendix A: Journal Publications

1. **Mustapha, M. A**., & Sim, S.-P. (2021). No Association of Latent Membrane Protein 1 (LMP1) Gene Expression of Epstein-Barr Virus and Oxidative Stress Production in Normal Nasopharyngeal Cell. *International Medical Journal*, *28*(2), 138-140.
2. **Mustapha, M. A**., & Sim, S.-P. (2021). Association of Epstein-Barr Virus Latent Membrane Protein 1 (LMP1) Gene Expression and Caspase Activity in Normal Nasopharyngeal Cell. *International Medical Journal*, *28*(5).
3. **Mustapha, M. A.,** & Sim, S.-P. (2024). Green Fluorescent Proteins (GFP) Induce Oxidative Stress in Normal Nasopharyngeal Cells Expression. *International Medical Journal*, 31(4), 108-110.
4. **Mustapha, M. A.,** & Sim, S.-P. (2024). Green Fluorescent Proteins (GFP) Induce Apoptosis in Normal Nasopharyngeal Cells Expression. International Medical Journal, 31(4), 117-119.
5. **Mustapha, M. A.,** & Sim, S.-P. (2024). The Role of Epstein-Barr virus LMP1 Gene Expression in Oxidative Stress in Nasopharyngeal Carcinoma. *International Medical Journal*, *31*(4), 83-86.
6. **Mustapha, M. A.,** & Sim, S.-P. (2024). The Role of Caspase Activity and Apoptosis in Nasopharyngeal Carcinoma. *International Medical Journal*, *31*(5), 141-143.
7. **Mustapha, M. A**., & Sim, S.-P. (2025). A Global Perspective on Nasopharyngeal Carcinoma: Clinical Features, Risk Factors, and Molecular. *International Medical Journal*, *34*(2).

8. **Mustapha, M. A**., & Sim, S.-P. (2025). Epstein-Barr Virus in Nasopharyngeal Carcinoma: Molecular Mechanisms and Oncogenic Drivers. *International Medical Journal*, *34*(2).

Appendix B: Data of oxidative stress measurement

No	Sample	pTracer	pcDNA	Sample	pTracer	pcDNA
1	Control	5636	56994	*LMP1*	6176	54907
2	Control	12942	58301	*LMP1*	6273	55427
3	Control	9003	53803	*LMP1*	6273	56335
4	Control	7836	53616	*LMP1*	6295	58857
5	Control	8736	55433	*LMP1*	6348	55520
6	Control	8335	57162	*LMP1*	6364	54450
7	Control	6520	58418	*LMP1*	6481	54533
8	Control	9400	57616	*LMP1*	6645	56442
9	Control	6701	56630	*LMP1*	6646	53970
10	Control	10648	52722	*LMP1*	6781	55648
11	Control	5015	64360	*LMP1*	6793	58411
12	Control	6604	58990	*LMP1*	6921	55594
13	Control	8432	58389	*LMP1*	7055	53054
14	Control	12033	61198	*LMP1*	7298	64195
15	Control	6698	60891	*LMP1*	7329	54970
16	Control	11909	57189	*LMP1*	7589	55912
17	Control	9718	58998	*LMP1*	7600	57638
18	Control	9742	58299	*LMP1*	7676	53649
19	Control	7040	56337	*LMP1*	7801	55022
20	Control	6136	54796	*LMP1*	7827	54949
21	Control	8985	58092	*LMP1*	7975	60226
22	Control	8723	59893	*LMP1*	8201	61568

23	Control	9687	56631	*LMP1*	8403	60163
24	Control	6949	61007	*LMP1*	8415	60236
25	Control	9515	58298	*LMP1*	8651	60872
26	Control	6851	58259	*LMP1*	9002	59163
27	Control	6121	57757	*LMP1*	9173	58925
28	Control	4844	61674	*LMP1*	9432	58388
29	Control	6658	63425	*LMP1*	9499	63619
30	Control	6348	60650	*LMP1*	9522	60732
31	Control	6693	58542	*LMP1*	9603	58148
32	Control	6365	57240	*LMP1*	9721	57532
33	Control	6401	57501	*LMP1*	9753	59531
34	Control	6409	58717	*LMP1*	9784	55005
35	Control	5825	55895	*LMP1*	9862	54295
36	Control	8355	61278	*LMP1*	9885	57970
37	Control	7813	57681	*LMP1*	10118	61569
38	Control	6326	58641	*LMP1*	10135	60179
39	Control	6833	65478	*LMP1*	10201	64996
40	Control	5889	61344	*LMP1*	10214	59545
41	Control	11700	480	*LMP1*	10435	39912
42	Control	12489	46699	*LMP1*	10482	43644
43	Control	14441	44665	*LMP1*	10669	43796
44	Control	9966	43878	*LMP1*	10844	43527
45	Control	13792	44477	*LMP1*	10885	44162
46	Control	15880	45708	*LMP1*	10975	42896

47	Control	17786	44162	*LMP1*	11011	44949
48	Control	9920	44170	*LMP1*	11108	43753
49	Control	15457	41816	*LMP1*	11132	43135
50	Control	11065	42081	*LMP1*	11462	45584
51	Control	11011	44040	*LMP1*	11726	42322
52	Control	10595	43280	*LMP1*	12329	43680
53	Control	11408	43075	*LMP1*	12367	41808
54	Control	11830	50007	*LMP1*	12590	42865
55	Control	8481	43580	*LMP1*	12660	46496
56	Control	9691	45388	*LMP1*	12852	47772
57	Control	8991	43885	*LMP1*	13211	56768
58	Control	11082	47611	*LMP1*	13516	51484
59	Control	10195	45116	*LMP1*	14190	44333
60	Control	11408	43982	*LMP1*	16188	43156
61	Control	28996	43877	*LMP1*	16645	42099
62	Control	25718	47177	*LMP1*	16839	44504
63	Control	30932	44303	*LMP1*	19869	44926
64	Control	30430	42616	*LMP1*	20173	41483
65	Control	33642	42358	*LMP1*	20423	42930
66	Control	30509	42841	*LMP1*	21146	42667
67	Control	30249	43378	*LMP1*	21918	49493
68	Control	29720	42780	*LMP1*	22318	53568
69	Control	32497	45665	*LMP1*	22367	45644
70	Control	37012	42989	*LMP1*	22449	45180

71	Control	26006	43740	*LMP1*	23236	45087
72	Control	30954	45444	*LMP1*	24359	43659
73	Control	37026	43279	*LMP1*	24878	44352
74	Control	30324	45607	*LMP1*	26264	42684
75	Control	29703	46235	*LMP1*	26367	42325
76	Control	31543	41372	*LMP1*	26935	44454
77	Control	27311	42172	*LMP1*	27509	42769
78	Control	24725	44545	*LMP1*	27822	43266
79	Control	31942	45522	*LMP1*	28503	43827
80	Control	23953	43747	*LMP1*	30138	44299
81	Control	41176	41959	*LMP1*	32034	39966
82	Control	33747	42628	*LMP1*	32468	41105
83	Control	33044	42167	*LMP1*	32595	42875
84	Control	31279	42946	*LMP1*	32958	43789
85	Control	36779	42941	*LMP1*	33212	44152
86	Control	35117	44082	*LMP1*	33293	44644
87	Control	38575	42766	*LMP1*	33711	41889
88	Control	35301	40770	*LMP1*	34526	42871
89	Control	39919	41698	*LMP1*	35380	42003
90	Control	31580	43166	*LMP1*	37287	42366
91	Control	34372	40953	*LMP1*	37310	41602
92	Control	31987	41555	*LMP1*	37430	41001
93	Control	38021	41976	*LMP1*	38543	41224
94	Control	26962	42470	*LMP1*	39881	39971

95	Control	30180	40366	*LMP1*	40221	42869
96	Control	22637	40998	*LMP1*	40539	42061
97	Control	26573	41853	*LMP1*	45736	42034
98	Control	30930	42653	*LMP1*	47578	42464
99	Control	26262	46137	*LMP1*	48078	44435
100	Control	28852	46902	*LMP1*	48822	41951

Appendix C: Data of Caspase Activity Measurement

No	Sample	pTracer	pcDNA	Sample	pTracer	pcDNA
1	Control	538	387	*LMP1*	347	340
2	Control	458	327	*LMP1*	283	401
3	Control	377	364	*LMP1*	308	404
4	Control	587	455	*LMP1*	245	355
5	Control	500	302	*LMP1*	443	563
6	Control	351	275	*LMP1*	94	416
7	Control	376	337	*LMP1*	280	296
8	Control	443	355	*LMP1*	302	345
9	Control	466	343	*LMP1*	217	368
10	Control	456	299	*LMP1*	278	392
11	Control	534	401	*LMP1*	328	442
12	Control	433	523	*LMP1*	233	417
13	Control	476	450	*LMP1*	258	330
14	Control	466	287	*LMP1*	382	400
15	Control	485	299	*LMP1*	232	433
16	Control	295	340	*LMP1*	194	528
17	Control	326	351	*LMP1*	189	389
18	Control	349	334	*LMP1*	253	525
19	Control	346	307	*LMP1*	227	358
20	Control	479	271	*LMP1*	232	394
21	Control	396	219	*LMP1*	383	369
22	Control	460	482	*LMP1*	367	434

23	Control	404	452	*LMP1*	420	302
24	Control	423	363	*LMP1*	278	285
25	Control	373	406	*LMP1*	354	281
26	Control	160	333	*LMP1*	195	506
27	Control	200	365	*LMP1*	200	541
28	Control	144	297	*LMP1*	144	745
29	Control	194	211	*LMP1*	146	445
30	Control	129	338	*LMP1*	169	435
31	Control	107	315	*LMP1*	179	471
32	Control	114	296	*LMP1*	160	430
33	Control	112	138	*LMP1*	133	406
34	Control	164	215	*LMP1*	120	414
35	Control	212	215	*LMP1*	128	511
36	Control	217	378	*LMP1*	251	344
37	Control	146	320	*LMP1*	132	419
38	Control	126	373	*LMP1*	126	277
39	Control	160	314	*LMP1*	118	311
40	Control	115	318	*LMP1*	133	284
41	Control	108	309	*LMP1*	129	177
42	Control	79	325	*LMP1*	124	314
43	Control	105	208	*LMP1*	94	245
44	Control	87	275	*LMP1*	145	239
45	Control	143	151	*LMP1*	77	246
46	Control	129	243	*LMP1*	128	389

47	Control	113	268	*LMP1*	123	445
48	Control	176	270	*LMP1*	128	410
49	Control	132	251	*LMP1*	83	412
50	Control	91	136	*LMP1*	107	339
51	Control	968	247	*LMP1*	361	216
52	Control	1191	236	*LMP1*	776	269
53	Control	484	309	*LMP1*	1076	286
54	Control	520	208	*LMP1*	758	326
55	Control	963	267	*LMP1*	532	243
56	Control	587	448	*LMP1*	540	285
57	Control	425	417	*LMP1*	305	270
58	Control	629	275	*LMP1*	348	243
59	Control	112	322	*LMP1*	677	264
60	Control	827	297	*LMP1*	470	185
61	Control	765	242	*LMP1*	1042	320
62	Control	1068	305	*LMP1*	527	250
63	Control	1333	372	*LMP1*	480	254
64	Control	1297	404	*LMP1*	372	239
65	Control	828	351	*LMP1*	113	201
66	Control	747	283	*LMP1*	423	254
67	Control	1075	227	*LMP1*	328	249
68	Control	1406	277	*LMP1*	572	238
69	Control	841	314	*LMP1*	616	263
70	Control	833	259	*LMP1*	663	204

71	Control	871	300	*LMP1*	851	346
72	Control	656	271	*LMP1*	1083	289
73	Control	638	358	*LMP1*	1027	243
74	Control	598	245	*LMP1*	842	249
75	Control	452	261	*LMP1*	979	328
76	Control	803	205	*LMP1*	710	147
77	Control	871	176	*LMP1*	2207	185
78	Control	825	187	*LMP1*	1563	237
79	Control	818	146	*LMP1*	1306	223
80	Control	804	179	*LMP1*	1018	155
81	Control	734	162	*LMP1*	1044	148
82	Control	801	151	*LMP1*	1484	131
83	Control	515	186	*LMP1*	821	124
84	Control	876	176	*LMP1*	1007	212
85	Control	659	146	*LMP1*	1367	119
86	Control	848	211	*LMP1*	1766	362
87	Control	805	174	*LMP1*	892	231
88	Control	912	172	*LMP1*	803	161
89	Control	882	144	*LMP1*	897	232
90	Control	502	109	*LMP1*	854	157
91	Control	976	147	*LMP1*	830	136
92	Control	931	95	*LMP1*	779	157
93	Control	551	134	*LMP1*	661	178
94	Control	566	145	*LMP1*	763	127

95	Control	493	93	*LMP1*	583	182
96	Control	563	171	*LMP1*	743	148
97	Control	905	119	*LMP1*	674	167
98	Control	896	148	*LMP1*	586	168
99	Control	700	104	*LMP1*	922	178
100	Control	389	186	*LMP1*	678	129

Appendix D: Number of Chromosome Break in SAR of *AF9* gene

No	Sample	pTracer	pcDNA	Sample	pTracer	pcDNA
1	Control	1	0	*LMP1*	4	3
2	Control	3	0	*LMP1*	4	3
3	Control	3	0	*LMP1*	4	3
4	Control	3	2	*LMP1*	4	3
5	Control	3	2	*LMP1*	5	3
6	Control	3	2	*LMP1*	5	3
7	Control	2	1	*LMP1*	5	1
8	Control	2	1	*LMP1*	5	1
9	Control	2	1	*LMP1*	4	1
10	Control	2	2	*LMP1*	4	1
11	Control	2	3	*LMP1*	4	1
12	Control	2	3	*LMP1*	3	1
13	Control	3	0	*LMP1*	3	1
14	Control	3	0	*LMP1*	3	1
15	Control	3	2	*LMP1*	1	1
16	Control	3	2	*LMP1*	1	3
17	Control	3	1	*LMP1*	1	3
18	Control	3	1	*LMP1*	2	3
19	Control	3	1	*LMP1*	3	3
20	Control	0	1	*LMP1*	3	3
21	Control	0	2	*LMP1*	3	3
22	Control	0	2	*LMP1*	3	3

23	Control	0	0	*LMP1*	3	3
24	Control	0	0	*LMP1*	3	1
25	Control	0	0	*LMP1*	3	1

Appendix E: Number of Chromosome Break in non-SAR of *AF9* gene

No	Sample	pTracer	pcDNA	Sample	pTracer	pcDNA
1	Control	1	1	*LMP1*	2	4
2	Control	1	1	*LMP1*	2	4
3	Control	1	1	*LMP1*	2	4
4	Control	1	4	*LMP1*	2	4
5	Control	2	4	*LMP1*	2	4
6	Control	2	4	*LMP1*	2	4
7	Control	2	4	*LMP1*	2	2
8	Control	2	4	*LMP1*	2	3
9	Control	3	2	*LMP1*	2	3
10	Control	3	2	*LMP1*	2	3
11	Control	3	2	*LMP1*	2	3
12	Control	3	1	*LMP1*	2	3
13	Control	1	1	*LMP1*	3	0
14	Control	1	1	*LMP1*	3	0
15	Control	1	1	*LMP1*	3	0
16	Control	1	1	*LMP1*	3	0
17	Control	3	1	*LMP1*	2	0
18	Control	3	4	*LMP1*	2	0
19	Control	3	4	*LMP1*	2	0
20	Control	1	1	*LMP1*	2	3
21	Control	1	1	*LMP1*	2	3
22	Control	2	1	*LMP1*	4	3

23	Control	3	1	*LMP1*	2	2
24	Control	2	1	*LMP1*	2	2
25	Control	2	1	*LMP1*	2	2

Appendix F: DNA Sequence of Region of Study in the SAR of the *AF9* Gene (Nucleotide position 236059-246297) [Ensembl: ENSG00000171843]

```
         10        20        30        40        50
GGATCCGAAG TAAGAAAGAC GTATTTTCTG TCCTCATCTT TATCTTTGGC    50
TAGTGTGACT GCAGGGAAAA GTGGCTAATA CAAACATTTA GATAACTTAC   100
ACACAAAGCA AAAAGTAGAA TAGTAGAATT TTACCAGGTA TGAAAAGGAT   150
ATGAAAAGTC AACTTAATAC ATTAAGACAA TTAGAATAAT AATTTCTATA   200
TTGGAGAGCA GTGGATATCT GGATAAAATA GTAATAGCAG GCCTATAACT   250
TATTATTATC TAGGTTTAGT TCAGAATGTT ATATATATAT CCAGAAAATA   300
GCCTGGTTTT TACCATAGGA GCAGACTGGC CAATACTAAT GGCCATAAAA   350
TCCGGTTAGC ATTTCTTACT GAGAATTTTT TGGGGTCTAG AATATTTGAT   400
GTAGTCACAT ATTCACTTGG CTGTCTCATT CCAGCACATC TTTAGAATAG   450
GAGTTTGGGG CATCCATGGG ATGGAGTCTG TCACTTCCAG TGATGACTGT   500
GACAGCTGAA AGCTTCTGTC AATAGGCTTT ATCCAATTCT GCAATGTTCA   550
TTACTCCAGC GACTGGGCCT CCAAACAAAG CTTGATGACA GGAGCCAGAT   600
CAACTCCTTG TTGGTCTTTC TTTAGCAAAG CTACAGCCAC GAAGGAAAAG   650
CCACACATAA TTCAGAGGCT AACTTGTCAG ACTTTGGGTT TATTGGCTTC   700
TGAGGTAGTA GGCTTGGGAC TGCTGACCAG AGAGGTATTT GAATGAATGA   750
CCCTTCTCCT TACTCTTCAA CATGCTGTGC CTTGGAAACT CTCCTTGAAA   800
CATTATTTTG TGCCACTTCA GGCACAGAAC AACACTTGCC CTAAACATAA   850
ACTGACTGGT AACACCATTC TTGTATTTAC TTACTCTAAT TGCAAACTCT   900
TTTTCATTTC AGATAACATT ATGTAAAGTA TGTGCTTTCA CATGGTACTT   950
TTTGGACCCT GTTAACCAAA TGGCACGGCA TCTCTACATG TTCTGTATAT  1000
GAAATACAGA AATGGCAGTG CCCATTGGAA ATGGGAATTC AGTTTCAACC  1050
TAAGGGACAA CAGGGCTGTC TCACCAGAAC CAGAGTGGAG TGATTGTCCT  1100
CTTCACTAAA GGGAGAGATG CAACACAAAT ATCTCAGAAG GCAGATTGCA  1150
GTTTACTTCA GAATACACCT TAGCATCAAT GCTAGAGAGA GACCATTGAA  1200
ATGGATCTTA TCAGACCAAA AGTATCATTT CATAATGGTC TTAAGATTAG  1250
AACATTATTA AAATGGAATC ATCTGTCAAT TAGTGGTATA TAATTATATG  1300
AGACTCTTGT AGGTAACTTC TTGAGAAAAC ATTATCCTAG GATAGTGTCT  1350
CCCCTACCCC ACCCCATTGG GCACTATCTG GAGACATTTT TAAATGTTGT  1400
GACTTGGGAG TGGGAGTGGT GTCAGCCAAA ATTTTCAGCT GAGATTTTAA  1450
CATTTTAGGA CAAATTATTA AGTCTTTCTA AAATCTTTTT GGACTTGGAA  1500
AAGTATTTCT TATTGCTTTA GGATGTTCTC GATTGAACCA AAAACAGACA  1550
TTTTGTTATT TTTATTCATA TATTTCAAAC ATAATTATTG TAAAAATTTT  1600
TGAAAGGTAC AGAATTGTTT TATAAAGAAG AAAATAAAAA GCACCATAAT  1650
CCTACCATTC ATTTGTACAT TTTCTTCGAG ATTCAAAATA TCATGGAACA  1700
GATGGTTTAT CAGAGTAGTC TCCTTAAGTA AATCAACTGG AAGTGGGATT  1750
TTTTCCTGGG ATATTGCTTA CATTATCTTT CCGTAGCCCC TTCTCTTTCA  1800
AATAATACTA CTGTCACACA AAGTTTAGGG TCAGTTTATC AGAACGTGAC  1850
TAATTATACT ATAAATACCT GAGTAGAATT TCCAGCTTTA GGAGTTCTTA  1900
CCGATATTTA AATAGACTCA GGAAAAAGTA TTGTAGACCC CTTTATTATT  1950
TTCAGTGTGT ATGTTAATCG GTTAGAGATT TTTATTAAGT ACCTTTACTT  2000
ATGGTAGGAT TTATTAATAA ATAAAAATAT TTTGTAACTT TTCAGTCTGA  2050
GTCTAGAGGT AACTTTTTAC TTTTCTGCAA GAATTTTTAG TTGATTATGA  2100
GATCTAAAAA GGCAAAGAAG AGGGAAAGAC CAGATAATCT AGACCACCAC  2150
ATAACAGTGT CTCTGGACTT CAGAAGCCCC AGGAGGGTGC GGCTGGGGAG  2200
TGCTATATTT TCTGTACACT TAGTGAGTCA AATTTGGTTT GGCAGAGAAT  2250
TTTTAGAAGT AGATCAGTCT CTGTAAGTAA AACTTTGCAG TTTTAGCCAA  2300
AGAAAACAGA CTGTAGATTT GATGACTCTA TTTGATTTAT TAGTGCCATA  2350
TGCAGGGCCT TTTGATCTTC AGGATTCAGT AAACCAATGA ATTGATATTT  2400
```

```
         10         20         30         40         50
ATCAGAACTC CATCACATGG GAAGAAATCT TAATTTGACT GTCGTGTTGA  2450
TACTAATATA ATTATCCAAC ACCTGCATCC CATTGATGAT AATTGTGTTA  2500
GATTATATTA AATTTTAATG ATAGGTACAA ACTGAGCATT AAGAACACGT  2550
TTTTGGATAA CATTAACATT GACAATTCTC TTATTTCTTA GGAAATGAGT  2600
ACATTAAACT TGCTTACTTC TATGTATAAC AAAACCTCTT ACTTAGGAGT  2650
GCAGGAACTG TGTATATGTA GTGATTGTAG TTTTTAAAGA AAAAAGAAAT  2700
AATTGGCATT GATGATATAC CACATGACCA GACGCTATAT AATAAAGTGA  2750
CTTACAATCT TAATTTGTCA GACAACACTA GAGAGTTTCA ATTAAGTTTG  2800
AGGTTTTGGC CCACCTCTGT ATTCCTTTGT AATATTTTAT AAAGTAGCAA  2850
AAACTACAGA CAAATATTTC TGCCATTAAG CAATCAGCAT CCATCTGTTT  2900
TTGTCTTTTT ACTACCAAGA CAAGTTAACA TGTGGCAGAA GCATATAAGT  2950
CTACAAAATT GAGATGAACT GTTAAGAACA GTATCATGAT ATTATTATTT  3000
TCTTTAAGTG GATCATTTTC TTGAGAGTCT TAGGGCCCAC TTTTTCTGCA  3050
AAAGGAAAAA CTTCTCAAAC AATGTGATGC CCTAAATAAG TTTCACTTTC  3100
TTTTTTCTAT GAATTCTCAA ATATTACTTG ACCAGGCAGT TCTTACCTAT  3150
AGATCATTTG TGCTTAGATT AACAATCACT TTTTAATTCT CATCTTTGTT  3200
TATTCTGTTT GTATAGAAGT CCTTGCTAGA ATGTCTTATA TATTGGGCAT  3250
TGAACTAGGT ACTTTATCAG CTGCCTTCTC CAGGAGAAGT TTCAGAGAAA  3300
ATAAATTTTG GATTAGTCTA AAGGAAGTTA CCTTATTTCT TGAAGTGTCA  3350
AGTGTTTTTC AGGGAAGACT ATCATGGGGG AGAGGGTGGA CAAGCCAAAT  3400
GGACTCAGAC ATAACTTTTT AATATCATAA TTTTATATTT ATTCCTATTG  3450
ATGATAATAG TAACTAATAT TTAGTCATCA TTTATTCAGT GTTCTGTCCT  3500
AACTGCTTTG CGTTGTCAAC ATTCTAACAT TAGCGATATG AATTGGTACC  3550
ACTATTACTT ATGTTTCCCC ATTAGAAATG AAGGAACTGA GGCCTGGAGA  3600
GATGATGTTA TGTGCTAAGG TCACATAGCT AGAAAGTGGT GGAACTGAGA  3650
TTCAAACCAG GTTTGATGTT TTCAGAACCT GAGCACCAAA TTATTATAGT  3700
GTTTCATTCT TTTCACCCTC ATACTAATCT AGATTTTTCA AAATATTTGA  3750
CATTAAGATT AAAATGCAAA GTTCATGTCA AAAGTAAAAT ATAATTATGT  3800
TAAAATTGTT TGACATTTTA AAATAACGTA CCCCCCATCA TAGAGCTGAT  3850
AGCTGCATTT GTTTTCTGCT CCTCACCAAC ACATCTTAAT ATATATAAAT  3900
CTCAATGTGT ATGCATATAT ACAACACAAT AGTGTTGTAG GTGTAGTTAA  3950
TAGCTGCCAG CAAACAGTGC CATCTGTGTT TTCTCTTGTT TCTCACTTAC  4000
TGGAAAATGT TACTGAAGCT CAAATTATGA GTAGGGGCTT AAAGTTCATG  4050
TTCAGTATCT CCACCTTTTA ACTCTACCAC CTTTTATTTG CACAAGAAAA  4100
TGTGATCCTC TATAGTGACT TTACCCTGTT TACAGAATAG GAGTGTGTGC  4150
TATTACTGAT AGGATGGGCA GGTCTTACAG TTTAAAAATA CCCATAAGAC  4200
AAATAGATCC TGCAATCTTT ACATAGTCAT TTCTTTACTT TGTACTTCAT  4250
AAAACCCTGC AGGCTAGAGC AAGGAATTGA TGTAAGTTGA TGTCAAACTG  4300
ACTTTACAAA AACATTCTGG AATTCTAAAT CAGTGTGATT TAAGATCCTG  4350
TTATACTAAG AGAGGTAGTT AAGCATTTGA CATCCGATCT CTCTAGAAGA  4400
ATGTCTGCTT ACCACACATA CAGTAGTATC AGATGGGAGA AGCTAGTTTC  4450
TAAAACAGTT CGGGGTGGAT GTTTGCCTCT GATTAGCCTG TGTATGTGTT  4500
AGATGAATGA GCAATTGGGA AGTCATGACC ATGTGTTTGC TGTCATTTAA  4550
ACTGCCAGTC TAATCTAAAT ATACATGTGG GTGCCTAAGG AGACTTCTTT  4600
CTTTTCAAGA AGATCATTCT TAACAGCTCT CCATTCCTTG CTGACACACA  4650
TGAACTTAGG GGCACTTCTG TTTTCCTGAG GGCTACACCG TTTTCTTTTG  4700
AAAGAGGTGA GAACAATTTC TTTGCTTGCT GTGTTTCTAA TTTTTAAAGA  4750
TAAACTATAT GTAAAGTTTA TTTACTATGA GTAGTAATTT CATAGGCAAG  4800
```

```
         10        20        30        40        50
ATAGCTTCTG GGATATTATG TAAGGTAATG AAAAATGGGG TGAGAGGGAT   4850
GGAACAAGGA AAAGCCACAG GAGGTTTTGA TGTATAGAAA AATTCAGGAA   4900
AATCTTACAA TTTTAGATGA TTAATGTGCA TATAACATGC ATTCATCTAA   4950
TTAAGAAAAA AATAGAGAAA AAGTTTTAAA ATAGGAGCCC CAGTGTGCAT   5000
TTGTTCTCCA GAGTGATTTA AAGTAGGAAA AAAGATACCC TTCTCCTCTG   5050
ACTTCAGCAC ATTTACAAGT ATTTTATCTA CCAGTGAATT TTTAAGTGCA   5100
TGTTAAAGAA ATAGAAGCAT TTTTTAACTC ATGATTATCG TATTAAATAA   5150
TTTCTCTTAT TTTCTGAGAG AAGCTAAGGA TTAGACCTCT AATTATTTTA   5200
CTAATCTGAA ATGAATCTTC TTTCAGAATA ATATTGTTCT TACATTTTTC   5250
TTAATTATAT AGCTTATCAT ACATTGAAAG TCGTTCATCC AAAATGATTT   5300
TCTCCATTCA CATATTTTTA AATATAAAAG TAAACACTGC AAGATTTAAC   5350
TTACCAGAAA AATAACAGAG TATCCCTTGT TATCTAACCA CTAGGGTATT   5400
TGGTAAAAAT TAGCCCTTAA TTATACATTA AAATGTTGTT TTATTAAATT   5450
AAAAATATAA CTATGTAAAT TCTCATGCAA GTACATGTAT TTCATTTTTA   5500
ACTTGTTTAA TAGTTATATG ACTGAATGGC ATTTAGATTG GTTTACAGGT   5550
TAATATACAT GGCAAATATA ATCTAATATA AGTAGTTACA AAATTTAAAG   5600
AACACAGCAT TTGTTCAGAG TTAGGTGCTT TTTTACACAT AGGCAATAAG   5650
TTCTGCCATT TTCTTCCTAG AGAAATATAA CAGTTATGGA GTTTTATATT   5700
TGGAATATCG GAAAGCGTGC ACATTCTTCA GCAGAAATAA ATTTTCCTGG   5750
AACTACATTC AAGGAGAATT TGTCATGTGG ATGTCTATAA ATTAAATGCA   5800
ATCACATGAT GTCCAAATAG TTAATATTAT ATGTAAATGT TGTATATGAA   5850
TGCTAATCCC GATAAAATTG TTTTTCTGAT GCACTAAAAC ATGTTGCTTC   5900
TACCCAAATG TGGAAATATC TAAAGGTTTA AAATACTTTG TTGGAGCATT   5950
TTCTTTTGAG ATTTCATTTC AGCACAATTT CAGAGAAAAC AGAAGCTGAT   6000
CATAATAAGT TCCTTTTATT CTAGGAAGAG TCATATTTAG TATATTATAT   6050
TTTAGAGCAA TAAGTGGAGC CTTAGTGTTT CTTTAAAGCT AGTCTCTCCC   6100
CTTCCTCCTG CCATTGCTCA GCCCATTCTC AGAGAATTCT GCAGTTTGCT   6150
CAGTTCATCA GGAAGATCAT GAGACTAATA ACATCAGTAT TTCTGGCTTT   6200
CTTTCGCATT TTTAGTTTAT AAGAAAACAG CACTTTCCAG TCTAATTTGA   6250
CTTTCAAGTT TCCAGCAGCC AGCTAGAAAA GAGAATGTAC TCTGTTACGC   6300
ATTTCACATT GTTCCTAAGA TAAAAACCAA TCTCTTAGGA GATGTTTAAG   6350
TCAGATGAAA TTAAGTGAAA AGAACGGTCT GTGTTTTCTC CTTGCTTTTC   6400
GGTCTATGGT GTTTAAGTGT AGTGGATAAA AGGTAATAAG CAGCCTTTAC   6450
TACATAGGAA GAAATTGGAA TGGCAGGTTA CAACTAAATG TGTTGAAGTT   6500
TTAAAAGCAT CCATATTTGA ATTGTACGTG CAATTGTGTG TGGATAGTAA   6550
GGTTTGTAGA ACTGAATAAT AATTCCTGAA TAATGATTCC ACAAGTAGGA   6600
ATTGCTCCTG TTTTATGAAG ATTAAAAACA AGTTAGGCCT TTAAATCATT   6650
CTTGAGATAT TTGAAATTGA ATGAGTTTCT GAGGATCAAT TCAGAAACGA   6700
TCCCTAACCA TCCCCTAAGA ATGCTGTAAA TCATGTGTGC TGTTGTTCTA   6750
ATTGAACAGT GTGGATCTGG ACCCTCACTA ACTTGTCTGT AGGAAGAGAA   6800
AATTGAAGAG AGAGAGGAAG AGGCTGAGGA GAAGACCTTT ACCGGTGCTG   6850
TCCTGAGCTG CCATTGTTGA AGGACATCCT GTGAGCTAAT TTCAGTCCCA   6900
TTTTAAAAAA AATCTGTTTA GTGATTAGTA ACCAGTGGAT AAGAGTAGAA   6950
AGCCCAAGTG AATAAAGGCC GTAAATACTG AAGTCTTTGG GATAACAGCC   7000
TGTAAAACTG TCAAAAACCA TTGCTTAAGT GACAAAAGAT AGCCTGAAGT   7050
AACAAGAATG TTAGCTCACT GGGCAATACT ACATATTCAT GAACGAGGAG   7100
ATGTAATTTA GTTTTCTTCA AAAATAGAGA TTCTTCTTGA ATCCAGTACC   7150
TGCCTCACTT CATTTATTTA TTTGTTTTGG ATGGTGTCTA TTTGTTTTAA   7200
```

```
         10        20        30        40        50
GTTTATGGTT AACTGCTTTT GCTAGTAGGA GTCATGGTGC CTGTTTAAAG   7250
ATAGCTTTGT TTCCAAAAGT GGGACTCTGA TTACTGAAAC TTCTGTTTAG   7300
ATCCATCAAG ATCTACTTAC TTGTTAAGAA TGTTAAGCCT AACTGGAGTG   7350
AATGGGTGAG TATGTGTGTG TGTATATTTT TAAAAGGCAG ATGGTTAGGA   7400
ATTAGATGCT GAAAAGTACC TTCTTATTGT TTTTAAATTA AAACCTAAAG   7450
TTTGTAATTC TTAAATATTT CTCCCTAAGG GGAAAAATTA AGTAAATAAA   7500
AACAGAGGAA TGTTAAGTCT ATATTGCAAA CTGTTTTGTT ATTCCTGCAC   7550
AATATTGAGA CTCAAATGAT TAAAGTCTTA CTAGCCAGTA ATTTAGATTT   7600
GTTTTTAGGG CTAGAGAAAA TTCAAGATGA TGAATCCAGT GTCTTTATTC   7650
TAAATTCCAA TTGTTCTCAC TATGATGATG ATTTGGGGGC TTTTGGGGGA   7700
CAATCTAAAT TCTGTGTTCT TTAGTTTTGT GGTATAACCT CTAAATCCCA   7750
AATGTATGTT TTAGAGATGG GACTTCAGTG TATTTAAAGT GTTGTGCTTC   7800
TTCAAGTTAG AGAATTTTAT CAAGTTAGAG TATTTGCTTT TCAACGTCCA   7850
GGCTGTGAGA ACTAGAGGTG TCATTCATTT CAAGGGAGCA GATGTTAAAG   7900
GTTCTTGCCC CTGCCTCCCT AATCCCAAAT ACCCTTGAAG AAGAGAAACG   7950
TTCTGTTACC CAAATAGGCA TCTTTCTGAT GGTGAGAAAA GGCGGTAAAA   8000
GGCAGTAAAT CTGCCTTCTC ATTATGAAGT GCAACAATAG ATTGATTTAT   8050
GAAGAAAAGC AACTGTTAGG CTTATATAAG ATGTGTGCAT GGAAATGCAG   8100
GTTCCTTTAG AGCCTTATAA GCAGCAGCAT TCCTCCTATT CTGGACAGAA   8150
ATCTAATATT ACTAATCAGG AATATGTGAT TCTGAATACC ACATTTTTAA   8200
AGGGGCTACA TTGAGAACAC AACAGAAGAA TCTACCTCCT AAGTTTTCAC   8250
ATGAGGGCTG TCTTGTAAAT GTAGCATGTT TTTGCTACTT TAGTTCATTT   8300
TCACTAAGAT AGAGGGACTT CTTAAAATAG ATATATTTTT ACAAACTTTT   8350
TATTAGGACT TAACATTTTA AATAGAAGTG AATGTAAGTA TAGCACATTT   8400
AGACTGATTA TAAACAATTT CACTTAATAT TGTATAATTT TCTGTAAACA   8450
CTATAAAAGC AAACTCTAAA ATAGAGCCAC AGATAATTTC AAACTTCTAG   8500
GGGGATTATG TAAAGTATTT TGTGTTGGGT TTTTAAGAAA AGCTATGTGC   8550
ACTTCAAAAT CTTTTAAATA AGTTTGTGTT ACCTCTTCTT CAATAAGGAA   8600
CTCTGTGTAC AGTATAAATA TAAAAAGTAA ACTTTTAAGC TCTGAGTTTG   8650
ATTGTTATAA AATACTGTAA AATGACAAAA ATCTGGATGC TCTCTGTGTG   8700
AGTTATAATG AGACTGTCCT GTTTATTTAG CATGTATCAC TCCAGAGCTT   8750
GAAATCATGT GATATAATTA AGATTTGAGC AGGTAAATTA CAGAATTAAA   8800
ACTTACAGAG TTGGTATTAC AACTCATCTC TCCCTGTAAA CCTGTATCCC   8850
CAATGTATGT ACATTTCCAA ATTATAGTAA GCCAACATAT TATGTACAAG   8900
AATAAGTTAT GCTGTACTTT CATATTATGT ACAAGAAAAC ACATACAAAA   8950
AATTGAAAAT TTAAAAGTAA AAGACTAATA TGAAAAGAGG TCTTGTTATT   9000
TTCTTTTACT ACTATGTATT ATCTTACACA CATCTATGTA AACATATACA   9050
TTCATACCCA CACCCAAAGA ACTTTGGAGT TCCTTGATCT AGATAGAGGA   9100
AAGAGCTCCA GTATTCTAAT TCTATTGAGG AAGATATAAG GCTTGTAGAG   9150
AAATAACCCT TGCAACTATC AGTAGTCCTT GTTTCATAAG CTGATAATGA   9200
GAAAATGTTG CATTTTTGAC CTGGTGAAGA TGAGGTAGTC TTAAGTCAAC   9250
ACTTTTTCCC CAGTGGATTT TGGTTATATT AAAAGTCAAA CTACCATGGA   9300
CATAAGTTCT TAGCAAATAA AATAAACTCT TAATGCCACT GCCATGATAG   9350
TAAATCCAGG TTTCACATTT AATCGGGAAA ATTCTGTAAG ATAGTCCAAA   9400
GATTAAAATT AATACTTCTT GACATCTGTG TAAATCTGTA AAGATTATAT   9450
TGGTGTCAAT CAAATGCTTT ATGAACATTT TAAAGTGACA CATTTTTTAA   9500
AAGGTTAAAG TTAATAAGGA AATGTAGTTA AATGTCTTGT CAAAATAATC   9550
TCTTCAGATA TATATCTTTT AGTTATTTCA TTTCATTATC ACCTTTATTT   9600
```

```
         10        20        30        40        50
CCTGAATTCA TTCTATTTAA TTTACTTTAT ACTTTTTTAT AATCCCAAAT      9650
ATACTTTAAT ATTTATAATC CAGTACTGAG CCCTGTAGCC ATAATGTACT      9700
TTTTATGACT TTGTTCAATT CAGTGTTCAG TTCAATGTTG CCGTTATAAA      9750
TGTTGAGAAG ACAAAAGTGA TTTTCCACTA TTTTAAATTT CTTATGCTCT      9800
GCACAGGAGG GTTACTTCAC AAGTAACTGA ATAGTTCTGA CAATAATCTC      9850
AAGATCTCAT TGTAAACTGA CAAGTAGTGG TGTTTCGATG TTCATTGGTC      9900
AACTCCAATA ACACTTAATT ATGGGCCAGT TCTATCAGGA AACTTTTTTT      9950
CCTTCCTGTT TCTCTTCTGT GAGGAATTAC AGCTCTGAGC CCTCTCACTT     10000
GTAACATTTT TCTTTCTCTA CCCGTCACTC TCAAAACACA CAAAAGAGAA     10050
AGAGAAAACA CAGCTAAAAC ATATGTCTTT GGCATCACAA CTAAGTGGTG     10100
TAGTTAAAGG TCAGTTGTGA AATTCAACTT CTGTACATGT GAACTTTGCT     10150
CAAAAATAGT ACTGTGTTTA TAGGTGGTGT GACTGTAAGT GGTCTGTTTT     10200
CAGATGATTT CTCATTCTCA AAGCTCTTAA CAAGGATCC                 10239
```

Appendix G: DNA Sequence Representing the 950 bp Intact Fragment of the *AF9* SAR

Forward primer used in the second round of IPCR of the *AF9* gene (AF9 245507 F)

*Bam*H I sites

```
                10        20        30        40        50
ATTGGTGTCA ATCAAATGCT TTATGAACAT TTTAAAGTGA CACATTTTTT    50
AAAAGGTTAA AGTTAATAAG GAAATGTAGT TAAATGTCTT GTCAAAATAA   100
TCTCTTCAGA TATATATCTT TTAGTTATTT CATTTCATTA TCACCTTTAT   150
TTCCTGAATT CATTCTATTT AATTTACTTT ATACTTTTTT ATAATCCCAA   200
ATATACTTTA ATATTTATAA TCCAGTACTG AGCCCTGTAG CCATAATGTA   250
CTTTTTATGA CTTTGTTCAA TTCAGTGTTC AGTTCAATGT TGCCGTTATA   300
AATGTTGAGA AGACAAAAGT GATTTTCCAC TATTTTAAAT TTCTTATGCT   350
CTGCACAGGA GGGTTACTTC ACAAGTAACT GAATAGTTCT GACAATAATC   400
TCAAGATCTC ATTGTAAACT GACAAGTAGT GGTGTTTCGA TGTTCATTGG   450
TCAACTCCAA TAACACTTAA TTATGGGCCA GTTCTATCAG GAAACTTTTT   500
TTCCTTCCTG TTTCTCTTCT GTGAGGAATT ACAGCTCTGA GCCCTCTCAC   550
TTGTAACATT TTTCTTTCTC TACCCGTCAC TCTCAAAACA CACAAAAGAG   600
AAAGAGAAAA CACAGCTAAA ACATATGTCT TTGGCATCAC AACTAAGTGG   650
TGTAGTTAAA GGTCAGTTGT GAAATTCAAC TTCTGTACAT GTGAACTTTG   700
CTCAAAAATA GTACTGTGTT TATAGGTGGT GTGACTGTAA GTGGTCTGTT   750
TTCAGATGAT TTCTCATTCT CAAAGCTCTT AACAAGGATC CGGATCCGAA   800
GTAAGAAAGA CGTATTTTCT GTCCTCATCT TTATCTTTGG CTAGTGTGAC   850
TGCAGGGAAA AGTGGCTAAT ACAAACATTT AGATAACTTA CACACAAAGC   900
AAAAAGTAGA ATAGTAGAAT TTTACCAGGT ATGAAAAGGA TATG         944
```

Reverse primer used in the second round of IPCR of the *AF9* gene (AF9 236211 R)

Appendix H: DNA Sequence of the Region of Study in the Non-SAR of the *AF9* Gene (Nucleotide position 71116-75282) [Ensembl: ENSG00000171843]

```
         10        20        30        40        50
GGATCCGGGT GGGTGGATCT GATGACTTGG CAACTCAGAG AGGGAAGCAA    50
GTGTAATTGT TGATTTTATC AGAGCAGGTT TGTGTATTTG TTTGTGCACG   100
TAGGGGGATG GGGGTTGAAT GAAGACTTAA CAACATGAAT TTCGTGAGGT   150
TACAAAAGTC TACTGCTGCA TAAATCGTAG GTAATATGTT CATGTCTATA   200
GGGTGTAATG TTAATGTTCA ATAAAAATTA ATATCAAAGT CTGTCATAGA   250
TTTTATGTTT TGGGGCCTCC AGGTAAATGG TTTGTATCAT CATTGGGAAT   300
TTATGGAAGT TGATTTAAGT ATACAGTAGT TAGCACTCTA TTAAATACTG   350
AAGTCATGCA CTGTGTCACC TTGATCAATG AGCAAGTTTT TTTTTCTTTT   400
CTGATTCTGG AATTTCAGCA CTTTCAGCTT TGAAGGTTGC CTTGGGTTTG   450
TTTTTGAATT AAAAGTGTCA CAGGAGAAAG GCTGATTGTT ACATTCATGA   500
CCAATCAGAA GGGTCTGTAG GACTCAAAAT GTTTGGTACA TTCCTTAAAT   550
TAGAAGATAT GCTTCTGGAC TTTTCCCAGG GAAAGCAGGT CCATGTAATA   600
ATTTACCAAA AGGAAGTTGG GAAAATGAAA CATCACCTTT TGTGAGAAAG   650
AGGATGGGGT GGGGAGAGTG CTTTAAAGTG ATCTTTGCTT TGTGGAATGC   700
ATTGGATTAA AGAGTCACAG AGGCCACAAA AAGCAGCTCA GTGAGTTATC   750
AAATATAGGA GTACTTGAGA AATTAAACCA GGTATGTTAT TATTCTTTCT   800
TGGGGAGCCT AAAAATAAAC TCAAGGATTG GGATTGTTCC TTTTACTTTC   850
ATATCTGCTG GCTGCTAAAG GCAATTGGAA ATTTTAGCCA ACATATTGAC   900
CTACTTGATT TTCACGATGA GTATTTTTTT CCTCAGCCGA ATTGGCCATT   950
TAAAGAAAAC TTTGTGGCTG GGTGCTGTGG CCCACGCCTG TAATCCCAGT  1000
CTTTGGAAGG CTAAGGTGGG AGGATTACTT GAGGCCAGGA GTTTGAGACT  1050
AGCCTGGTTA ATGTAGTGAG ACCCCATTTC TACCAAGAAA AAAATGAAAA  1100
AAATTTAGCT GGGCATGATG GCACATGCCT GTAGTCCTAG CTACTTGGGA  1150
AGCTGAGGTG GGAGGATTGC TTGAGCCCAG GAGGTTGTAG TTATAGTGAG  1200
CTATGATCAT GCCACTGCAG TCCAGCCTGG ACAACAGAGT GAGACCCTGT  1250
CTCTTAAAAA ACTAAATAAA TAAAATAAAT AAGACTTTGT AACATATTAA  1300
ACAAGCTTTA TCATATTAAA AAATACACAC AGCTGGGCAC AGTGGCTCAT  1350
GCCTGTAATC CCAACATTTT AAAGAGGCCG AGGAAGTTCA AGACTAGCCT  1400
GGGCAACATA GTGAGACTCT GTCTATAAAA AAAAATTTAG TCATGTGGTG  1450
GTGGTGTGTG CTTGTGGTCC CAGTTACTCG TGGGGTGGGA GGATCACTTG  1500
AGTCTGGGAG GCCGAGGTTA TGGTGAGCTG TGATTGTGCC ACTGCATTCC  1550
AGCCTGGGTG ACAAAGCAAG ACCTGTCTCA AAAATAAAAA ACAACAACAC  1600
ACACACACAG TGAAACGAAT GGCACCAAGG AAGAAATCCA AGTGCAAGAA  1650
AAGAGGGAAT CATACAATAA GAACAATAGG GTGGAAAGCC AGGGTGAAAT  1700
ACACATCTCT GATGAAGGAA GGGCTGTGTA TCTGGGTCCT TACTTTTCCC  1750
CTTTTACTTG GGTGCTTCTG TCATTATTAT TTGTAAGTCA CTTCTCTCAT  1800
GAAGTGCTTT GAGCTTTTGG GGCAAGTAAG AGTAGAGAAA CTAGAAATAG  1850
TTGGCAATTC ATTATCAACC TTGTTACGTG GGATTAGAAC TATTGTCTAA  1900
GATGGGATTG TGAGTTTTTG AATTGTAGGT GAGATTTTAA TGCGTGTGTC  1950
CATGCCCCCA TTATTGAAGT GGGTTTGGAG TTCATGATAT TAGAAGCAAT  2000
TCCTTAGTTA GCAGTTCTTG TAGATGAGTG ATGAGCCACA AGATTTCTGT  2050
TGAGATCAAC ATTCAGTTGT ATGAAATTCT TGAGGTGCAG AGGGGAGTAG  2100
AATAGGAATT TCCTAATTTG GTTGGATTGA CTCCATTTTG GAACTAGATT  2150
TTCTGCACGC ATGTGCTAAG AAGAGAGATG GTGGTCACAG TGAATGGATC  2200
ACTGCGCAAA GATCTTTGGT GAGTAAATGT GGAAAGAAGT TTTCATGTAT  2250
TTGTATTTCT GTCAGGGTGG GATTAAAAGG TCTGAGTAGG TAGTGGTCTT  2300
TATAATTCTC AGACCAGCCA AGGTGAATTT GATAGTCTTG GCTGTACAGC  2350
GTAGTGGTTA AAAGCATGCC ATTTGGAGCC AGACTGTTTA GGTTCAAACC  2400
```

```
         10         20         30         40         50
TTAGGTCTTC CGCCAGTCGC CTGTGTGACC TTTGTTCATT AACTGTCATT 2450
TCCATACTTT CTTCATCTTC AGAATGGGTC GCATAAAGTA CCTGCTTTCT 2500
AGAGTTGTTG GGCAGATTAA ACGAGTTAAC TACGTAAAAT GCTTAGGACA 2550
GTTCCTGGCA CGGAGTAAAT GGCATATATG TGTTAGCTAT TATTATTACT 2600
ACTAATATTA TTAGTAGGAA CTATTTGAGT TAGTCGTTGT TGTTGTTATA 2650
AACCATGCTG GGTTGGTAAC AGAGATTGTT ACCAACAATT GTTGGTGTTT 2700
CTTTGTAGAG AAAGCAAAGC CTGATGTCAT TTTTATCTTT GTGTCCAGTG 2750
TTTTGGCTCA CACAAATTAT AGCTACTTGG CTTTTCTCAT TCTGATGTTG 2800
CCAAGGAATC TTATAATTTA ATCAGTGATT TGAGTGGCCT GTGCCTTGTC 2850
AAATGAAGGC ACTATTCCTA ATGCCTCATT TTATTATTAT TGTTATTTTT 2900
GTTAATATCT TTGAGGCAGG GCGTGATGGC TAATGCCTGT AACCCCAGCA 2950
CTTTGGCAGG CCCGGGTGTG TGGATTGCTT GAGCCCAGGA GTTCGGGACC 3000
CACCTGGGCA ACATGGTGAA ACCCCCTCTC TACAAAAAAT ACTAAAATTA 3050
GCCGGGTATG GAGATGCATG GCTGTAGTCC TAGCTACTTG GGAGGCAAAA 3100
GTGGGAGGAT CAGCTGAGCC GGGGGAGGTC GAGGATGCAG TGATGGTGCC 3150
ACTGCACTCT AGCCTGGCTG ACACAGTGAG ACCCTGTTTG GAAAAAAAAA 3200
AAAAATCTTT GAGCAAAATA TGTGTTTTGG TAAGTTTGAG AATATTCAAA 3250
TTCCCAACAA ATATTGCCCA TCAGTCTTTC AAAGGCATTC AGGTGAGTAG 3300
TTTATTCTTT TTGCAGCTCC TTGCTGCCTT TTTATCAGCC TGCTGCCCAG 3350
AGCCCGCAGA AGAGATTACA CCTGGGACAG GGGATGACTT TTCTTCAATC 3400
TTACTCTTAA CAGGGCCATT CCAAATGGAT TTTAGAACAA TTTCACATAA 3450
TTAGGGTAAG CAGGTTTTAC CCTTGGTGTT TCCAATAATT AATTATTACA 3500
ATTAAAATCC ATTTACTGTA TGTACTCCAA TGGGATGTCA AATGTAATCT 3550
GTATTTAGCA CCCCCCTGGA AAACTTAAAA TATTTTCATT TACCATTGTT 3600
CAGACTCCAA ATGCTGATGG CTTTTCAGAG AACAAAAATG TCTTCCTTGC 3650
GGAGGAGGTT AATCTCAGCC TGTACCAGTT AGATACTACG AGAGCCATCA 3700
TAAACGCAAT GTCAGAGAAC TGGAGAAGAT AGAGCCAACT GATGAGATAG 3750
ATGCTTTTTC CTGTTCCTGT TAGACATATG CTGTCCAATA TGGCAGTCAC 3800
TAGGCAAATT TGACTGTTGA GTACTTGAAA TGTGGTTAGA TTAGAGAATT 3850
GAATTCTAAA TTTCATTTAA TTTTAAACTT AAACAAGAGG CTTGATTCAA 3900
TTAATGGAAA ATTGTAAATA TGTTTAGAGC AACTTGGGTA TGTGAATCTA 3950
CCTTTTCAAC TGTAAAGTTT TTTTGGACTC TAGATACAAG TCAAGTAGTC 4000
AAATGAAAAT TTAGCATCTG AATTGAGATG TGCTATAGGT GTAAAATATG 4050
TACCAACTTT TATTTTGTTT TATTTTTGAG ATAAGATCTT GCTGCGTTGT 4100
CCAGATGGGA GTGCAGTCAT GTGATCTCAG TTTATTGTAG CCTTGAGCTC 4150
CTGGGCTTGA GGGATCC                                     4167
```

Appendix I: DNA Sequence of the 950 bp Intact Fragment of the *AF9* Non-SAR

Forward primer used in the second round of IPCR of the *AF9* gene (AF9 74494 F)

*Bam*H I sites

```
         10         20         30         40         50
AGGGGATGAC TTTTCTTCAA TCTTACTCTT AACAGGGCCA TTCCAAATGG    50
ATTTTAGAAC AATTTCACAT AATTAGGGTA AGCAGGTTTT ACCCTTGGTG   100
TTTCCAATAA TTAATTATTA CAATTAAAAT CCATTTACTG TATGTACTCC   150
AATGGGATGT CAAATGTAAT CTGTATTTAG CACCCCCCTG GAAAACTTAA   200
AATATTTTCA TTTACCATTG TTCAGACTCC AAATGCTGAT GGCTTTTCAG   250
AGAACAAAAA TGTCTTCCTT GCGGAGGAGG TTAATCTCAG CCTGTACCAG   300
TTAGATACTA CGAGAGCCAT CATAAACGCA ATGTCAGAGA ACTGGAGAAG   350
ATAGAGCCAA CTGATGAGAT AGATGCTTTT TCCTGTTCCT GTTAGACATA   400
TGCTGTCCAA TATGGCAGTC ACTAGGCAAA TTTGACTGTT GAGTACTTGA   450
AATGTGGTTA GATTAGAGAA TTGAATTCTA AATTTCATTT AATTTTAAAC   500
TTAAACAAGA GGCTTGATTC AATTAATGGA AAATTGTAAA TATGTTTAGA   550
GCAACTTGGG TATGTGAATC TACCTTTTCA ACTGTAAAGT TTTTTTGGAC   600
TCTAGATACA AGTCAAGTAG TCAAATGAAA ATTTAGCATC TGAATTGAGA   650
TGTGCTATAG GTGTAAAATA TGTACCAACT TTTATTTTGT TTTATTTTTG   700
AGATAAGATC TTGCTGCGTT GTCCAGATGG GAGTGCAGTC ATGTGATCTC   750
AGTTTATTGT AGCCTTGAGC TCCTGGGCTT GAGGGATCCG GATCCGGGTG   800
GGTGGATCTG ATGACTTGGC AACTCAGAGA GGGAAGCAAG TGTAATTGTT   850
GATTTTATCA GAGCAGGTTT GTGTATTTGT TTGTGCACGT AGGGGGATGG   900
GGGTTGAATG AAGACTTAAC AACATGAATT TCGTGAGGTT ACAAAAGTCT   950
ACTGCT                                                   956
```

Reverse primer used in the second round of IPCR of the *AF9* gene (AF9 71282 R)

Appendix J: DNA Sequences of IPCR Bands Representing Chromosomal Breaks within the *AF9* SAR Identified in *LMP1* Transfected Cells.

DNA sequencing was performed by using the reverse primer that was used in the second round of IPCR of the *AF9* gene. The arrow above the sequence shows the direction of the forward primer (*AF9* 245507 F). The double arrow indicates the breakpoint. The nucleotide positions of the chromosomal breaks were mapped to the *AF9* sequence accessed from Ensembl database [Ensembl: ENSG00000171843].

1. Code SE1R1: DNA sequence representing chromosomal break detected within the *AF9* SAR region in *LMP1* transfected cells. The breakpoint is at coordinate 245527.

→ Direction of forward primer (*AF9* 245507 F).

```
         10         20         30         40         50
AGTCGATTGG TGTCAATCAA ATGCTTGATC CGAAGTAAGA AAGACGTATT    50
TTCTGTCCTC ATCTTTATCT TTGGCTAGTG TGACTGCAGG GAAAAGTGGC   100
TAATACAAAC ATTTAGATAA CTACACACAA GCAAAAAGAG AAATGG       146
```

2. Code SE1R4: DNA sequence representing chromosomal break detected within the *AF9* SAR region in in *LMP1* transfected cells. The breakpoint is at coordinate 245575.

→ Direction of forward primer (*AF9* 245507 F)

```
         10         20         30         40         50
GTTTATTGGT GTCAATCAAA TGCTTTATGA ACATTTTAAA GTGACACATT    50
TTTTAAAAGG TTAAAGTTAA TAAGATCCGA AGTAAGAAAG ACGTATTTTC   100
TGTCCTCATC TTTATCTTTG GCTAGTGTGA CTGCAGGGAA AAGTGGCTAA   150
TACAAACATT TAGATAACTA CACACAAAGC AAAAGAGAAG CTGC         194
```

3. Code SETR1: DNA sequence representing chromosomal break detected within the *AF9* SAR region in in *LMP1* transfected cells. The breakpoint is at coordinate 245596.

→ Direction of forward primer (*AF9* 245507 F)

```
         10         20         30         40         50
TTTATTGGTG TCAATCAAAT GCTTTATGAA CATTTTAAAG TGACACATTT    50
TTAAAAGGTT AAAGTTAATA AGGAAATGTA GTTAAATGTC TTGATCCGAA   100
GTAAGAAAGA CGTATTTCTG TCCTCATCTT TATCTTTGGC TAGTGTGACT   150
GCAGGAAAAG TGGCTAATAC AAACATAGAT AACTACACAC AAGCACAGGA   200
ATAAC                                                    205
```

4. Code SE1R9: DNA sequence representing chromosomal break detected within the *AF9* SAR region in *LMP1* transfected cells. The breakpoint is at coordinate 245649.

→ Direction of forward primer (*AF9* 245507 F)

```
         10         20         30         40         50
TATTATTGGT GTCAACAAAT GCTTTATGAA CATTTTAAAG TGACACATTT    50
TTTAAAAGGT TAAAGTTAAT AAGGAAATGT AGTTAAATGT CTTGTCAAAA   100
TAATCTCTTC AGATATATAT CTTTTAGTTA TTTCATTTCA TTATCAGATC   150
CGAAGTAAGA AAGACGTATT TTCTGTCCTC ATCTTTATCT TTGGCTAGTG   200
TGACTGCAGG GAAAAGTGGC TAATACAAAC ATTTAGATAA CTACACACAA   250
GCAAAAGAGA AAGACT                                        266
```

5. Code SE1R2: DNA sequence representing chromosomal break detected within the AF9SAR region in *LMP1* transfected cells. The breakpoint is at coordinate 245725.

→ Direction of forward primer (*AF9* 245507 F)

```
         10         20         30         40         50
ATTATTGGTG TCACAAATGC TTTATGAACA TTTTAAAGTG ACACATTTTT    50
TAAAAGGTTA AAGTTAATAA GGAAATGTAG TTAAATGTCT TGTCAAAATA   100
ATCTCTTCAG ATATATATCT TTTAGTTATT TCATTTCATT ATCACCTTTA   150
TTTCCTGAAT TCATTCTATT TAATTTACTT TATACTTTTT TATAATCCCA   200
AATATACTTT AATATTTATA GATCCGAAGT AAGAAAGACG TATTTTCTGT   250
CCTCATCTTT ATCTTTGGCT AGTGTGACTG CAGGGAAAAG TGGCTAATCA   300
ACATTAGAAA CTCACACAAG CAAAAAGTGA ATTGT                   335
```

6. Code SETR4: DNA sequence representing chromosomal break detected within the AF9 SAR region in *LMP1* transfected cells. The breakpoint is at coordinate 245979.

→ Direction of forward primer (AF9 245507 F)

```
         10         20         30         40         50
TATTGGTGTT CATCAAATGC TTTATGAACA TTTTAAAGTG ACACATTTTT    50
TAAAAGGTTA AAGTTAATAA GGAAATGTAG TTAAATGTCT TGTCAAAATA   100
ATCTCTTCAG ATATATATCT TTTAGTTATT TCATTTCATT ATCACCTTTA   150
TTTCCTGAAT TCATTCTATT TAATTTACTT TATACTTTTT TATAATCCCA   200
AATATACTTT AATATTTATA ATCCAGTACT GAGCCCTGTA GCCATAATGT   250
ACTTTTTATG ACTTTGTTCA ATTCAGTGTT CAGTTCAATG TTGCCGTTAT   300
AAATGTTGAG AAGACAAAAG TGATTTTCCA CTATTTTAAA TTTCTTATGC   350
TCTGCACAGG AGGGTTACTT CACAAGTAAC TGAATAGTTC TGACAATAAT   400
CTCAAGATCT CATTGTAAAC TGACAAGTAT TGGTGTTTCG ATGTTCATTG   450
GTCAACTCCA ATAACACTTA ATTAGATCCG AAGTAAGAAA GACGTATTTT   500
CTGTCCTCAT CTTTATCTTT GGCTAGTGTG ACTGCAGGGA AAAGTGGCTA   550
ATACAAACAT TTAGATAACT ACACACAAAG CAAAAGTGAA TTGGG        595
```

Appendix K: DNA Sequences of IPCR Bands Representing Chromosomal Breaks within the *AF9* non-SAR identified in *LMP1* Transfected Cells.

DNA sequencing was performed by using the reverse primer that was used in the second round of IPCR of the *AF9* gene. The arrow above the sequence shows the direction of the forward primer (*AF9* 74494 F). The double arrow indicates the breakpoint. The nucleotide positions of the chromosomal breaks were mapped to the *AF9* sequence accessed from Ensembl database [Ensembl: ENSG00000171843].

1. Code NSE1R11: DNA sequence representing chromosomal break detected within the *AF9* SAR region in *LMP1* transfected cells. The breakpoint is at coordinate 74653.

   → Direction of forward primer (*AF9* 74494 F)

```
         10         20         30         40         50
ATCTTGGCAT TCAAATGGAT TTTAGACAAT TTCACATAAT TAGGGTAAGC     50
AGGTTTTACC CTTGGTGTTT CCAATAATTA ATTATTACAA TTAAAATCCA    100
TTTACTGTAT GTACTCCAAT GGGATGTGAT CCGGGTGGGT GGATCTGATG    150
ACTTGGCAAC TCAGAGAGGG AAGCAAGTGT AATTGTTGAT TTTATCAGAG    200
CAGGTTTGTG TATTTGTTTG TGCACATAGG GGGATGGGGG TTGAATGAAG    250
ACTTAACAAC ATGAATTTCG TGAGGTTACA AAAGAAATAC TGCTA         295
```

2. Code NSE1R3: DNA sequence representing chromosomal break detected within the *AF9* SAR region in *LMP1* transfected cells. The breakpoint is at coordinate 74752.

   → Direction of forward primer (AF9 74494 F)

```
         10         20         30         40         50
AATCCAGTCA TTCAATGGAT TTTAGACAAT TTCACATAAT TAGGGTAAGC     50
AGGTTTTACC CTTGGTGTTT CCAATAATTA ATTATTACAA TTAAAATCCA    100
TTTACTGTAT GTACTCCAAT GGGATGTCAA ATGTAATCTG TATTTAGCAC    150
CCCCCTGGAA AACTTAAAAT ATTTTCATTT ACCATTGTTC AGACTCCAAA    200
TGCTGATGGC TTTTCAGAGA ACAAAAGATC CGGGTGGGTG GATCTGATGA    250
CTTGGCAACT CAGAGAGGGA AGCAAGTGTA ATTGTTGATT TTATCAGAGC    300
AGGTTTGTGT ATTTGTTTGT GCACATAGGG GGATGGGGGT TGAATGAAGA    350
CTTAACAACA TGAATTTCGT GAGGTTACAA AAGTTCTACT GCTA          394
```

3. Code NSE2R2: DNA sequence representing chromosomal break detected within the AF9 SAR region in *LMP1* transfected cells. The breakpoint is at coordinate 74786.

→ Direction of forward primer (*AF9* 74494 F)

```
         10         20         30         40         50
TTCCTGTCAT TCAATGGATT TTAGACATTT CACATAATTA GGGTAAGCAG    50
GTTTTACCCT TGGTGTTTCC AATAATTAAT TATTACAATT AAAATCCATT   100
TACTGTATGT ACTCCAATGG GATGTCAAAT GTAATCTGTA TTTAGCACCC   150
CCCTGGAAAA CTTAAAATAT TTTCATTTAC CATTGTTCAG ACTCCAAATG   200
CTGATGGCTT TTCAGAGAAC AAAAATGTCT TCCTTGCGGA GGAGGTTAAT   250
CTCAGCCTGA TCCGGGTGGG TGGATCTGAT GACTTGGCAA CTCAGAGAGG   300
GAAGCAAGTG TAATTGTTGA TTTTATCAGA GCAGGTTTGT GTATTTGTTT   350
GTGCACATAG GGGGATGGGG GTTGAATGAA GACTTAACAA CATGAATTTC   400
GTGAGGTTAC AAAAATCTAC TGCTA                              425
```

4. Code NSE2R9: DNA sequence representing chromosomal break detected within the *AF9* SAR region in *LMP1* transfected cells. The breakpoint is at coordinate 75013.

→ Direction of forward primer (*AF9* 74494 F)

```
         10         20         30         40         50
AAATCCAGGC ATTCAATGGA TTTTAGACAT TTCACATAAT TAGGGTAAGC    50
AGGTTTTACC CTTGGTGTTT CCAATAATTA ATTATTACAA TTAAAATCCA   100
TTTACTGTAT GTACTCCAAT GGGATGTCAA ATGTAATCTG TATTTAGCAC   150
CCCCCTGGAA AACTTAAAAT ATTTTCATTT ACCATTGTTC AGACTCCAAA   200
TGCTGATGGC TTTTCAGAGA ACAAAAATGT CTTCCTTGCG GAGGAGGTTA   250
ATCTCAGCCT GTACCAGTTA GATACTACGA GAGCCATCAT AAACGCAATG   300
TCAGAGAACT GGAGAAGATA GAGCCAACTG ATGAGATAGA TGCTTTTTCC   350
TGTTCCTGTT AGACATATGC TGTCCAATAT GGCAGTCACT AGGCAAATTT   400
GACTGTTGAG TACTTGAAAT GTGGTTAGAC TAGAGAATTG AATTCTAAAT   450
TTCATTTAAT TTTAAACTTA AACAAGAGGC TTGATTCGAT CCGGGTGGGT   500
GGATCTGATG ACTTGGCAAC TCAGAGAGGG AAGCAAGTGT AATTGTTGAT   550
TTTATCAGAG CAGGTTTGTG TATTTGTTTG TGCACATAGG GGGATGGGGG   600
TTGAATGAAG ACTTAACAAC ATGAATTTCG TGAGGATACA AAATTCCTAC   650
TGCT                                                     654
```

5. Code NSE2R1: DNA sequence representing chromosomal break detected within the *AF9* SAR region in in *LMP1* transfected cells.

→ The breakpoint is at coordinate 75081.

```
         10         20         30         40         50
TATTCCAAGG TCATTTCAAT GGATTTTAGA CAATTTCACA TAATTAGGGT     50
AAGCAGGTTT TACCCTTGGT GTTTCCAATA ATTAATTATT ACAATTAAAA    100
TCCATTTACT GTATGTACTC CAATGGGATG TCAAATGTAA TCTGTATTTA    150
GCACCCCCCT GGAAAACTTA AAATATTTTC ATTTACCATT GTTCAGACTC    200
CAAATGCTGA TGGCTTTTCA GAGAACAAAA ATGTCTTCCT TGCGGAGGAG    250
GTTAATCTCA GCCTGTACCA GTTAGATACT ACGAGAGCCA TCATAAACGC    300
AATGTCAGAG AACTGGAGAA GATAGAGCCA ACTGATGAGA TAGATGCTTT    350
TTCCTGTTCC TGTTAGACAT ATGCTGTCCA ATATGGCAGT CACTAGGCAA    400
ATTTGACTGT TGAGTACTTG AAATGTGGTT AGACTAGAGA ATTGAATTCT    450
AAATTTCATT TAATTTTAAA CTTAAACAAG AGGCTTGATT CAATTAATGG    500
AAAATTGTAA ATATGTTTAG AGCAACTTGG GTATGTGAAT CTACCTTTTC    550
AACTGTAAAG ATCCGGGTGG GTGGATCTGA TGACTTGGCA ACTCAGAGAG    600
GGAAGCAAGT GTAATTGTTG ATTTTATCAG AGCAGTTTGT GTATTTGTTG    650
TGCACATAGG GGATGGGGGT GATGAGACTT AACAACATGA ATTTCGTGAG    700
GTTACAAAAG TTCTACTGCT AA                                  722
```