Department of Microbiology, Immunology and Parasitology, College of Health sciences

School of Medicine, Addis Ababa University

Molecular Epidemiology and Genetic Diversity of Epstein - Barr virus among lymphoma patients in Ethiopia

By

Seifegebriel Teshome

BSc., MSc. in Medical Microbiology

A thesis submitted to the department of Microbiology, Immunology and Parasitology, College of Health Sciences, Addis Ababa University in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Medical Microbiology

June, 2024

Department of Microbiology, Immunology and Parasitology, College of Health sciences

School of Medicine, Addis Ababa University

Molecular Epidemiology and Genetic Diversity of Epstein - Barr virus among lymphoma patients in Ethiopia

By

Seifegebriel Teshome

BSc., MSc. in Medical Microbiology

A thesis submitted to the department of Microbiology, Immunology and Parasitology, College of Health Sciences, Addis Ababa University in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Medical Microbiology

Advisors:

Dr. Tamrat Abebe; MSc., PhD, Assistant Professor of Medical Microbiology & Immunology, Department of Microbiology, Immunology & Parasitology School of Medicine, College of Health Sciences, Addis Ababa University

Prof. Robert Baiocchi; MD, PhD., Professor of internal Medicine, Division of Hematology, Department of Internal Medicine, College of Medicine, The Ohio State University, USA

Collaborators:

Dr. Fisihatsion Tadesse; MD, Specialist, Associate Professor of Internal Medicine, Division of Hematology, department of Internal Medicine, School of Medicine, College of Health Sciences, Addis Ababa University

Dr. Abdulaziz Abubeker; MD, Specialist, Assistant Professor of Internal Medicine, Division of Hematology, department of Internal Medicine, School of Medicine, College of Health Sciences, Addis Ababa University

June, 2024

Declaration

I Seifegebriel Teshome Feleke, declare that this PhD Thesis is the result of my original work and has not been submitted for a degree at any other university. I further affirm that all sources of material used in this thesis have been appropriately acknowledged and cited.

Name: Seifegebriel Teshome Feleke

Signature: _____ **Date:** 27/06/2024

Address: Department of Microbiology, Immunology & Parasitology, School of Medicine, College of Health Sciences, Addis Ababa University

Phone: +251 910104699

Email: Seifegebriel.teshome@aau.edu.et /yegebrielseif@gmail.com

Advisors				
No.	**Name**	**Email**	**Telephone**	**Signature**
1	Dr. Tamrat Abebe	Tamrat.abebe@aau.edu.et	0911447227	
2	Prof. Robert Baiocchi	Robert.baiocchi@osumc.edu	1614915208	

June, 2024

Abstract

Background: The Epstein-Barr virus (EBV) is a member of the herpesvirus family and is highly prevalent among the global population. It is estimated that more than 90% of adults have been exposed to this virus, making it one of the most common human viruses worldwide. After initial infection, infectious mononucleosis (IM), which typically occurs during childhood or adolescence, EBV establishes a long-lasting latent infection within B lymphocytes. EBV is associated with various diseases such as Burkitt's lymphoma, Hodgkin's lymphoma, nasopharyngeal carcinoma, and certain types of Gastric cancer. Studying the overall viral molecular epidemiology and genetic diversity of EBV can provide valuable insights into the pathogenesis, diagnosis, and management of EBV-associated diseases. The objectives of the study were on various aspects of EBV among lymphoma patients. It encompasses the detection of EBV, determining the genotypes, analysis of its methylation status, and identification of sequence variations.

Methods: The study was conducted at Tikur Anbessa Specialized Hospital in Addis Ababa, Ethiopia, spanning from June 2021 to August 2023. A total of 305 participants were enrolled retrospectively and prospectively, and formalin-fixed paraffin-embedded (FFPE) tissues, blood, and lymph node biopsy samples were collected. From the whole blood samples, serum plasma and polymorphonuclear blood mononuclear cells (PBMCs) were isolated. Additionally, lymphomononuclear cells (LMNCs) were isolated from the lymph node biopsy samples. DNA was extracted from FFPE samples, PBMCs, and LMNCs. Serological tests were conducted using Viral capsid antigen (VCA) IgG to detect antibodies, and molecular detection and quantification were carried out using quantitative PCR (qPCR) by targeting EBV EBNA1 gene. Genotyping was performed using conventional PCR targeting the EBNA3C gene. The viral methylation profile was assessed using the IPLEX assay, and sequencing of the viral genome was performed using next-generation sequencing technology (Illumina). We utilized statistical analysis tools such as descriptive statistics (mean, median, frequencies, percentages) and employed techniques like chi-square test, bioinformatics, and phylogenetic tree analysis for in-depth data examination.

Results: The mean age of the study participants was 38 years (± SD, 17.5). Among the lymphoma patients in the study, 32% (n=91) were diagnosed with Hodgkin lymphoma, while the remaining (68%) patients had non-Hodgkin lymphoma (n=197). Within the non-Hodgkin lymphoma cases, the highest percentage (18%) was attributed to small lymphocytic lymphoma, followed by diffuse

large B-cell lymphoma (13%). Serological testing revealed that 96% of the study participants had evidence of previous exposure to EBV. Using qPCR, EBV was detected in 99% of the lymphoma patients. The quantification of EBV copies/ml ranged from 10^2 to 10^9, with a median viral load count of approximately 1.1×10^3 EBV copies/ml. Among the study participants, 56.3% had a high EBV viral load (>10,000 EBV copies/mL), 16% had a low EBV viral load (ranging from 5,000 to 10,000 EBV copies/mL), and 27.8% had a very low EBV viral load (<5,000 EBV copies/mL). With regards to EBV genotyping, it was found that 52.2% of the patients had EBV genotype 1. EBV genotype 2 was detected in 38.2% of the patients, and 9.7% of the patients were coinfected with both EBV genotypes. Looking specifically at Hodgkin's lymphoma (HL) patients, 52.8% showed the presence of genotype 1, while 42.8% had genotype 2 and 2.3% had both genotypes. Among non-Hodgkin's lymphoma (NHL) patients, 51.8% had genotype 1, 35.8% had genotype 2, and 12.4% had both. The average methylation status of the lymphoma samples ranged from 45% to 50%. Most samples exhibited an intermediate/polyclonal-like methylation pattern, indicating a non-clonal origin of EBV (i.e., derived from non-tumor cells). The correlation between EBV load and EBV methylation status suggests that the release of EBV is predominantly non-lytic in nature. The analysis of EBV capture sequencing data revealed contrasting degrees of genetic variation between EBV type 1 and type 2. Specifically, EBV type 1 exhibited variations in the amino acid sequences of coding regions like LMP1, whereas EBV type 2 showed no variation in coding regions such as EBNA LP, EBNA 2, and LMP2B and variations on other latent genes. Moreover, phylogenetic tree analysis demonstrated the presence of sequence similarities among multiple samples within each genotype.

Conclusion: This study presents the first evidence-based data in Ethiopia, providing valuable insights into the epidemiology and virological characteristics of lymphoma patients with EBV infection. These findings significantly contribute to our understanding of the relationship between EBV and lymphoma, emphasizing the importance of viral load, genotyping, methylation status, and genetic variations in different EBV genotypes. However, further research across diverse populations is necessary to broaden our knowledge of this complex virus and its association with lymphoma development. Such knowledge holds great potential for the development of region-specific diagnostic and therapeutic strategies targeting EBV-associated lymphoma.

Acknowledgements

I would like to express my heartfelt appreciation to my esteemed supervisors, Dr. Tamrat Abebe, Prof. Robert Baiocchi, Dr. Abdulaziz Abubeker, and Dr. Fisihatsion Tadesse, for their invaluable intellectual input and guidance throughout my candidacy. Their expertise and support have been instrumental in shaping the direction and quality of this thesis. I am sincerely grateful for their time and effort in editing and proofreading this work, as well as their assistance in preparing manuscripts and providing constructive feedback whenever needed.

I would also like to extend my gratitude to the members of Prof. Baiocchi's lab groups. Their collaboration, insights, and camaraderie have played a significant role in the development and execution of this research. I am particularly thankful to Dr. Elshafa Ahmed and Dr. Christoph Weigl for their invaluable supervision and guidance in the laboratory.

My deepest appreciation goes to the study participants whose cooperation and willingness made this research possible. Without their involvement, this work would not have come to fruition. I am also grateful to the research nurses who diligently collected samples and managed the associated data, contributing to the smooth progress of the study.

I am indebted to the funding institutions, Addis Ababa University and The Ohio State University, USA, for their financial support throughout this research. Their investment has been crucial in enabling the completion of this study.

I would like to express my heartfelt thanks to my families and colleagues for their unwavering support and encouragement throughout this journey. Their belief in me and their constant encouragement have been a source of strength and motivation.

Lastly, I am profoundly grateful to the almighty God for guiding me every step of the way and providing me with the strength and resilience to overcome challenges and achieve my goals.

Once again, I extend my sincere gratitude to all those who have contributed to this work, directly or indirectly. Your support and guidance have been invaluable, and I am truly grateful for the opportunity to have worked with each one of you.

Table of contents

Contents

List of figures

List of tables

List of Appendices

Nomenclature

The viruses in this thesis are named in accordance with the International Committee on Taxonomy of Viruses (ICTV) guidelines. Nucleic acids, organic and biochemical nomenclature, symbols and terminology have been written in accordance with the recommendations set by the International Union of Pure and Applied Chemistry (IUPAC).

Glossary of terms/words

AIDS	Acquired Immunodeficiency Syndrome
BARTs	BamHI A rightward transcripts
BL	Burkitt's lymphoma
CD	Cluster of differentiation
CTLs	Cytotoxic T lymphocytes
DLBCL	Diffuse large B cell lymphoma
DNA	Deoxyribonuclease acid
DNMT	DNA methyltransferases
EBERs	EBV-encoded RNAs
EBNA	EBV nuclear antigens
EBV	Epstein-Barr virus
FFPE	Formalin-Fixed Paraffin-Embedded
GM-CSF	Granulocyte Macrophage Colony-Stimulating Factor
GSV	Ganciclovir
H3K27me3	H3 lysine 27 trimethylation
HAART	Highly Antiretroviral Therapy
HD	Hodgkin's disease
HIV	Human Immunodeficiency virus
HL	Hodgkin's Lymphoma
HLA	Human leukocyte antigen
IE	Immediate early
IFN-a	Interferon-a
ILs	Interleukins
IM	Infectious Mononucleosis
IL	2- Interleukin 2
IRFs	Interferon-regulatory factors
LMPs	Latent member proteins

LPD	Lymphoproliferative Disease
MHC	Major Histocompatibility Complex
MIP	Macrophage inflammatory protein
miRNAs	EBV-encoded microRNAs
ncRNAs	noncoding RNAs
NK	Natural Killer
NK	Natural killer
NPC	Nasopharyngeal carcinoma
OHL	Oral hairy leukoplakia
PCR	Polymerase Chain Reaction
PTCL	Peripheral T-cell lymphoma
PTLD	post-transplant lymphoproliferative disorders
RNA	Ribonuclease acid
TFs	Transcription factors
TGF-b	Transforming growth factor b
TLRs	Toll-like receptors
TNFR	Tumor necrosis factor receptor
TNFs	Tumor necrosis factors
TPA	12-O-tetradecanoyl-phorbol- 13-acetate
TRADD	TNFR-associated death domain
VCA	Viral capsid antigen
ZEBRA	BamHIZ Epstein-Barr virus replication activator

Publications from this work

Manuscripts published in peer-reviewed journals in sequential order of appearance:

Publication 1: Seifegebriel Teshome, Kidist Zealiyas, Abdulaziz Abubeker, Fisihatsion Tadesse Jayalakshmi Balakrishna, Christoph Weigel, Elshafa Hassan Ahmed, Tamrat Abebe and Robert A. Baiocchi. **Detection and Quantification of the Epstein–Barr Virus in Lymphoma Patients from Ethiopia: Molecular and Serological Approaches.** ***Microorganisms*** **2023*, 11*(10), 2606; https://doi.org/10.3390/microorganisms11102606**

Publication 2: Seifegebriel Teshome, Kidist Zealiyas, Abdulaziz Abubeker, Fisihatsion Tadesse, Christoph Weigel, Elshafa Hassan Ahmed, Robert A. Baiocchi and Tamrat Abebe. **Genotypes distribution of Epstein - Barr virus among lymphoma patients in Ethiopia.** ***Int. J. Mol. Sci. 2023, 24(18), 13891; https://doi.org/10.3390/ijms241813891***

Manuscript in preparation:

Manuscript 3: Seifegebriel Teshome, Kidist Zealiyas, Abdulaziz Abubeker, Fisihatsion Tadesse Jayalakshmi Balakrishna, Christoph Weigel, Elshafa Hassan Ahmed, Tamrat Abebe and Robert A. Baiocchi. **Profiling the CpG methylation patterns of Epstein-Barr Virus DNA in lymphoma patients from Ethiopia.**

Co-authored publications:

Publication 1: Kidist Zealiyas, **Seifegebriel Teshome,** Nega Berhe, Wondwossen Amogne, Aklilu Feleke Haile, Ebba Abate, Getnet Yimer, Christoph Weigel, Elshafa Hassan Ahmed, Tamrat Abebe, Robert Baiocchi. The Burden of Epstein–Barr Virus (EBV) and Its Determinants among Adult HIV-Positive Individuals in Ethiopia. *Viruses* 2023, *15*, 1743. https:// doi.org/10.3390/v15081743

Publication 2: Zealiyas K, **Teshome S**, Haile AF, Weigel C, Alemu A, Amogne W, Yimer G, Abebe T, Berhe N, Ahmed EH and Baiocchi RA (2023) Genotype characterization of Epstein–Barr virus among adults living with human immunodeficiency virus in Ethiopia. Front. Microbiol. 14:1270824. doi: 10.3389/fmicb.2023.1270824

Scientific presentations

- Profiling the CpG methylation patterns of Epstein-Barr Virus DNA in lymphoma patients from Ethiopia. American Association for cancer Research (AACR) annual meeting from April 5-10, San Diego, USA. (Poster presentation)
- Molecular Epidemiology and Genetic Diversity of Epstein-Barr Virus Among Lymphoma Patients in Ethiopia: American Society for Microbiology (ASM) annual meeting, June 13-17 2024, Atlanta, USA. (Poster presentation)
- Epigenetic Insights into Epstein-Barr Virus-Associated Lymphomas among Ethiopian Patients. International Federation of Clinical Chemistry and Laboratory Medicine (IFCC) annual meeting. Dubai, UAE, May 26-30, 2024. (Poster presentation)
- Unveiling the Molecular Detection and Quantification of Epstein-Barr Virus in Lymphoma Patients from Ethiopia. Ethiopian Medical Association (EMA) annual meeting at Addis Ababa Ethiopia from March 8-10,2024. (Oral presentation)
- Detection of Epstein Barr Virus using Epstein-Barr encoding region (EBER) in situ hybridization in Lymphoma Patients in Ethiopia. Ethiopian Public Health Association (EPHA) annual meeting at Addis Ababa Ethiopia from March 17-19,2024. (Poster presentation)

CHAPTER 1. INTRODUCTION

1.1. Background

Epstein-Barr virus (EBV), a human tumor virus, was first identified in 1964 by Epstein and colleagues in a cell line derived from Burkitt lymphoma (Epstein, 1964). EBV is known as human herpesvirus 4, is a member of the Herpesviridae family and the Gamma Herpesviridae subfamily. It is a double-stranded DNA virus. As a γ-1 herpesvirus, EBV has the capability to infect B-lymphocytes and establish a latent infection within the host. This latent infection has the potential to transform B-lymphocytes (C. A. Alford et al., 1993). The prevalence of EBV is exceptionally high, with over 90% of the global population being carriers of the virus (Tzellos & Farrell, 2012).

EBV is primarily transmitted through oral routes, such as saliva. The virus can be detected in the oropharyngeal secretions of individuals with Infectious Mononucleosis (IM), as well as immunosuppressed individuals. Additionally, EBV can be found to a lesser extent in the oropharyngeal secretions of healthy individuals who have previously been infected with the virus (Tao et al., 2006). Despite its high prevalence and transmission through saliva, EBV typically causes minimal adverse effects in most infected individuals. The infection persists throughout life, with primary infection often occurring without symptoms during childhood. The pathogenicity of EBV is generally limited unless there is an underlying compromise in the immune system (Williams & Crawford, 2006).

EBV-associated cancers are characterized by the presence of viral DNA or RNA and the expression of viral genes within tumor tissues. These cancers encompass a range of lymphocytic diseases, including Hodgkin's disease (HD), Burkitt's lymphoma (BL), and post-transplant lymphoproliferative disorders (PTLD). Although EBV primarily infects B-lymphocytes, it can also infect T lymphocytes or epithelial cells. As a result, EBV is associated with T-cell lymphomas, oral hairy leukoplakia (OHL), nasopharyngeal carcinoma (NPC), and undifferentiated gastric carcinoma (Maeda et al., 2009; Santpere et al., 2014).

Following primary infection, the immune response, including antigen-specific adaptive cellular and humoral components, usually effectively controls EBV. However, the virus eventually enters a latent state within the memory B cell compartment, resulting in lifelong persistence. The mechanisms of oncogenesis by EBV vary depending on the specific type of tumor. Nevertheless, different EBV-associated cancers share fundamental characteristics of EBV biology. The initial stage of EBV-associated tumorigenesis involves the establishment of a persistent (latent) infection (Kutok & Wang, 2006).

Approximately 20% of human tumors are associated with latent viral infections, including EBV. The viral latency observed in these tumors is controlled through a combination of viral genetic and epigenetic regulatory mechanisms (Speck & Ganem, 2010). During latent infection of EBV, a highly active and dynamic process occurs, involving the orchestrated interaction between viral DNA and host-cell fate decisions such as proliferation, differentiation, and survival. Within this state, the viral genome selectively expresses a restricted set of latency-associated genes, undergoes controlled replication, and faithfully distributes to newly divided daughter cells. The expression programs of latency-associated genes are influenced by host-cell-specific factors and environmental signals, contributing to the dynamic nature of EBV latency (Lieberman, 2015).

During latency, EBV expresses multiple genes. Epstein-Barr nuclear antigens (EBNAs) are expressed through alternative splicing of primary transcripts, resulting in the production of EBNA1, EBNA2, EBNA3A, EBNA3B, EBNA3C, and EBNA leader protein (EBNA-LP). Latent membrane proteins (LMPs), specifically LMP1, LMP2A, and LMP2B, are expressed individually from distinct promoters. Additionally, EBV expresses non-coding RNAs, including EBV-encoded small RNAs (EBER)-1 and EBER-2, as well as several viral microRNAs (miRNAs), during latency (Ansari et al., 2013).

EBV demonstrates three distinct types of latencies known as latency I, II, and III. Each latency program is distinguished by the specific expression of a unique set of latency-associated genes (Kempkes & Robertson, 2015). During latency I, EBV expresses a limited number of genes, primarily EBNA1 and non-coding RNAs (EBERs and viral miRNAs). Latency II is characterized by the expression of EBNA1, LMP1, LMP2A, and LMP2B,

along with non-coding RNAs. Latency III involves the expression of all latency-associated genes, including EBNA1, EBNAs 2, 3A, 3B, 3C, LMP1, LMP2A, LMP2B, as well as non-coding RNAs. (Yin et al., 2019). The presence of distinct latency programs in EBV enables the virus to adjust its gene expression profile according to different cellular environments. This adaptability allows EBV to establish persistent infection and potentially contribute to tumorigenesis under specific circumstances (Tempera & Lieberman, 2014). EBV demonstrates variations in latency programs between B cells and epithelial cells, resulting in the production of distinct viral proteins and transcripts specific to each type of latency (Calderwood et al., 2007).

EBV can be categorized into two primary genotypes: type 1 (A) and type 2 (B). These genotypes are differentiated by variations in the EBV nuclear antigen 2 (EBNA-2) gene, which displays only 54% similarity between the two types. The divergence in the sequence of the EBNA-2 gene serves as a crucial marker for distinguishing between type 1 and type 2 EBV genotypes (Münz, 2015). Among the genes of EBV, the latent membrane protein 1 (LMP-1) oncogene displays higher degree of polymorphism. Different variants of LMP-1 have been categorized into seven main groups: B95-8, Alaskan, China 1, China 2, Med+, Med−, and NC. These various LMP-1 variants represent distinct genetic variations within the population of EBV (H.-P. Li & Chang, 2003; Yakovleva et al., 2015).

The detection of EBV in various EBV-associated malignancies varies depending on geographic regions, as well as factors such as age and sex (Khan et al., 2020). Geographically, populations from Europe, America, China, and South Asia have a higher prevalence of EBV type 1, while EBV type 2 is less commonly found in these regions. Conversely, EBV type 2 is more frequently observed in African and Papua New Guinean populations. These regions with a higher prevalence of EBV type 2 also tend to have a higher incidence of immunocompromised individuals who are more susceptible to acquiring both types of EBV infections (Bouvard et al., 2009). EBV type 1 and type 2 differ in their abilities to transform B cells, with EBV type 1 demonstrating a higher efficiency in B cell transformation compared to type 2. This discrepancy in transforming abilities between the two types contributes to the distinct characteristics and outcomes observed in EBV-associated diseases (Brooks et al., 2000; Dolan et al., 2006; Tiwawech et al., 2008).

Research conducted in various geographic regions has indicated a correlation between specific socioeconomic factors, such as overcrowded living conditions with poor sanitation and low income, and a higher likelihood of younger-aged children testing positive for EBV antibodies (seropositivity) (Bouvard et al., 2009; Chabay & Preciado, 2013). In countries where poor hygiene standards are prevalent, EBV infection is typically acquired during early childhood. Consequently, the majority of children in these developing countries are seropositive for EBV by the age of 6 years (Papesch & Watkins, 2001).

Various techniques can be employed to detect EBV, including Epstein-Barr virus-encoded RNA-in situ hybridization (EBER-ISH), serology using viral IgG and IgM ELISA, and molecular techniques such as polymerase chain reaction (PCR). Each method has its own strengths and limitations. EBER-ISH is commonly regarded as the gold standard for detecting EBV in pathology specimens. It enables the visualization of viral RNA within cells, offering direct evidence of viral presence. PCR is valuable for assessing viral load and is extensively utilized in both research and clinical settings (Dowd et al., 2013; Gulley, 2001; Hassan et al., 2006; Nystad & Myrmel, 2007).

1.2. Statement of the problem

Between 1990 and 2010, there has been a significant increase in the overall global burden of malignancies attributed to EBV. Over this 20-year period, the burden has risen from 5.779 million cases in 1990 to 7.978 million cases in 2010. Analysis of EBV-attributed malignancies in 21 world regions revealed that the highest mortality was in East Asia. 47% of all EBV-attributed malignancies occurred in East Asia (Khan & Hashim, 2014) (**Figure 1.1**).

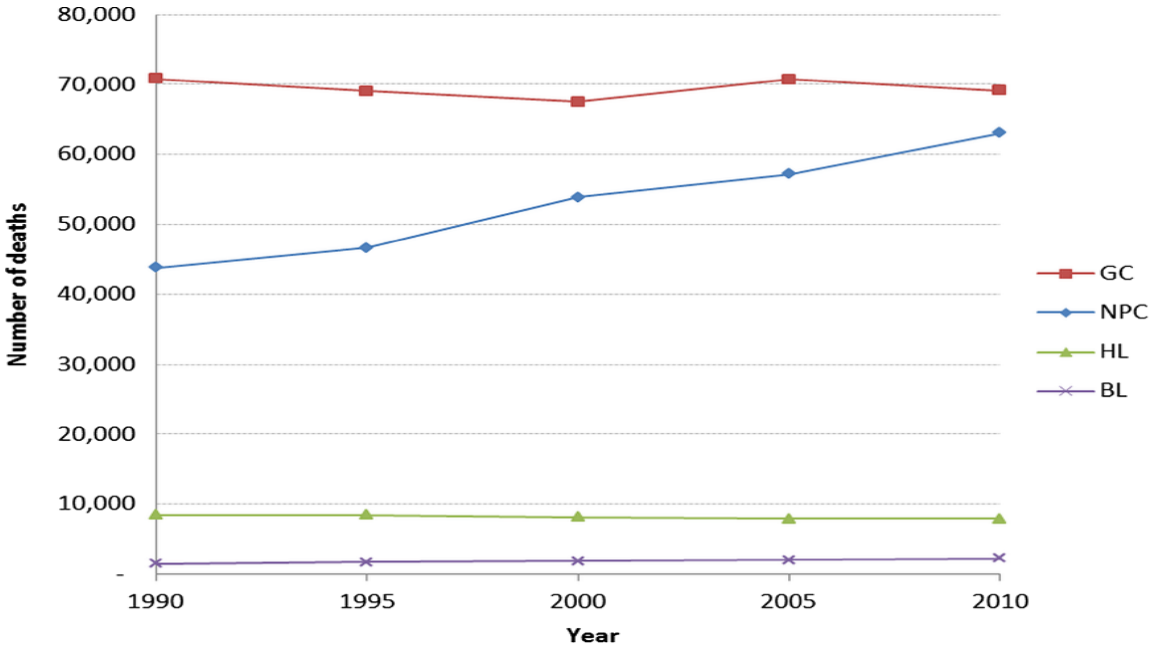


Figure 1.1: Global burden of death from EBV attributed malignancies, 1990-2010

EBV is associated with the development of several malignancies, including the endemic form of BL, NPC, 30 to 50% of Hodgkin's disease, and approximately half of AIDS-associated lymphomas. EBV is present in over 95 %, 5–10 %, and 3–40 % of endemic, sporadic, and HIV-associated Burkitt's lymphoma cases, respectively. EBV infects approximately 60% to 80% of PTLD patients and is the virus also associated with approximately 40% of cHL cases (Ocheni et al., 2010; Shannon-Lowe & Rickinson, 2019).

In 2017, there were 1.442 million new cases and 973,000 deaths attributed to HL, BL, NPC and GC combined. The global burden of EBV-associated fraction in these four malignancies accounted for over 265,000 (18%) new cases, 164,000 (17%) deaths, and 4.6 million (20%) disability-adjusted life years (DALYs) (Khan et al., 2020). Another estimate in 2020 by Wong *et al.* also showed that EBV was associated with approximately 239,700 to 357,900 new cases of cancer and 137,900 to 208,700 cancer-related deaths worldwide. These figures represent approximately 1.3% to 1.9% of the global cancer burden (Wong et al., 2022).

In 2020, NHL accounts for over 254,400 to 265,200 deaths and 536,000 to 552,800 new cases of lymphoma globally and this comprise more than 75% of all lymphomas, while HL accounts for the remaining 10-25% (Ervik et al., 2021; Sung et al., 2021). According to

global Burden of Disease (GBD) 2017 estimates, there were 1,526 incident cases of non-Hodgkin lymphomas (NHLs) in the age group of 1-4 years in East Sub-Saharan Africa. In this specific age group and region, it has been estimated that 90.5% of NHL cases correspond to BL and 95% of the BL cases in this age group are EBV associated (De Martel et al., 2012; Parkin, 2006).

Ethiopia is one of the countries in SSA that has been most severely affected by HIV infections, with about 720,000 PWH and 27,104 newly diagnosed cases by 2020 (Misganaw et al., 2022) and EBV is a major hurdle in managing HIV/AIDS-infected individuals, since it is one of the most frequent causes of morbidity and mortality in this population (Sachithanandham et al., 2014).

In Ethiopia, it was estimated that in 2019, there were approximately 53,560 new incident cases of cancer, resulting in 39,480 deaths, and causing 1.42 million DALYs. Out of the total new incident cases reported, there were 650 cases of HL and 480 cases of NHL. Between 2010 and 2019, there was an observed increase in cancer incidence by 32%, deaths by 29%, and DALYs by 19% (Awedew et al., 2022). Lymphoma is the leading cancer diagnosed in the Gondar hospital, Ethiopia which accounts for 17.2% of the total cancer cases (Tefera et al., 2016) Additionally, according to Ethiopian cancer registry report, NHL is the 2nd most common cancer in the men population (Timotewos et al., 2018).

International collaboration is crucial to combat B-cell malignancies in Ethiopia and Sub-Saharan Africa. Many countries in this region still face challenges in terms of diagnostic capabilities and lack the necessary resources for expensive chemotherapy treatments. To address these obstacles, a proactive approach to preventing the development of EBV-associated malignancies becomes paramount.

Studies conducted globally have identified variations in the molecular epidemiology and genetic diversity of EBV strains among different populations. However, the knowledge regarding the molecular characteristics of EBV strains and their association with lymphomas in Ethiopia is limited. Additionally, investigating possible risk factors and transmission patterns of EBV in this context is crucial to understanding the underlying mechanisms of lymphomagenesis. The specific issue this research aims to address is the

lack of understanding regarding the molecular epidemiology and genetic diversity of EBV among lymphoma patients in Ethiopia. More specifically, we need to determine the prevalence and distribution of different EBV strains, identify potential genetic variations, and assess their association with lymphoma subtypes in the Ethiopian population.

Understanding the molecular epidemiology and genetic diversity of EBV in lymphoma patients in Ethiopia is significant for several reasons. Firstly, it provides crucial insights into the pathogenesis and etiology of EBV-associated lymphomas, which can contribute to the development of targeted diagnostic and therapeutic strategies. Additionally, identifying specific EBV strains and genetic variations associated with lymphoma subtypes can aid in risk stratification, prognosis assessment, and personalized treatment approaches. Furthermore, knowledge about the prevalence and transmission patterns of EBV strains in Ethiopia can inform public health initiatives and preventive measures to reduce the burden of EBV-associated lymphomas. Therefore, conducting research on the molecular epidemiology and genetic diversity of EBV in Ethiopian lymphoma patients will make a significant contribution. This study will provide a vital baseline of data for future investigations, policymakers, and other stakeholders involved in preventive measures.

1.3. Significance of the study

The International Agency for Research on Cancer (IARC) classifies EBV infection as a Group I carcinogen (Niedobitek et al., 2001). Although EBV is prevalent globally, there is observed geographic variability in the incidence of EBV-related tumors. This means that the occurrence of tumors associated with EBV can vary across different regions of the world (Smatti et al., 2018).

The comprehensive study of molecular epidemiology and genetic diversity of EBV is crucial for gaining insights into the pathogenic mechanisms underlying EBV-associated malignancies. Such research can help identify effective therapeutic targets for the clinical management of these diseases. By extensively characterizing the properties of pathogenic EBV strains, a deeper understanding of the role of EBV in various EBV-associated malignancies can be achieved. Moreover, it has been discovered that the genomic diversity of EBV is more extensive than previously acknowledged. (C. M. Chang et al., 2009). Our

understanding of the specific characteristics and variations of EBV strains in this geographic area is currently limited. This knowledge gap highlights the need for further research and investigation to elucidate the molecular epidemiology and genetic diversity of EBV in sub-Saharan Africa, specifically in East Africa. Therefore, exploring and recognizing this greater genomic diversity can provide novel perspectives and knowledge regarding the behavior and impact of EBV in different EBV-associated malignancies

The understanding of genetic variation in different geographic locations is of paramount importance for disease monitoring and evaluations. However, the lack of studies investigating the distribution of EBV genotypes in sub-Saharan Africa, specifically in Ethiopia, is a notable gap in our understanding about virus-disease interactions. In addition, there is a lack of knowledge regarding the prevalence, distribution, and specific EBV strains associated with different lymphoma subtypes in the Ethiopian population. This study is significant as it aims to fill this knowledge gap and provide valuable insights.

This research represents a significant and valuable contribution to our understanding of the relationship between viral variability and geography-related demographic events, particularly regarding the underrepresentation of circulating EBV genomes in Africa, with a focus on Ethiopian lymphoma populations. Furthermore, this study has the potential to shape future research, inform healthcare policies, and guide public health interventions, ultimately contributing to the overall understanding and management of EBV-associated lymphomas in Ethiopia and potentially in other similar populations. Additionally, the findings have the potential to impact future research, policy, and practice. It can also serve as a foundation for further research on EBV-associated lymphomas in Ethiopia, prompting comprehensive studies on the molecular mechanisms of lymphomagenesis, genetic determinants of disease progression, and potential therapeutic targets. From a policy perspective, the results can inform healthcare policies and guidelines for the diagnosis and management of EBV-associated lymphomas in Ethiopia, while also influencing public health practice by guiding preventive measures such as vaccination campaigns or educational programs to reduce the incidence and impact of EBV-associated lymphomas in the population.

CHAPTER 2. LITERATRE REVIEW

2.1. Structure and genome

The EBV genome is composed of a linear, double-stranded DNA molecule that spans approximately 172 kilobases (kb) in length. It encodes over 85 genes, each serving specific roles in the viral lifecycle (Kieff, 2001). It possesses a protein core with a toroid shape, housing the viral DNA that is tightly wrapped around it. The nucleocapsid of EBV is composed of 162 capsomers, with a protein tegument situated between the nucleocapsid and the outer envelope. The outer envelope of EBV features virus-encoded glycoprotein spikes (C. Alford et al., 1993). The EBV genome contains more than 80 protein-coding open reading frames (ORFs) and approximately 30 distinct non-coding RNAs (ncRNAs) (Skalsky & Cullen, 2015) **(Figure 2.1).**

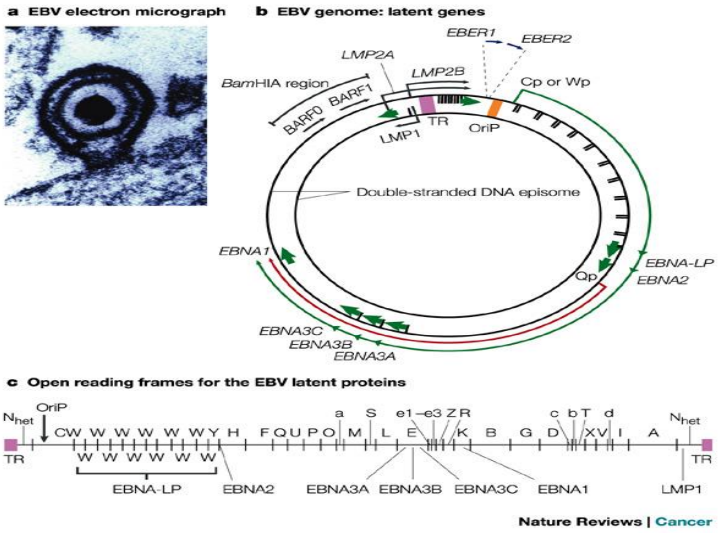


Figure 2.1: Structure of EBV genome

2.2. EBV types and strain

EBV is classified into two types: EBV-1 (B95-8) and EBV-2 (AG876), which are differentiated based on genetic variations in the EBNAs (EBNA-2, -3A, -3B, and -3C) (Sample et al., 1990). Specifically, when comparing EBV type 1 and type 2, the EBNA-2 gene exhibits 70% homology at the gene level and 54% homology at the protein level.

These differences in the EBNA-2 gene sequence contribute to the distinction between EBV type 1 and type 2 (Farrell, 2015). The B95-8 type of EBV was isolated from a patient in North America, while the AG876 type was obtained from patients with infectious mononucleosis and Burkitt lymphoma in central Africa. EBNA-3C, along with EBNA-2 and other EBNA-3 proteins, plays a crucial role in regulating the expression of B lymphocyte genes and is essential for the transformation of B cells. Thes two EBV types may exhibit varying abilities to transform B cells. In particular, EBV type 1 is known to have a higher tendency for B-cell transformation compared to type 2 (Dolan et al., 2006; Tiwawech et al., 2008).

2.3. EBV lifecycle and transmission

EBV can infect both lymphoid and epithelial cells. It undergoes replication in the oropharynx, but establishes a latent infection in long-lived memory B lymphocytes. During primary infection, which is typically transmitted through saliva, EBV often goes unnoticed clinically. However, in some cases, it can lead to IM. IM is a self-limiting lymphoproliferative disease characterized by an overactive but ineffective T-cell response to EBV-infected B cells (Callan et al., 1996). EBV infects both lymphoid and epithelial cells through distinct attachment and entry mechanisms. In the case of B lymphocytes, viral attachment protein (VAP), also known as gp350, binds to the cluster of differentiation (CD)21 receptor on the cell membrane. Additionally, the viral protein gp42 interacts with MHC class II molecules present on the cell surface. These interactions facilitate the fusion of the viral envelope with the cell membrane, enabling EBV to enter B cells (Johannessen et al., 2002; Moore et al., 1991).

Once EBV enters the cell, an uncoating process occurs in the cytoplasm, during which the viral protein coat is shed, and the viral DNA is transported to the cell nucleus. In both lymphoid and epithelial cells, EBV can undergo a lytic replication cycle, leading to the production of infectious virions. In B cells, the lytic replication of EBV is typically observed following the reactivation of latent virus. On the other hand, in epithelial cells, lytic replication has been observed to occur directly during primary infection after the virus enters the cell (Odumade et al., 2011).

During the lytic cycle, EBV reproduces and releases viral particles, which can infect other mucosal cells and lymphocytes. Infected B cells that enter the latent cycle migrate back to lymphoid tissue. There, they undergo proliferation and a portion of them are eliminated by cytotoxic T cells. The remaining infected B cells enter a latency program, ceasing the expression of antigen molecules that trigger cytotoxic responses. These cells become resting memory B cells, allowing EBV to persist in a lifelong infection. Infected memory B cells act as a reservoir for the virus, occasionally releasing infectious viruses into saliva. Some infected memory B cells differentiate into plasma cells, releasing EBV and entering the lytic cycle in peripheral tissues. EBV released from B cells in the lytic cycle is associated with EBV-related epithelial neoplasms. Burkitt lymphoma is linked to activated c-myc oncogenes, while Hodgkin's lymphoma arises from EBV-infected B cells blocked at the germinal center due to cellular mutations. In immunosuppressed individuals, the absence of cytotoxic T cells can lead to lymphoproliferative diseases **(Figure 2.2)** (Maeda et al., 2009).

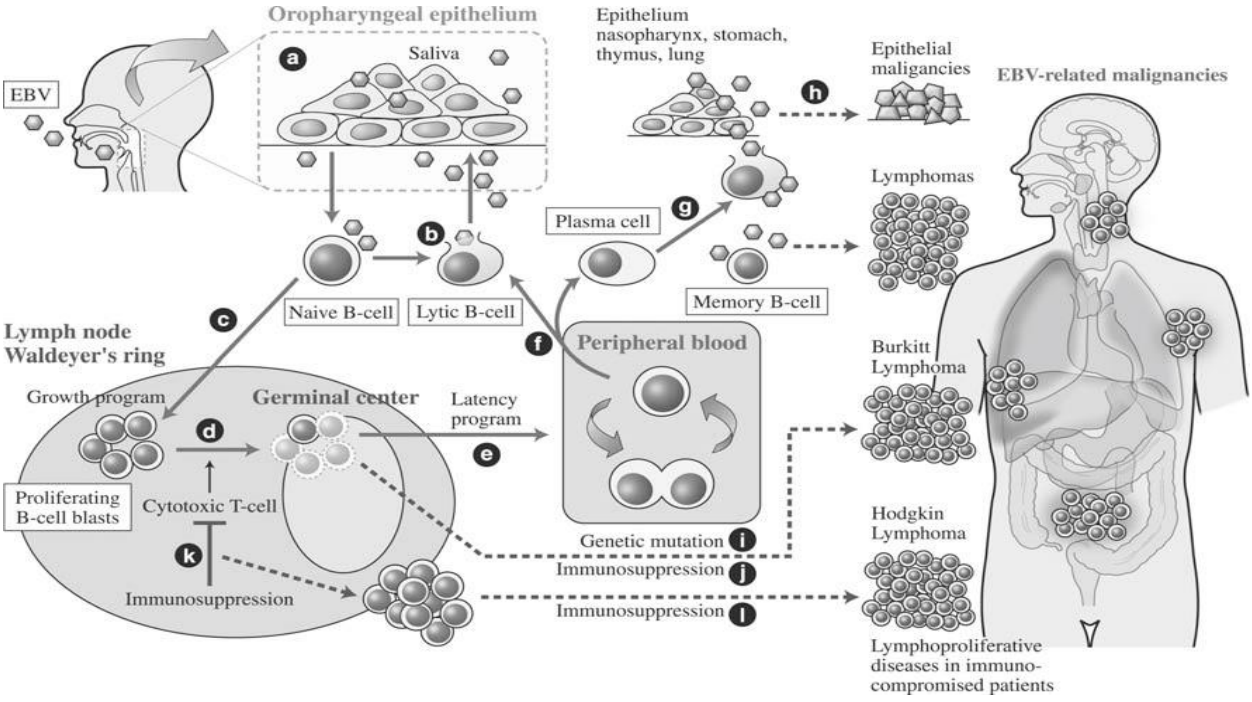


Figure 2.2: Epstein-Barr virus (EBV) infection and pathogenesis of EBV-related diseases. Adopted from (Maeda et al., 2009).

2.4. EBV epidemiology

EBV is carried by more than 90% of individuals worldwide, with the majority being asymptomatic carriers (Young & Rickinson, 2004). EBV has been linked to various lymphoproliferative diseases, particularly in individuals with compromised immune systems. In immunocompromised patients, such as those with impaired cell-mediated immunity, acute EBV infection can lead to the development of lymphoproliferative diseases. These conditions are associated with high mortality rates ranging from 50% to 80% (Paya et al., 1999).

Certain EBV-related tumors exhibit geographic variability in their incidence. NPC and BL are endemic in southern China and equatorial Africa, respectively, but uncommon in other parts of the world. Extra nodal nasal-type natural killer (NK)/T-cell lymphomas also demonstrate a geographic predilection for regions in Asia, South America, and Central America, while being rarely observed in the United States and Europe (Neparidze & Lacy, 2014).

Studies indicate that BL is more common in areas where malaria is endemic. However, when examining incidence data at the country level, it is observed that BL is more prevalent in Eastern Africa compared to other African countries where malaria is also endemic. This indicates that additional factors beyond malaria may contribute to the higher incidence of BL in Eastern Africa (Orem et al., 2007). In equatorial and central Africa, where BL has become endemic, the reported annual incidence rate of BL in children is approximately 5 to 10 cases per 100,000 (Van den Bosch, 2004).

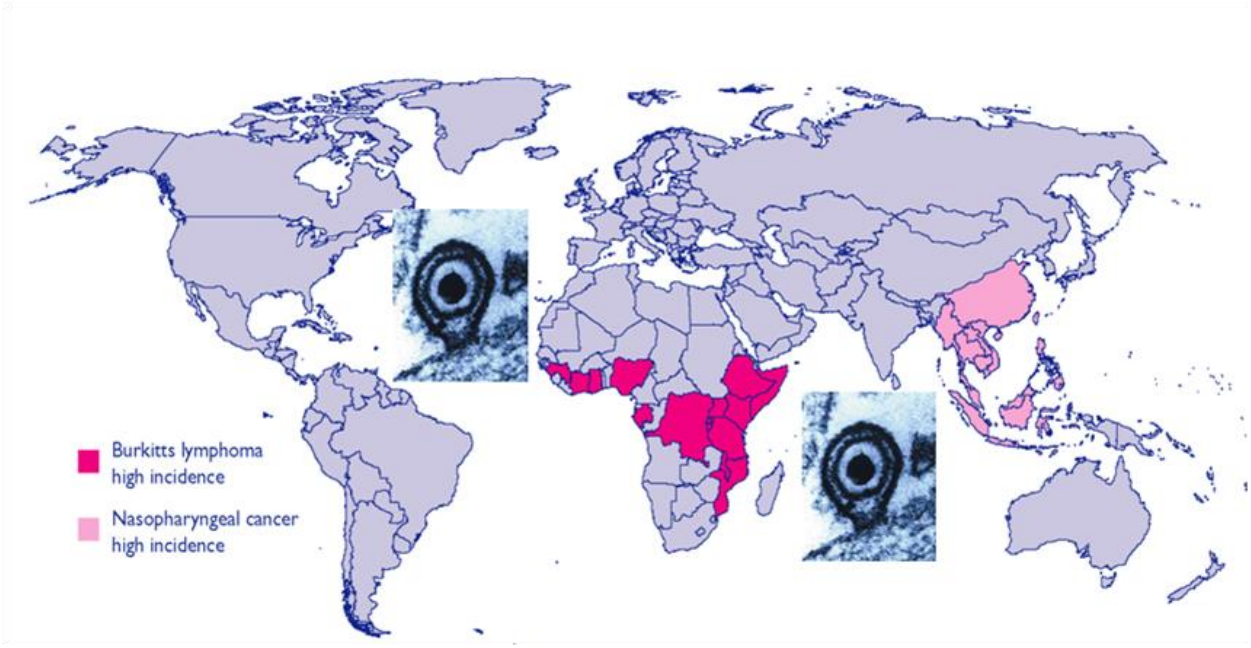


Figure 2.3: Epidemiology of EBV associated carcinoma.

2.5. EBV persistence and latency

Persistent EBV infection is a continuous cycle involving infection, differentiation, persistent infection, reactivation, and reinfection, which takes advantage of various aspects of mature B cell biology. The expansion of the virus is balanced by the immune response. This ongoing cycle, combined with the quiescent infection of peripheral memory B cells, allows the virus to be maintained at very low and stable levels typical of persistent infection. Importantly, EBV does not disrupt the normal progression of latently infected cells into memory cells (D. A. Thorley-Lawson, 2015).

During latency, EBV gene expression is significantly limited compared to the active lytic cycle. The extent of gene expression during latency varies depending on the specific tissue or tumor type from which the EBV-positive cell line originated (Babcock et al., 2000).

The EBV life cycle consists of four latency stages (III, II, I, and 0), during which the virus modulates its gene expression program. In each latency stage, up to 9 viral proteins are expressed, including EBNA1, EBNA2, EBNA3A, EBNA3B, EBNA3C, EBNALP, latent membrane protein 1 (LMP1), LMP2A, and LMP2B. Additionally, two non-coding RNAs

(ncRNAs), EBER1 and EBER2, are expressed consistently throughout all latency stages. While the expression of EBV-encoded proteins is regulated across the four latency stages, EBER1 and EBER2 are continuously present (Speck & Ganem, 2010).

Latency III is characterized by the unrestricted expression of all EBNAs, EBERs, and LMPs. However, the prolonged expression of these proteins is typically observed only when T cell immune surveillance is impaired, as many of these proteins elicit strong immune responses. Latency II is more limited, involving the expression of EBNA-1, EBERs, LMP-1, and LMP-2A and LMP-2B. Latency 0-I represents the most restricted program observed in resting B cells. During this phase, EBV-infected B cells may either exhibit no expression of any EBV gene product (latency 0) or only express EBNA-1 (latency I), along with non-coding RNAs (Tempera & Lieberman, 2014).

EBV undergoes replication during latency I, II, and III by proliferating activated B cells. However, it is primarily during latency 0 and I, following extensive methylation of the viral genome, that efficient induction of lytic replication occurs. In the lytic cycle, over 80 viral genes are expressed. This induction is facilitated by the immediate early transcription factor BZLF1, which works in cooperation with the BRLF1 transcription factor to initiate the production of infectious viral particles. Notably, BZLF1 prefers methylated CpG sequences. Stimulation of the B cell receptor in EBV-infected B cells expressing latency 0 or I triggers lytic reactivation (Woellmer, Arteaga-Salas and Hammerschmidt, 2012).

Distinct EBV transcription programs during specific latency stages have been observed in various human tumors, such as immunoblastic lymphoma in immunosuppressed patients, BL, HD, and NPC. These tumors originate from specific stages within the EBV life cycle and appear to be linked to immune system dysregulation (D. A. Thorley-Lawson & Gross, 2004). Latency I is primarily associated with EBV-associated Burkitt lymphoma. Latency II is associated with Hodgkin disease, T-cell non-Hodgkin lymphoma, as well as nasopharyngeal and gastric carcinoma. Latency III is commonly observed in immunocompromised individuals, including those with posttransplant lymphoproliferative disorders (PTLD) and HIV-associated lymphoproliferative disorders (Klein & Ernberg, 2007).

Table 2.1: EBV latency types and associated malignancies

Latency types	EBV genes expressed	EBV-associated malignancies
Latency 0	EBER-1, EBER-2, EBNA-1, LMP-2A, BART	Dormant B cells in healthy carriers
Latency I	EBER-1, EBER-2, EBNA-1, BART, BARF0	Burkitt's lymphoma, gastric carcinoma
Latency II	EBER-1, EBER-2, EBNA-1, LMP-1, LMP-2A, LMP-2B, BART	Hodgkin lymphoma, nasopharyngeal carcinoma
Latency III	EBER-1, EBER-2, EBNA-1, EBNA-2, EBNA-3 (A, B and C), LMP-1, LMP-2A, LMP-2B, EBNA-LP, BART, BARF0	Post-transplant lymphoproliferative disease, infectious mononucleosis, AIDS-associated lymphomas

2.6. The role of EBV latent genes on the pathogenesis of EBV

Different types of lymphomas are associated with distinct patterns of latency programs. In the first pattern (type I), BL is characterized by the selective expression of EBNA 1. The second pattern (type II), observed in HL and peripheral T-cell lymphoma, involves the expression of EBNA-1, LMP-1, and LMP-2. The third pattern (type III), known as the "growth program," is commonly seen in post-transplantation lymphoproliferative disorders and is characterized by the expression of all nine latent-cycle EBV antigens (Roschewski & Wilson, 2012).

LMP 1

It's a membrane protein which plays a crucial role in signal transduction by engaging in numerous interactions through its C-terminal region. These interactions have a significant impact on the regulation of NF-κB and cell survival through various mechanisms. Therefore, it is conceivable that sequence variations in LMP1 can influence these processes. A significant advancement in comprehending the extensive polymorphism in LMP1 occurred with the categorization of LMP1 variants into seven primary groups (Mainou & Raab-Traub, 2006). LMP1 exhibits sequence variants in comparison to the B95-8 strain, which have been assigned specific names to distinguish them. These variants include the Alaskan variant, China 1 variant, China 2 variant, Med+ variant, Med- variant, and NC variant (W. E. Miller et al., 1994; Walling et al., 2004).

LMP-1 mimics the activity of the CD40 receptor, a cell surface receptor typically found on B cells, dendritic cells, and other immune cells. When LMP-1 is expressed, it acts as a constitutively activated CD40 receptor, meaning it is continuously turned on and transmitting signals within the cell. This constitutive activation leads to the activation of downstream signaling pathways, including the NF-κB pathway, MAPK/ERK pathway, and PI3K/AKT pathway. These signaling pathways are involved in various cellular processes, including the differentiation and survival of B cells. In the context of memory B lymphocytes, LMP-1 signaling promotes their differentiation, leading to the generation of long-lived memory B cells that can provide a rapid and robust immune response upon re-exposure to the same pathogen (Breda et al., 2010; Petrara et al., 2015). In addition to its role in signaling pathways and cellular processes, LMP-1 has been found to have implications for genomic stability. LMP-1 has been shown to induce genomic instability by interfering with DNA repair mechanisms and suppressing DNA damage checkpoints (Q. Cai et al., 2015).

LMP2A

LMP2A has been found to play a role in enhancing cell survival. One of the mechanisms through which LMP-2A exerts this effect is by inhibiting apoptosis induced by TGF-β1. In addition to inhibiting TGF-β1-induced apoptosis, LMP-2A has been implicated in other

mechanisms that enhance cell survival. For instance, LMP-2A can activate signaling pathways, such as the PI3K/AKT pathway, which promotes cell survival and inhibits apoptosis. LMP-2A can also modulate the expression of anti-apoptotic proteins, such as BCL-2, which further contributes to cell survival (Liu et al., 2016; Roughan & Thorley-Lawson, 2009).

Additionally, LMP-2 can activate the Lyn/Syk signaling pathway. This pathway is a tyrosine kinase pathway that plays a critical role in the survival of hematopoietic malignancies, making it an essential component for tumor cell survival (Frappier, 2012). Activation of this pathway can promote the survival of cancer cells by enhancing anti-apoptotic signaling and promoting cell growth and proliferation. By hijacking this pathway, LMP-2 can contribute to the survival and proliferation of cancer cells, facilitating tumor progression and evasion of cell death mechanisms (M. Cohen et al., 2013). Additionally, LMP-2A possesses a unique capability of inducing epigenetic changes. One of the mechanisms through which LMP-2A accomplishes this is by promoting the phosphorylation of STAT3, a transcription factor that regulates gene expression. This phosphorylation event triggered by LMP-2A leads to the activation of DNA methyltransferases (DNMTs)(Grömminger et al., 2012; Shareena & Kumar, 2023).

EBNA-1

EBNA-1 is the only viral protein that is consistently expressed in all latently infected cells. EBNA-1 plays a crucial role in the maintenance of the viral genome and in the regulation of viral gene expression (Shinozaki-Ushiku et al., 2015). It plays a pivotal role in ensuring the persistence of the viral genetic material within infected cells and serves as a critical protein for the replication and maintenance of the EBV. Additionally, emerging evidence suggests that EBNA-1 may act as an oncogene, contributing to the development and progression of EBV-associated malignancies. (Swerdlow et al., 2016). EBNA-1 is involved in protecting infected cells from undergoing apoptosis. It achieves this by suppressing the expression of the myc oncogene while promoting the expression of anti-apoptotic proteins such as Bcl-2 and surviving (Q. Cai et al., 2015). In addition, EBNA-1 has been found to induce the generation of reactive oxygen species (ROS) within infected cells. The exact

mechanisms by which EBNA-1 triggers ROS production are not fully understood, but it is believed to involve the modulation of cellular signaling pathways and alterations in cellular metabolism (Gruhne et al., 2009).

EBNA2

EBNA-2 plays a pivotal role in both promoting the proliferation of transformed B cells and preventing their apoptosis. Working in collaboration with EBNA-LP, EBNA-2 is directly responsible for initiating the transcription of several key viral and cellular proteins (Petrara et al., 2015). EBNA-2 and EBNA-LP form a complex that interacts with specific DNA sequences in the viral genome, known as enhancer elements. This interaction leads to the activation of transcription of viral genes, as well as cellular genes that play critical roles in B cell biology. By initiating the transcription of these genes, EBNA-2 and EBNA-LP contribute to the establishment of a cellular environment favorable for B cell immortalization and transformation (Ojeniyi, 2017; Q. Peng et al., 2020).

EBNA3 Family

EBNA-3 exerts multiple effects that contribute to the transformation and survival of infected cells. Its actions include preventing the accumulation of cyclin-dependent kinase (CDK) inhibitors, degrading the tumor suppressor protein Rb (retinoblastoma protein), stabilizing the c-myc oncogene, and suppressing pro-apoptotic proteins (Fukuda et al., 2001). EBNA-3 interferes with the accumulation of CDK inhibitors, such as p16 and p27, thereby promoting cell cycle progression and facilitating the proliferation of infected cells. By preventing the accumulation of CDK inhibitors, EBNA-3 supports uncontrolled cell growth, a hallmark of transformed cells(Saha & Robertson, 2013).

EBER RNAs

During latent infection, cells infected with EBV express a significant amount of viral RNA transcripts known as EBERs. These EBERs have been found to exert diverse effects on various cellular processes, including cell proliferation, apoptosis, production of growth factors, and cellular signaling (Shinozaki-Ushiku et al., 2015). One prominent effect of EBERs is their influence on cell proliferation. Studies have revealed that EBERs can

enhance the proliferation of infected cells by promoting cell cycle progression and increasing the rate of cell division (Z. Li et al., 2021). EBERs play a crucial role in protecting cells from apoptosis by modulating signaling pathways such as IRF-3 and NF-κB, as well as suppressing IFN-α-mediated apoptosis. In addition, EBERs are capable of inducing the production of growth-promoting cytokines, including IL-6, IL-9, IL-10, and IGF-1(Q. Cai et al., 2015).

BZLF1

BZLF1 is a transcription factor that plays a crucial role in initiating the reactivation of the lytic cycle in B cells infected with EBV. The promoter of BZLF1 is tightly regulated, as it serves as the key mediator for switching from viral latency to the lytic cycle. Understanding the variants of BZLF1 that could potentially influence its function or expression has attracted significant interest in research (Imajoh et al., 2012; Jin et al., 2010; Lorenzetti et al., 2009). The activation of BZLF1 is a critical step in the transition from viral latency, where the virus remains dormant within the host cell, to the lytic cycle, where the virus replicates and produces new viral particles. BZLF1 binds to specific DNA sequences in the viral genome and initiates the transcription of a cascade of lytic genes, leading to viral replication and the release of infectious virions (Murata et al., 2021; Swaminathan & Kenney, 2009).

Table 2.2: EBV latent antigens on B-cell transformation and subsequent lymphoma development

EBV latent protein	Function related to B-cell lymphomagenesis
EBNA1	Regulates viral DNA replication and transcription of numerous viral and cellular genes; facilitates p53 degradation, thereby promoting overall oncogenesis.
EBNA2	Serves as a key viral transcription factor; in conjunction with EBNALP, regulates transcription of multiple viral and cellular genes; essential for B-cell transformation.

EBNALP	Acts as a transcriptional coactivator for EBNA2-mediated transcription of both viral and cellular genes; bypass cell innate immune response; essential for B-cell transformation.
EBNA3A	Together with EBNA3C, represses transcription of BIM, p14, p15, p16, and p18 genes through epigenetic regulation; inhibits B-cell-to-plasma cell differentiation; essential for B-cell transformation.
EBAN3B	Functions as a virus-encoded tumor suppressor protein.
EBNA3C	Alongside EBNA3A, represses transcription of BIM, p14, p15, p16, and p18 genes through epigenetic regulation; facilitates cell cycle transitions from G1 to S and G2 to M; hijacks the ubiquitin-proteasome pathway; inhibits p53-, E3F1-, and Bim-mediated apoptosis; activates autophagy; essential for B-cell transformation.
LMP1	Functionally mimics the CD40 signaling pathway; acts as a major transcriptional regulator; constitutively activates NF-κB, JAK/STAT, ERK MAPK, IRF, and Wnt signaling pathways; stimulates expression of bcl-2 and a20 to block apoptosis; essential for B-cell transformation.
LMP2A	Functionally mimics the B-cell receptor (BCR) signaling pathway; blocks apoptosis; regulates EBV latency.
LMP2B	Regulates the functions of LMP2A.
EBERs	Most abundant noncoding viral RNAs found in all forms of EBV latency programs; influences innate immune response and gene expression; blocks PKR-dependent apoptosis.
miRNAs	Transcribed from BART and BHRF1 loci; maintain latently infected B cells by blocking cellular apoptosis.

2.7. EBV Lytic infection and oncogenesis

Previously, it was believed that EBV-induced cancers solely resulted from latent infection with the virus, and the lytic cycle did not contribute to oncogenesis. Latency was considered a critical state for cancer development, as most EBV-positive cancer cells exhibited latent infection, and initiation of the lytic cycle led to cell cycle arrest. However, recent studies utilizing lytic replication-incompetent EBV have demonstrated that the lytic cycle of EBV enhances the efficiency of B cell transformation (Katsumura, Maruo and Takada, 2012).

Lytic infection is now recognized to contribute to the oncogenicity of EBV. It has been observed that lytic infection precedes the development of EBV-associated Hodgkin's lymphoma, nasopharyngeal carcinoma, Burkitt's lymphoma, and post-transplant lymphoproliferative disorders (C. K. Chan et al., 1991). Therefore, it is highly probable that lytic infection plays a crucial role in tumor development by expanding the population of latently infected cells. This expansion increases the chances of a latently infected cell undergoing neoplastic transformation. Additionally, lytic infection may contribute to the oncogenicity of EBV by stimulating the expression of proinflammatory cytokines, growth factors, and cellular signaling molecules (Arvey et al., 2012). In a humanized mouse model, it has been demonstrated that the induction of the lytic cycle is essential for the development of B cell lymphoma (Ma et al., 2011).

The process of lytic replication leads to the efficient secretion of viral and cellular cytokines and growth factors, with IL-13 playing a significant role. These factors can potentially contribute to oncogenesis. Notably, the virus encodes functional IL-10, and cellular cytokines and growth factors such as IL-10, IL-8, TGF-b, and vascular endothelial growth factor are released from cells undergoing lytic replication. These released factors have the ability to promote cell cycling in neighboring cells that harbor latent EBV infection, further augmenting the potential for oncogenesis (Hong et al., 2005; Rosemarie & Sugden, 2020).

2.8. EBV associated diseases

EBV association refers to cases where the viral genome or gene products are detected within malignant cells (Razzouk et al., 1996). The mechanisms by which EBV causes

tumors differ depending on the specific type of tumor. However, there are certain fundamental characteristics of EBV biology that are present in various EBV-associated cancers. The initial step in the process of EBV tumorigenesis is the establishment of a persistent infection (Kutok & Wang, 2006). Different types of lymphomas are associated with distinct patterns of latency programs. In the first pattern (type I), BL is characterized by the selective expression of EBNA 1. The second pattern (type II), observed in HL and peripheral T-cell lymphoma, involves the expression of EBNA-1, LMP-1, and LMP-2. The third pattern (type III), known as the "growth program," is commonly seen in post-transplantation lymphoproliferative disorders and is characterized by the expression of all nine latent-cycle EBV antigens (Roschewski & Wilson, 2012).

2.7.1. Hematologic malignancies

2.7.1.1. Hodgkin’s Lymphoma

This particular lymphoma is characterized by the presence of Reed-Sternberg cells, which are large mononucleated or multinucleated neoplastic cells (Ayee, Ofori, Wright, et al., 2020).. The presence of EBV can influence the expression of cytokines and chemokines and these cytokines play a critical role in the immunopathogenesis of HL, acting as both autocrine growth factors and initiators/sustainers of the reactive infiltrate (Massini et al., 2009).

There are two variants of Hodgkin's lymphoma (HL) are known as classical Hodgkin's lymphoma (cHL) and nodular lymphocyte-predominant Hodgkin's lymphoma (NLPHL). These variants differ in terms of their clinical presentation, histological characteristics, and prognosis (Swerdlow et al., 2008; Vockerodt et al., 2014). cHL can be categorized into four distinct histological subtypes: mixed cellularity (MC), nodular sclerosis (NS), lymphocyte-rich (LR), and lymphocyte-depleted (LD). These subtypes differ in their histological characteristics and have unique clinical and prognostic implications (Pileri et al., 2002).

In developed countries, the rate of Epstein-Barr virus (EBV) association in classical Hodgkin's lymphoma (cHL) typically ranges between 30% and 50%. However, significantly higher rates of EBV positivity have been reported in developing countries.

Within the context of cHL, EBV-positive cases are more commonly observed in two age groups: children under 10 years of age and adults over 80 years of age. Furthermore, EBV positivity is usually associated with the less common subtypes of mixed cellularity and lymphocyte-depleted cHL (Kutok & Wang, 2006; Shannon-Lowe et al., 2017). In contrast, nodular sclerosis Hodgkin's lymphoma is infrequently associated with EBV, with a prevalence of approximately 15% to 20%. Additionally, EBV is never detected in lymphocyte-rich Hodgkin's lymphoma (Grywalska & Rolinski, 2015; S. Wang et al., 2014).

2.7.1.2. Non-Hodgkin's lymphoma

2.7.1.2.1. Burkit's Lymphoma

BL is an aggressive B-cell lymphoma derived from germinal centers, frequently characterized by MYC translocation. It gained significance as the first malignancy attributed to EBV and was initially identified in central Africa (Epstein & Achong, n.d.). The strong correlation of BL with malaria endemicity suggested the involvement of an infectious agent, which was later confirmed to be EBV (D. Thorley-Lawson et al., 2016). In BL a type 1 pattern of latency is observed, characterized by the expression of EBNA-1 and EBER without the expression of latent membrane proteins(Tsao et al., 2017).

BL can be categorized into three clinical variants: endemic BL, sporadic BL, and immunodeficiency-associated BL, each exhibiting notable differences in epidemiology, clinical presentation, and biology(Gulley et al., 1992). Endemic Burkitt lymphoma (eBL) has an estimated annual incidence ranging from 5 to 10 cases per 100,000 individuals. This form of lymphoma represents a significant portion, accounting for approximately 50% of all pediatric cancers in malaria-endemic regions such as Equatorial Africa and Papua New Guinea. In these areas, the Epstein-Barr virus (EBV) is detected in approximately 95% of all diagnosed eBL cases (Bornkamm, 2009; Lenze et al., 2011). Sporadic Burkitt's lymphoma (sBL) exhibits a broad geographic distribution worldwide but occurs at a lower frequency compared to endemic Burkitt's lymphoma. It primarily affects children and young adults. In the United States and Western Europe, sBL accounts for approximately 30-50% of all pediatric neoplasms, whereas its incidence in adults is less than 1% (Blum et al., 2004). Immunodeficiency-associated Burkitt's lymphoma (iBL) is characterized by

its occurrence in individuals with HIV infection who develop the lymphoma prior to progressing to AIDS. Compared to the sporadic form of the disease, iBL has a significantly higher incidence, ranging from 10 to 100 times greater. Furthermore, approximately 30-40% of iBL cases are found to be positive for EBV(Gloghini et al., 2013).

In endemic BL, the primary role of EBV may be to protect B-cells already harboring a c-myc translocation from apoptosis. The presence of EBV contributes to telomere dysfunction and genomic instability in these cells. Patients with acute malarial infection exhibit an increased EBV viral load, possibly due to viral reactivation triggered by malarial antigens. The immunophenotype of BL resembles that of germinal center cells and persistent malarial infection could promote the hyperactivation of these germinal centers, thereby heightening the risk of somatic hypermutation and c-myc translocations (Chene et al., 2009).

2.7.1.2.2. Diffuse Large B-Cell Lymphoma

DLBCL is a subtype of non-Hodgkin's lymphoma that represents 30% to 40% of all NHL cases. The latency pattern of EBV infection in EBV-positive DLBCL cases varies depending on the immune status, with the virus exhibiting either type II or type III latency (Menon et al., 2012). Among immunocompromised patients, particularly those with HIV, approximately 30-35% of DLBCL cases are associated with EBV. In these cases, there is an increased diversity in the latency programs exhibited by the virus (Cesarman, 2013). However, in immunocompetent hosts, EBV infection is associated with DLBCL in approximately 10% of cases. (Roschewski & Wilson, 2012).

DLBCL can be classified into two subsets based on the cell-of-origin mode: germinal center B cell-like (GCB) and activated B cell-like (ABC). Initially, ABC DLBCLs were identified in older patients and tentatively referred to as EBV-positive DLBCLs of the elderly. However, as more evidence has emerged regarding the occurrence of EBV-positive DLBCL in younger immunocompetent patients, the disease has been reclassified as DLBCL-NOS (‘not otherwise specified’) to encompass this broader patient population (M. Cohen et al., 2014). DLBCL is the most prevalent subtype of NHL among people with HIV (PWH). The introduction of combined antiretroviral therapy (cART) has led to substantial

improvements in immune function and better management of infectious complications during lymphoma treatment in PWH (Huguet et al., 2023). As a result, the composition of HIV-related DLBCLs has undergone changes in the post- cART era. However, the overall incidence and the rate of association with EBV in these tumors still remain significantly higher than those observed in DLBCL cases within the general population (Arvey et al., 2015).

2.7.1.2.3. Post-transplantation lymphoproliferative disease

PTLD encompass a diverse range of diseases observed in immunocompromised patients who have undergone solid organ or hematopoietic stem cell transplantation. The 2016 WHO classification categorizes PTLD into several subtypes, including non-destructive PTLD (such as plasmocytic hyperplasia, florid follicular hyperplasia, and infectious mononucleosis), polymorphic PTLD, monomorphic PTLD (consisting of B-cell or T-cell lymphoma), and lesions resembling classic Hodgkin lymphoma (Swerdlow et al., 2008)

PTLD can be seen in up to 15-20% of lung transplant cases, which involve the transplantation of larger amounts of lymphoid tissue. In PTLD cases arising from lung transplants, the malignant cells typically originate from the host rather than the donor. Another significant risk factor for PTLD development is when the transplant recipient is EBV seronegative and receives an organ from an EBV-positive donor (Paya et al., 1999). EBV infection is detected in a significant proportion of PTLD patients, ranging from 60% to 80% overall. In the case of early-onset PTLD, EBV infection is found in 100% of affected individuals (Thompson & Kurzrock, 2004).

2.7.1.2.4. Extranodal NK/T-cell lymphoma

Extranodal NK/T-cell lymphoma (ENKTL) is classified as a rare subtype of non-Hodgkin lymphoma (NHL). Compared to other types of NHL, ENKTL is less commonly encountered in clinical practice and accounts for a smaller proportion of lymphoma cases(Teras et al., 2016). Epidemiological studies have indicated that ENKTL is more commonly diagnosed in individuals of Asian descent and is particularly prevalent in certain regions of Asia(Fox et al., 2020). ENKTL primarily affects adults, with the median age of onset falling between 40 and 50 years, and it is more commonly observed in males than

females(Shannon-Lowe et al., 2017). ENKTLs display a Latency II gene expression profile characterized by the presence of LMP1 expression, which is a defining feature of these malignancies. Gene expression profiling of ENKTL tumors has revealed the activation of multiple oncogenic pathways, including NF-kB, MAPK, and JAK-STAT (Jiang et al., 2015). The persistent activation of NF-kB and its downstream pathways has been suggested to play a crucial role in the development and advancement of T/NK cell diseases (Chuang et al., 2005).

EBV is consistently associated with ENKTL, including nasal NK/T-cell lymphoma. While the involvement of EBV in the pathogenesis of this specific subtype is evident, the precise mechanisms by which EBV contributes to the development and progression of nasal NK/T-cell lymphoma are still under investigation (Harabuchi et al., 1990). The level of circulating EBV DNA serves as a measure of tumor burden and activity in NK/T-cell lymphoma, as it is released into the bloodstream from apoptotic tumor cells. Increased quantities of EBV DNA copies are associated with adverse clinical characteristics, diminished response to treatment, and inferior overall clinical outcomes. Regular monitoring of EBV DNA levels can assist in prognostication, treatment planning, and assessing disease progression in patients with this aggressive lymphoma (Au et al., 2004; Kwong et al., 2014).

2.7.2. Epithelial malignancies

The precise molecular mechanism by which EBV infects epithelial cells is still not fully understood. However, several key factors involved in viral attachment, membrane fusion, and cell-to-cell contact-mediated viral transmission have been identified (Wu et al., 2005; Wu & Hutt-Fletcher, 2007). Gastric cancer and nasopharyngeal carcinoma are the exclusive epithelial tumors currently known to be associated with EBV infection.

2.7.2.1. Nasopharyngeal Carcinoma

NPC is the first known epithelial carcinoma to be associated with EBV infection (Klein G et al., 1974). MP1 and LMP2 play a crucial role in influencing the genetic expression and promoting the proliferation of cells, leading to the aggressive and malignant growth seen in NPC. The establishment of EBV latency and transformation in epithelial cells is considered a major contributing factor in the development of NPC (Raab-Traub, 2002).

Nearly all cells in NPC have been found to be infected with EBV. In NPC, EBV establishes type II latency, characterized by the expression of multiple latent genes. The presence of EBV infection and the expression of latent genes, particularly LMP1, can contribute to global hypermethylation, leading to the dysregulation of various cancer-related genes and increased tumor heterogeneity in NPC. As the tumor progresses, additional genetic alterations are acquired (Tsao et al., 2017).

BV infection plays a significant role in driving the pathogenesis of NPC through multiple pathways. One key mechanism is the promotion of a hypermethylation phenotype in the host, leading to the inactivation of tumor suppressor genes. This hypermethylation can be a result of the host's defense response or the direct action of viral genes, such as LMP1, which activates DNMTs. Both lytic and latent EBV genes contribute to NPC development. The lytic genes induce genomic instability and stimulate the release of immune-suppressive cytokines to evade immune surveillance. The expression of latent genes drives tumorigenesis and the acquisition of stemness characteristics in NPC cells. LMP1 is well-documented for its transformation properties, many of which are mediated through NF-kB activation. Additionally, emerging evidence highlights the important role of BART-miRNAs in NPC pathogenesis, particularly in anti-apoptosis and immune evasion (Su et al., 2023; Tsao et al., 2017).

2.7.2.2. Gastric carcinoma

Gastric cancer is among the top five most commonly diagnosed cancers worldwide and ranks as the third leading cause of cancer-related deaths (Jemal et al., 2011). Approximately 10% of gastric carcinomas worldwide are characterized by monoclonal proliferations of tumor cells carrying EBV (Shibata & Weiss, 1992). EBV-positive gastric cancer predominantly manifests in the proximal stomach, affecting the cardia and gastric body. It typically presents as lumps or ulcers accompanied by lymphocyte infiltration. An important characteristic of EBV-positive gastric cancer is its tendency to invade the submucosa easily, while exhibiting a lower rate of lymph node metastasis (Sun et al., 2020).

EBV-associated gastric cancer (EBVaGC) is classified as latency type I or II, characterized by the expression of EBERs, EBNA-1, BARTs, and BART miRNAs. Additionally, approximately half of EBVaGC cases exhibit the expression of LMP-2A (Luo et al., 2005; Sugiura et al., 1996). Global CpG island hypermethylation, which leads to the epigenetic silencing of tumor suppressor genes, is a distinctive characteristic of EBVaGC. This hypermethylation is considered to play a crucial role in the carcinogenesis of EBVaGC (Shinozaki-Ushiku et al., 2015).

2.7.3. Infectious mononucleosis

Infectious mononucleosis (IM), also known as glandular fever or kissing disease, is a manifestation of EBV infection that typically occurs in individuals between the ages of 15 and 25. However, the EBV infection is frequently asymptomatic in healthy individuals and, in young children, often leads to acute but self-limiting disease (Hess, 2004). When primary EBV infection is delayed until adolescence, it can result in infectious mononucleosis. This condition is characterized by a range of symptoms, including severe fatigue, sore throat, swollen lymph nodes, fever, and enlarged spleen. Additionally, individuals with infectious mononucleosis may experience headaches, muscle aches, loss of appetite, and a general feeling of malaise (Ascherio & Munger, 2010; J. I. Cohen, 2000). It is estimated that approximately 30% to 50% of individuals who experience late primary EBV infection will develop IM. In these cases, the delayed onset of EBV infection, typically occurring in adolescence or early adulthood, increases the likelihood of developing IM compared to primary infection in younger individuals (Ascherio & Munger, 2010). In most cases, the infection typically resolves within a period of 3 to 6 weeks without the need for medical intervention. Most individuals who contract the virus during this time frame experience a self-limiting course of the disease (Vetsika & Callan, 2004). While the diagnosis of infectious mononucleosis is primarily based on clinical symptoms and the detection of heterophile antibodies, it is important to note that the heterophile antibody test can be unreliable and prone to false negatives. Healthcare providers consider multiple factors, including clinical presentation and other laboratory tests, to make an accurate diagnosis of IM (Ravanfar et al., 2016).

Table 2.3: Prevalence of EBV among EBV associated malignancies

EBV- associated diseases	EBV association	Infected cell	High risk population
Infectious mononucleosis	>99%	B cells or plasma cells	Children and young individuals
Hodgkin lymphoma	20-80%	B cells	
Burkitt's lymphoma-endemic	100%	B cells	Equatorial Africa, New Guiana
Burkitt's lymphoma-sporadic	30%	B cells	
Burkitt's lymphoma-HIV-related	30-40%	B cells	HIV patients
Post-transplant lymphoproliferative disease	60-80%	B cells	Recipients with heart, lung, or intestine transplantation
Diffuse large B cell lymphoma-HIV related	20-60%	B cells	HIV patients
Extranodal NK/T-cell lymphoma	>90%	NK/T cells	
Nasopharyngeal carcinoma	>99%	Epithelial or squamous cells	East Asia
Gastric carcinoma	5-25%	Epithelial cells	

2.8. Immune response for EBV associated malignancies

Host immune responses are considered crucial in both limiting primary EBV infection and controlling the persistent carrier state throughout a person's life. In healthy individuals, EBV infections are typically well-controlled by the immune system, particularly through the activity of antigen-specific T lymphocytes. When humans are infected with EBV, both humoral (antibody-based) and cellular (T-cell-based) immune responses are generated against the virus. In cases of EBV-related malignancies, nearly all patients test positive for

EBV antibodies. Similar to many viral infections, primary exposure to EBV is characterized by an early IgM antibody response (Münz, 2015).

2.8.1. Innate immunity

The primary target cells of EBV include B cells and oropharyngeal epithelial cells. The oropharyngeal epithelium serves as the initial innate barrier against infection. The interaction between EBV and the innate immune system involves various pathways and steps, including recognition by cell surface, endosomal, and intracellular sensors, as well as production and signaling of interferons (IFNs). The interplay between EBV and innate immunity plays a significant role in determining the outcome of the infection (Pontejo et al., 2018).

The mechanism by which EBV gains access to B cells through the epithelial barrier is not fully understood, and there are ongoing debates regarding this process. According to one current model, EBV released from epithelial cells has an affinity for B cells, while EBV released from B cells shows a preference for epithelial cells (Shannon-Lowe & Rowe, 2014). However, further research is needed to elucidate the precise mechanisms and clarify this aspect of EBV infection.

During primary infection, EBV triggers the innate immune response, primarily by activating Toll-like receptors (TLRs). The infected cells recognize viral components containing pathogen-associated molecular patterns (PAMPs) through pattern recognition receptors (PRRs). This recognition leads to the induction of an innate antiviral immune response. As a result, the infected cells produce and release various cytokines, including interleukins (ILs), tumor necrosis factor (TNF), and interferons (IFNs). The type I IFN response is a crucial defense mechanism employed by host cells to combat viral infections (Stephenson et al., 2016).

Innate immune responses, involving natural killer (NK) cells and natural killer T (NKT) cells, may be involved during the primary acquisition of EBV. NK cells can potentially contribute to either protecting against EBV infection or the pathogenesis of IM (Williams et al., 2005). Despite conflicting findings, a substantial body of research supports the role of NK cells in the immune response against EBV. Specifically, NK cells have been shown

to exert a suppressive effect on the EBV-mediated growth transformation of B cells (Strowig et al., 2008). To evade innate immune sensing and the subsequent activation of antiviral responses, EBV has developed various effective countermeasures. These countermeasures can interfere with different pathways and steps involved in the recognition of EBV by cell surface, endosomal, and intracellular sensors, as well as the production and signaling of IFNs. The interplay between EBV and innate immunity plays a significant role in determining the outcome of the infection (Tabtieng & Gaglia, 2018).

Classical dendritic cells (cDCs) can produce IL-12, which activates CD56bright NK cell subsets, as well as produce IFN-α/β, which can directly inhibit EBV infection and transformation within a short timeframe. In EBV-infected B cells, the production and secretion of Epstein-Barr virus-encoded small RNAs (EBERs) occur. EBERs can be recognized by cDCs through Toll-like receptor 3 (TLR3), leading to the activation of cytokine production. Plasmacytoid dendritic cells (pDCs) also produce IFN-α/β, which can recruit and activate NK cells and neutrophil granulocytes. The specific role of neutrophil granulocytes in the context of EBV infection is still to be determined. Macrophages can sense EBV and EBERs through TLR3, but their precise role remains unclear. Monocytes recognize unmethylated EBV DNA through TLR9, although their significance as non-target cells and the in vivo relevance of this interaction are subjects of ongoing debate. NK cells have demonstrated efficient inhibition of B-cell transformation and targeting of lytic replicating cells, with these effects being dependent on distinct subsets of NK cells (Chijioke et al., 2013; Christian, 2014).

During EBV infection, the expression of various lytic EBV gene proteins can hinder the secretion of antiviral cytokines. Notably, immediate early lytic proteins such as BZLF1 and BRLF1 have been found to inhibit interferon response genes and the production of type I interferons. Moreover, these lytic EBV proteins can impact the host's innate immunity by modulating the expression of Toll-like receptors (TLRs) on the surface of infected cells, thereby interfering with intracellular NF-kB signaling (Albanese et al., 2022; Kenney, 2007).

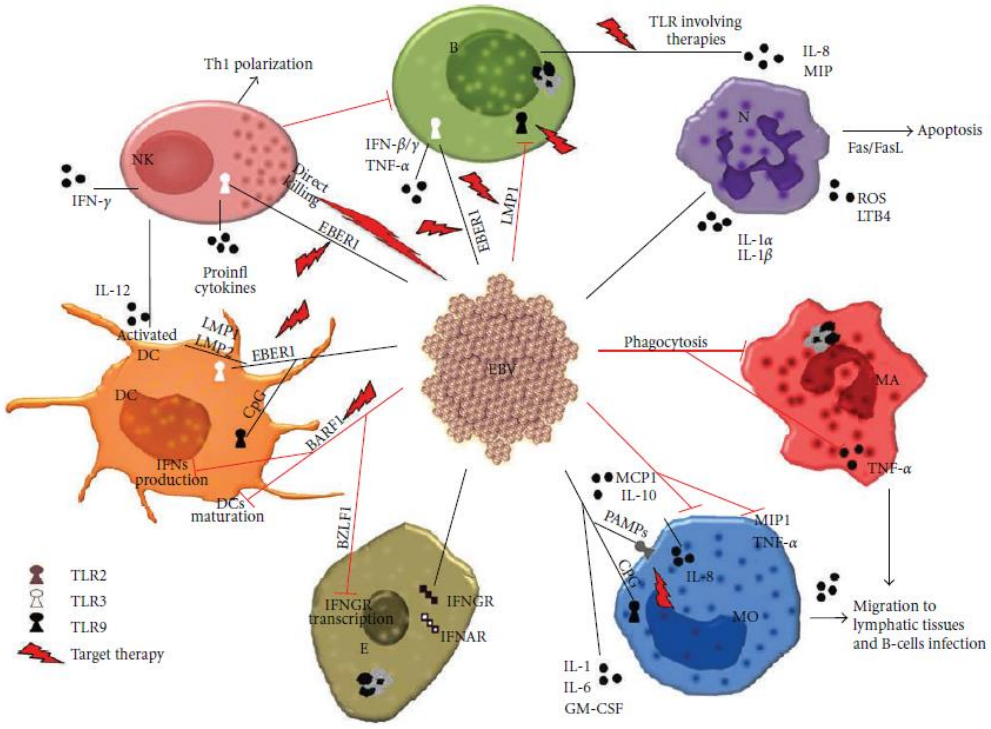


Figure 2.4: Interplay between EBV and innate immunity effectors

2.8.2. Adaptive immunity

2.8.2.1. Cell mediated immunity

Cellular immunity plays a crucial role in controlling EBV infection during both the primary and persistent phases. CD8+ and CD4+ T lymphocytes can distinguish between EBV-infected or EBV-transformed B cells and, as a result, can inhibit their growth. EBV-specific T lymphocytes recognize antigens in the form of molecular complexes consisting of viral peptide epitopes bound to Major Histocompatibility Complex (MHC) molecules, which are expressed on the surface of infected or transformed B lymphocytes (Callan, 2004; Landais et al., 2005).

In the context of cellular immune control of EBV latency I and II malignancies, CD8+ and CD4+ T cells primarily function as cytotoxic T cells capable of lysing target cells that present EBNA1 and/or LMP1 on MHC class II. However, in the case of EBV latency III, it is probable that CD4+ T cells play a crucial role in supporting CD8+ T cell responses

that specifically target the EBNA3 and LMP2 proteins. While CD8+ T cells have limited recognition of LMP1, CD4+ T cells frequently target this protein (Sohn et al., 2015).

CD4+ T cell responses are generated against various epitopes within latent cycle antigens of EBV. Among these responses, a subset of CD4+ T cells is capable of recognizing latently infected target cells that express HLA class II molecules (Hislop et al., 2007). CD4+ T cells specific to the EBNA1 protein can exert direct antiviral effects on EBV-transformed cells through their production of cytokines and their lytic function. Additionally, these CD4+ T cells can support and sustain the CD8+ cytotoxic T lymphocyte (CTL) response against other EBV antigens associated with lymphoma, such as LMP1 and LMP2 (Münz et al., 2000).

2.8.2.2. Humoral immunity

During primary EBV infection, the virus elicits a humoral immune response characterized by robust antibody production against certain lytic cycle proteins. When individuals present with IM, most of them demonstrate strong IgM responses and developing IgG responses to nucleocapsid and envelope proteins, also known as "viral capsid" antigens. Additionally, IgG responses to specific immediate early and early lytic cycle proteins, referred to as "early" antigens, as well as to the latent protein EBNA 2, are typically detectable (Rowe & Zuo, 2010). Subsequently, IgG antibodies targeting the latent protein EBNA 1 begin to develop. However, neutralizing antibodies, which are specifically directed against gp350 (a component of the "membrane" antigen), emerge relatively late during primary infection. As a result, these neutralizing antibodies are unlikely to have a significant role in preventing the spread of the virus during the early phase of infection. (Callan, 2004).

The evaluation of the humoral response to EBV has primarily focused on its diagnostic role and the development of EBV vaccines. During natural infection, the initial appearance of antibodies is directed against the viral capsid antigen (VCA), which likely plays a role in viral neutralization. VCA is expressed late during the lytic phase of infection and typically follows earlier T cell responses targeting immediate early (IE) and some early (E) antigens in terms of the adaptive immune response timeline. Serologic testing for EBV is

considered the gold standard for assessing exposure to the virus. Antibodies against viral structural proteins, specifically IgM antibodies, indicate a primary acute infection, while IgG antibodies against VCA emerge later, persist for a lifetime, and indicate recent or past infection. Antibodies to EBNA, a latent cycle protein, appear later after infection and persist for life, indicating a prior infection. The presence of IgG antibodies against early antigen diffuse (EA-D) can suggest viral reactivation. Therefore, evaluating the combined serostatus of these various antigens can aid in distinguishing whether an individual is experiencing an early primary infection, an ongoing active infection, a prior infection, or even viral reactivation (Martinez & Krams, 2017).

2.9. EBV strategies against immune evasion

Immune evasion is a prominent characteristic of human cancer cells, and the expression of foreign antigens in virally infected cells makes them vulnerable to immune detection. In the context of EBV infection, the virus has developed various strategies to evade host immune responses. The role of EBV lytic proteins in immune evasion during lytic reactivation of the virus in B cells is widely acknowledged (Bentz et al., 2010).

EBV employs immune evasion strategies that involve both lytic and latent cycle gene products, as well as virally encoded microRNAs. These strategies operate at multiple levels to thwart immune responses, including inhibiting the production of antiviral proteins, disrupting antigen presentation by infected cells, interfering with cell death pathways used by virus-specific effector T cells or NK cells, and producing molecules that dampen the function of innate or adaptive immune cells (Ressing et al., 2015).

Host immune responses are modulated by latent EBV gene products. Among them, the EBV-encoded gene product called EBER is expressed at high levels in all types of EBV latency. EBER exerts its immune-modulating effects by binding to PKR, a double-stranded RNA-dependent protein kinase, and inhibiting interferon-stimulated gene activity (Nanbo et al., 2002). EBV also codes for BARF1, a soluble form of the CSF-1 receptor. BARF1 can neutralize the effects of CSF-1, resulting in a decrease in IFN-alpha secretion by mononuclear cells. Additionally, BARF1 is referred to as viral IL-10 due to its capacity to

suppress the production of IFN-gamma, IL-2, and IL-6 by CD4+ T cells that are involved in antiviral responses (Jochum et al., 2012; Strockbine et al., 1998).

Various types of EBV proteins are involved in targeting the processing and presentation of viral antigens, facilitating immune evasion. For instance, EBNA-1 possesses a glycine-alanine repeat that hinders the processing and presentation of antigens by HLA class I molecules. During the early lytic cycle, the protein BNLF2a disrupts peptide loading onto HLA class I molecules by interacting with the transporter associated with antigen processing (TAP). Additionally, BGLF5, a DNAse/exonuclease with host shutoff function, inhibits the synthesis of new HLA class I and class II molecules. BILF1, on the other hand, downregulates cell surface class I molecules by targeting them for lysosomal degradation. In its effort to promote viral persistence, EBV employs various strategies to prevent apoptosis of infected cells. One such tactic involves the functional bcl-2 homolog encoded by BHRF1, which can inhibit apoptosis induced by different stimuli, partly by binding to the pro-apoptotic protein Bim (Hatton et al., 2014).

2.10. EBV epigenetics

EBV-induced epigenetic changes play a critical role in the pathogenesis of EBV-associated malignancies and LPDs (Kaneda et al., 2012; H. H. Niller et al., 2016; Tsao et al., 2012). Epigenetic programming encompasses a range of processes, such as DNA methylation, nucleosome positioning, histone tail modifications, and the formation of higher-ordered structures like promoter-enhancer loop interactions(Tempera & Lieberman, 2014). These epigenetic changes play a crucial role in regulating gene expression and are involved in various biological processes, including development, differentiation, and disease pathogenesis.

During the development of EBV-associated cancers, the infection with EBV can alter the formation of heterochromatin and influence the transcriptional regulation of host genes through DNA methylation. These epigenetic modifications induced by EBV play a significant role in the pathogenesis of EBV-associated malignancies (L. Li et al., 2018). Moreover, the viral products of EBV could directly impact gene expression by mimicking the epigenetic regulators found in the host cells. This allows EBV to manipulate the

epigenetic landscape and exert control over the expression of various genes (Buschle & Hammerschmidt, 2020). The latent transcripts of EBV, including EBNA2, EBNALP, EBNA3A, EBNA3C, and LMP1, engage in interactions with various host epigenetic modifiers. These interactions are critical for the transformation of B cells, a process associated with EBV-induced oncogenesis. The engagement of these viral transcripts with host epigenetic modifiers plays a pivotal role in shaping the epigenetic landscape and gene expression patterns that contribute to B cell transformation (Kang & Kieff, 2015).

The viral genome is typically devoid of histone modifications and CpG methylation during the virion stage. However, upon infection and establishment of latency in host cells, the EBV genome can undergo significant epigenetic modifications, including histone modifications and CpG methylation, which play crucial roles in regulating viral gene expression and the host-virus interaction (Fernandez et al., 2009). Upon initial infection, both latent and lytic genes can be expressed for a brief period of time due to the absence of epigenetic inhibition (Wen et al., 2007). However, as epigenetic modifications occur, the viral genome undergoes changes that ultimately lead to the suppression of immediate early gene transcription when the viral genome remains intact. These epigenetic modifications play a crucial role in regulating the switch from lytic to latent infection and maintaining viral latency (Paulson et al., 2002; Scott, 2017)

2.10.1. CpG methylation of EBV genome

DNA methyltransferases are enzymes responsible for catalyzing the addition of a methyl group to cytosine residues, resulting in the conversion of cytosine to 5-methylcytosine. This process predominantly occurs at CpG-rich regions of the genome, with the exception of CpG islands located near promoters of actively transcribed genes (Smith & Meissner, 2013). CpG methylation plays crucial roles in various aspects of cancer development, including the silencing of tumor suppressor genes, loss of imprinting, impaired expression of DNA repair enzymes, and the loss of expression of viral immunodominant antigens necessary for immune surveillance. These epigenetic modifications contribute significantly to the pathogenesis of cancer (Robertson & Wolffe, 2000; Tao & Robertson, 2003).

Methylation of the EBV genome plays a crucial role in controlling viral gene expression in both healthy and malignant cells. This epigenetic modification leads to the silencing of the majority of latent EBV promoters. By silencing these promoters, infected cells become more resistant to antiviral nucleoside analogues and can evade immune surveillance. Methylation of the EBV genome provides a mechanism for infected cells to regulate viral gene expression and promote their survival and persistence (Ambinder et al., 1999). Frequent inactivation of tumor suppressor genes through focal hypermethylation can contribute to the development of cancer. This process involves the abnormal increase in DNA methylation at specific regions of the genome, leading to the silencing or reduced expression of tumor suppressor genes (Niller et al. 2012).

Upon primary infection in B lymphocytes, the regulation of LMP1/2 promoters and the sites where lytic gene transcription initiates is influenced by the gradual methylation of viral DNA (Paulson et al. 2002; Bergbauer et al. 2010). CpG methylation is more abundant in the C promoter (Cp), which is essential for EBNA transcription, as well as in the LMP1 promoters. This higher level of CpG methylation plays a crucial role in regulating the establishment of different latency types, specifically Type I latency versus Type III latency. The differential CpG methylation patterns in these promoters contribute to the regulation of EBNA expression and the expression of LMP1, ultimately influencing the specific latency program adopted by the virus. (Tempera & Lieberman, 2010) **(Figure 2.5A)**.

Even though methylation of EBV DNA is typically associated with the repression of transcription, methylation of viral promoter regions is also necessary for the reactivation of the lytic cycle. This paradoxical effect is due to the ability of the virally expressed immediate early protein Zta to bind specifically to the region of methylated DNA that contains viral lytic gene promoters. The interaction between Zta and methylated DNA facilitates the initiation of the lytic cycle by promoting the expression of genes required for viral replication and release. (Bergbauer et al., 2010; Sinclair, 2013; Zalani et al., 1996) **(Figure 2.5B)**. Thus, methylation of specific viral promoter regions serves as a regulatory mechanism for the controlled reactivation of the lytic cycle in EBV

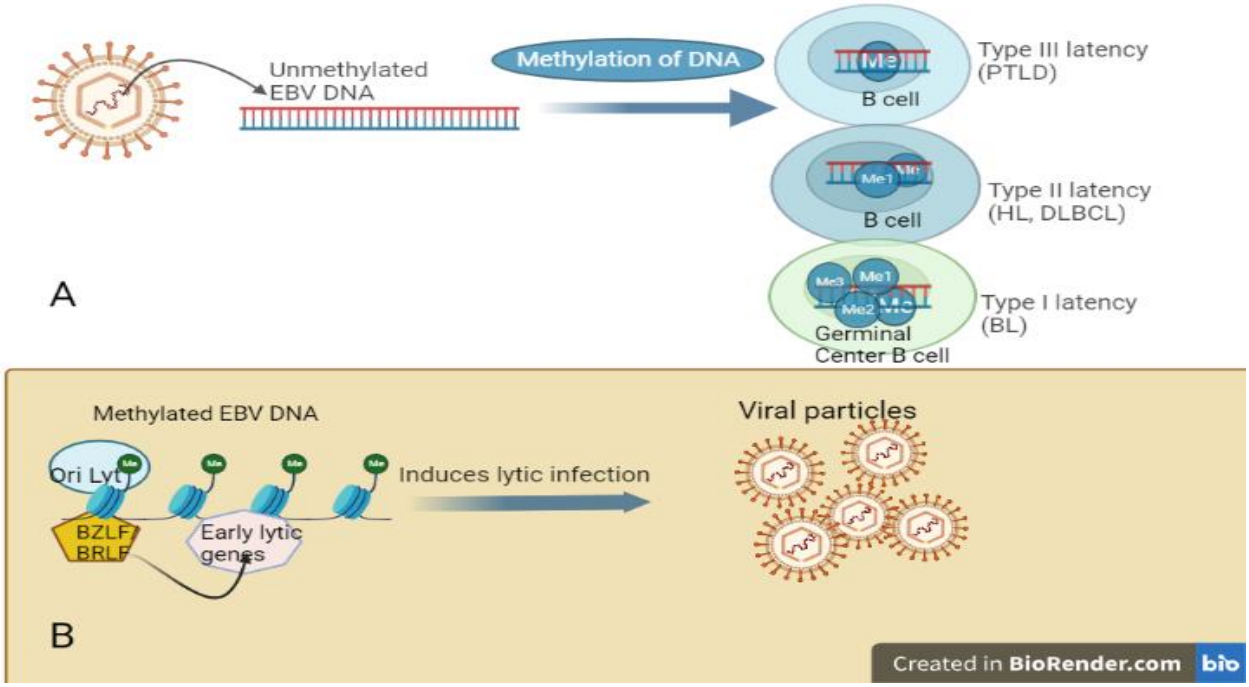


Figure 2.5: Effects of CpG DNA methylation of EBV genome on the viral life cycle and disease progression. **(A)** Methylation of the EBV DNA will silence most of the latent genes and keep in different latent stages. **(B)** Reactivation of the EBV genome from latent stage via the binding of BZLF/BRLF into the methylated region of the viral DNA origin of lytic replication, ori-Lyt and initiates the lytic phase of EBV which then starts to replicate and infects new B cells. **Abbreviations:** DLBCL: Diffuse large B cell lymphoma; BL: Burkitt lymphoma; HL: Hodgkin lymphoma

2.10.2. Histone modification

Histone modifications play a vital role in regulating chromatin structure and gene transcription. The acetylation and deacetylation of histones have a significant impact on the accessibility of DNA to transcription, replication, and repair elements (Fuks et al., 2003). Generally, histone deacetylation is associated with gene inactivation, while histone hyperacetylation promotes transcription. Dysregulation of chromatin structure contributes to EBV-induced cellular transformation and cancer progression (H. H. Niller et al., 2014). he chromatin composition of the EBV genome varies depending on the latency type, and histone modifications frequently coincide with expected marks at regions involved in transcriptional gene activation and repression. Notably, there are substantial epigenetic

differences between Type I and Type III latency at the transcription start sites of viral genes exclusive to Type III infection. Acetylated histones are particularly abundant in the LMP1/LMP2 and Cp promoters (Gerle et al., 2007; J. M. Niller et al., 2007)

Histone modifications also play a critical role in controlling the reactivation of the EBV genome from its latent phase to the lytic phase by inducing a de-repressive epigenetic modification of the Zp promoter, which is essential for immediate early transcription of Zta. In latent cells, the Zp region exhibits suppressive histone markers, including H3K27me3, H3K9me2/3, and H4K20me3. Conversely, high levels of histone acetylation, phosphorylation, and histone H3 lysine 4 trimethylation markers are associated with viral reactivation (Murata et al., 2012). Among these modifications, the repressive histone marker H3K27me3 appears to be particularly crucial for maintaining latency. Interestingly, a point mutation in the MEF2 binding sites within Zp leads to a significant increase in H3K27me3 modifications. This increase in H3K27me3 correlates with a loss of Zta expression, subsequent loss of viral latency, and the induction of the lytic phase (Murata et al., 2013).

2.10.3. microRNAs

EBV was the first virus in which microRNAs (miRNAs) were discovered. These small RNA molecules serve as an additional layer of epigenetic regulation and play a critical role in the development of EBV-associated neoplasms. Two primary groups of miRNAs are implicated: EBV-encoded viral miRNAs and EBV-regulated host-cell miRNAs. The EBV-encoded miRNAs can be categorized into two clusters: the BHRF1 miRNA cluster (miR-BHRF1-3) and the BART miRNA cluster (miR-BART1-22) (Giudice et al., 2016; M. Wang et al., 2018).

Approximately 44 distinct miRNA species have been identified to play a role in both the development of carcinogenesis and the regulation of viral latency (Barth et al., 2008). These miRNAs contribute to the intricate network of molecular interactions within infected cells, influencing key cellular processes involved in cancer development (L. Li et al., 2018). The coordinated regulation of both viral and host-cell miRNA types plays a crucial role in aiding the virus in evading host immune surveillance and contributing to EBV-

associated tumorigenesis. The miRNA families involved in this process have distinct impacts on different types of tumors. The BART miRNA family, for example, influences the development of lymphomas and epithelial tumors, while the BHRF1 miRNA family predominantly contributes to a subset of lymphocyte malignancies (Yang et al., 2013).

Viral-encoded miRNAs play a crucial role in downregulating the expression of viral immediate-early genes, which facilitates the establishment of persistent EBV infection and contributes to host cell carcinogenesis (Barth et al., 2011). These viral miRNAs also assist the virus in evading the host's immune response (Stern-Ginossar et al., 2007). Through their regulatory effects on gene expression, these miRNAs participate in the molecular mechanisms underlying tumor initiation, progression, and immune evasion mediated by EBV (Albanese et al., 2016; L. Cai et al., 2015).

2.11. Detection of EBV

EBV detection is crucial for the diagnosis and monitoring of patients affected by EBV-related diseases, which encompass a broad range of clinical manifestations, from transient benign infections to aggressive malignancies (Gulley, 2001). In the field of diagnostics, a combination of new molecular tests alongside traditional serological or histochemical assays has proven to be valuable. The choice of diagnostic approach depends on the specific clinical context and the types of samples available for testing. For instance, the use of EBER in situ hybridization on biopsy samples has been established as a reliable method. Additionally, more recent advancements include EBV viral load testing of blood samples, which provides an accurate assessment of the patient's clinical status. These diagnostic tools aid clinicians in accurately diagnosing and monitoring EBV-related diseases, enabling appropriate management and treatment strategies (Gulley & Tang, 2008).

2.11.1. Serological methods

Recent advancements have led to the development of several serological tests that aid in diagnosing EBV infection. By utilizing viral capsid antigen (VCA) IgG, VCA IgM, and EBV nuclear antigen (EBNA)-1 IgG, it is now possible to differentiate between acute and past infections. Acute infection is indicated by the presence of VCA IgM and VCA IgG without EBNA-1 IgG, while past infection is characterized by the presence of VCA IgG

and EBNA-1 IgG without VCA IgM. However, interpreting serological findings can sometimes be challenging. VCA IgG may be present without VCA IgM or EBNA-1 IgG in cases of both acute and past infections. Additionally, all three parameters can be detected simultaneously in recent infections or during reactivation (De Paschale & Clerici, 2012).

To accurately interpret these patterns, additional tests are necessary. These include determining IgG avidity, identifying anti-EBV IgG and IgM antibodies through immunoblotting, and searching for heterophile antibodies, anti-EA (D) antibodies, or viral genomes using molecular biology methods (Abusalah et al., 2020).

2.11.2. EBER in Situ Hybridization

EBER in situ hybridization offers a significant advantage in its capability to precisely localize EBV within the context of cytological and histopathological features of the tissue. Typically, EBER is expressed uniformly across all tumor cells in EBV-associated malignancies, indicating a consistent presence of the virus (J. W. Sixbey, 2000). EBER in situ hybridization is widely regarded as the gold standard for the detection and localization of latent EBV in tissue samples. This technique plays a crucial role in various clinical scenarios. One common application is its routine use in confirming a diagnosis of EBV-driven PTLD (Gulley & Tang, 2010).

EBER in situ hybridization can be performed on paraffin sections or cytology preparations, providing flexibility in sample types (Braz-Silva et al., 2008). This technique serves as an excellent marker for EBV infection due to its high sensitivity and specificity. It can be detected in nearly all EBV-infected cells, and its expression levels are typically quite high. In cases where the differential diagnosis involves infectious mononucleosis, Hodgkin's disease, and/or non-Hodgkin's lymphoma, EBER hybridization is often invaluable in facilitating an accurate diagnosis (Ambinder & Mann, 1994; Gulley & Tang, 2008).

2.11.3. Molecular methods

Multiple molecular techniques have been developed and utilized for the detection and quantification of EBV DNA, enabling the diagnosis and monitoring of primary EBV infection, reactivation, and EBV-related diseases. Among these techniques are quantitative

real-time PCR (qRT-PCR), Southern blotting, and Dot blotting (Gulley & Tang, 2008; Kimura et al., 2008). qRT-PCR has proven to be highly sensitive in detecting and quantifying EBV DNA, enabling early and accurate diagnosis of primary EBV infection and the detection of silent reactivation of the virus. This molecular technique provides a quantitative assessment of viral DNA levels, allowing for precise monitoring of the viral load in patient samples over time (Bauer et al., 2005; K. H. Chan et al., 2001; Gartzonika et al., 2012; Maurmann et al., 2003).

qRT-PCR assays can be performed on various types of samples, including serum, whole blood, tissue biopsy, and peripheral blood mononuclear cells (PBMCs). However, determining the optimal specimen to use depends on several factors, such as the patient's condition and the stage of the disease. While there is ongoing debate surrounding this topic, the selection of the best sample type aims to maximize diagnostic accuracy and clinical relevance (Gan et al., 1994; Kimura et al., 2008; Ryan et al., 2004). EBV DNA detected in serum can serve as a valuable indicator for both primary EBV infection and reactivation. Studies have shown that patients with active infection or those with EBV-associated cancers tend to exhibit higher levels of viral load in their cell-free blood (Gulley & Tang, 2008). Moreover, there is evidence to suggest that the viral load in serum correlates with the severity of EBV-associated malignancies and lymphoproliferative diseases. Higher viral loads have been associated with more severe disease outcomes. Consequently, the quantification of EBV DNA in serum can serve as a prognostic marker, providing insights into the progression and prognosis of these conditions (Gärtner & Preiksaitis, 2010; Holman et al., 2012)

Table 2.4: Laboratory test for EBV detection

Laboratory tests	Purpose
***In situ* hybridization**	Identify EBER transcripts or EBV DNA in specific cell types within histologic lesions. It provides information about the presence and localization of the virus in tissues.

EBV clonality assay by Southern blot analysis	This assay assesses the clonality of lesions by analyzing the structure of EBV DNA. It can distinguish between latent and replicative infection based on the episomal (latent) or linear (replicative) structure of the EBV genome
EBV DNA amplification	Detects the presence of viral DNA in patient tissues. However, it lacks disease specificity and may not differentiate between different EBV-associated conditions.
EBV viral load	This test quantitates the amount of EBV DNA in blood or body fluids. It is used to monitor the disease status over time, providing insights into the viral replication and response to treatment.
Immunohistochemistry (LMP1, EBNA1, EBNA2, LMP2A, BZLF1)	Identify EBV protein expression in specific cell types within histologic lesions; distinguish latent from replicative infection based on expression profiles
Culture of EBV or of EBV-infected B lymphocytes	This method detects and semi-quantitatively measures infectious virions or latently-infected B lymphocytes. However, it is not routinely used in clinical practice due to practical limitations.
Electron microscopy	Used to identify whole virions that represent replicative viral infection. However, it is not practical for routine clinical use due to technical challenges and limitations.
Serology (VCA, EBNA, EA, heterophile antibodies)	Measure antibody response to viral proteins in serum sample. They helped to distinguish acute from remote infection and monitor disease status over time

2.12. Therapeutic strategies for EBV-associated malignancies

The expression of viral antigens in tumors that are positive for EBV presents potential targets for immune-based therapies. T cells specific to each of the latency-associated antigens have been identified in patients with EBV-related malignancies. In immunocompromised patients, inhibiting lytic infection has been proposed as a prophylactic strategy. However, currently available drugs do not effectively block lytic infections in patients. The discovery of new anti-EBV drugs depends on a comprehensive understanding of the diverse mechanisms utilized by the virus to promote lytic infection (Gottschalk & Rooney, 2015; Pei et al., 2020).

Clinical trials have extensively investigated immunotherapeutic approaches for the treatment of EBV-associated diseases, with promising results observed in certain cases. During primary infection, the host's CD4+ and CD8+ cytotoxic T cells, along with NK cells, play a crucial role in eliminating EBV-infected cells (W. Li et al., 2023). One effective immunotherapeutic strategy is adoptive immunotherapy, which involves the ex vivo expansion of virus-specific cytotoxic T lymphocytes (CTLs) for subsequent administration. Adoptive immunotherapy has demonstrated safety and efficacy in the treatment of various EBV-associated malignancies, including HL, extranodal NK/T-cell lymphoma, PTLD, and NPC (Heslop et al., 2021).

Another promising strategy for targeting EBV and infected cells involves a combination therapy approach. This approach focuses on inducing the lytic phase gene expression of EBV, followed by exposing the tumor cells to anti-herpesvirus drugs. To achieve this, adenoviral vectors containing the genes BZLF1 or BRLF1 are injected. These genes are responsible for initiating the lytic phase of EBV. Once the lytic phase is induced, anti-herpesvirus drugs are administered. These drugs are specifically designed to target and inhibit the replication of herpesviruses, including EBV (Lv et al., 2022; Pei et al., 2020).

HYPOTHESIS

Hypothesis1: *The prevalence and viral load of Epstein-Barr virus among lymphoma patients in Ethiopia will exhibit a significant elevation compared to other populations.*

Hypothesis 2: *The genomic variability of EBV within the lymphoma patient population in Ethiopia will demonstrate distinguishable patterns when compared to EBV variations observed in other populations.*

Hypothesis 3: *The methylation profile of EBV DNA in lymphoma patients from Ethiopia will exhibit distinctive characteristics, differing from the methylation profiles observed in other populations.*

OBJECTIVES

General objective

- To determine molecular epidemiology and genetic diversity of Epstein - Barr virus among lymphoma patients in Ethiopia.

Specific objectives

- To determine the Sero- prevalence and molecular detection of EBV among patients with different lymphoma.
- To characterize the distributions of EBV genotypes isolated from patients with lymphoma.
- To assess DNA methylation profiles of EBV from EBV associated lymphomas.
- To determine the genetic variations of EBV among different lymphoma types

CHAPTER 3. MATERIALS AND METHODS

3.1. Study area

The study was conducted at Tikur Anbesa Specialized Hospital (TASH) located Lideta sub city in Addis Ababa Ethiopia under the College of Health Sciences, Addis Ababa University **(Figure 3.1).** It is one of the referrals and oldest public hospitals serving the entire nation in providing all forms of medical service. It is the main teaching hospital for both clinical and preclinical training of most disciplines. It is the largest of the public teaching hospitals having more than 26 departments. It is also an institution where specialized clinical services that are not available in other public or private institutions are rendered to the whole nation. The TASH has 200 doctors, 379 nurses and 115 other health professionals dedicated to providing health care services. The various departments, faculties and residents under specialty training in the School of Medicine provide patient care in the hospital. The hospital also has 950 permanent and contract administrative staff to support the hospital activities. The annual Cancer patients diagnosed at TASH is more than five thousand excluding the blood cancer coming from all over the country. The blood cancer patients that are under follow-ups at haematology clinic are more than one thousand per year.

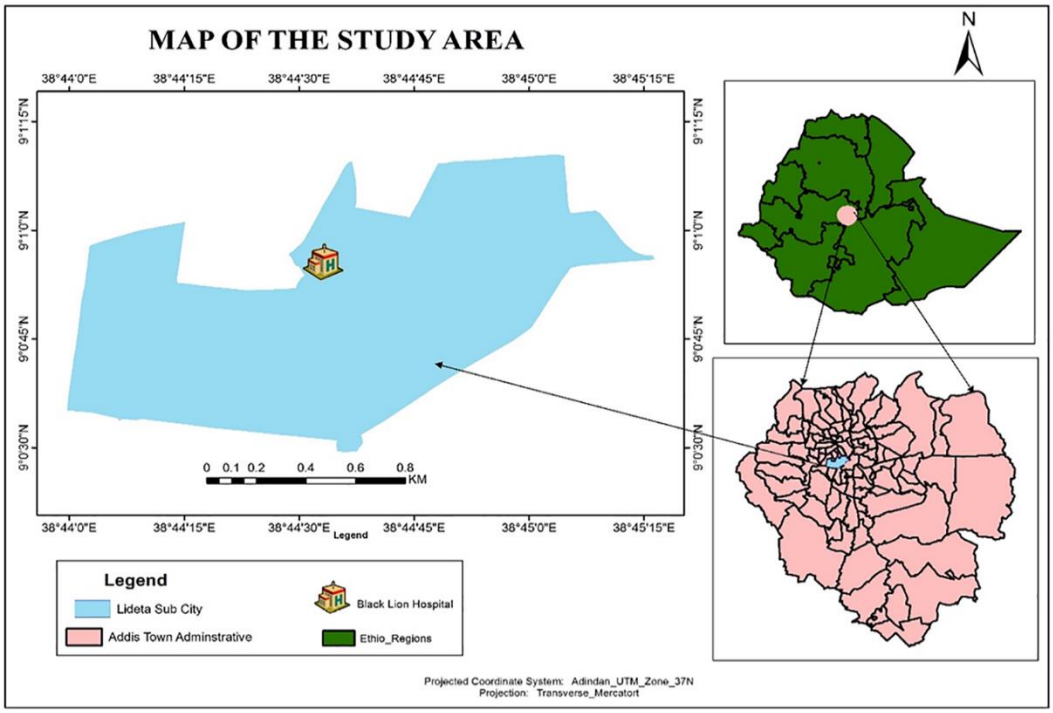


Figure 3.1: Map of Tikur Anbessa Specialized Hospital, Lideta Sub city, Addis Ababa Ethiopia.

3.2. Study design and period

We have conducted a hospital-based study at Tikur Anbessa Specialized Hospital in Addis Ababa, Ethiopia, combining retrospective and prospective cross-sectional approaches. For the retrospective component, we have collected formalin-fixed paraffin-embedded (FFPE) lymphoma tissue blocks from the Pathology department at the hospital. This data collection focuses on FFPE blocks diagnosed within a two-year period, specifically from 2019 to 2020. In addition, for the prospective portion of the study, we recruited two groups of participants. Firstly, we enrolled lymphoma patients who were already under follow-up in the hematology clinic at the hospital. Secondly, we included suspected lymphoma patients who were referred to the hospital for lymph node biopsy to confirm the presence of lymphoma.

3.3. Sample size calculation

A total of 133 Lymphoma confirmed FFPE blocks were collected from pathology department from the two years (2019-2020). The prospective sample size was determined based on the global prevalence of EBV, which has been reported to be more than 90% according to the epidemiological studies conducted by (Tzellos & Farrell, 2012). Considering an assumed prevalence of 90% for EBV, the sample size was calculated using the following formula:

$$n = \frac{\left(\frac{Z\alpha}{2}\right)^2 pq}{d^2}$$

$$\text{where, } n = \text{sample size}$$

$$Z = \text{confidence interval}$$

$$p = \text{proportion}$$

$$q = 1 - p$$

$$\text{By using 95\% confidence interval i. e. } Z = 1.96$$

$$P=0.9$$

d= 0.05 as 95% confidence interval is taken.

Therefore, to achieve the desired sample size, a total of 172 study participants were prospectively recruited, considering a non-lymphoma rate of 20% and accounting for potential non-responses.

3.4. Inclusion and exclusion criteria

3.4.1. Inclusion criteria

- FFPE blocks diagnosed as lymphoma between the years 2019 and 2020 were included.
- Selected blocks needed to contain enough tissue for staining and genomic DNA isolation.
- Patients who expressed willingness to participate in the study and provide oral and written consent were included in the prospective cohort.

3.4.2. Exclusion criteria

- We excluded FFPE blocks that had insufficient tissue samples and lacked sociodemographic and clinical data.
- Pregnant women, patients with other hematologic malignancies, and those undergoing chemotherapy were also not included in the study.

3.5. Data and sample collection

For the retrospective study, we conducted a thorough review of pathology department reports from 2019 to 2020. Any cases of lymphoma identified in these reports were included in our FFPE block collection. In addition to the FFPE blocks, we obtained sociodemographic data and Hematoxylin and Eosin (H&E) reports for each included case. To analyze the genomic DNA, we collected two sections, each measuring 5µm, from the FFPE blocks. These sections were carefully preserved in 1.5ml Eppendorf tubes until we performed nucleic acid extraction.

In the prospective study, we enrolled two distinct groups of participants to gather data for our research. The first group consisted of lymphoma patients, from whom we collected 20ml of peripheral blood to isolate peripheral blood mononuclear cells (PBMCs) and extract plasma and serum. The second group of participants comprised individuals with suspected lymphoma who underwent lymph node biopsy. From these patients, we removed a $5mm^3$ block of fresh lymph node tissue for further analysis. Additionally, we collected a 10ml blood sample from this group. Blood samples were collected in Acid Citrate Dextrose (ACD) and serum separator tubes (SST) in accordance with the manufacturer's instructions for the plasma and serum separation respectively. While the tissue sample was stored in 1% phosphate-buffered saline (PBS) immediately after surgery until single cell suspension was performed. Lymphomononuclear cells were isolated from the lymph node biopsy specimen. In addition to the biological samples, we also collected clinical and sociodemographic data from both groups of participants.

3.6. Laboratory procedure

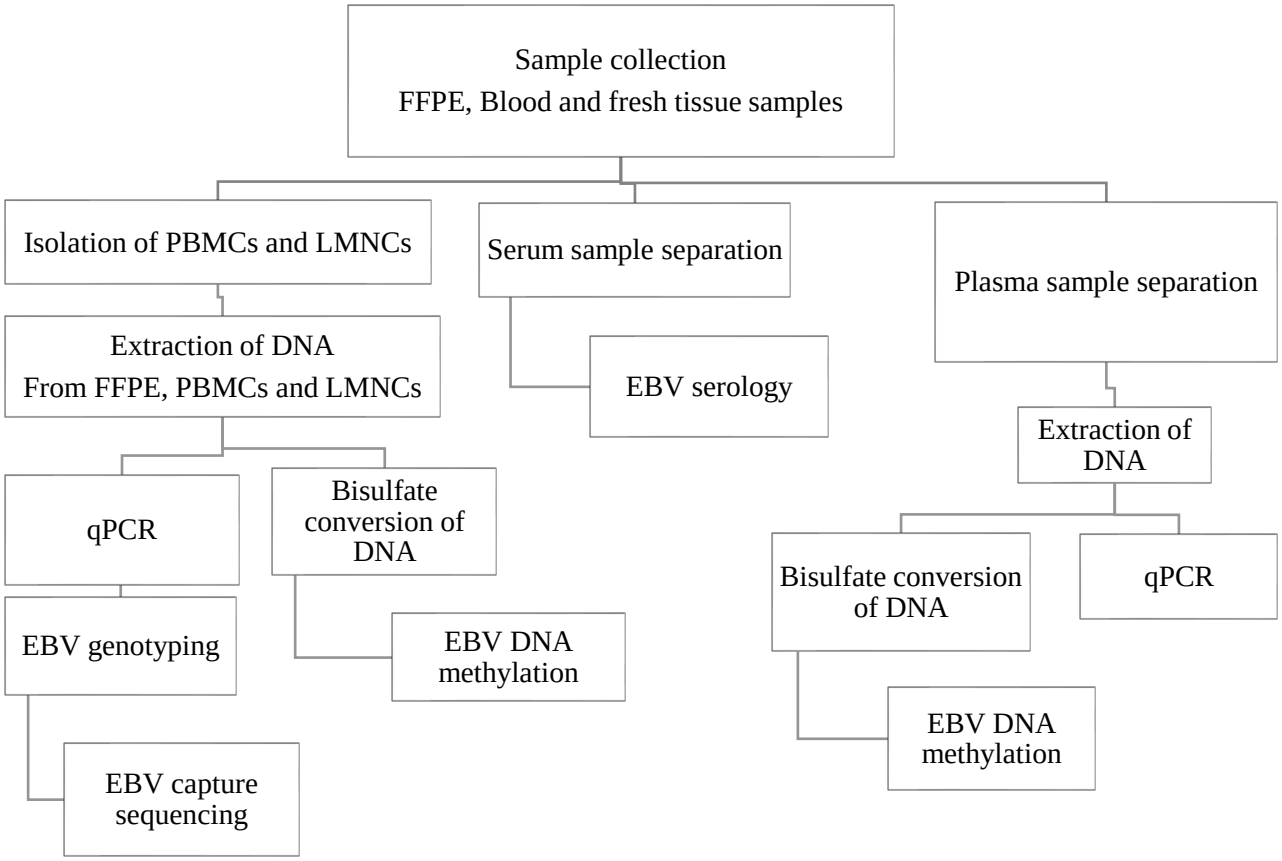


Figure 3.2: The flowchart illustrating the sequential steps and processes of the study.

3.7. Sample processing

3.7.1. Isolation of peripheral blood mononuclear cells

To isolate peripheral PBMCs from the collected blood samples, we employed the Ficoll–Paque technique (Global Life Sciences Solutions, Marlborough, MA, USA). This method utilizes a density gradient medium, Ficoll–Paque, which facilitates the separation of different cellular components based on their density. The blood samples were first diluted 1:1 in 1X phosphate buffered saline (PBS), and the diluted blood was then layered on top of Ficoll-Paque PLUS solution. After centrifugation at 2000rpm for 30 minutes at 20°C, the blood cells underwent differential migration, leading to the formation of distinct layers consisting of various cell types. (**Figure 3.3).** PBMCs were collected from the interface layer between the PBS and Ficoll-Paque PLUS. Then PBMCs were washed with a PBS

twice to eliminate platelets, Ficoll-Paque PLUS, and any remaining plasma. After determining the viable count of PBMCs, the PBMCs were transferred to cryovials along with freezing medium containing 10% dimethyl sulfoxide (DMSO). The cryovials were placed in a Mr. Freezie and after 24 hours they were stored at -80°C for further analysis.

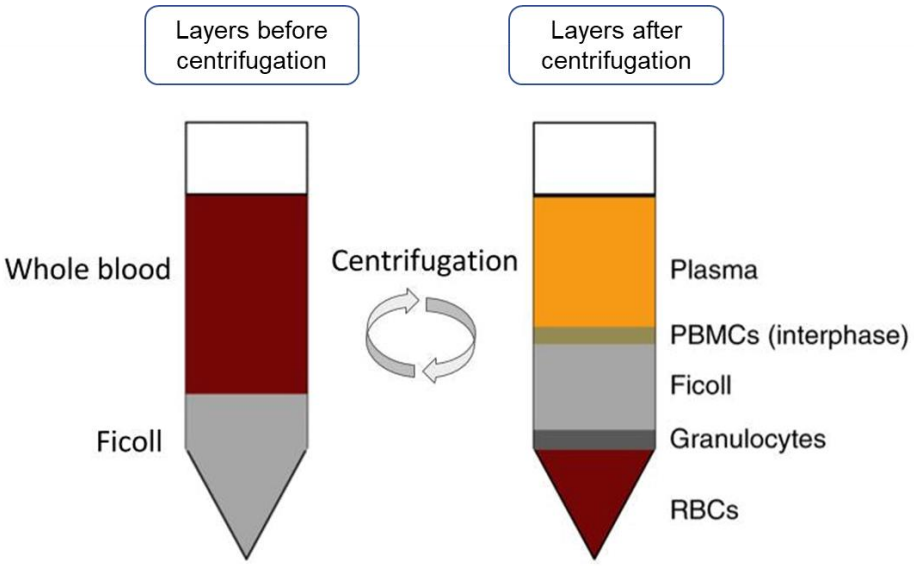


Figure 3.3: Isolation of PBMCs using Ficoll-Paque

3.7.2. Isolation of Lymphomononuclear cells

For the isolation of lymph node mononuclear cells from fresh lymph node biopsies, we utilized a single cell suspension technique. This technique involved dissociating the tissue into individual cells, enabling the separation and isolation of specific cell populations (LMNCs). To sterilely harvest the lymph node, it was placed in a conical tube with PBS. In a hood, a 70uM cell strainer was placed in a petri dish, and the lymph node and PBS were poured into the strainer. Using the backside of a syringe plunger, the lymph node was mashed and passed through the strainer, ensuring fluid flow into the petri dish. Once fully dissociated, the strainer was placed back on the conical tube, and the liquid was filtered through a 70uM cell strainer using a pipette. Additional rinsing with PBS was done to maximize recovery. The tube was then filled to 50mL with PBS. For cell counting, 10uL

of the solution was taken and the remaining 50mL conical tube with lymph node cells was placed on ice. The tube was centrifuged at 1200 RPM for 10 minutes, the supernatant was removed, and freezing media was added. The pellet was gently resuspended by pipetting, and 1mL of the solution was transferred into each cryovial. The cryovials were placed in a Mr. Freezie for at least 24 hours before storing them in -80°C.

3.7.3. Trypan blue staining

To ensure the quality of the isolated cells, viability and cell count assessments were performed. To perform Trypan blue staining for cell counting, we started by preparing a cell suspension through cell harvesting. The suspension was then appropriately diluted to achieve a suitable concentration for counting. Next, an equal volume of the diluted cell suspension was mixed with Trypan blue dye using a 1:1 dilution. The stained cell suspension was loaded onto a counting chamber or hemocytometer, allowing the cells to settle and distribute evenly. Under a microscope at 100X magnification, the viable (unstained) and non-viable (stained) cells were observed and counted within the designated grids or squares of the counting chamber. The total cell count and viability were calculated based on the obtained numbers. The viability percentage was determined by dividing the total number of viable cells per ml of the aliquot by the total number of cells per ml of the aliquot and multiplying by 100.

$$\text{Viable cells (\%)} = \frac{\text{total number of viable cells per ml of aliquot}}{\text{total number of cells per ml of aliquot}} \times 100$$

Most of the samples demonstrated viability exceeding 90%, indicating that the isolated cells were of good quality and suitable for subsequent analysis. Following isolation, both the PBMCs and LMNCs were stored at -80°C in 10% dimethyl sulfoxide (DMSO) (Thermo Fisher Scientific, Waltham, MA, USA) until DNA extraction was conducted.

3.7.4. Serum and plasma isolation

To isolate serum and plasma, blood samples were collected using appropriate tubes or containers. For serum collection, the samples allowed to clot, resulting in the separation of serum from the clot. In contrast, for plasma isolation, the collected blood samples were directly subjected to centrifugation. After clot formation, centrifugation was carried out at

1500× g for 10 minutes. Following centrifugation, the serum or plasma was carefully separated and transferred into separate vials using pipettes. Finally, the isolated serum and plasma samples were stored at -80°C in a freezer for future use and analysis.

3.8. Extraction of Genomic DNA

Two 5µm FFPE tissue samples were deparaffinized using a ready-to-use deparaffinization solution (QIAGEN, Hilden, Germany). The solution was added to the samples, facilitating the removal of the paraffin from the tissue. Subsequently, DNA extraction from the FFPE samples was conducted using the QIAamp DSP DNA FFPE tissue kit (QIAGEN, Hilden, Germany). For the extraction of genomic DNA from the PBMCs and LMNCs, we followed the manufacturer's instructions and employed the QIAamp mini-DNA kit (QIAGEN, Hilden, Germany). This kit provided the necessary reagents and protocols for efficient DNA extraction.

The final elusion volume for the extracted DNA differed depending on the sample type. For the FFPE DNA, the elusion volume was 40µL, while for the DNA extracted from PBMCs and LMNCs, the elusion volume was 100µL. These elusion volumes ensured enough DNA for subsequent analysis. To assess the concentration and quality of the extracted DNA, we utilized both the Qubit (Invitrogen, Waltham, MA, USA) and Nanodrop (Thermo Scientific, Waltham, MA, USA) spectrophotometers. The DNA concentration was measured using the Qubit, and the range varied from 8.6 ng/µL up to 908.4 ng/µL. The quality of the DNA was determined by measuring the absorbance at 260/280 nm using the Nanodrop spectrophotometer, with values ranging between 1.7 and 2.2. A 2µL volume of elution was used to determine the quality and quantity of the extracted DNA.

3.9. Serology of EBV using VCA IgG antibody using ELISA

To determine the presence of EBV IgG antibodies, we employed the BioPlex 2200 EBV IgG kits on a BioPlex 2200 Analyzer (Bio-Rad, Hercules, CA, USA) following the manufacturer's instructions. Serum samples collected from the study participants were used for the serological detection of EBV. The BioPlex 2200 Analyzer, along with the EBV IgG kits, allowed us to measure the levels of EBV IgG antibodies present in the serum samples.

The results of the EBV IgG antibodies were categorized based on the antibody index (AI). The interpretation criteria for the results were as follows: AI ≤ 0.8 indicated a negative result, AI 0.9-1 indicated an equivocal or intermediate result, and AI ≥ 1.1 indicated a positive result. To ensure the accuracy and reliability of the test, we included internal positive and negative controls as part of the analysis. These controls served as reference points to validate the performance of the test and ensure the accuracy of the results.

3.10. Detection of the EBV *EBNA1* gene using real time qPCR

The presence of EBV was detected using real-time quantitative PCR (qPCR) using a Viia7 qPCR machine (Applied Biosystems, Waltham, MA, USA). For each qPCR reaction, 10 ng of DNA was used as the starting material. The reaction mixture consisted of 5 µL of 2× Fast SYBR green master mix (Applied Biosystems), 0.25 µL of forward primer (10 µM), 0.25 µL of reverse primer (10 µM), 2.5 µL of distilled water, and 2 µL of DNA with a concentration of 5 ng/µL, resulting in a total reaction volume of 10 µL. The qPCR reaction consisted of 40 cycles, with denaturation at 95°C for 1 second, annealing at 60°C for 20 seconds, and extension at 70°C for 30 seconds.

To quantify the EBV viral load, we used 12 standards for the *EBNA1* gene and the *ACTB* gene. The qPCR reaction for the standards used a total volume of 10 µL, including 5 µL of 2× Fast SYBR green master mix (Applied Biosystems), 0.25 µL of forward primer (10 µM), 0.25 µL of reverse primer (10 µM), and 2.5 µL of distilled water, along with 2 µL of the ACTB standards. The thermal profile for the reaction was denaturation at 95°C for 1 second, annealing at 60°C for 20 seconds, and extension at 70°C for 30 seconds, with a total of 40 cycles.

3.11. Reference strains and cell lines

To validate the accuracy of the assay, we used the Raji cell line as a positive control and the K562 cell line as a negative control. These cell lines, obtained from the American Type Culture Collection (ATCC), were stored in the laboratory (ATCC numbers CCL-86 and CCL-243, respectively). For genotyping, the B95-8 cell line (ATCC VR-1492) was used as a representative of EBV type 1, and the Jijoye cell line (ATCC CCL-87) was used as a representative of EBV type 2.

The *EBNA1* gene standard was derived from the *EBNA1* amplicon with an accession number of *EBNA1*-NC_007605.1 (**Figure 3.4**). The stock concentration of the standard was 1.98 ng/µL, and the standards were prepared by twofold dilutions, resulting in a range of concentrations from 2.6 × 106 copies/µL to 1.2 × 103 copies/µL.

Figure 3.4: *EBNA1* amplicon primer sequence sites (sequence number; Forward-85,519-85,538; Reverse 85,566-85,586)

The *ACTB* gene, which served as a housekeeping gene for normalizing the host genome DNA, was obtained from Homo sapiens with an accession number of NM_001101.5 (**Figure 3.5**). The stock concentration of the *ACTB* standard was 4.3 ng/µL, and the standards were prepared through twofold serial dilutions, ranging from 8.2 × 105 copies/µL to 4 × 102 copies/µL.

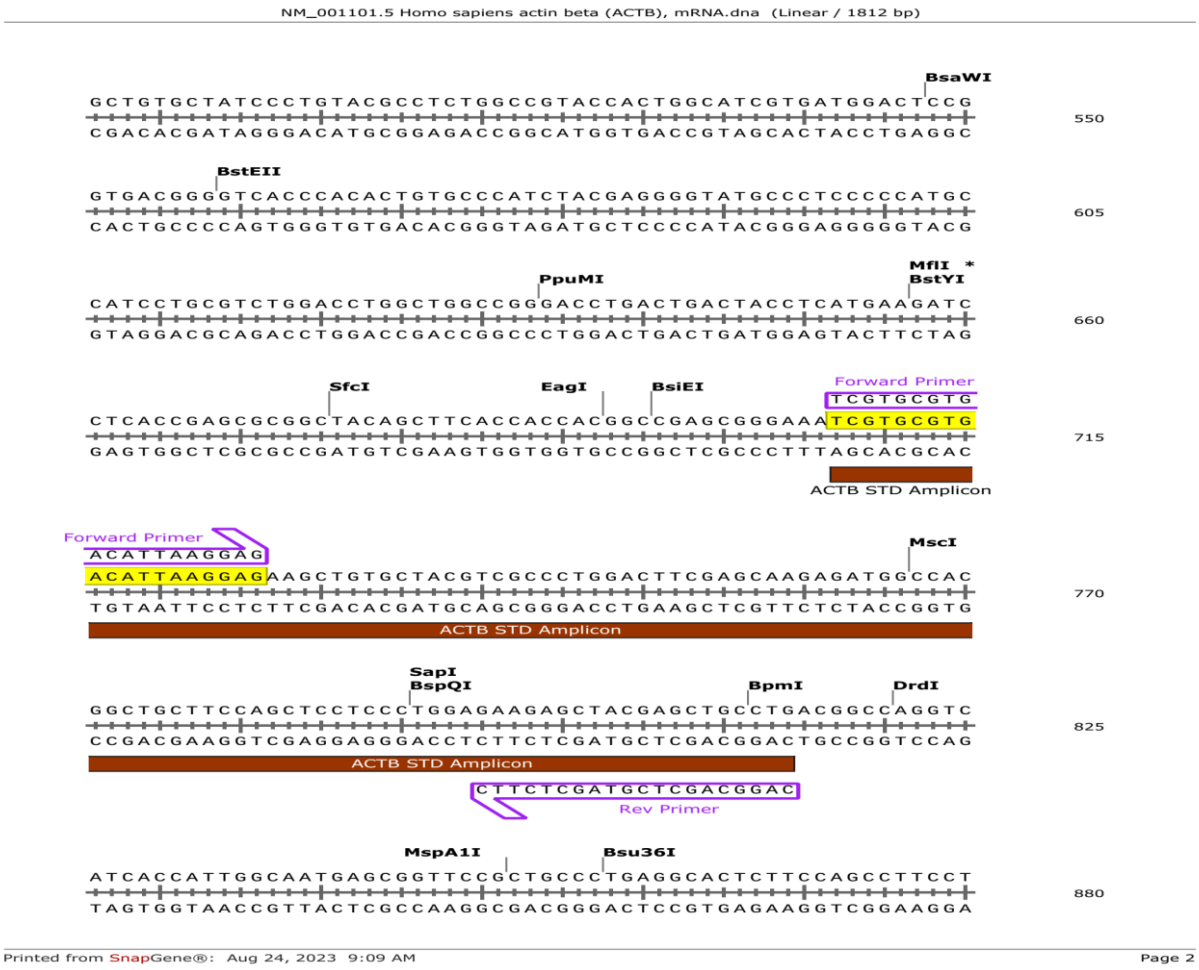


Figure 3.5: *ACTB* amplicon primer sequence sites (sequence number; forward 707-726; reverse 794-814)

3.12. Standard curves design

For the design of the EBV standard curve, a reference standard with known concentrations of EBV for both *EBNA1* and *ACTB* genes was obtained. Serial dilutions of the reference standard were prepared to establish the dynamic range of real-time quantitative PCR. The Ct values obtained from the analysis of 12 different concentrations of the standards were then plotted against the logarithm of the known concentrations, resulting in the generation of a standard curve. Subsequently, the standard curve was utilized to quantify EBV levels in the study samples by interpolating the Ct values. All the standards, samples, and controls were run in triplicate to ensure reproducibility. The CT values obtained from the qPCR were used

to generate a standard curve for quantification. The CT values of each standard were calculated and converted into EBV copies/mL to obtain the EBV viral load using the ΔΔCt method (**Figure 3.6**).

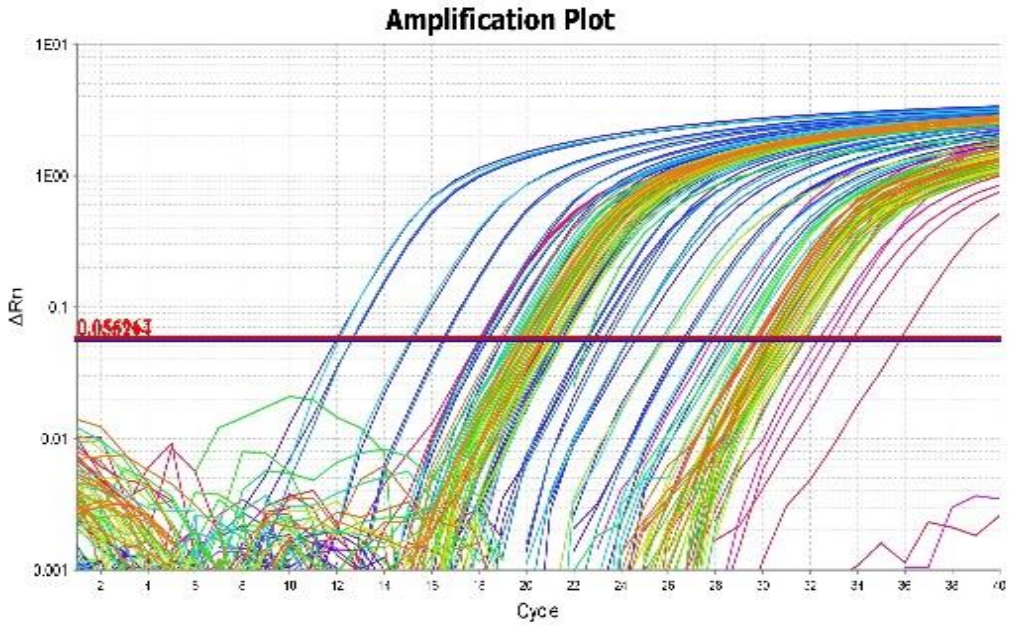


Figure 3.6: The amplification curve for EBV detection and quantification by targeting the *EBNA1* gene.

3.13. EBV typing using *EBNA3C* gene

EBV typing was determined by targeting EBNA3C gene using a specific primer designed to distinguish between EBV type 1 and type 2. The primer sequences are provided in **Table 3.1** and were designed to target the region of divergence between EBV type 1 (B95 coordinate 87,651–87,669) and EBV type 2 (B95-8 coordinate 87,803–87,783) (Kafita et al., 2018). These coordinates correspond to the nucleotide sequences in the reference genome (accession number: NC_007605.1). The PCR reaction was performed in a total volume of 20 µL, which included 10 µL of NEBNext Q5 Hot Start HiFi PCR Master Mix (New England Biolabs, Ipswich, MA, United States), 1 µL of forward primer (10 µM), 1 µL of reverse primer (10 µM), 3 µL of distilled water, and 5 µL of DNA sample with a total concentration of 50 ng.

The PCR reaction consisted of an initial denaturation step at 98°C for 30 seconds, followed by 50 cycles of denaturation at 98°C for 10 seconds, annealing at 68°C for 30 seconds, and extension at 72°C for 10 seconds. After PCR amplification, 2% agarose gel electrophoresis was performed to visualize the PCR products. For reference, the B95-8 cell line (ATCC VR-1492) was used as a representative of EBV type 1, and the Jijoye cell line (ATCC CCL-87) was used as a representative of EBV type 2. These cell lines were obtained from the American Type Culture Collection (ATCC) and stored in the laboratory. Differentiation between EBV type 1 and type 2 genotypes was achieved by observing the size of the amplification products, with a product size of 153 bp indicating EBV type 1 and a product size of 246 bp indicating EBV type 2.

Table 3.1: Primer sequences for EBV detection and genotyping

PCR Reactions	**Forward**	**Reverse**
EBNA1	TCATCATCATCCGGGTCTCC	CCTACAGGGTGGAAAAATGGC
ACTB	TCGTGCGTGACATTAAGGAG	CAGGCAGCTCGTAGCTCTTC
EBNA3C	AGAAGGGGAGCGTGTGTTG	GGCTCGTTTTTGACGTCGG

3.14. Bisulfite conversion of DNA for methylation assay

For the bisulfite conversion of DNA, we utilized the EZ DNA Methylation kit (Zymo research, CA, USA) according to the manufacturer's instructions (**Figure 3.7**). Each bisulfite treatment involved using an input amount of DNA ranging from 500 ng to 2 µg.

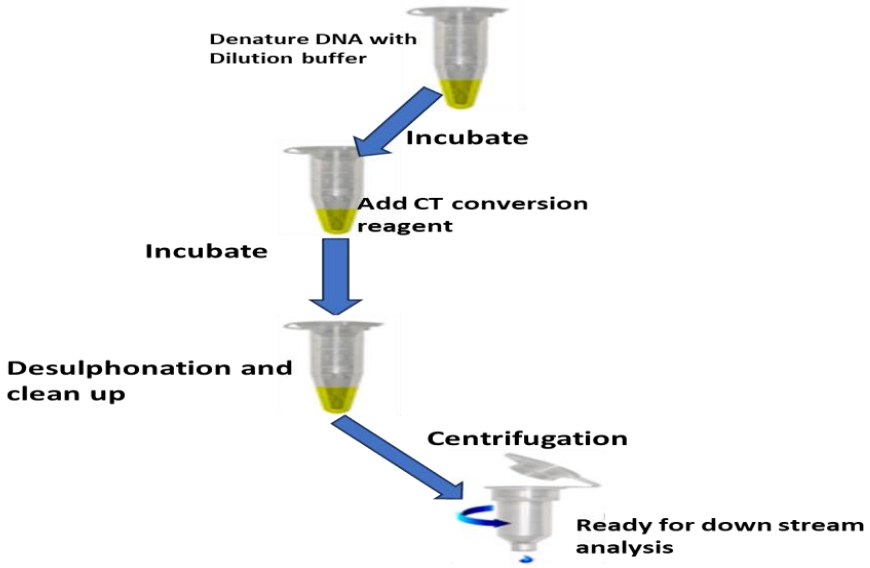


Figure 3.7: The general procedure for DNA bisulfite conversion using the EZ DNA Methylation Kit

3.15. Region-specific quantitative DNA methylation analysis using methylation-iPLEX

EBV DNA methylation analysis was carried out using methylation-iPLEX (Giacopelli et al., 2019). iPLEX capture and extension primers were designed to amplify and target EBV CpGs from bisulfite-converted DNA sequences using the Typer4.0 software (Agena Biosciences, Sandiago, USA). Capture primers were designed to overlap at least 3 non-CpG cytosines to ensure amplification of bisulfite-converted DNA only. For extension primer design, CpGs of interest were treated as a C/T nucleotide polymorphism, reflecting methyl cytosine and cytosine after bisulfite conversion **(Table 3.7)**.

For the methylation-iPLEX workflow, we followed the iPLEX method (Agena Biosciences, Sandiago, USA). Briefly, regions around CpGs of interest are amplified in a multiplexed capture PCR reaction followed by treatment with Shrimp alkaline phosphatase to dephosphorylate unincorporated dNTPs. A multiplexed single base extension PCR was then performed using mass modified terminator nucleotides and primers that anneal immediately 5’ to the CpGs of interest. Following desalting, samples are dispensed onto the Spectro CHIP array for analysis by MALDI-ToF (Shimadzu instruments, Maryland, USA). The ratio of the extension products is reported directly as the percent methylation.

Samples were dispensed using the RS1000 nano dispenser and analyzed using the Mass ARRAY Analyzer4 system in 384 sample chip formats. For data quality control, single CpG measurements displaying less than 80% primer extension or combined extended peak areas (methylated + unmethylated) <10 intensity units in the mass spectrum were excluded (Giacopelli et al., 2019).

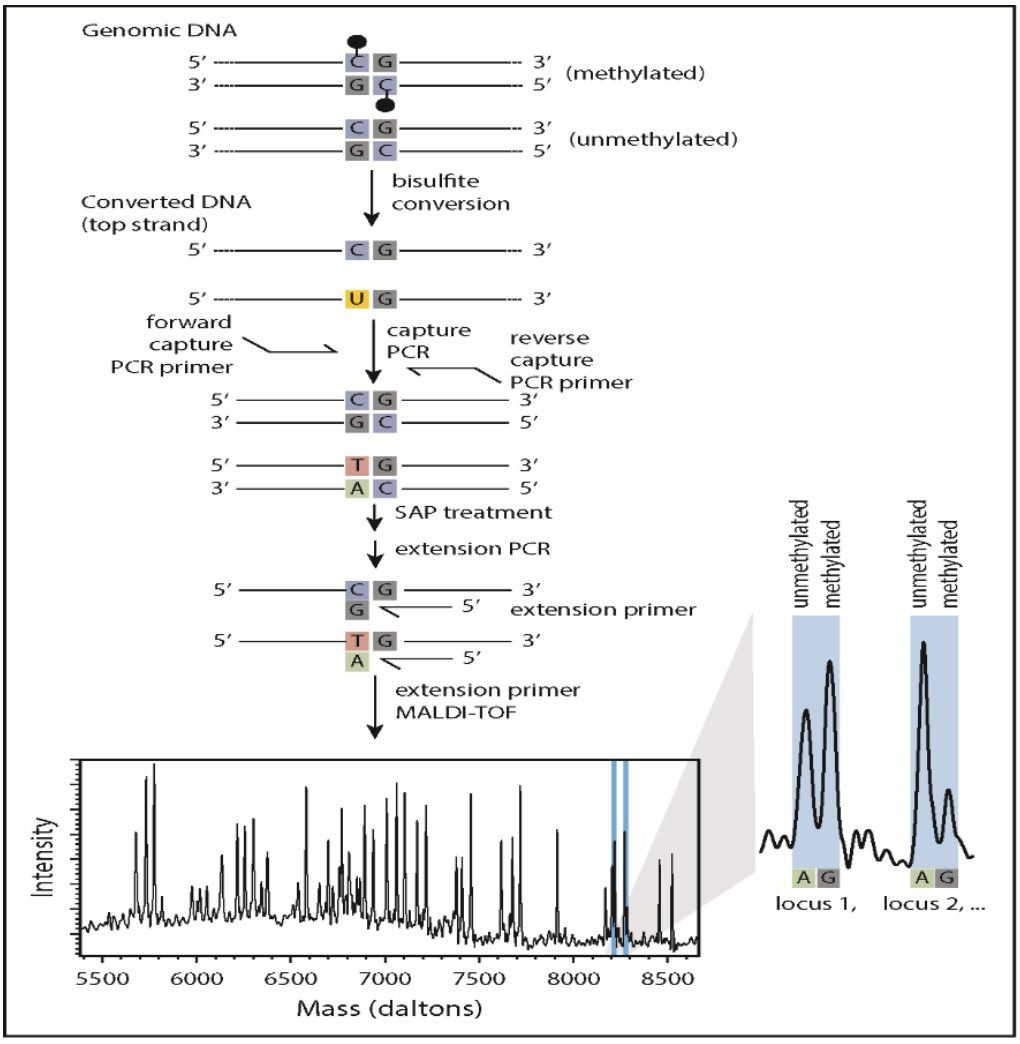


Figure 3.8: IPLEX assay procedure for determination of EBV DNA methylation status

3.16. EBV capture sequencing

3.16.1. Library preparation

We have used KAPA HyperPrep Kit (Roche diagnostics, IN, USA) for performing the library preparation for EBV capture sequencing using Illumina. The protocol begins with

input recommendations, where libraries are prepared from fragmented DNA between 150-350 bp using third-party library preparation kits. For this capture sequencing, we have used 500 ng of DNA as a starting material. The protocol requires specific consumables and equipment, including xGen Lockdown Probe Pools, xGen Hybridization and Wash Kit, blocking oligos, xGen Universal Blockers, and other items from IDT and other suppliers.

3.16.2. EBV genome capture

The workflow consists of several steps. The steps include performing the hybridization reaction by combining the library and probe pool, followed by incubation and capturing of target DNA fragments. Buffers are prepared for the hybridization and wash steps, and Streptavidin beads are washed to remove unbound DNA fragments and contaminants. Bead capture is performed by incubating the Streptavidin beads with the captured DNA, followed by additional washes to remove non-specifically bound DNA fragments. Post-capture PCR is then conducted to amplify the captured DNA fragments, and the resulting PCR fragments are purified to remove unwanted artifacts and contaminants. The captured library is validated and quantified using appropriate methods, and finally, sequencing of the captured library is performed using next-generation sequencing platforms. **Figure 3.9** illustrates the overall workflow of EBV capture sequencing. The detailed procedure for Library preparation **(Annex 10)** and Illumina capture sequencing can be found in the **Annex 11.**

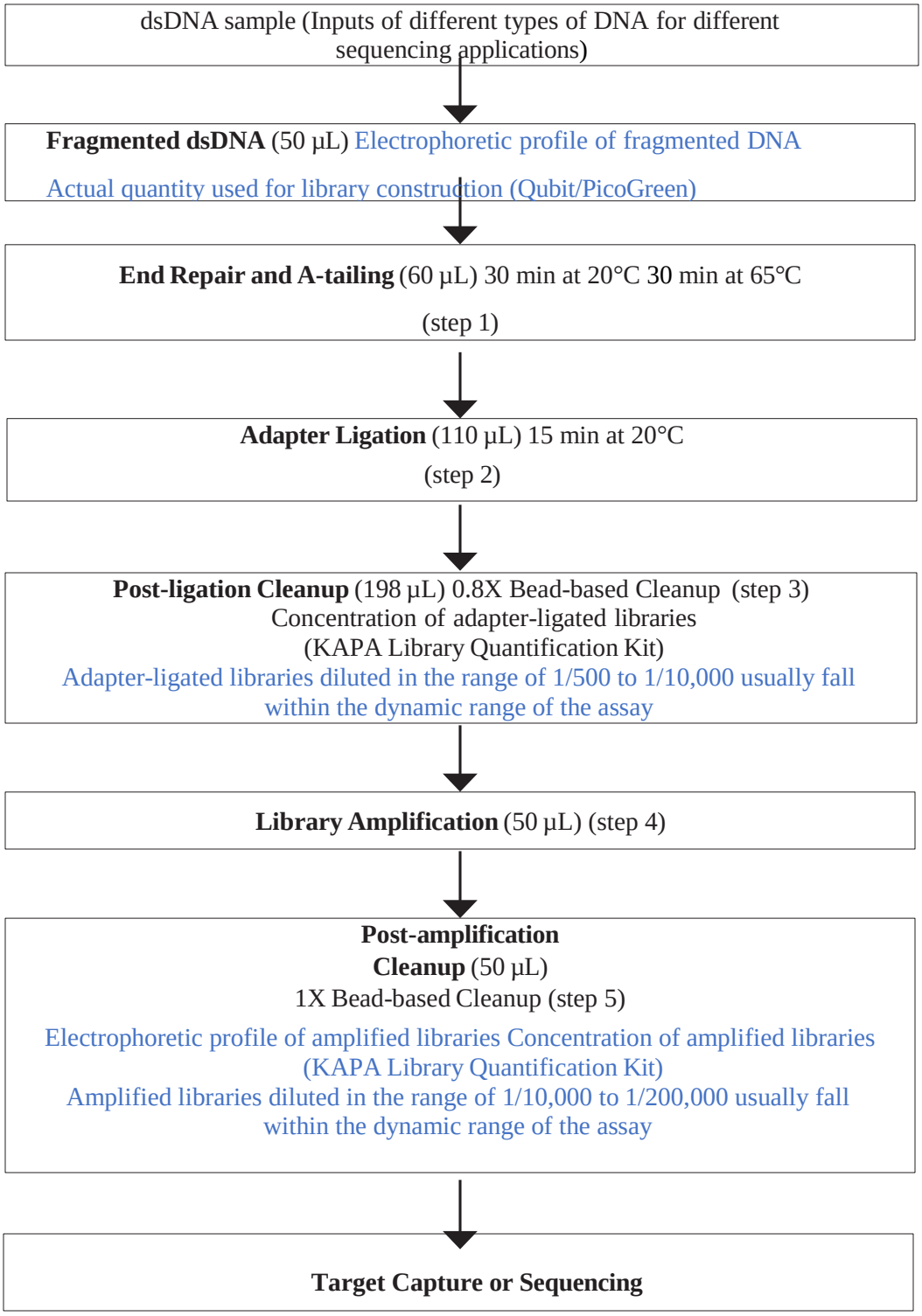


Figure 3.9: The overall work flow of EBV capture sequencing using KAPA HyperPrep Kit.

3.17. Phylogenetic tree analysis

Phylogenetic trees were built using the identified variants as input for the *IQTREE2* tool. IQTREE2 is a fast and effective stochastic algorithm to infer phylogenetic trees by maximum likelihood. IQTREE2 has been run with the following parameters: **B 1000**: to enable the ultrafast bootstrap (**UFBoot**) approximation with 1000 replicates, which provides approximately unbiased branch support values. **m MFP**: to enable **ModelFinder**, which automatically selects the best-fit model to infer the maximum-likelihood tree**. bnni**: to reduce the risk of overestimating branch supports with UFBoot due to severe model violations. These phylogenomic trees have been generated using two layouts: a classic layout and a semi-circular one.

3.18. Quality assurance

The research quality was assured through three main phases, with each phase characterized by specific measures:

Phase I: Pre-Analytical

- Ethical approval for the research proposal was obtained from the relevant department, college, and national ethical review committees, ensuring compliance with ethical standards.
- The laboratory setup was prepared, and specific training was undertaken to ensure proper handling of samples and data.
- Study participants were recruited based on predefined inclusion criteria, ensuring a representative sample population.
- Sample and data collection were performed in collaboration with nurses, phlebotomists, and surgeons, following standard protocols.
- Samples and isolated DNA were preserved at -80 degrees Celsius to maintain sample integrity.

Phase II: Analytical

- Qualified laboratory experts and senior advisors were involved in the analysis, specifically for epigenomic and genomic analysis.
- Standard Operating Procedures (SOPs) were strictly followed during all study protocols to maintain consistency and accuracy.
- Bioinformaticians played a crucial role in analyzing the genomic data, ensuring accurate interpretation and analysis.

Phase III: Post-Analytical

- All clinical, demographic, and laboratory findings were meticulously entered into computer software, ensuring accurate data management.
- Manuscript writing was undertaken and the results were disseminated to the scientific community through conferences.

3.19. Data analysis

Data entry and analysis was performed using Statistical Package for the Social Sciences (SPSS) version 26 software. Descriptive statistics were computed to provide a summary of the data. To assess the association between EBV detection, viral load, EBV typing and other sociodemographic and clinical variables, the Pearson Chi-square test was employed. Statistical significance was determined for associations with p-values below 0.05.

DNA methylation data from iPLEX were quality-filtered based on number of obtained methylation values per samples, with the lowest quality sample tertile removed (amounting to samples with >50% of data excluded due to technical peak intensity threshold). Hierarchical clustering was carried out using **Cluster 3.0 V1.52** (De Hoon et al., 2004) with Euclidean Distance metric and TreeView (Saldanha, 2004) for visualization. Principal Component analysis (PCA) was generated with **Qlucore Omics Explorer V3.8**. Box- and Violin plots were generated using BoxPlotR (Spitzer et al., 2014) and **Vioplot** (Adler &

Kelly, 2005), respectively. Statistical analysis of inter-group differences was carried out using Fisher exact test and Student's t-test.

Analysis of the capture sequencing involves trimming adapters and filtering low-quality bases from the sequencing reads. Adapters are sequences that are ligated to the DNA fragments before sequencing and need to be removed to avoid interference in subsequent analysis. Low-quality bases, which can arise due to sequencing errors or technical factors, also need to be filtered out to ensure the accuracy of downstream analysis. Once the preprocessing steps are completed, the reads are aligned against the reference genomes of the host (human) and other potential contaminants, such as the Illumina PhiX spike-in. This alignment helps identify and remove reads that originate from the host or other sources, ensuring that subsequent analysis focuses on EBV. For the unmapped reads that do not align to the host or other contaminants, a secondary alignment step is performed against the reference genomes of EBV Types I (NC_007605.1) and II (NC_009334.1). This step aims to identify and categorize reads that potentially originate from the EBV virus. To avoid counting duplicate reads multiple times, a marking process is implemented to flag and remove duplicated reads. These duplicated reads are not considered in the subsequent variant calling analyses.

The analysis of the sequencing data involved several software tools. First, **fastqc v0.12.1** was used to conduct a quality check of the FASTQ data, evaluating the overall sequencing data quality. Subsequently, **fastp v0.23.4** was applied for trimming and quality filtering of the FASTQ files. This step aimed to remove low-quality reads and adapter sequences from the data. To address potential contamination in the FASTQ data, a combination of **bowtie2 v2.4.4, samtools 1.17**, and **picard v3.1.0** was employed. Bowtie2 performed the alignment of the reads against a contaminant database, while samtools and picard facilitated the removal of the aligned reads corresponding to contaminants. For the extraction and alignment of reads associated with the EBV, **bwa-mem2 v2.2.1** was utilized. This specific tool conducted the alignment of the reads against the EBV reference genome, allowing for the identification of EBV-associated sequences.

In the subsequent steps, variant calling was performed using **lofreq v2.1.5**, normalizing the identified variants with **vt v0.57721--heae7c10_3**. Finally, variant annotation was accomplished using **snpeff v5.1d** and **snpsift v5.1d**, providing functional annotations for the identified variants.

3.20. Ethical clearance

Ethical clearance was secured from Department of Microbiology, Immunology and Parasitology (Meeting number: DRERC/004/2021), College of Health Sciences Institutional Review Board (Protocol number: 056/21/DMIP), Addis Ababa University and National Ethical Review Committee (Reference number: 8/182/552/22).

Both verbal and written consent were obtained prior to collecting samples from each patient. Patient confidentiality was strictly maintained throughout the process, as their identification was not disclosed. Instead, laboratory IDs were assigned for data and sample collection purposes. To ensure the security of the collected information, it was stored confidentially on a password-protected computer.

CHAPTER 4. RESULTS

4.1. Sociodemographic and Clinico-pathological characteristics

A total of 305 participants were enrolled in this study. Among them, 133 individuals were included based on retrospectively collected FFPE lymphoma blocks, while 120 participants were confirmed lymphoma patients. The remaining 52 participants were suspected to have lymphoma.

In terms of gender distribution, n = 197 (64.6%) of the study participants were male, with a male-to-female ratio of 1.8 to 1. The age of the participants varied from 3 to 85 years, with a mean age of 38 ± 17.5 years (standard deviation). The age group between 41 and 50 years had the highest number of participants n = 69 (22.6%). Sociodemographic and clinical profiles of the study participants are provided in **Table 4.1**.

Table 4.1: Demographic and clinical characteristics of the study participants.

Characteristics	Number	Percent
Study participants (n = 305)		
FFPE blocks	133	43.6
Confirmed lymphoma subjects	120	39.3
Suspected lymphoma subjects	52	17
Sex		
Male	197	64.6
Female	108	35.4
Age at diagnosis		
<20	50	16.4
21–30	58	19.0
31–40	59	19.3
41–50	69	22.6
51–60	34	11.1
>61	32	10.5
Not applicable	3	1

HIV status (*n* = 172)		
Positive	29	16.9
Negative	122	70.9
Unknown	21	12.2
Lymphoma types		
Hodgkin's lymphoma	91	29.8
Non-Hodgkin's lymphoma	197	64.6
Non-lymphoma (from suspected subjects)	17	5.6

4.2. Histopathological Profiles

Out of the total 305 study participants, 288 individuals were diagnosed with lymphoma based on confirmed pathology reports. On the other hand, 17 patients initially suspected to have lymphoma were found to be negative for the disease after obtaining a pathology report. Consequently, these non-lymphoma patients were excluded from further analyses.

Among the 288 lymphoma patients, *n* = 91 (32%) were classified as Hodgkin's lymphoma patients and *n* = 197 (68%) were categorized as non-Hodgkin's lymphoma patients. Within the NHL subgroup, most cases were small lymphocytic lymphomas (SLLs), accounting for 18% (*n* = 52) of the total NHL patients. Other NHL subtypes included diffuse large B cell lymphoma (DLBCL) with 13% (*n* = 37), Burkitt lymphoma (BL) with 4% (*n* = 10), follicular lymphoma (FL) with 2% (*n* = 6), mantle cell lymphoma (MCL) with 1% (*n* = 3), T cell lymphoma (TCL) with 1% (*n* = 3), low grade lymphoma with 0.7% (*n* = 2), mucosa-associated lymphoid tissue lymphoma (MALT L) with 0.3% (*n* = 1), marginal zone B cell lymphoma (MZBCL) with 0.3% (*n* = 1), extranodal natural killer/T cell lymphoma (ENKTCL) with 0.3% (*n* = 1), and the remaining cases were NHL subtypes not otherwise classified(*n* = 81.28%) **(Figure 4.1).**

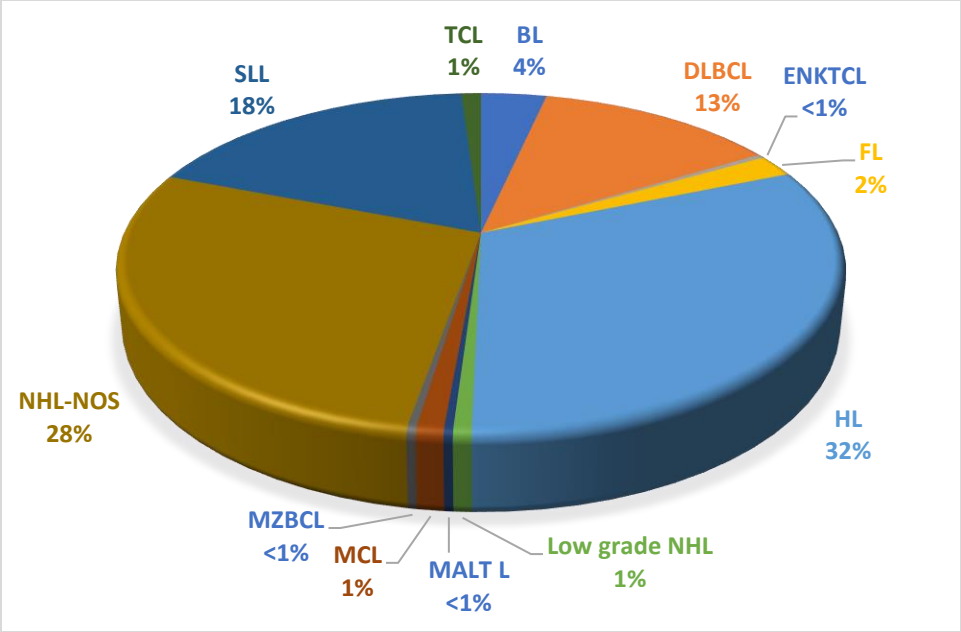
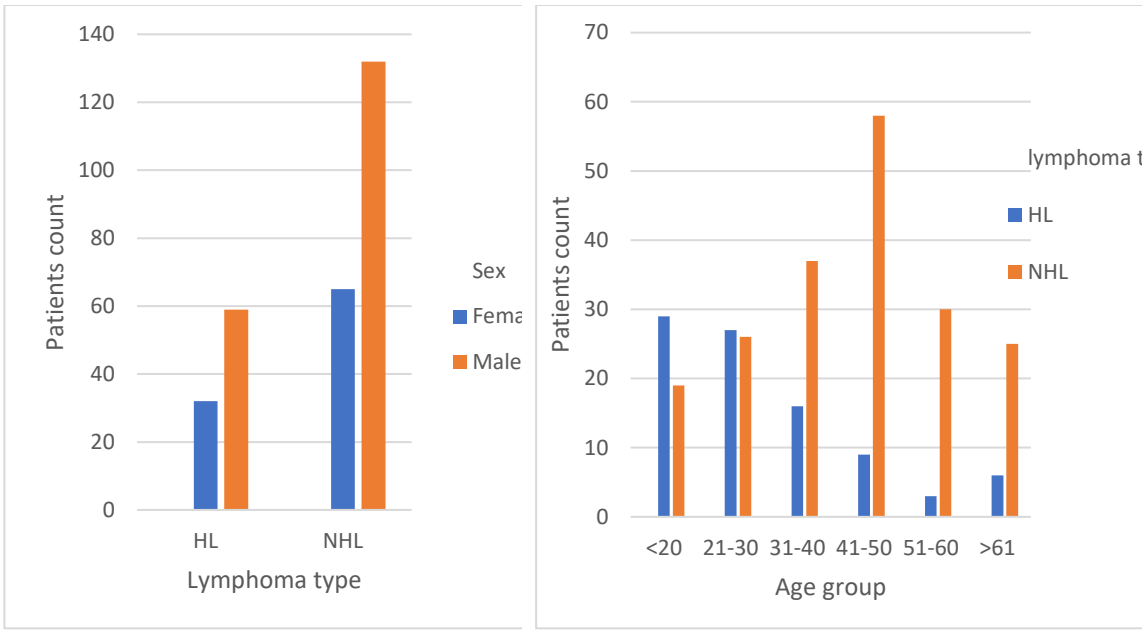


Figure 4.1: Histologic subtypes of lymphoma in all study participants (*n* = 288). Abbreviations: DLBCL: diffuse large B cell lymphoma; BL: Burkitt lymphoma; HL: Hodgkin's lymphoma; NHL: non-Hodgkin's lymphoma—not otherwise specified; SLL: small lymphocytic lymphoma; FL: follicular lymphoma, MCL: mantle cell lymphoma; MALT L: mucosa-associated lymphoid tissue lymphoma; MZBCL: marginal zone B cell lymphoma; ENKTCL: extranodal natural killer/T cell lymphoma; TCL: T cell lymphoma.

There was a higher prevalence of male patients in both HL and NHL groups. Among the 91 HL patients, 64.8% (*n* = 59) were male, while among the 197 NHL patients, 67% (*n* = 132) were male **(Figure 4.2a).** HL was more predominant in the age group younger than 20 years old, accounting for 60.4% (*n* = 29) of cases within that age range. On the other hand, a higher proportion of NHL patients were found in the age group of 51–60 years old, comprising 91% (*n* = 30) of cases within this age group **(Figure 4.2b).** Out of the total HL patients, 9.9% (*n* = 9) were found to be HIV-positive. Similarly, among the NHL patients, 8.6% (*n* = 20) were HIV-positive. Among the NHL patients who were HIV-positive, 45% (*n* = 9) of them were diagnosed with DLBCL.



(**a**) (**b**)

Figure 4.2: Distribution of sex and age groups with the lymphoma types. (**a**) The distribution of different lymphoma types with sex; (**b**) the distribution of different lymphoma types with age group. Abbreviations: HL: Hodgkin's lymphoma; NHL: non-Hodgkin's lymphoma.

4.3. Seroprevalence of EBV

A total of 164 serum samples were collected from both suspected and confirmed lymphoma patients to test for EBV VCA IgG. Among these samples, 150 were subjected to serological analysis (49 from HL patients and 101 from NHL patients). The remaining 14 samples were excluded from further analysis due to non-lymphoma reports obtained from pathologies of suspected patients cohort groups.

Out of the 150 serum samples analyzed, 96% (n = 144) of the study participants were tested as positive for the EBV IgG antibody. However, 3.3% (n = 5) tested negative and 0.07% (n = 1) showed an intermediate result for EBV IgG antibodies, respectively. From the different lymphoma types, EBV IgG was detected in the serum of 98% (n = 48) of HL patients and 95% (n = 96) of NHL patients. Among the different types of NHL, it was

observed that some patients with T cell lymphoma (TCL), mantle cell lymphoma (MCL), and small lymphocytic lymphoma (SLL) were EBV seronegative. Specifically, out of the total number of TCL patients, one out of two individuals were found to be EBV seronegative. Similarly, among MCL patients, one out of three patients lacked EBV seropositivity. Among SLL patients, 2 out of 13 individuals were EBV seronegative. There was no statistical significance between the HL and NHL types and EBV serology was observed (*p*-value = 0.64) **(Figure 4.3).**

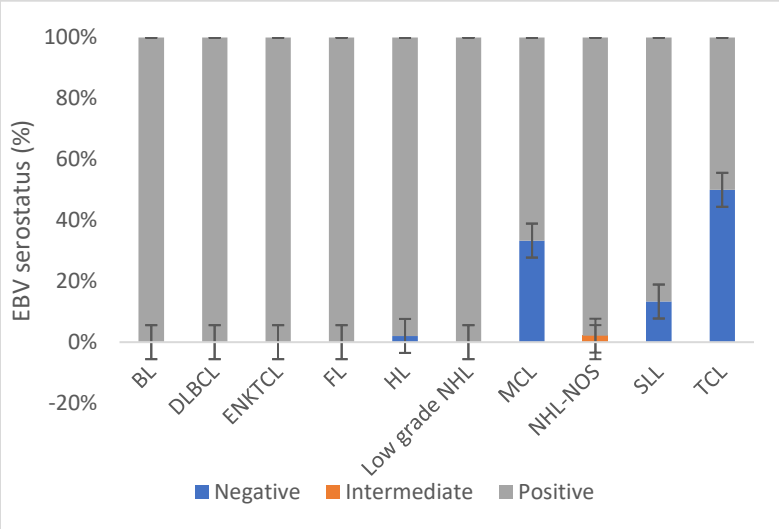


Figure 4.3: Serological detection of the EBV among different lymphoma types in Ethiopia. Abbreviations: DLBCL: diffuse large B cell lymphoma; BL: Burkitt lymphoma; HL: Hodgkin's lymphoma; NHL-NOS: non-Hodgkin's lymphoma—not otherwise classified; SLL: small lymphocytic lymphoma; FL: follicular lymphoma, MCL: mantle cell lymphoma, ENKTCL: extranodal natural killer/T cell lymphoma; TCL: T cell lymphoma; Low grade NHL: Low grade non-Hodgkin's lymphoma.

Regarding gender distribution, 95.8% of male participants and 96.3% of female participants were seropositive for the EBV and there was no statistically significant association between sex and EBV serology (*p*-value = 0.31). Among the 28 HIV-positive study participants, 26 had their EBV serology tested, and all of them (100%) were found to be EBV seropositive. Two serum samples were not collected from HIV-positive patients from the suspected groups due to unwillingness to give their blood sample. Among the 106 HIV-negative patients, 94.3% (*n* = 100) were seropositive for the EBV. All 18 patients

who were unaware of their HIV status were seropositive for the EBV and there was no significant association between HIV status and EBV serological positivity (*p*-value = 0.63). **Table 4.2** summarizes the serological findings related to EBV infection within various datasets.

Table 4.2: Associations of sociodemographic and clinical data with EBV serology.

Characteristics (*n* = 150)	**VCA IgG Negative (*n*)**	**VCA IgG Intermediate (*n*)**	**VCA IgG Positive (*n*)**	**Total (*n*)**	***p*-Value**
Age					0.56
<20 years	1	0	15	16	
21–30 years	0	0	37	37	
31–40 years	0	1	32	33	
41–50 years	2	0	36	38	
51–60 years	1	0	12	13	
>60 years	1	0	12	13	
Sex					0.31
Male	4	0	92	96	
Female	1	1	52	54	
HIV status					0.63
HIV positive	0	0	26	26	
HIV negative	5	1	100	106	
Unknown	0	0	0	18	
Lymphoma type					0.64
Hodgkin's lymphoma	1	0	48	49	
Non-Hodgkin's lymphoma	4	1	96	101	

4.4. Molecular detection and viral load of EBV DNA

Among the 288 lymphoma patients, the EBV was detected in 99% (*n* = 285) of the cases. The quantification of EBV copies/mL ranged from 10^2 to 10^9 copies/mL, with a median viral load count of approximately 1.1×10^3 EBV copies/ml. Based on the limit of detection of the assay and using the negative control, the EBV viral copy numbers were categorized into three groups: high, low, and very low. A high EBV viral load (>10,000 EBV copies/mL) was observed in 56.3% (*n* = 162) of the study participants, while 16% (*n* = 46) had a low EBV viral load (5000–10,000 EBV copies/mL) and 27.8% (*n* = 80) had a very low EBV viral load (<5000 EBV copies/mL).

When analyzing different lymphoma types, a high EBV viral load was found in 59.3% (*n* = 54) of Hodgkin's lymphoma patients and 54.8% (*n* = 108) of non-Hodgkin's lymphoma patients. Among NHL patients, those with DLBCL had the highest prevalence of high EBV viral loads, with 67.6% (*n*= 25) of the 37 DLBCL patients exhibiting high EBV viral loads. This was followed by BL, with 50% (*n* = 5) of 10 cases, SLL with 46% (*n* = 24) of 52 cases, and FL with 33.3% (*n* = 2) of 6 cases. However, no significant association was observed between different lymphoma types and EBV viral loads (*p*-value = 0.11).

In terms of sample types, tissue biopsy samples collected from suspected lymphoma patients showed a higher EBV viral load compared to retrospectively collected FFPE blocks and blood samples obtained from confirmed lymphoma patients. Among the samples, 94.3% of DNA extracted from LMNCs had a high EBV viral load. However, 72.5% of the DNA extracted from PBMC samples and 31.6% of DNA from FFPE lymphoma block samples exhibited a high EBV viral load (>10,000 EBV copies/mL). There was a significant association between different sample types and EBV viral loads (*p*-value < 0.001). Specifically, DNA isolated from lymph node biopsies had a higher EBV viral copy number compared to DNA samples extracted from blood and FFPE block samples.

The analysis also revealed that high EBV viral copy numbers were detected in 56% (*n* = 107) of male participants and 56.7% (*n* = 55) of female participants. However, no significant association between sex and EBV viral load was found (*p*-value = 0.99).

Furthermore, among the different age groups, 73.5% (*n* = 39) of individuals in the young age group (21–30 years) exhibited a higher EBV copies/mL value compared to the other age groups. Nevertheless, there was no significant association between age groups and EBV viral loads (*p*-value = 0.49). Among the 28 lymphoma patients who were HIV-positive, 85.7% (*n* = 24) had a high EBV viral load. Similarly, among the 109 lymphoma patients who were HIV-negative, approximately 73.4% (*n* = 80) had a high EBV viral load. There was a significant association between HIV positivity in lymphoma patients and a higher EBV viral load (*p*-value < 0.01) **(Table 4.3).**

Table 4.3: The distribution of EBV viral load among lymphoma patients in Ethiopia.

Characteristics	**<5000 EBV Copies/mL**	**5000–10,000 EBV Copies/mL**	**≥10,000 EBV Copies/mL**	**Total**	***p*-Value**
Study participants (*n* = 288)					<0.01
FFPE blocks	77 (57.8%)	14 (10.5%)	42 (31.6%)	133	
Lymphoma patients	3 (2.5%)	30 (25%)	87 (72.5%)	120	
Suspected lymphoma patients	0 (0%)	2 (5.7%	33 (94.3%)	35	
Age					0.78
<20 years	18 (37.5%)	9 (18.7%)	21 (43.8%)	48	
21–30 years	8 (15.1%)	6 (11.3%)	39 (73.6%)	53	
31–40 years	15 (28.3%)	11 (20.7%)	27 (51%)	53	
41–50 years	16 (23.9%)	7 (10.4%)	44 (65.7%)	67	
51–60 years	12 (36.4%)	7 (21.2%)	14 (42.4%)	33	
>60 years	10 (32.2%)	6 (19.3%)	15 (48.4%)	31	
Not applicable	1 (33.3%)	0 (0%)	2 (66.7%)	3	
Sex					0.97
Male	53 (27.7%)	31 (16.2%)	107 (56%)	191	
Female	27 (27.8%)	15 (15.5%)	55 (56.7%	97	
HIV status					<0.01

HIV positive	0 (0%)	4 (14.3%)	24 (85.7%)	28	
HIV negative	3 (2.7%)	26 (23.9%)	80 (73.4%)	109	
Unknown	77 (51%)	16 (10.6%)	58 (38.4%)	151	
Lymphoma type					0.65
Hodgkin's lymphoma	22 (24.2%)	15 (16.5%)	54 (59.3%)	91	
Non-Hodgkin's lymphoma	58 (29.4%)	31 (15.7%)	108 (54.8%)	197	

4.5. EBV genotypes among lymphoma patients

4.5.1. Sociodemographic and clinical characteristics

A total of *n* = 207 DNA samples was collected from FFPE, PBMCs, and LMNCs for EBV detection and subtyping. A total of 63 (30.4%) DNA samples were isolated from FFPE lymphoma blocks, 118 (57%) from PBMCs, and 26 (12.6%) of the DNA samples were extracted from LMNCs. Most of the study participants were male patients (n = 135, 65.2%). The study participants age ranged from 3 to 80 years, with a mean age of 38.91 ± 16.94 SD years. The overall clinical and sociodemographic data of the study participants are listed in **Table 4.4.**

Table 4.4: Clinical and sociodemographic profile of the study participants for EBV typing (*n*=207)

Characteristics	**Number**	**Percent**
Study participants (n = 207)		
FFPE blocks	63	30.4
Lymphoma patients	144	69.6
Sex		
Male	135	65.2
Female	72	34.8
Age		
<20 years	32	15.5

21–40 years	79	38.2
41–60 years	75	36.2
>61 years	20	9.7
Not applicable	1	0.5
HIV status		
Positive	28	13.5
Negative	98	47.3
Unknown	81	39.2
Lymphoma types		
Hodgkin’s lymphoma	70	33.8
Non-Hodgkin's lymphoma	137	66.2

Out of the 137 non-Hodgkin’s lymphoma (NHL) study participants, the majority (*n* = 35, 25.5%) were small lymphocytic lymphoma (SLL) patients, followed by diffuse large B cell lymphoma patients (*n* = 33, 24.1%). The distribution of the different lymphoma types in the study participants showed in **Figure 4.4**. Male patients comprised 65.7% (*n* = 46,) and 65% (*n* = 89) of HL and NHL patients, respectively. The predominant age group in HL patients was the younger population. A total of 32 (45.7%) of HL patients were categorized between the age of 20 and 40. However, in NHL patients, the majority (*n* = 63, 46.3%) of patients were categorized under the age group between 40 and 60 years. Among NHL patients, HIV prevalence was higher (*n* = 9, 27.3%) in DLBCL patients than in the other NHL types.

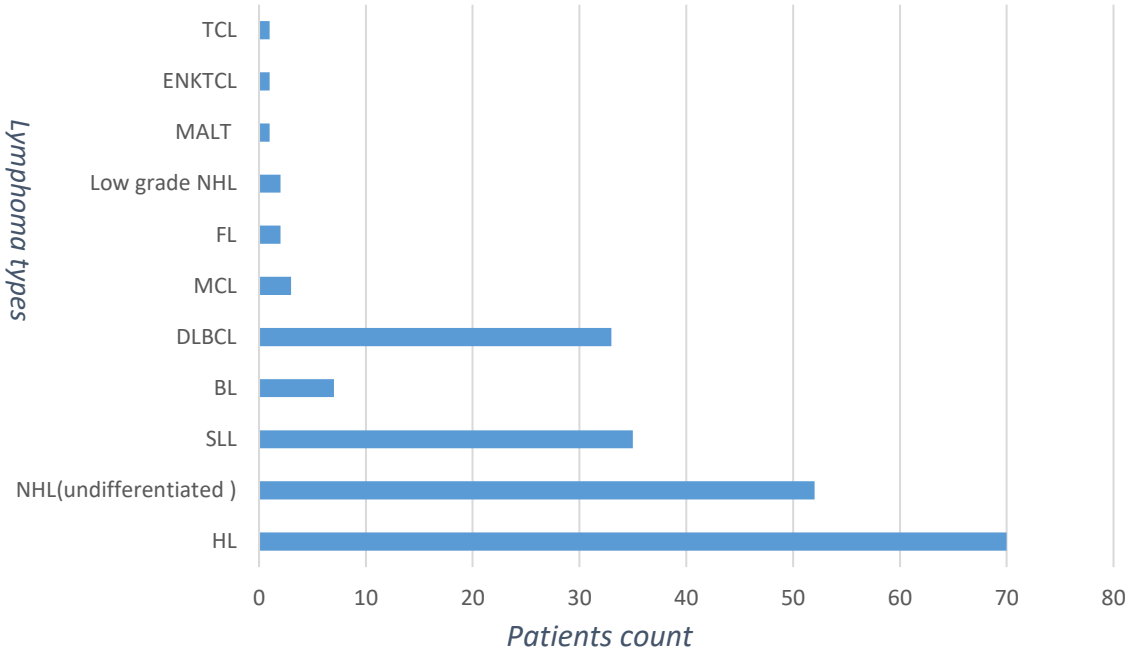


Figure 4.4: Distribution of lymphoma types among the study participants (*n* = 207). Abbreviations: DLBCL: diffuse large B cell lymphoma; BL: Burkitt lymphoma; HL: Hodgkin's lymphoma; NHL: non-Hodgkin's lymphoma; SLL: small lymphocytic lymphoma; FL: follicular lymphoma; MCL: mantle cell lymphoma; ENKTCL: extranodal natural killer/T-cell lymphoma.; MALT: mucosa-associated lymphoid tissue lymphoma.

4.5.2. EBV genotyping using EBNA3C gene typing

Out of the 207 lymphoma patients, 52.2% (*n* = 108) had EBV genotype 1, while 79 (38.2%) had EBV genotype 2, and 20 (9.7%) were coinfected with both genotypes **(Figure 4.5).**

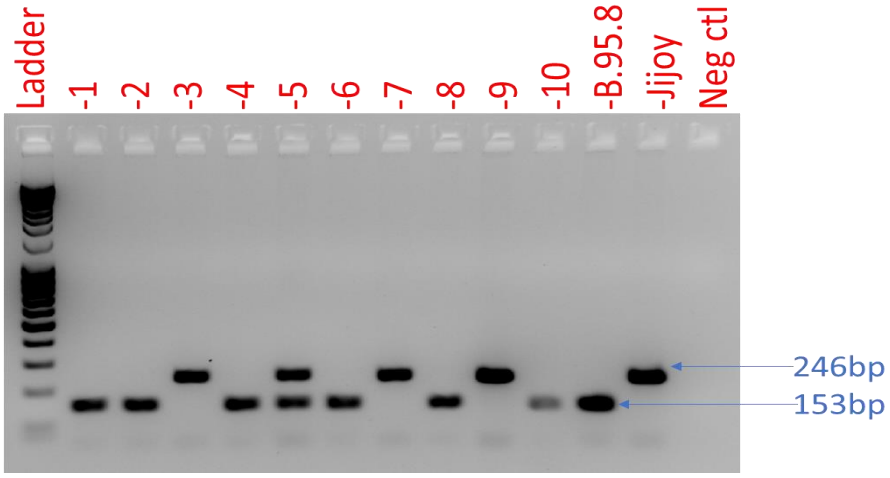


Figure 4.5: Representation of EBV genotyping using *EBNA3C* gene typing using conventional PCR. Numbers 1 to 10 refer to examples for patient sample PCRs, B95.8 refers a positive control for EBV genotype I, Jijoy refers to positive control for EBV genotype 2, and Neg Ctl refers to EBV negative cell line used as control.

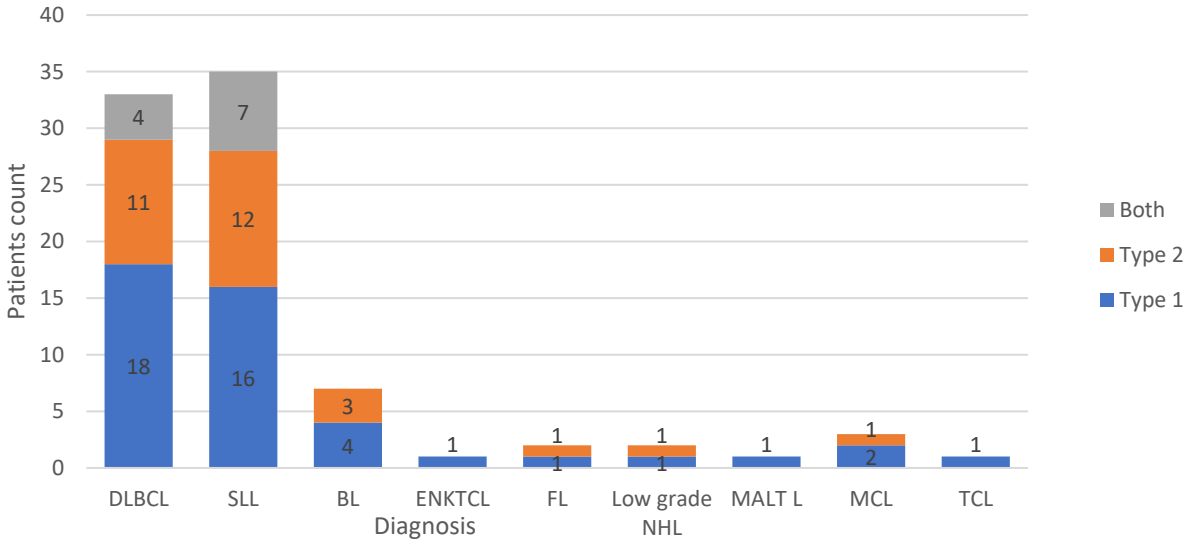
Next, we assessed if the type of EBV infection varied per sex, HIV status, and type of lymphoma. Our results showed no difference between males and females in the type of EBV genotype infection. However, a significant association between the age groups and EBV genotypes was observed ($p = 0.027$). EBV genotype 1 was higher in age groups less than 20 years ($n = 20/32$, 62.5%), while EBV genotype 2 was higher within the age group of 21 to 40 ($n = 37/79$, 46.8%). Among HIV-positive and EBV-positive lymphoma patients ($n = 28$), EBV genotype 1, genotype 2, and coinfection was detected in 13 (46.4%), 12 (42.8%), and 3 (10.7%) of HIV-positive lymphoma patients, respectively **(Table 4.5).**

Table 4.5: Distributions of EBV genotypes across sex, HIV status, and disease phenotype.

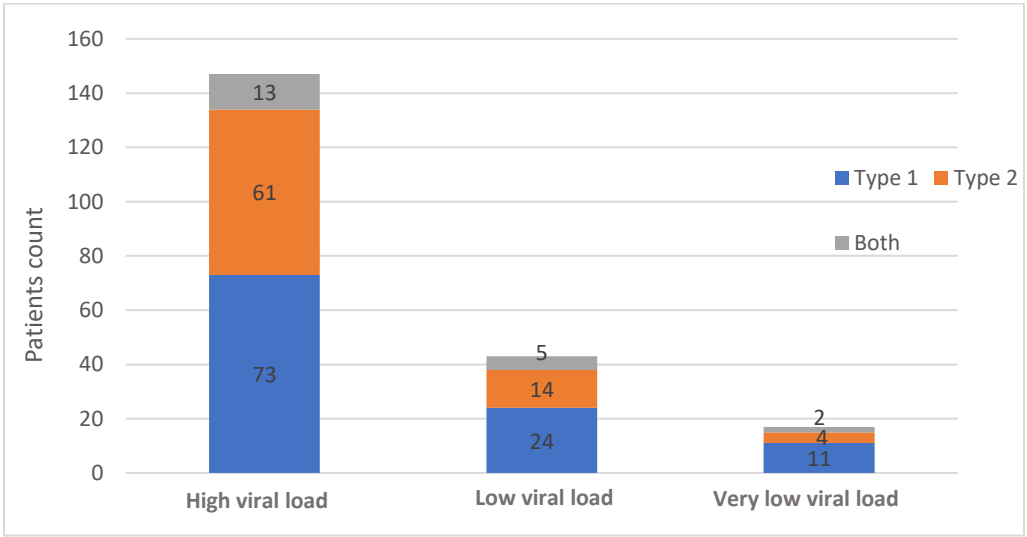
Characteristics	EBV1	EBV2	Coinfections	Total	*p*-value
Age					
<20 years	20 (62.5%)	11 (34.4%)	1 (3.1%)	32	
21–40 years	39 (49.4%)	37 (46.8%)	3 (3.8%)	79	
41–60 years	36 (48%)	25 (33.3%)	14 (18.7%)	75	0.027
>60 years	12 (60%)	6 (30%)	2 (10%)	20	
Not applicable	1			1	
Sex					
Male	70 (51.8%)	52 (38.5%)	13(9.6%)	135	0.99
Female	38 (52.8%)	27 (37.5%)	7(9.7%)	72	
HIV status					
HIV positive	13 (46.4%)	12 (42.9%)	3 (10.7%)	28	
HIV negative	45 (45.9%)	42 (42.9%)	11 (11.2%)	98	0.169
Not applicable	50 (61.7%)	25 (30.9%)	6 (7.4%)	81	
Lymphoma type					
HL	37 (52.9%)	30 (42.6%)	3 (4.3%)	70	0.15
NHL	71 (51.8%)	49 (35.7%	17 (12.4%)	137	

We next assed if the EBV genotype varies with the clinical form of the lymphoma disease. Of the n = 137 NHL types, SLL (n = 35, 25.5%) and DLBCL (n = 33, 24.1%) comprise most of the study population. From n = 35 SLL cases, EBV genotype 1 was detected in n = 16 (45.7%), EBV genotype 2 was detected in n = 12 (34.3%), and n = 7 (20%) represented coinfection with both genotypes. Of the n = 33 DLBCL cases, EBV genotype 1, EBV genotype 2, and both types were detected in n = 18 (54.5%), n = 11 (33.3%), and n = 4 (12%) patients, respectively. There was no significant association between the different NHL types and the distribution of EBV genotypes (p value = 0.92). The overall genotype

distribution of EBV among different NHL types is indicated in **figure 4.6a**. Of the samples with high EBV viral load (*n* = 147), EBV type 1 was detected in *n* = 73 (49.7%), EBV type 2 in n = 61 (41.5%), and both types in *n* = 13 (8.8%). There was no significant association between EBV viral load and EBV genotypes (*p* value = 0.57). **(Figure 4.6b).**



(A)



(A)

Figure 4.6: The magnitude of EBV genotypes. (**A**) EBV genotype distribution within different non-Hodgkin's lymphoma types; (**B**) EBV viral load status vs. EBV genotypes. Abbreviations: DLBCL: diffuse large B cell lymphoma; BL: Burkitt lymphoma; NHL: non-Hodgkin's lymphoma; SLL: small lymphocytic lymphoma; FL: follicular lymphoma, MCL: mantle cell lymphoma, ENKTCL: extranodal natural killer/T-cell lymphoma. MALT L: mucosa-associated lymphoid tissue lymphoma.

4.6. Methylation profiles of EBV DNA

4.6.1. Sociodemographic and clinical characteristics

A total of 115 FFPE and plasma samples had enough methylation data for clustering. Out of these 79 of them were FFPE samples collected from first cohort (2021) and the rest 36 samples were FFPE and plasma samples collected during the second cohorts. 65.2% (n=75) were male patients' majority (n=35, 30.4%) of the study participants were categorized within the age group of 41-60 years. HL patient accounts 16.5% (n=19) and NHL patients accounts 83.5% (n=96) of the study participants. Out of n=96 NHL patients, SLL accounts 28.1% (n=27) followed by DLBCL (n=24, 25%). The overall clinical and sociodemographic characteristics of the study participants were showed in **table 4.6**.

Table 4.6: Sociodemographic and clinical characteristics of the study participants (n=115)

Characteristics	**Number**	**Percent**
Study cohorts		
First batch (2021)	79	68.7%
Second batch (2022)	36	31.3%
Age		
<20 years	19	16.5%
21-40 years	34	29.6%
41-60 years	35	30.4%
>60 years	26	22.6%

Not applicable	1	0.9%
Sex		
Male	75	65.2%
Female	40	34.8%
Lymphoma type		
HL	19	16.5%
NHL	96	83.5%
NHL type		
BL	5	5.2%
DLBCL	24	25%
ENKTCL	1	1%
FL	12	12.5%
MCL	7	7.3%
SLL	27	28.1%
Other NHL types	20	20.8%

4.6.2. EBV status

We performed EBER-ISH for the first 79 FFPE batch of samples. Of these, 6.3%(n=5) of the tested samples showed EBV positivity confirming EBV associated lymphoma. Whereas, 15.2 %(n=12) of the samples had rare or few EBER positive cells which was concluded to be due to latent EBV infection

(Figure 4.7).

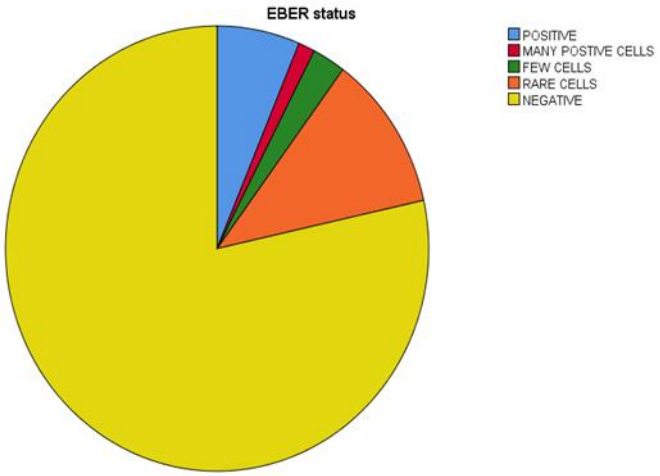
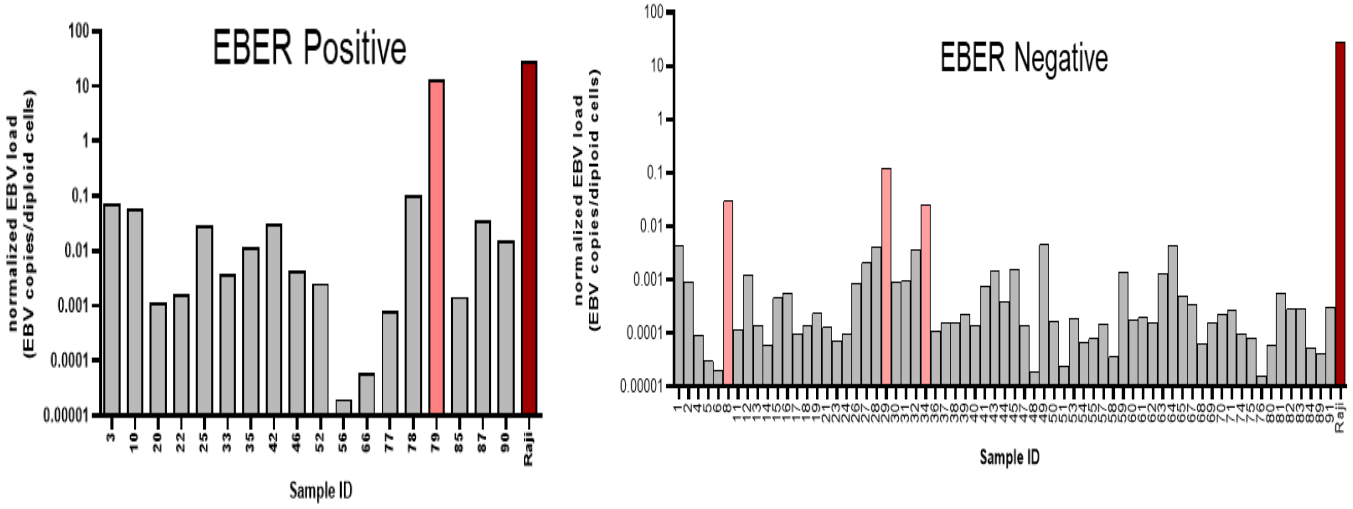
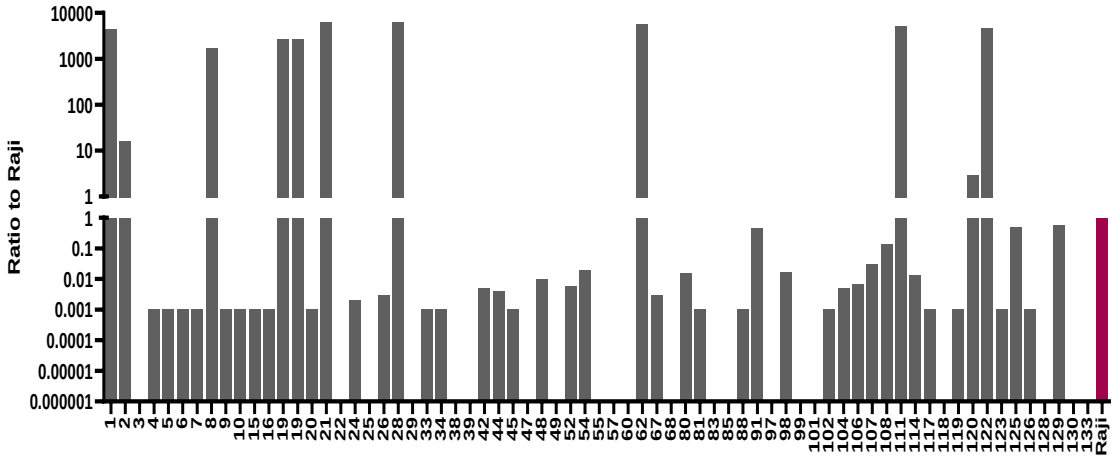


Figure 4.7: EBV positivity among FFPE samples. **EBER status:** Five of EBV positive lymphoma were EBER positive; Twelve with rare to some EBV positive cells in the background; The rest were EBER negative

Furthermore, qPCR was employed to determine the EBV viral load in both batches of samples. The analysis involved comparing the number of EBV copies per diploid cell with the Raji cell line, which serves as a reference. This approach allowed us for the quantification and comparison of the viral load across the samples in relation to the Raji cell line, providing insights into the relative abundance of EBV within each batch **(Figure 4.8)**.



(A) **(B)**

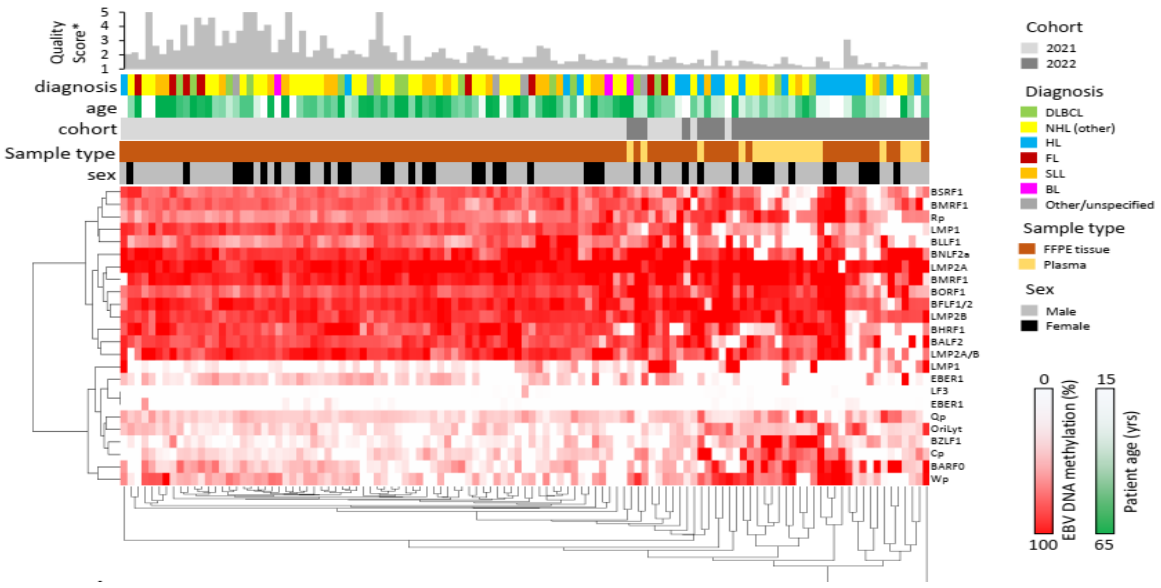


(C)

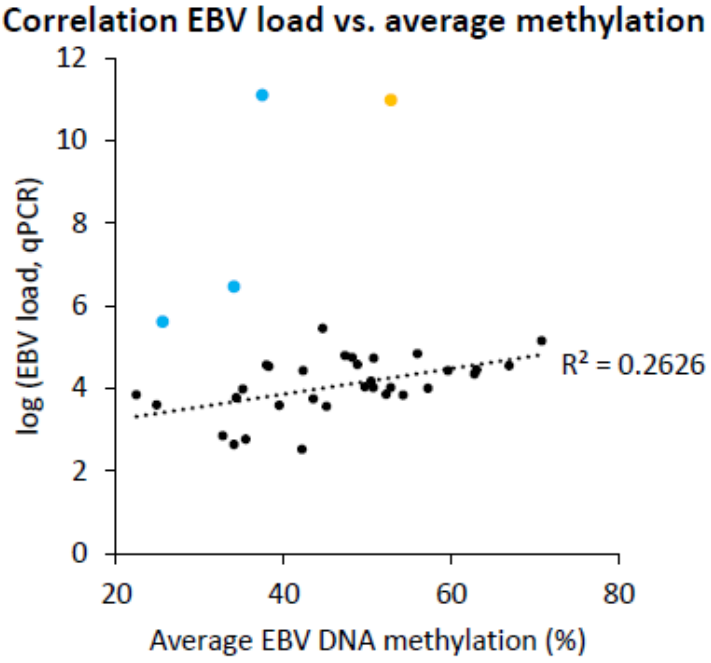
Figure 4.8: EBV viral load among different sample collection batch. **(A)** EBER positive group and **(B)** EBER negative group from the first batch. (**C)** EBV viral load status of the second batch.

4.6.3. CpG methylation of EBV genome

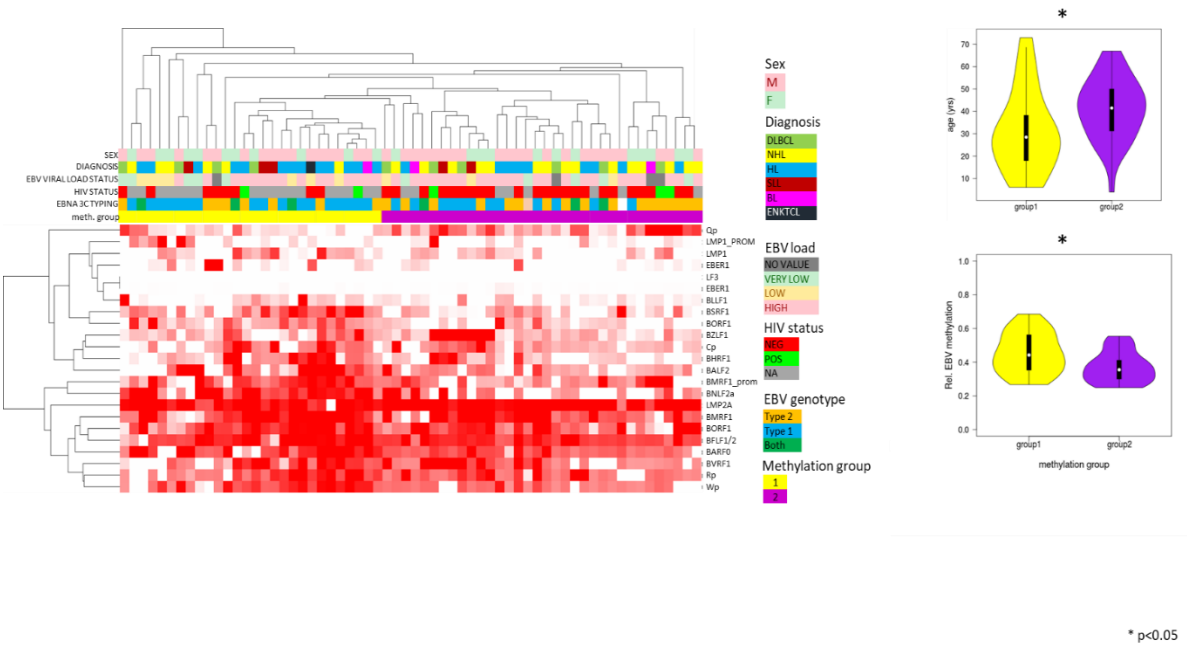
Of the different EBV genome methylation sites, we prefreed 24 EBV genes to assess the methlation status of the EBV genome based on the the quality of data generated during the analysis of the CpG cites. We excluded some CpG locus due to lower quality of data generated for comaring the individual cpGs. According to our analysis of the 24 refrence EBV genes the avarage methylation status of the lymphoma samples lie between 45% to 50%. Most of the samples had intermediate/polyclonal-like methylation and indicator of non-clonal EBV origin (i.e. non-tumor cell derived). Further analysis of the correlation of EBV viral load and EBV methylation suggests mostly non-lytic EBV release **(Figure 4.9)**.


Quality Score*
diagnosis
age
cohort
Sample type
sex
BSRF1
BMRF1
Rp
LMP1
BLLF1
BNLF2a
LMP2A
BMRF1
BORF1
BFLF1/2
LMP2B
BHRF1
BALF2
LMP2A/B
LMP1
EBER1
LF3
EBER1
Qp
OriLyt
BZLF1
Cp
BARF0
Wp
Cohort
2021
2022
Diagnosis
DLBCL
NHL (other)
HL
FL
SLL
BL
Other/unspecified
Sample type
FFPE tissue
Plasma
Sex
Male
Female
EBV DNA methylation (%)
Patient age (yrs)

(A)


Correlation EBV load vs. average methylation
log (EBV load, qPCR)
Average EBV DNA methylation (%)
R² = 0.2626

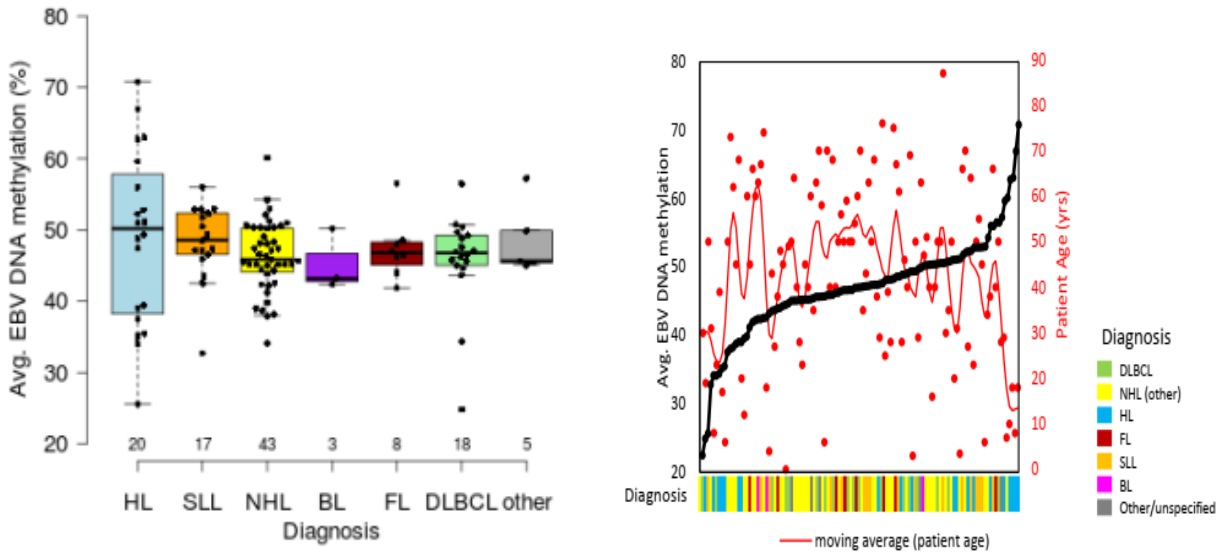
(B)



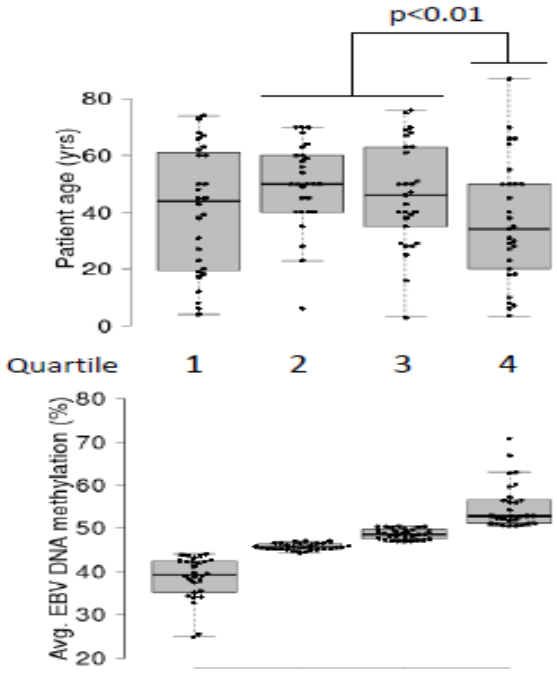
(C)

Figure 4.9 CpG methylation status of EBV genome from lymphoma patients in Ethiopia; **(A)** Methylation status of EBV vs different sociodemographic and clinical characteristics. **(B)** Average methylation profiles of the samples in association with EBV viral load (Outliers: mostly HL + SLL with low methylation, potentially lytic **(C)** Analysis of methylation status with the viral characteristics.

When we compare the methylation status, we observed a striking methylation heterogeneity in patient diagnosed with HL (low and high methylation subtypes). However similar CpG methylation status of EBV genes was observed among patients with NHL subtypes. Similar to the case of HL, extremely high and low methylation status was associated with patient age (**Figure 4.10**).



(A) **(B)**



(C)

Figure 4.10: Average CpG methylation status across different lymphoma types **(A)** and patients age **(B)**. The interquartile range of DNA methylation status among patient age **(C)**.

Inorder to assess the diffrences in in the EBV methylation status across the different factors such as sample batchs, sample types, lymphoma types and patient age, a t-SNE plot was created to visualize the methylation status of EBV samples. Each data point in the plot represents a sample, and the position of the data points reflects their similarity in terms of methylation patterns. Samples with similar methylation profiles are positioned closer to each other, while samples with dissimilar profiles are farther apart (**Figure 4.11**). The t-SNE plot was manually annotated to highlight the similarities and differences in CpG methylation among the samples (group 1-4). The analysis revealed a consistent methylation profile pattern across the 2021 cohorts, as well as in patients with Hodgkin's lymphoma belonging to group 1. Additionally, older patients displayed a similar methylation status, indicating a potential association between age and methylation patterns.

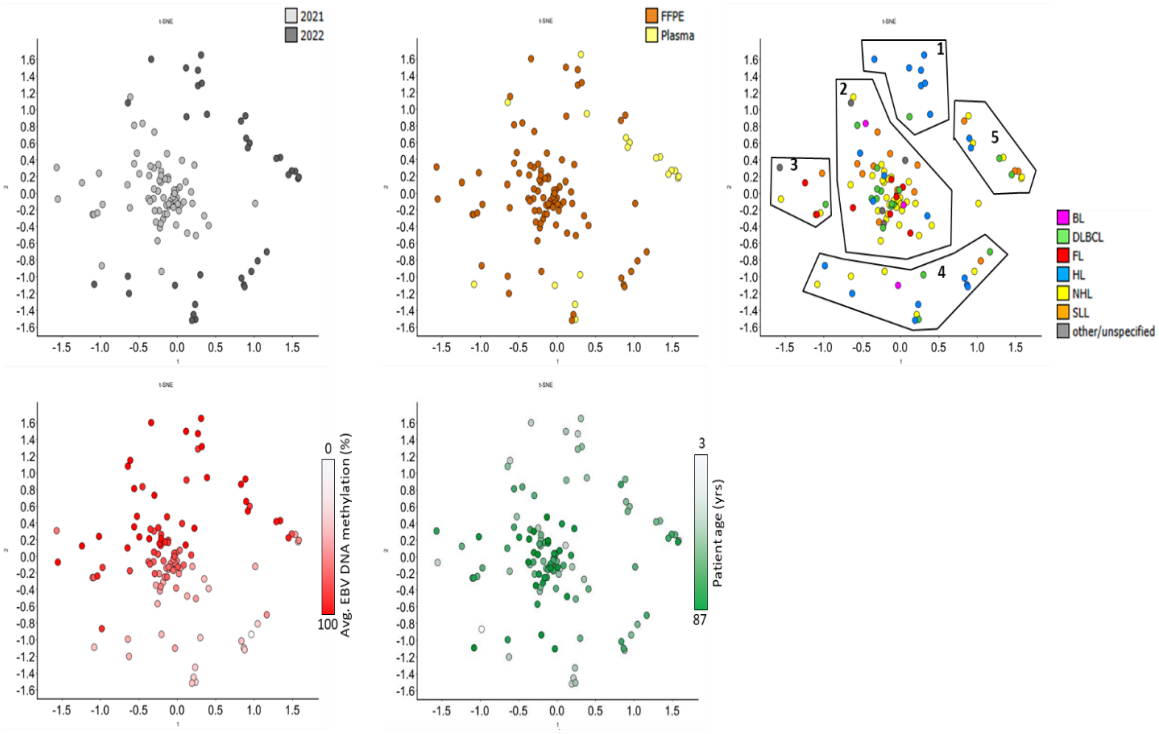


Figure 4.11: t-SNE plot for EBV methylation vs. other sample characteristics

4.7. EBV capture sequencing

4.7.1. Genomic characteristics

For EBV capture sequencing, a total of 110 EBV positive samples with a high viral load were selected. These selected samples exhibited a higher viral load, and their EBV quantity ratio with the Raji cell line was also determined. The minimum and maximum EBV viral loads observed in these samples were 7,892 EBV copies/mL and 4,037,445 EBV copies/mL, respectively. Among the samples, 58 (52.7%) were identified as EBV type 1, 40 (36.4%) as type 2, and the remaining 12 (10.9%) samples exhibited coinfection.

4.7.2. Quality control

Capture sequencing was conducted on all samples, and subsequent quality control was carried out based on the depth of coverage. Based on the quality control analysis conducted on the samples, it was found that out of the 110 samples subjected to capture sequencing, 30 of them exhibited low depth of coverage. These samples did not meet the required criteria for further bioinformatics and phylogenetic tree analysis and only 80 samples were used. To ensure accuracy, only reads with a mapping quality score of 20 or higher were included in the calculation of the depth of coverage during the quality control analysis. Co-infected samples were termed Type 1 (Co) and Type 2 (Co). These samples were aligned against both Type 1 and Type 2 variants of the virus. The mean depth of coverage for the samples were illustrated in **figure 4.12.**

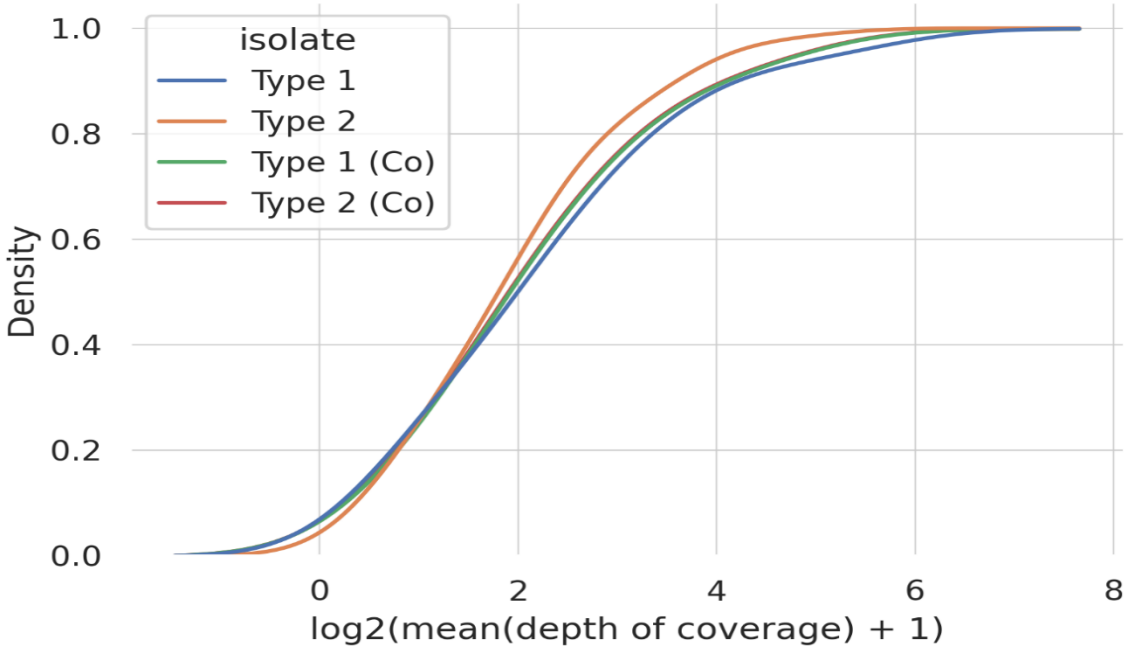


Figure 4.12: The mean depth of coverage across the samples

During the analysis of coinfected EBV samples, alignment was performed against both the EBV type 1 reference genome (NC_007605.1) and the EBV type 2 reference genome (NC_009334.1) to determine the depth of coverage for each EBV type. The analysis of the samples, including different lymphoma samples, revealed that the overall depth of coverage for EBV was low. The box plot displays the mean depth of coverage for each sample across conditions and EBV strains. For a given sample, the mean depth of coverage is calculated by averaging the number of reads aligned (minimum. mapping quality (MQ) score >= 20) at each genomic position **(Figure 4.13).**

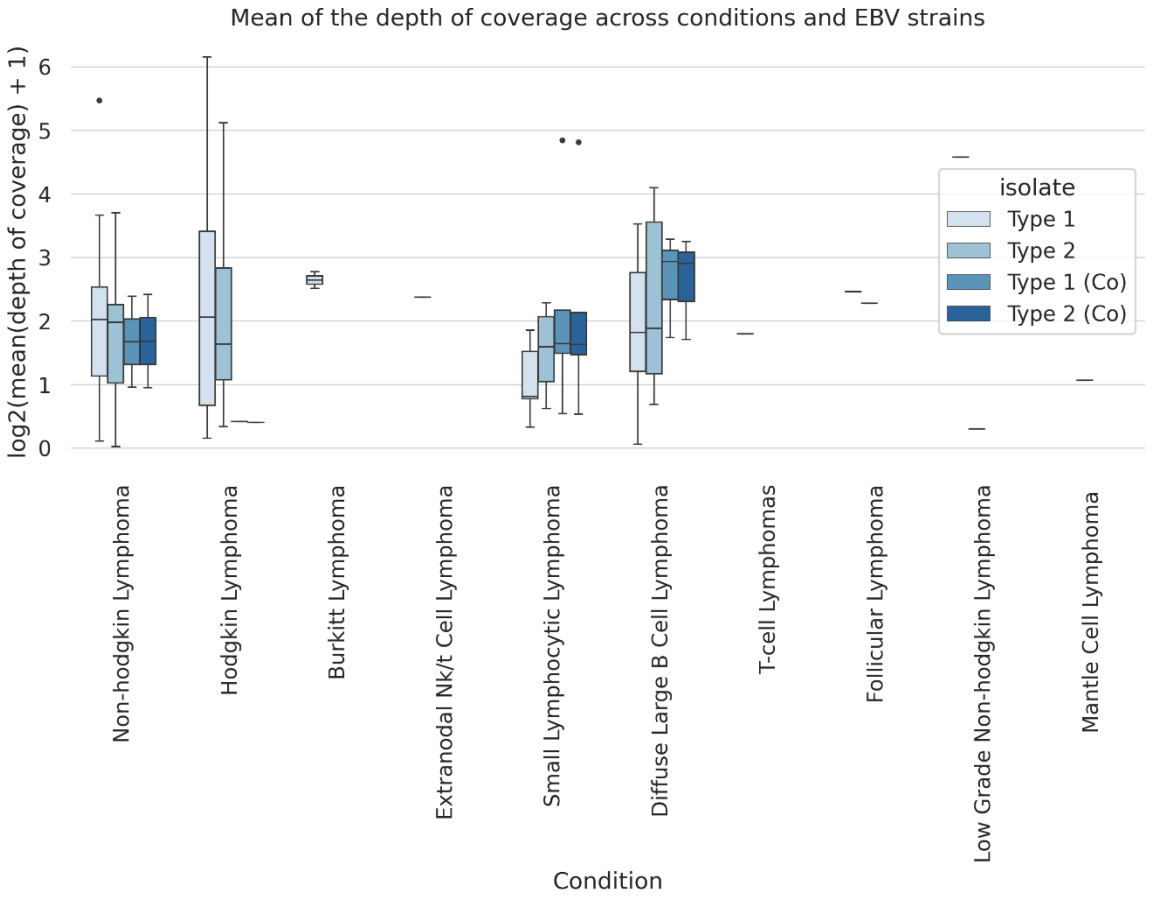


Figure 4.13: The mean depth of coverage across the different lymphoma types

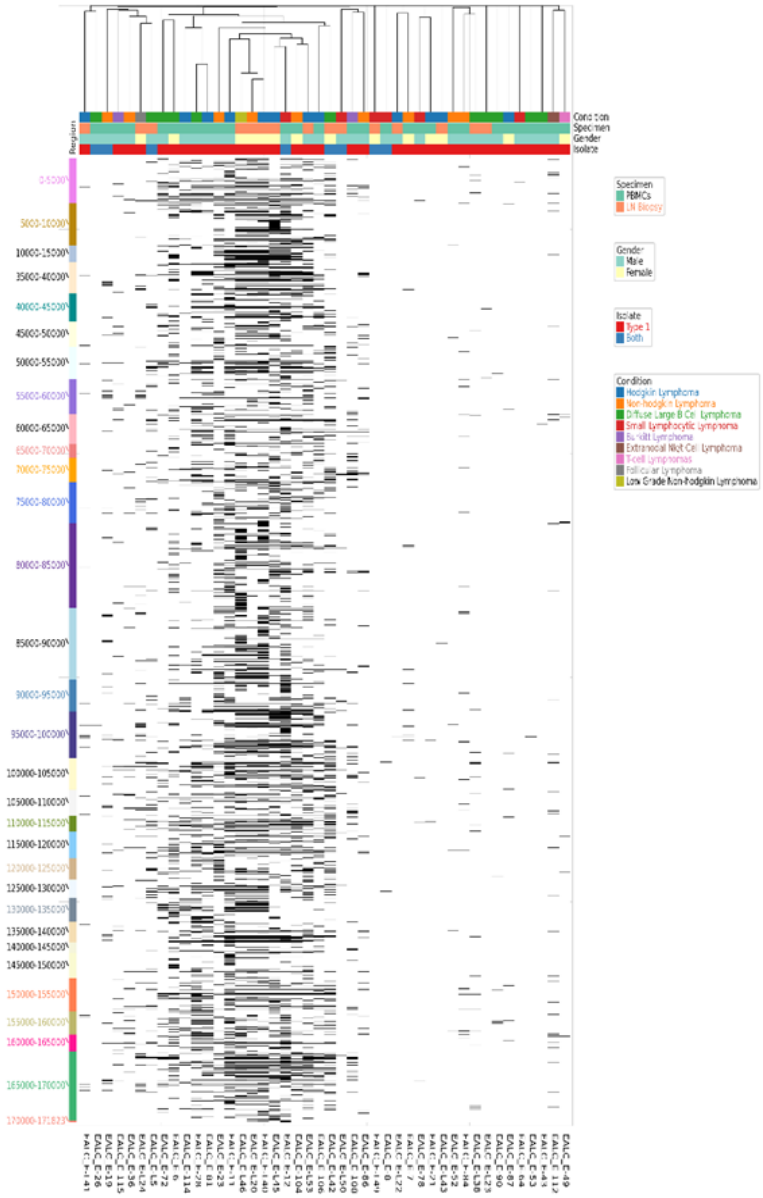
4.7.3. Variant calling

We have performed the variant allele frequency analysis for each sample, considering different allele frequency (AF) thresholds, including 0.01, 0.5, and 0.9. In this context, the values 0.01, 0.5, and 0.9 represent the proportions of aligned reads that support the alternative allele. Specifically, a value of 0.01 corresponds to 1% of the aligned reads supporting the alternative allele, a value of 0.5 indicates that 50% of the aligned reads support the alternative allele, and a value of 0.9 signifies that 90% of the aligned reads support the alternative allele.

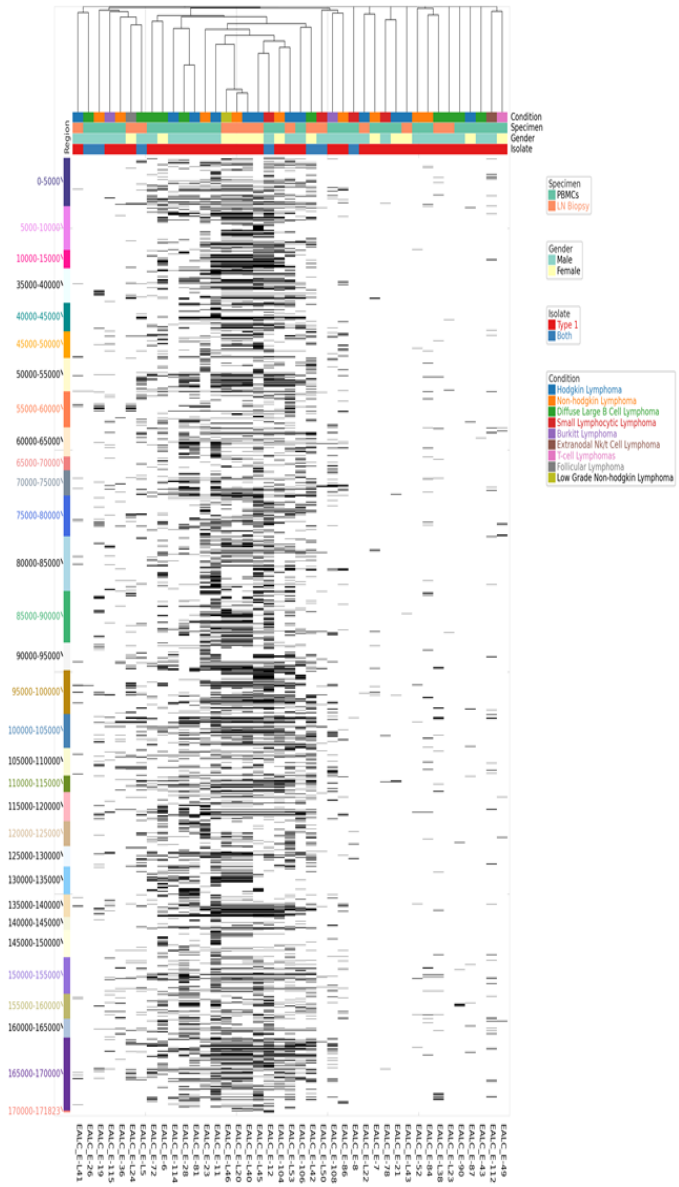
We generated heatmap plots to visualize the distribution of variants across samples, based on three different allele frequency (AF) thresholds: 0.01, 0.5, and 0.9. We utilized LoFreq, a variant calling tool, to identify variants in our samples. The following parameters were

employed during the variant calling process: Variants were considered only if the read depth (number of aligned reads at a given position) was equal to or greater than 10 reads, variants were included if they had a minimum of 5 reads supporting the alternative allele, variants were required to have a mapping quality of at least 20 and variants were considered only if the base quality (quality score of the base call) was at least 20. A minimum of 500-600 single nucleotide variants (SNVs) were identified per each sample.

The identified variants were represented in a binary format across samples, where the presence of a variant was denoted by 1 (black), and the absence of a variant was denoted by 0 (white) across the samples and conditions. To gain a better understanding of the relationships between samples based on their variant profiles, we performed hierarchical clustering. The samples (columns) in the analysis were clustered using a complete linkage approach and Jaccard distance metric. The variants (rows) were organized into groups or windows of 5,000 nucleotides. The reported information for each sample includes the specimen, gender, isolate, and condition. In the analysis, only samples that had at least one variant were included in the reporting. Variants were reported based on a minimum allele frequency threshold of >=0.01(**Figure 4.14 A and B**).


Condition
Specimen
Gender
Isolate
Region
Specimen
PBMCs
LN Biopsy
Gender
Male
Female
Isolate
Type 1
Both
Condition
Hodgkin Lymphoma
Non-hodgkin Lymphoma
Diffuse Large B Cell Lymphoma
Small Lymphocytic Lymphoma
Burkitt Lymphoma
Extranodal NkT Cell Lymphoma
T-cell Lymphomas
Follicular Lymphoma
Low Grade Non-hodgkin Lymphoma

AF = 0.01


Condition
Specimen
Gender
Isolate
Region
Specimen
PBMCs
LN Biopsy
Gender
Male
Female
Isolate
Type 1
Both
Condition
Hodgkin Lymphoma
Non-hodgkin Lymphoma
Diffuse Large B Cell Lymphoma
Small Lymphocytic Lymphoma
Burkitt Lymphoma
Extranodal NkT Cell Lymphoma
T-cell Lymphomas
Follicular Lymphoma
Low Grade Non-hodgkin Lymphoma

AF = 0.5

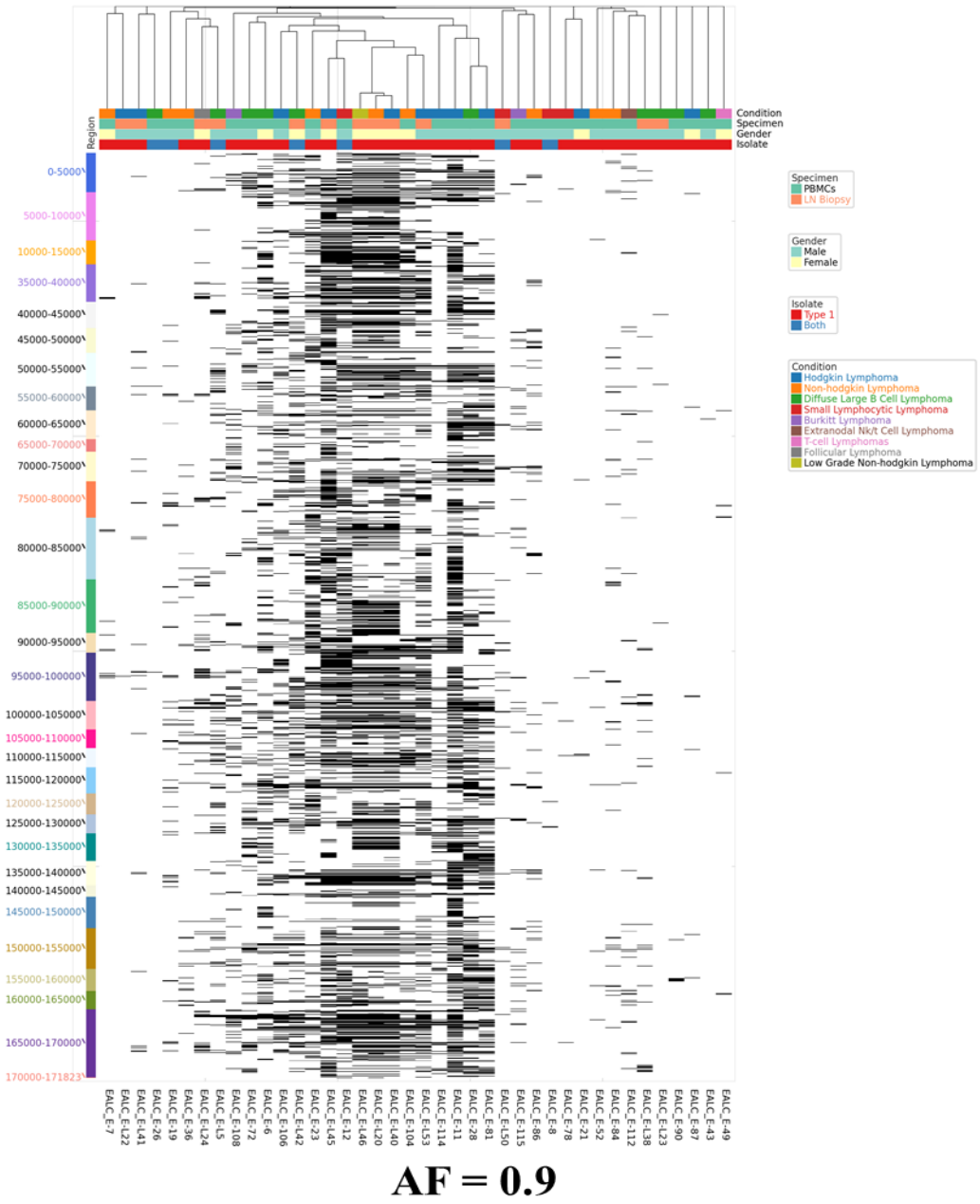


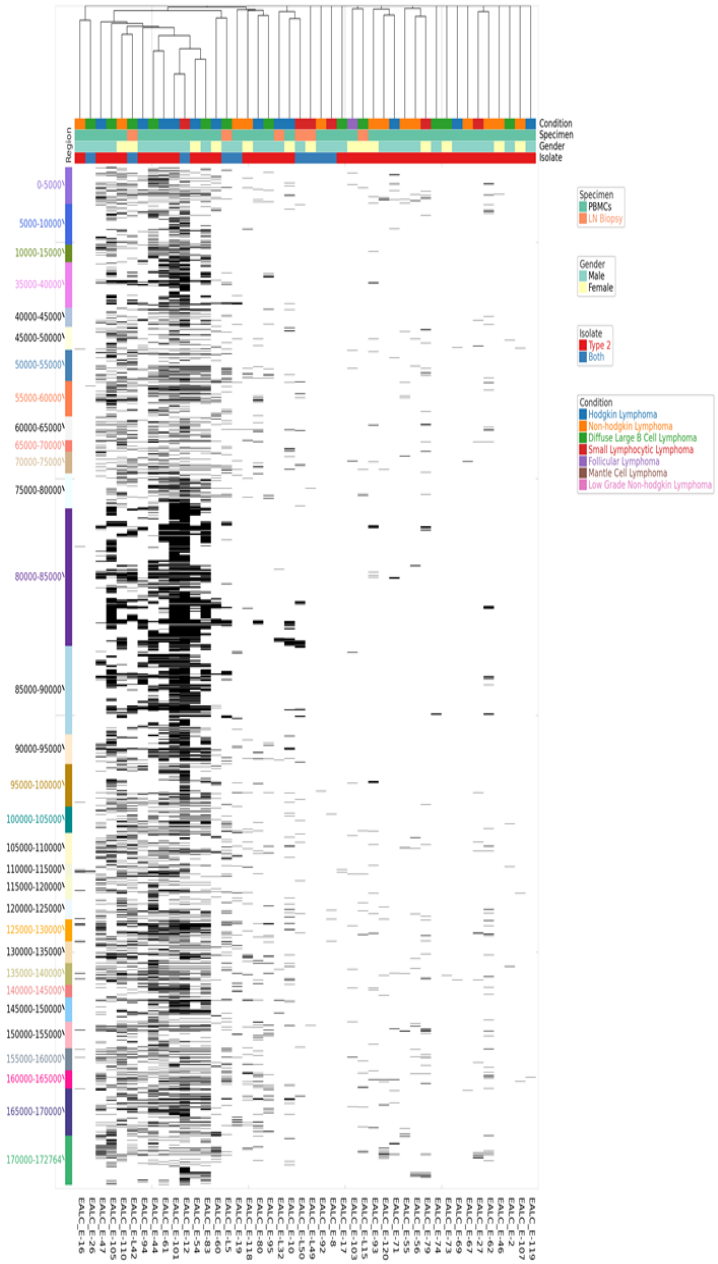
Figure 4.14A: Heatmap plot visualizing variants across EBV Type 1 at different allele frequencies: 0.01 (1% aligned reads supporting the alternative allele), 0.5 (50% aligned reads supporting the alternative allele), and 0.9 (90% aligned reads supporting the alternative allele, almost monoclonal).

Out of the 44 samples aligned for EBV type l, 45% (n=20) exhibited a gene variation exceeding 50% from the reference gene, indicating a substantial level of genetic diversity. This subset represents the highest degree of gene variation within the studied population. Approximately 23% (n=10) of the samples displayed a gene variation ranging between 25% and 50%, indicating a moderate level of genetic variation compared to the reference gene. The remaining 27% (n=12) of the samples had a gene variation of less than 25%, signifying relatively minor genetic variations in comparison to the reference gene.

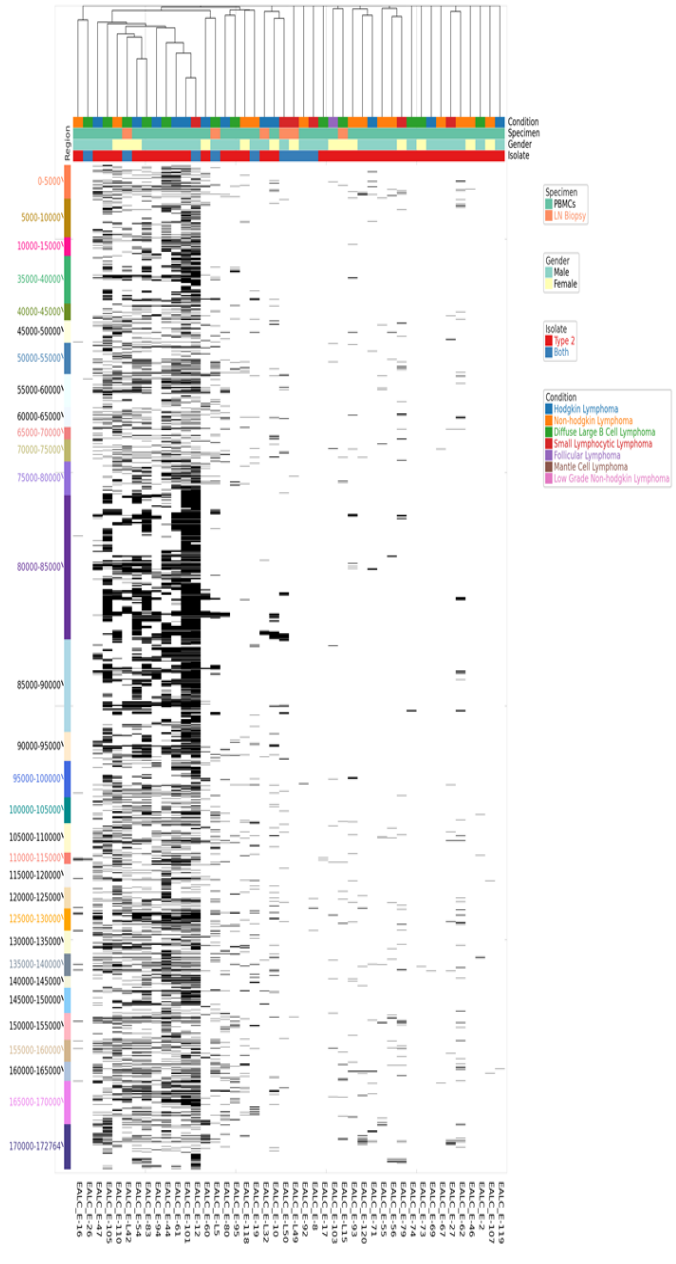
Further analysis revealed that the highly variable samples exhibited notable variations within specific nucleotide sequences. These variations were predominantly observed between nucleotide positions 5,000 and 40,000, as well as between nucleotide positions 90,000 and 110,000. These nucleotide segments are known for harboring coding regions associated with latent proteins, specifically EBERs and EBNA LP genes. The presence of variability in these coding regions implies potential functional implications and highlights the importance of studying these specific nucleotide segments in understanding the genetic characteristics of EBV type l.

Out of the 44 samples aligned to the EBV genotype 2 reference strain, only 25% (n=11) samples exhibited a gene variation exceeding 50% when compared to the reference strain. These samples demonstrate a substantial level of genetic divergence from the reference. A subset of 20% (n=9) samples displayed a gene variation ranging between 25% and 50%, indicating a moderate level of divergence from the reference genome. The remaining 55% (n=24) samples had a gene variation of less than 25%, suggesting relatively minor genetic variations compared to the reference strain.

Notably, a higher variation rate was observed within the nucleotide sequence region spanning 75,000 to 95,000. This specific region is well-known for coding EBERs (**Figure 4.14B**). Additionally, these nucleotide sequence regions have been identified as coding regions for several latent proteins, particularly Epstein-Barr nuclear antigens.


Region
Condition
Specimen
Gender
Isolate
Specimen
PBMCs
LN Biopsy
Gender
Male
Female
Isolate
Type 2
Both
Condition
Hodgkin Lymphoma
Non-hodgkin Lymphoma
Diffuse Large B Cell Lymphoma
Small Lymphocytic Lymphoma
Follicular Lymphoma
Mantle Cell Lymphoma
Low Grade Non-hodgkin Lymphoma

AF = 0.01


Region
Condition
Specimen
Gender
Isolate
Specimen
PBMCs
LN Biopsy
Gender
Male
Female
Isolate
Type 2
Both
Condition
Hodgkin Lymphoma
Non-hodgkin Lymphoma
Diffuse Large B Cell Lymphoma
Small Lymphocytic Lymphoma
Follicular Lymphoma
Mantle Cell Lymphoma
Low Grade Non-hodgkin Lymphoma

AF = 0.5

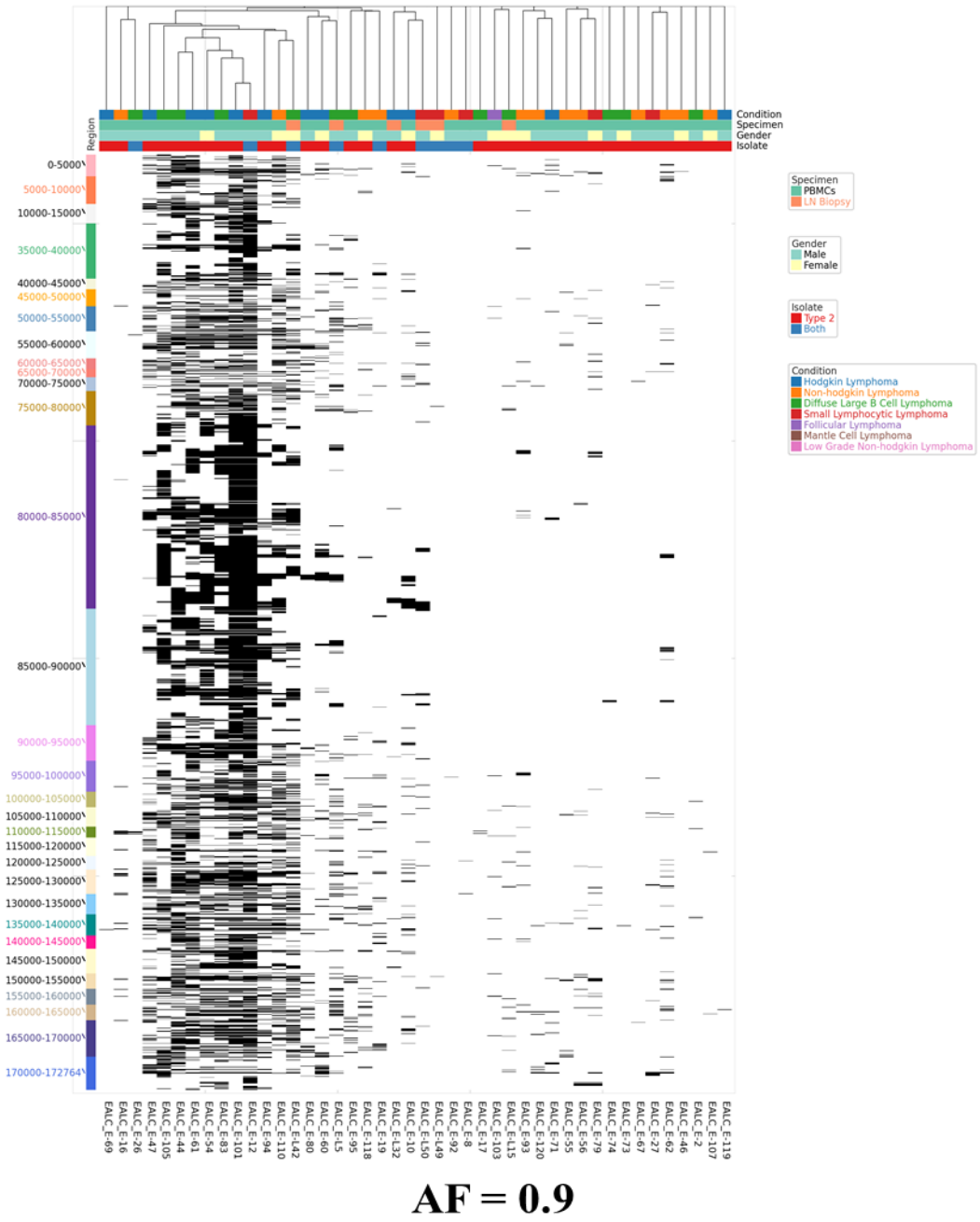
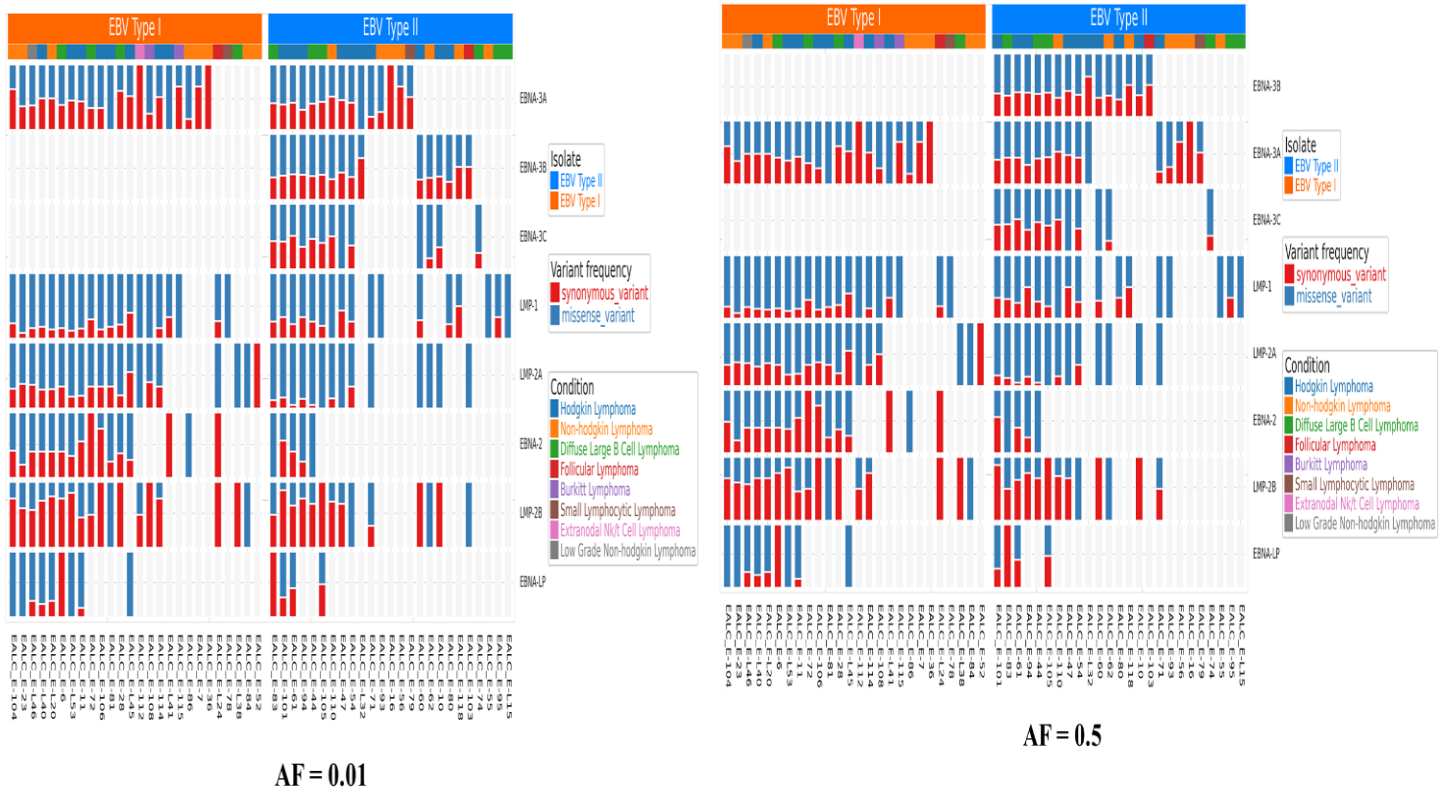


Figure 4.14B: Heatmap plot visualizing variants across EBV Type 2 at different allele frequencies: 0.01 (1% aligned reads supporting the alternative allele), 0.5 (50% aligned reads supporting the alternative allele), and 0.9 (90% aligned reads supporting the alternative allele, almost monoclonal).

4.7.4. Functional annotation

We have analyzed the functional annotation of the identified EBV variants to assess their potential impact on the viral proteins. This annotation process involved predicting the functional consequences of the variants and their potential effects on the structure, function, and interactions of viral proteins. Building upon the variant calling process, we focused on identifying mutations within specific latent protein regions of EBV. We performed this analysis by considering three different allele frequency thresholds: 0.01, 0.5, and 0.9 **(Figure 4.15).** The plot provides a visual representation of the normalized distribution (range: 0-100%) of synonymous (shown in red) and missense (shown in blue) variants in key coding genes, such as EBNA-1 and EBNA-3C, across different samples. Synonymous variants are those that do not result in a change in the amino acid sequence of the encoded protein, while missense variants introduce an alteration in the amino acid sequence. In the plot, gray bars represent samples where no synonymous or missense variants were found in the coding region of the key latent genes. These bars indicate the absence of any detectable genetic variations in those specific regions.

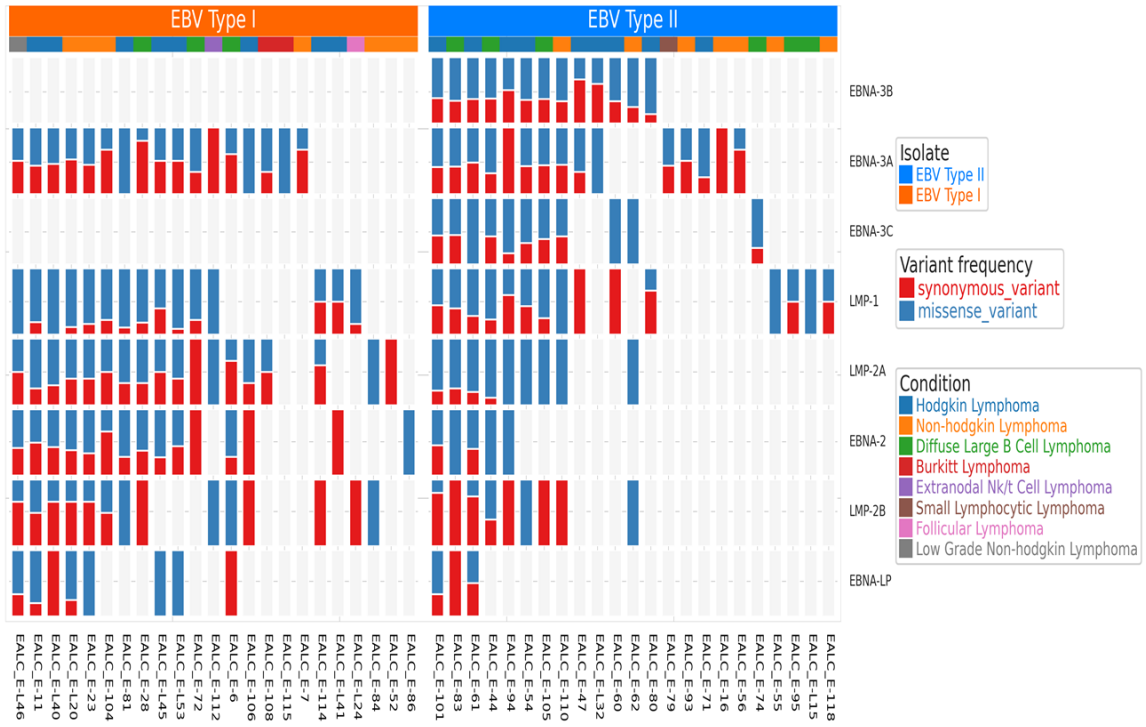


Comparing the latent genes of EBV, it has been observed that EBV type 2 exhibits more variable regions in its latent genes compared to type 1. Among the EBV type 1 samples, no changes in the amino acid sequences were identified in the coding regions for EBNA 3B and EBNA 3C. However, there was a variation in the amino acid sequences of LMP1 in EBV type 1 samples. Furthermore, EBV type 1 samples exhibited more conservation in the coding regions for latent genes.

On the other hand, the EBV type 2 samples did not show any variation in the amino acid sequences of coding regions such as EBNA LP, EBNA 2, and LMP2B. These regions remained constant across the EBV type 2 samples. These findings suggest that EBV type 2 exhibits more variable regions in its latent genes compared to type 1, indicating more genetic diversity within the type 2 samples. In contrast, EBV type 1 samples show a higher level of conservation in certain coding regions for latent genes, indicating a more stable and consistent genetic makeup within these regions.

Extensive analysis of various lymphoma types has revealed that all lymphoma types exhibit amino acid changes in at least one of the EBV latent genes. However, notable differences have been observed between non-Hodgkin lymphoma and Hodgkin lymphoma samples. Specifically, NHL samples commonly demonstrate no change in their amino acid sequences compared to HL types.

Among the NHL types, Diffuse Large B-cell Lymphoma, samples stand out with a higher proportion of missense variants in most latent genes compared to other NHL types. DLBCL samples exhibit a greater number of these missense variants in the latent genes associated with EBV infection. In contrast, Hodgkin lymphoma types and other NHL types show a relatively lower frequency of missense variants in their EBV latent genes. This suggests a lower degree of genetic variation in these regions among these lymphoma types compared to DLBCL.



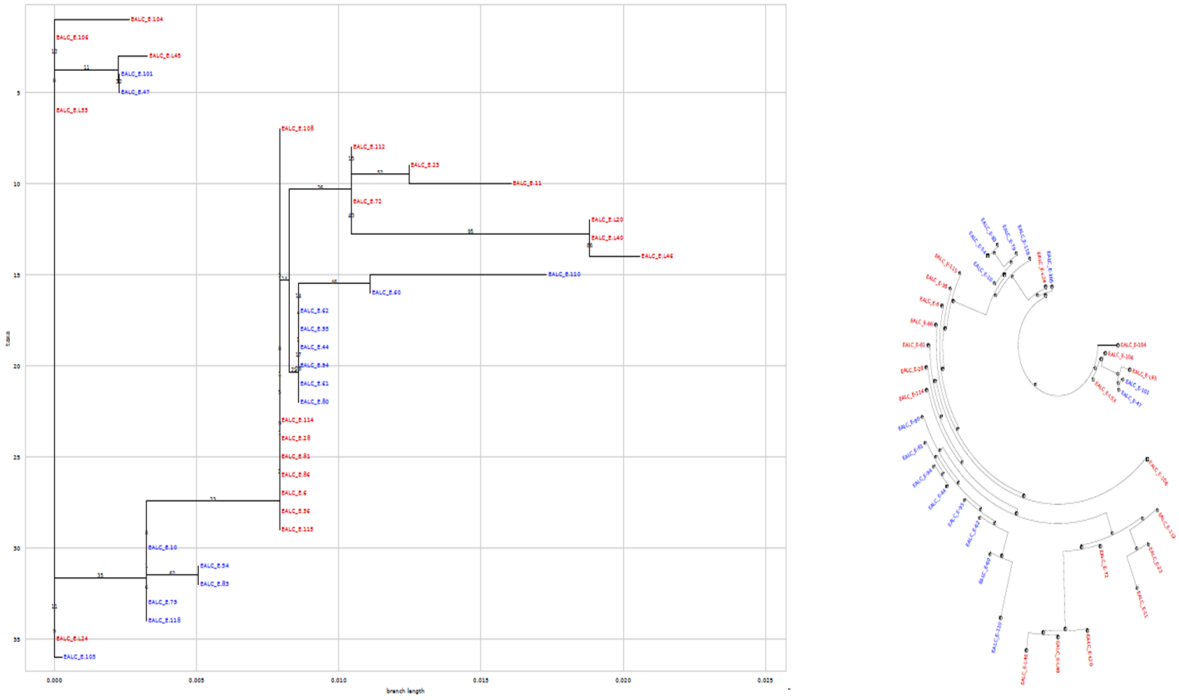
AF = 0.9

Figure 4.15: An alternative plot showing synonymous and missense variants frequency (normalized, between 0 and 100%) is used to highlight the variability of central genes at different allele frequencies: 0.01 (1% aligned reads supporting the alternative allele), 0.5 (50% aligned reads supporting the alternative allele), and 0.9 (90% aligned reads supporting the alternative allele, almost monoclonal).

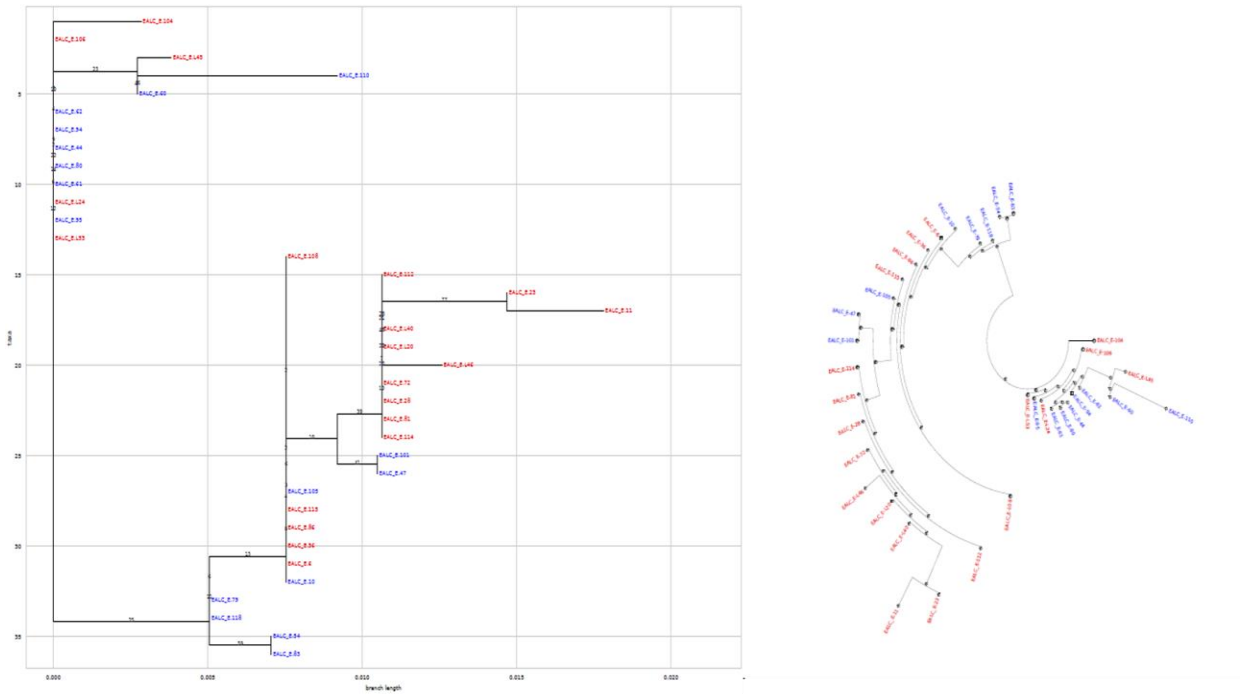
4.7.5. Phylogenetic tree analysis

The phylogenetic analysis was performed using EBV isolates sequenced by this study. To construct the tree, we utilized the variants of the samples, which were filtered based on a minimum allele frequency threshold of >=0.01. Phylogenomic trees have been generated using two layouts: a classic layout and a semi-circular one. We have observed the genetic variation of our samples within each variant. There are three phylogenomic trees for each layout, according to a minimum AF of **0.01**, **0.5**, and **0.9**. For the generation of the phylogenomic tree, we excluded samples that had less than 50 variants across the entire genome. Phylogenetic tree analysis included 36 samples, excluding co-infected samples, and utilized the maximum likelihood method. The generated phylogenetic trees

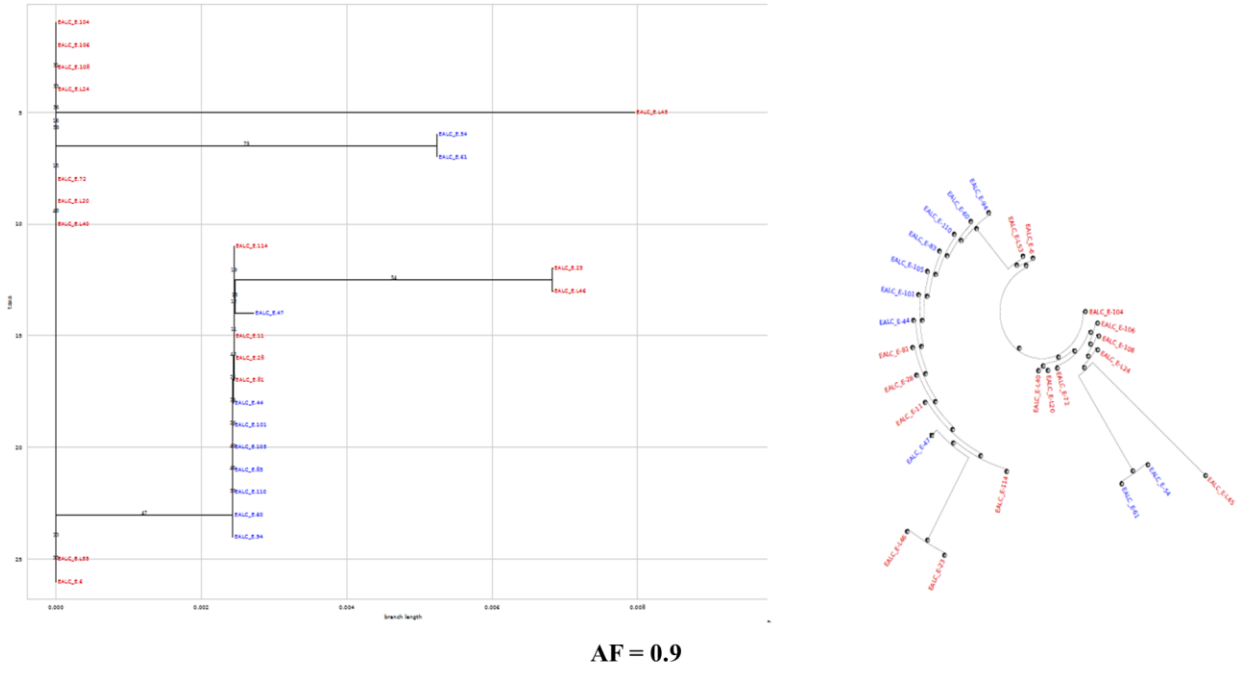
demonstrated the genetic variation within each EBV type and highlighted the evolutionary relationships between samples. To infer the tree structure, we employed IQTree2 (http://www.iqtree.org), a software tool that utilizes the maximum likelihood method (Minh et al., 2020). (**Figure 4.16**).



AF = 0.01



AF = 0.5



AF = 0.9

Figure 4.16: Phylogenetic tree analysis of EBV variants with different allele frequencies: 0.01,0.5 and 0.9. EBV Type I in red and EBV Type II in blue.

A wide-range analysis of the whole EBV genome involved conducting phylogenetic analyses to assess the evolutionary relationships among different sequences. To critically evaluate the phylogenetic signal, the dataset was segregated into EBV type 1 and type 2 genotypes.

Remarkably, similarities in sequences were observed across multiple samples within each genotype. Specifically, four samples exhibited remarkable sequence similarity. Three of these samples were classified as EBV type 1, while the remaining sample belonged to EBV type 2.

These findings highlight the presence of distinct phylogenetic clusters within the dataset, corresponding to the different EBV genotypes. The observed sequence similarities within each genotype suggest a closer evolutionary relationship among the samples. This implies a shared ancestry and potentially indicates a common origin or transmission pathway for the respective EBV genotypes.

CHAPTER 5. DISCUSSION

5.1. EBV detection and quantification

EBV infection is highly prevalent worldwide and is associated with a range of malignancies linked to B or T lymphocytes as well as epithelial cells (Shannon-Lowe & Rickinson, 2019). Detecting EBV using serological and molecular techniques is of utmost importance in monitoring diseases, assessing prognosis, and implementing preventive measures. These detection methods play a crucial role in identifying and understanding the presence of EBV in individuals, enabling effective disease management, and facilitating the development of targeted preventive strategies (Gulley, 2001). Therefore, it is crucial to investigate the prevalence and impact of EBV among lymphoma patients and other population groups across diverse regions worldwide. This research is particularly significant for developing countries like Ethiopia, where evidence-based studies on this topic are often lacking. Conducting this study holds great importance as it will be the first to determine the molecular and serological prevalence of EBV among lymphoma patients in Ethiopia.

Previous studies have established that over 90% of the global population is infected with EBV (Oh & Weiderpass, 2014; Rickinson, 2007). Our study similarly found a serological prevalence of EBV among Ethiopian lymphoma patients to be 96%. This finding aligns with other studies and reports conducted worldwide (Trottier et al., 2012). Moreover, our results indicated a higher prevalence of EBV VCA IgG antibodies in male individuals, HIV-positive patients, and those diagnosed with HL.

The pathological findings in our study indicated a higher prevalence of both Hodgkin lymphoma and non-Hodgkin lymphoma among male patients, which is consistent with findings from other studies conducted globally (Mushtaq et al., 2008; Olu-Eddo & Omoti, 2011; Yeole, 2008). Additionally, our study revealed that HL was more predominant in the age group of less than 20 years. This finding aligns with observations made in studies conducted in Zambia (Kafita et al., 2018) and Iraq (Al-Mudallal & Al-Sinjery, 2012). However, in developed countries, studies have shown that the incidence of HL is highest in individuals above the age of 60 (Cartwright & Watkins, 2004). This disparity is primarily

attributed to factors such as poor socioeconomic status in developing countries, hygienic conditions, and genetic variations (Maggioncalda et al., 2011; Olu-Eddo & Omoti, 2011).

Our study also demonstrated the molecular presence and quantification of EBV in different lymphoma types. High viral load (>10,000EBV copies/ml) EBV was detected in 56.3% of the study participants which is comparable with similar studies conducted in Nigeria (54.5%) (Iliyasu et al., 2014), Zambia (51.8%) (Kafita et al., 2018) and Qatar (52.6%) (Smatti et al., 2017) in lymphoma patients. Higher EBV viral load detection was observed in the study conducted in Pakistan (75.3%) (Salahuddin et al., 2018) and Libya (77.7%) (Shaklawoon et al., 2020) on different types lymphoma. Studies conducted in African countries like Egypt (El-Naby et al., 2017) and Eretria (Fessahaye et al., 2017) on FFPE samples had shown a lower EBV prevalence. In these study sites, EBV was detected in 38% and 27.8%, respectively. Another study in China (W. Zhang et al., 2017) also revealed a 24% EBV prevalence. These differences could be attributed to geographical variation, difference in the study participant group and sample types (FFPE, blood and tissue biopsy).

Furthermore, our study revealed that 59.3% of HL patients and 54.8% of NHL patients exhibited a higher EBV viral load. Similar findings were observed in Zambia (Kafita et al., 2018). However, variations in EBV viral load among HL and NHL patients were observed in other studies conducted in different regions such as France (Donzel et al., 2022), Iran (Habibian et al., 2018; Tabibzadeh et al., 2020), Libya (Shaklawoon et al., 2020), and Thailand (Mitarnun et al., 2002). These discrepancies across geographic locations may be attributed to differences in EBV genotypes (Sample et al., 1990), as well as environmental and dietary factors (Bakkalci et al., 2020). The lack of standardized EBV viral load reporting could also contribute for the difference in viral loads across different population.

Additionally, our findings indicated that 67.6% of DLBCL patients had a higher EBV viral load compared to the other NHL groups. This prevalence is consistent with a study conducted in Libya (66.7%) (Shaklawoon et al., 2020) but higher than the rates reported in studies from France (5%) (Donzel et al., 2022), Turkey(5.3%) (Uner et al., 2011) and Korea (14.3%). The variations in these findings could be attributed to differences in the

methodologies used for detection (e.g., immunohistochemistry vs. RT qPCR) as well as variations in sample types utilized in the studies.

In our study, we made noteworthy observations regarding associations between EBV viral load and specific variables, including HIV status and sample type. These findings align with similar research conducted in Switzerland (Telenti et al., 1993), Argentina (Fellner et al., 2007), and USA (Ling et al., 2003). However, no significant associations were observed between EBV viral load and various sociodemographic characteristics, including age, sex, and lymphoma types. These findings align with comparable research finding conducted in Qatar (Smatti et al., 2017), Iran (Habibian et al., 2013), and Zambia (Kafita et al., 2018). However, studies conducted in England revealed an association between EBV status and age groups (Winter et al., 2019). Furthermore, a study conducted in the United Kingdom revealed notable variations in EBV prevalence concerning age and gender. The research findings indicated that the distribution of EBV differed across different age groups and between males and females (Kuri et al., 2020). These observed differences in EBV prevalence can be attributed to several factors, including variations in study populations, sample types, and geographic locations.

5.2. EBV genotype characterization

Analyzing the distribution of EBV genotypes plays a pivotal role in unraveling the intricate disease pathogenesis and comprehending the broader impact of EBV in lymphoma. In our study, we reported the first-ever dataset on the genotype distribution of EBV among lymphoma patients in Ethiopia. To achieve this, we employed the characterization of EBV genotypes using *EBNA 3C* gene detection. These findings hold immense value in terms of enhancing our understanding of EBV-associated lymphomas in the Ethiopian context, aiding in accurate diagnosis, treatment optimization, and potential advancements in personalized care strategies.

EBV genotyping can be made by targeting various genes such as *EBNA 2* and *EBNA 3C* genes. Variation in the EBV *EBNA3C* gene has been reported to classify EBV types (Palser et al., 2015). To determine the EBV genotype in our study, we followed the suggested approach by Sample et al., which identified significant genetic divergence at four specific

loci and maintained distinct characteristics for each genotype at these loci. By utilizing this methodology, we aimed to accurately classify and differentiate between the two EBV genotypes based on the identified genetic variations (Sample et al., 1990). Using EBNA3C gene typing, our study determined the distributions of EBV genotypes in the lymphoma patient. We found that EBV genotype 1 was present in 52.2% of cases, EBV genotype 2 in 38.2% of cases, and coinfections with both genotypes in 9.7% of cases.

A consistent trend observed in the majority of studies conducted worldwide is the higher predominance of EBV type 1. However, in sub-Saharan Africa and Papua New Guinea, a distinct pattern emerges, with EBV genotype 2 being predominantly found in these regions (Bouvard et al., 2009; Kwok et al., 2015). EBV type 1 was detected in other studies conducted on different lymphoma types in Pakistan by Salahuddin S et al., (90.7%) (Salahuddin et al., 2018), 91.2% from Iran on hematologic malignancies by Tabibzadeh et al. (Tabibzadeh et al., 2020), 76.3% from China on NPC by Cui Y et al., (Cui et al., 2011), and 72.5% from Qatar on healthy individuals by Smatti M et al., (Smatti et al., 2017). Additionally, EBV type 1 was isolated from 71.1% of the study populations in Brazil by Monteiro TAF et al. (Monteiro et al., 2020). The collective findings, including those from other countries, consistently indicate a higher prevalence of EBV type 1 compared to our study (52.2%). This discrepancy may be attributed to various factors specific to our study methodology, particularly the use of the EBNA 3C gene for determining EBV genotypes. In their review, Smatti et al. (Smatti et al., 2018) showed that ENBA 3 shows less difference than EBNA 2. Besides, we are reporting data from only 207 patients, which may not represent the lymphoma patients in Ethiopia or the population in sub–Saharan Africa.

Our findings align with results from studies conducted in the USA, by Sixbey JW et al. In that study, EBV type 1 was identified in 50% of participants, type 2 in 41%, and both genotypes coexisting in 9% of the study population (J. Sixbey et al., 1989). Another study on lymphoma patients revealed a 56% prevalence of EBV type 1, 13% with type 2, and 31% with dual viral infections (Lin et al., 1993). A study conducted in Brazil on HIV patients showed a 47.37% EBV type 1 prevalence (Pereira et al., 2022).

Our study revealed that EBV type 2 infection was 38.2%, and such high EBV genotype 2 predominance was observed in some sub-Saharan African countries. A study conducted in Ghana showed a 52% EBV type 2 prevalence (Ayee, Ofori, Tagoe, et al., 2020). Other studies in other regions also revealed a high predominance of EBV type 2. A study conducted in China on lymphoma patients assessed a 98.6% EBV type 2 prevalence (Y.-J. Wang et al., 2021). In HL patients in Mexico, EBV type 2 was discovered in 47.6% of children and 69.2% of adults (Palma et al., 2013). Another study in Turkey showed a 44.4% EBV type 2 prevalence (Tinguely et al., 2000). The observed disparities in the prevalence of EBV genotypes could be attributed to multiple factors, including variations in their oncogenic properties, the immune status of the study populations, and geographic and strain differences (Chabay & Preciado, 2013; Ibrahim et al., 2010; Renzette et al., 2014).

In our study, we observed a similar prevalence of genotypes in both HL and NHL patients. Interestingly, this finding aligns with a study conducted in Zambia, which focused on HL patients and reported the identification of EBV type 1 in 55.6% of cases. In addition, 33.3% were EBV type 2, and 11.1% were type 1 and 2 coinfected cases (Kafita et al., 2018). On the other hand, the prevalence of EBV genotype 1 in our study in both HL and NHL patients was lower compared to other European studies, such as those conducted in Belgium (Trimèche et al., 2007) and Croatia (Begić et al., 2022). These disparities could be mainly due to the geographic variations (C. M. Chang et al., 2009; Neves et al., 2017) and immune status (Pereira et al., 2022; Zealiyas et al., 2023) of the populations within each country.

There were no significant associations between the EBV genotypes and the different lymphoma types. Similar reports were revealed in other studies where lymphoma type did not affect the distribution of EBV genotypes (Salahuddin et al., 2018; Tabibzadeh et al., 2020). However, we found significant differences in EBV types with age groups, similar to a study conducted in Mexico and Brazil respectively (Monteiro et al., 2019; Palma et al., 2013).

According to certain reports, it has been suggested that coinfection with both EBV type 1 and type 2 is feasible in immunocompromised individuals, implying that acquiring EBV type 2 may occur during the patient's immunocompromised condition (Pereira et al., 2022;

Yao et al., 1996). Studies conducted in different part of the world revealed higher rates of coinfection with both EBV types in individuals living with HIV. Studies in Argentina (32.6%) (Correa et al., 2007), China (12.96%) (Wan et al., 2023), and Brazil (26.32%) (Pereira et al., 2022) reported higher rates of coinfections to our report (9.7%).

5.3. EBV methylation profiling

The findings of our study on the methylation profiles of EBV DNA in lymphoma patients provide valuable insights into the epigenetic regulation of EBV and its potential implications in lymphomagenesis. In recent years, there have been several studies investigating the methylation profiles of EBV DNA in various malignancies, providing valuable insights into the epigenetic regulation of EBV and its implications in the development of cancer cancers associated with EBV.

DNA methylation is one of the key epigenetic modifications that regulate gene expression and cellular processes (J. L. Miller & Grant, 2012; N. Zhang, 2018). Alterations in DNA methylation patterns can have profound effects on gene regulation and contribute to tumorigenesis (Wachsman, 1997). Therefore, studying the methylation profiles of EBV DNA in lymphoma patients can provide valuable insights into the epigenetic dysregulation associated with EBV-associated lymphomagenesis.

In an unpublished study conducted at the Ohio State Comprehensive Cancer Center, researchers performed additional comparisons among different lymphoma types within various cohorts. The primary objective was to investigate the variations in methylation profiles across lymphoma types, focusing on samples obtained from the western countries and Ethiopia. The analysis revealed a methylation loss specifically in the lytic genes BZLF1 and at the lytic origin of replication. This finding suggests that there may be significant differences in the methylation profiles between the Ethiopian lymphoma samples and those from Western cohorts. **Figure 5.1** visualizes the methylation profiles of the lymphoma samples in Ethiopia compared to the Western cohorts.

The comparison of methylation profiles between these two populations is important for understanding potential variations in the epigenetic regulation of lymphoma development and progression. The observed methylation loss in the lytic genes BZLF1 and at the lytic

origin of replication indicates potential dysregulation in these regions in the Ethiopian lymphoma samples.

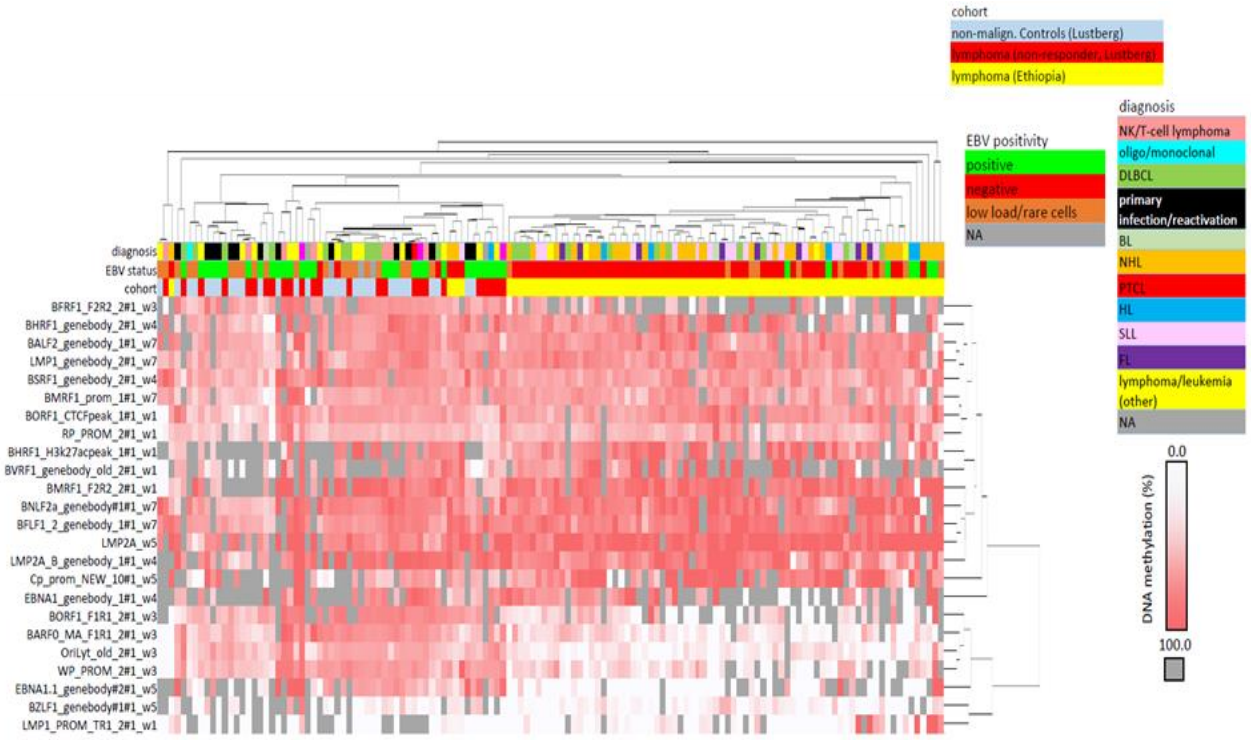


Figure 5.1: Methylation profiles of lymphoma samples in Ethiopia Vs western cohorts. It provides a graphical representation of the differences in methylation patterns between the two groups, highlighting the specific regions where methylation loss was observed.

Several studies have focused on investigating the methylation patterns of EBV DNA in different lymphoma subtypes. For instance, in Burkitt lymphoma, a highly aggressive B-cell lymphoma, studies have shown that the viral genome undergoes specific DNA methylation changes (Kretzmer et al., 2015). Methylation of certain viral genes, including LMP1 and EBNA1, has been observed, which can impact their expression and contribute to the pathogenesis of Burkitt lymphoma (Takacs et al., 2009).

In Hodgkin lymphoma, DNA methylation studies have revealed distinct methylation patterns associated with EBV infection (Dhiab et al., 2015; Leonard et al., 2011). For example, DNA hypermethylation of specific genes such as *SHP1* gene promoter has been

observed in EBV-positive Hodgkin lymphoma cases (Dhiab et al., 2015). These methylation changes can influence the expression of tumor suppressor genes and contribute to the development and progression of the disease (Massini et al., 2009).

Additional comparison was also made between the methylation status of lymphoma samples in Ethiopia and OSU cohorts. The findings are summarized in **Figure 5.2**, which illustrates the disparities in methylation profiles across different lymphoma types within the OSU and ETH cohorts. As it shown in **Figure 5.2**, differences in EBV methylation patterns among the lymphoma subtypes observed in the OSU and ETH cohorts. These differences provide valuable insights into potential variations in EBV-associated lymphoma biology between these two cohorts, suggesting that there may be distinct epigenetic regulation of EBV in different lymphoma subtypes across these populations.

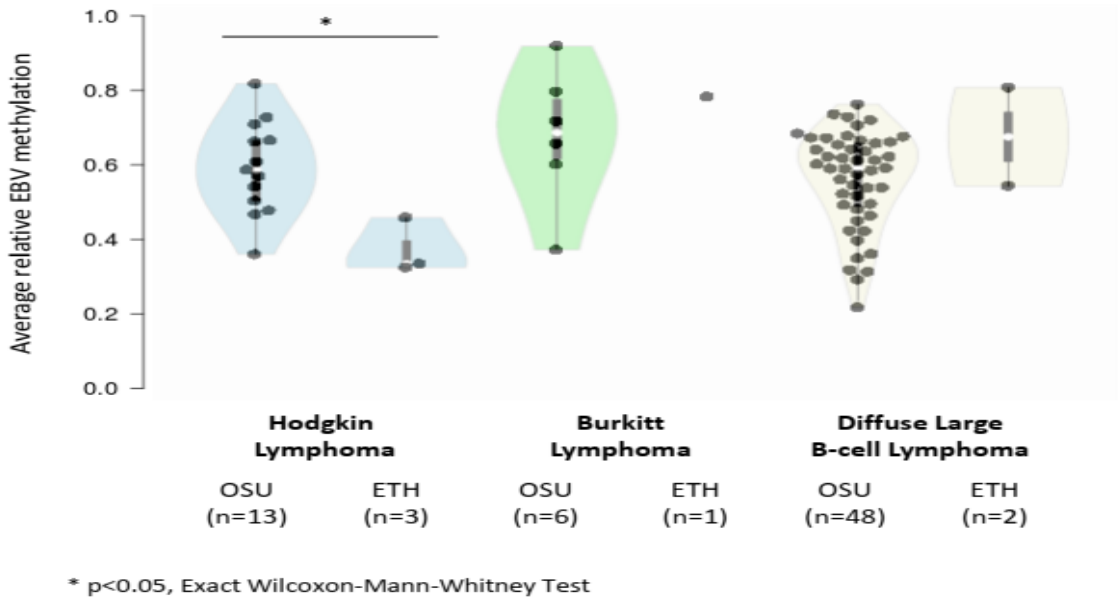


Figure 5.2: Methylation profiles of EBV among different lymphoma types

In recent years, several studies have also focused on identifying specific genes or loci that undergo methylation changes in EBV-associated lymphomas. For instance, *Pike BL et al*. conducted DNA methylation profiles in diffuse large B-cell lymphoma and their relationship to gene expression status. They identified the DNA methylation levels observed for specific CpG islands suggest there is considerable epigenetic heterogeneity

within the tumors analyzed (Pike et al., 2008). Similarly, *Stanland LJ et al.* investigated the role of EBV-induced hypermethylation in gastric cancer tumorigenesis, providing insights into the epigenetic dysregulation induced by EBV in this malignancy(Stanland & Luftig, 2020).

Another relevant study conducted by Tang CL et al., which investigated the genomic methylation profiles of various EBV-associated tumors and identified specific genes that exhibited differential methylation patterns. These differences in methylation modifications suggest potential variations in gene expression levels among the tumors (Tang et al., 2024).

5.4. EBV sequencing

The present study utilized EBV capture sequencing to investigate the genomic characteristics of EBV in a selected group with high viral load of 110 samples. The analysis focused on viral load, type identification, quality control, depth of coverage, variant calling, and functional annotation. The findings shed light on the genetic diversity, variations in specific nucleotide sequences, and the impact of these variations on selected viral proteins.

Previous research groups have successfully generated a substantial number of complete EBV genome sequences, reaching up to 1000 in total, from diverse geographic locations(Correia et al., 2018; Wegner et al., 2019). However, each research group analyzed their respective datasets independently, without adhering to standardized criteria. While this approach may not pose significant issues for individual analyses, it becomes a significant concern when attempting to integrate the independently generated consensus sequences for further analysis. The lack of standardized criteria introduces potential inconsistencies and challenges in comparing and combining the data across different studies.

The absence of a universally accepted gold-standard bioinformatic pipeline for obtaining the final consensus sequence of the EBV genome is a significant challenge (Blazquez et al., 2021). This is particularly crucial due to the presence of repetitive regions within the EBV genome, which can lead to variations in the obtained consensus genomes, even when analyzing the same dataset (Ba abdullah et al., 2017). To ensure higher quality and more reliable results, we have demonstrated the critical importance of analyzing the raw fastq

data files using a consistent and standardized pipeline. By employing the same pipeline across multiple research groups, we can minimize discrepancies and enhance the comparability of results, ultimately improving the accuracy and consistency of the obtained consensus sequences (Ebert & Brutlag, 2006). This approach ensures greater confidence in the genomic variations identified and facilitates more meaningful and comprehensive analyses of EBV genomes.

The samples exhibited different EBV types and coinfection. The unique patterns of gene expressions associated with the EBV are crucial for distinguishing between different subtypes of EBV associated diseases (Neves et al., 2017; Smatti et al., 2018). Numerous research studies conducted on EBV have demonstrated that these expression patterns play essential roles in initiating oncogenic signaling pathways, activating the immune system, and facilitating the transformation of B-cells (M. S. Chang & Kim, 2005; Di Pietro, 2020; Price & Luftig, 2014).

EBV genomic variations have been found to be influenced by various factors, including geographic distributions, environmental influences, disease subtypes, and the phase of EBV latency, as indicated by previous study (Palser et al., 2015). To accurately determine the dissemination of EBV, it is necessary to trace EBV genomic variations phylogenetically (Correia et al., 2018). This involves analyzing the genetic makeup of different EBV types and comparing them to establish evolutionary relationships and patterns of transmission. (Palser et al., 2015; Tzellos & Farrell, 2012). The results of the variant calling analysis from our study revealed a substantial level of genetic diversity within the EBV types. The highly variable EBV types exhibited notable variations within specific nucleotide sequences, particularly in coding regions associated with latency, proteins such as EBERs and EBNA LP genes. These findings suggest potential functional implications and highlight the importance of studying these specific nucleotide segments in understanding the genetic characteristics of EBV.

Several previous studies have highlighted the presence of numerous genomic variations within the EBV genome. These variations play crucial roles in viral regulation and cellular transformation processes. The genomic variations in EBV can affect various aspects of

viral biology, including viral gene expression, immune evasion mechanisms, and interactions with host cells (Rüeger et al., 2021; Umakanthan & Bukelo, 2021; Xiao et al., 2016). These contribute to the adaptability and persistence of EBV in the host, as well as its ability to establish latent infections and reactivate under specific conditions. Furthermore, these genomic variations have been associated with the development of EBV-associated diseases, such as certain types of lymphomas and carcinomas. These can influence viral oncogenic potential, altering the expression of oncogenic proteins and modulating cellular signaling pathways (Agwati et al., 2022; Tzellos & Farrell, 2012; Zhou et al., 2017). Indeed, several environmental factors have been found to significantly influence EBV genomic evolution and polymorphisms. It is of great importance to investigate whether these genomic variations exhibit ethnic or geographical correlations and whether they have a direct impact on the development of specific diseases (Neves et al., 2017).

In this study, we have conducted phylogenetic analyses for the first time in Ethiopia on EBV. Our findings revealed the existence of multiple distinct EBV strains within the region. Remarkably, our study challenges the prevailing notion that genomic divergences within EBV taxonomy are predominantly influenced by similar subtypes of EBV-associated malignancies. The genomic divergences observed in EBV genomes may not be dependent on the subtypes of EBV-associated diseases. This conclusion is supported by previous studies that have highlighted the limitations of using disease subtypes as a sole criterion for classifying EBV genomes. Instead, it has been suggested that multiple co-factors can contribute to the genomic variation observed in EBV. These findings emphasize the complex nature of EBV genomic diversity and suggest the involvement of additional factors beyond disease classification in shaping the evolutionary trajectories of EBV strains (Choi et al., 2018).

Our findings suggest that the phylogenetic study of EBV genomes may reveal fine-grained genomic variations within or between EBV-associated malignancies. Moreover, the phylogenetic analysis allowed us to discern broader disparities among EBV genomes, providing a comprehensive understanding of the overall genomic diversity within the virus. Therefore, the phylogenetic relatedness of EBV genomes is probably better inferred from

EBV genomic similarities. The examination of EBV genomic variations through phylogenetic analysis holds significant implications for multiple areas, including the development of vaccines and immune therapies to prevent EBV infection (Correia et al., 2018; Wegner et al., 2019). Additionally, it provides a valuable tool for investigating the connections between EBV genomic variations and EBV-associated diseases, enabling the development of advanced technologies for studying these relationships (Neves et al., 2017). This study significantly expands the pool of fully characterized EBV genomes, particularly from the region of East Africa, where representative data was previously lacking in the NCBI database. By including EBV genomes from this previously underrepresented region, our study addresses an important gap in the available knowledge. The identification and reporting of different genotypes and the observed divergence within EBV genomes hold significance. These findings contribute to a more comprehensive understanding of the genetic diversity and evolutionary dynamics of EBV.

Limitations of the study

This study has certain limitations that should be acknowledged. One limitation is the loss of follow-up and the absence of pathology results for some participants suspected of having lymphoma. As a result, data analysis for a subset of cases was incomplete. Additionally, another limitation is the lack of other serologic assays, such as EBV-specific IgM, and limited serological assay results. Only the EBV VCA IgG test was conducted, potentially missing out on additional serological markers or antibody responses associated with EBV infection. Moreover, there was a notable absence of standardized reporting guidelines for EBV viral load recommended by the WHO, which hampers the ability to compare research findings. Additionally, the availability of studies exploring the methylation profile of EBV in lymphoma cases was scarce, making it challenging to discuss and contextualize the results in relation to other research. Our study primarily focused on viral methylation profiles, but it would have been beneficial to include the methylation profiles of the host as well. Unfortunately, limited funding restricted our ability to incorporate host methylation profiles into the study.

CHAPTER 6. CONCLUSION AND RECOMMENDATIONS

6.1. Conclusion

This study on the molecular epidemiology and genetic diversity of Epstein-Barr virus among lymphoma patients in Ethiopia has provided valuable insights into the characteristics and distribution of EBV genotypes in this population. The findings highlight the importance of understanding the genetic diversity of EBV in relation to lymphoma development. The study sheds light on the prevalence and viral load of EBV, as well as the genotypic variations observed in Ethiopian lymphoma patients. These findings contribute to the growing body of knowledge on EBV-related malignancies and have implications for both research and clinical management.

The study underscores the need for incorporating EBV viral load assessments and pathological detection techniques, such as EBER ISH, in the evaluation and identification of EBV-associated lymphomas. By considering these factors, healthcare professionals can gain a better understanding of disease progression, devise effective preventive strategies, and develop targeted therapeutic approaches. Moreover, the study's outcomes have broader implications, providing insights into EBV genotype distribution patterns worldwide and their potential correlations with disease outcomes and therapeutic responses.

Additionally, the study's findings shed light on the methylation profile of EBV, offering insights into the epigenetic regulation of the virus and its implications for viral biology, disease development, and clinical outcomes. However, further research is needed to deepen our understanding of the role of EBV epigenetics in the pathogenesis of EBV-associated malignancies.

Furthermore, this study represents a significant milestone as it provides the first genomic and epigenomic analysis of the Epstein-Barr virus genome in Ethiopia and, more broadly, in sub-Saharan Africa. By conducting this analysis, we have contributed to expanding the number of genomic data represented in the NBCI database. The inclusion of genomic data from underrepresented regions such as Ethiopia is crucial for addressing global health disparities. By expanding the available genomic data from underrepresented regions, we

are able to gain insights into regional variations in EBV genomics and their implications for disease management. This inclusive approach is essential for addressing global health disparities and advancing our knowledge of EBV-related diseases on a global scale.

Overall, this study contributes to the growing body of knowledge in the field of EBV research, advancing our understanding of EBV-related malignancies and guiding efforts to combat these diseases in Ethiopia and beyond. The insights gained from this study can inform future research endeavors, support the development of region-specific diagnostic and therapeutic strategies, and ultimately improve patient outcomes.

6.2. Recommendations

Based on our findings, the following recommendations can be made to advance research and clinical understanding of Epstein-Barr virus and lymphoma in Ethiopia:

Establish a National Cancer Registry: It is recommended to establish a comprehensive national cancer registry in Ethiopia to systematically collect and analyze data on lymphoma and other EBV associated malignancy patients. This registry would include clinical information, demographic data, and molecular characterization of EBV infections. The registry will contribute to a better understanding of the epidemiology, incidence, prevalence, and outcomes of lymphoma in Ethiopia, providing a foundation for future research and evidence-based decision-making.

Implement Routine EBV Testing: Given the association between EBV and lymphoma, it is recommended to implement routine EBV testing for lymphoma patients in Ethiopia. In addition, the inclusion of Epstein-Barr virus-encoded small RNA in situ hybridization (EBER-ISH) in routine diagnostic activities is crucial. By incorporating EBER-ISH into the diagnostic process, healthcare providers can enhance their ability to accurately detect and diagnose lymphoma cases. This will help identify the proportion of lymphoma cases associated with EBV infection and provide valuable data for further research.

Expand Molecular Epidemiology Studies: Further molecular epidemiology studies focusing on EBV and lymphoma in Ethiopia should be conducted. These studies should involve larger sample sizes and include diverse geographic regions and ethnic groups

within the country. By investigating the genetic diversity and distribution of EBV strains associated with lymphoma, researchers can identify potential risk factors, understand disease progression, and develop targeted interventions.

Enhance Collaboration and Data Sharing: Encouraging collaboration among researchers and institutions working on EBV globally is crucial. It is recommended to establish collaborative networks that facilitate data sharing, exchange of samples, and joint analyses. This collaboration will promote a more comprehensive understanding of EBV genomics and its implications for disease outcomes. To ensure a comprehensive understanding of EBV genomics and its implications, it is crucial to actively involve researchers and institutions from LMICs in vaccine trials, clinical studies, and research collaborations related to EBV. Additionally, sharing data and resources with researchers in LMICs will foster equitable access to genomic research and strengthen capacity building efforts

Investigate Host-Virus Interactions: Further studies should explore the host-virus interactions in EBV-associated lymphomas among Ethiopian patients. This research should focus on understanding the genetic and immunological factors that influence the development and progression of EBV-related lymphomas. By unraveling the intricate interplay between the virus and the host immune response, researchers can identify potential biomarkers, therapeutic targets, and personalized treatment options.

Strengthen Genomic Surveillance: Building upon the success of this study, it is recommended to establish ongoing genomic surveillance programs focused on EBV in Ethiopia and other underrepresented regions within sub-Saharan Africa. By actively monitoring the genetic diversity and evolution of the virus, we can gain insights into emerging strains, identify potential virulence factors, and anticipate changes in disease patterns. This information will aid in the development of targeted interventions and preventive strategies.

REFERENCES

Abusalah, M. A. H., Gan, S. H., Al-Hatamleh, M. A. I., Irekeola, A. A., Shueb, R. H., & Yean Yean, C. (2020). Recent advances in diagnostic approaches for epstein–barr virus. *Pathogens*, *9*(3), 226.

Adler, D., & Kelly, S. T. (2005). vioplot: Violin plot. R package version 0.2. *Zugriff Am*, *2*, 2017.

Agwati, E. O., Oduor, C. I., Ayieko, C., Ong'echa, J. M., Moormann, A. M., & Bailey, J. A. (2022). Profiling genome-wide recombination in Epstein Barr virus reveals type-specific patterns and associations with endemic-Burkitt lymphoma. *Virology Journal*, *19*(1), 1–12.

Al-Mudallal, S. S., & Al-Sinjery, G. M. (2012). Immunohistochemical Expression of Epstein Barr Virus Antigen Latent Membrane Protein-1 and Bcl-2 in Classical Hodgkin Lymphoma. *Iraqi Journal of Medical Sciences*, *10*(3).

Albanese, M., Tagawa, T., Bouvet, M., Maliqi, L., Lutter, D., Hoser, J., Hastreiter, M., Hayes, M., Sugden, B., & Martin, L. (2016). Epstein–Barr virus microRNAs reduce immune surveillance by virus-specific CD8+ T cells. *Proceedings of the National Academy of Sciences*, *113*(42), E6467–E6475.

Albanese, M., Tagawa, T., & Hammerschmidt, W. (2022). Strategies of Epstein-Barr virus to evade innate antiviral immunity of its human host. *Frontiers in Microbiology*, *13*, 955603.

Alford, C. A., Britt, W. J., Roizman, B., Whitley, R. J., & Lopez, C. (1993). *The human herpesviruses*.

Alford, C., Britt, W. B., Whiteley, R., & Lopez, C. (1993). Cytomegalovirus: The human herpesviruses. *Roizman, Raven Press Ltd, New York*, *227*.

Ambinder, R. F., & Mann, R. B. (1994). Epstein-Barr-encoded RNA in situ hybridization: diagnostic applications. *Human Pathology*, *25*(6), 602–605.

Ambinder, R. F., Robertson, K. D., & Tao, Q. (1999). DNA methylation and the Epstein–Barr virus. *Seminars in Cancer Biology*, *9*(5), 369–375.

Ansari, M. A., Singh, V. V., Dutta, S., Veettil, M. V., Dutta, D., Chikoti, L., Lu, J., Everly, D., & Chandran, B. (2013). Constitutive interferon-inducible protein 16-inflammasome activation during Epstein-Barr virus latency I, II, and III in B and epithelial cells. *Journal of Virology*, *87*(15), 8606–8623.

Arvey, A., Ojesina, A. I., Pedamallu, C. S., Ballon, G., Jung, J., Duke, F., Leoncini, L., De Falco, G., Bressman, E., & Tam, W. (2015). The tumor virus landscape of AIDS-related lymphomas. *Blood, The Journal of the American Society of Hematology*, *125*(20), e14–e22.

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.

Ascherio, A., & Munger, K. L. (2010). Epstein–Barr virus infection and multiple sclerosis: a review. *Journal of Neuroimmune Pharmacology*, *5*, 271–277.

Au, W.-Y., Pang, A., Choy, C., Chim, C.-S., & Kwong, Y.-L. (2004). Quantification of circulating Epstein-Barr virus (EBV) DNA in the diagnosis and monitoring of natural killer cell and EBV-positive lymphomas in immunocompetent patients. *Blood*, *104*(1), 243–249.

Awedew, A. F., Asefa, Z., & Belay, W. B. (2022). National burden and trend of cancer in Ethiopia, 2010–2019: a systemic analysis for Global burden of disease study. *Scientific Reports*, *12*(1), 12736.

Ayee, R., Ofori, M. E. O., Tagoe, E. A., Languon, S., Searyoh, K., Armooh, L., Bilson-Amoah, E., Baidoo, K., Kitcher, E., & Wright, E. (2020). Genotypic characterization of Epstein Barr virus in blood of patients with suspected nasopharyngeal carcinoma in Ghana. *Viruses*, *12*(7), 766.

Ayee, R., Ofori, M. E. O., Wright, E., & Quaye, O. (2020). Epstein Barr virus associated lymphomas and epithelia cancers in humans. *Journal of Cancer*, *11*(7), 1737.

Ba abdullah, M. M., Palermo, R. D., Palser, A. L., Grayson, N. E., Kellam, P., Correia, S., Szymula, A., & White, R. E. (2017). Heterogeneity of the Epstein-Barr virus (EBV) major internal repeat reveals evolutionary mechanisms of EBV and a functional defect in the prototype EBV strain B95-8. *Journal of Virology*, *91*(23), 10–1128.

Babcock, G. J., Hochberg, D., & Thorley-Lawson, D. A. (2000). The expression pattern of Epstein-Barr virus latent genes in vivo is dependent upon the differentiation stage of the infected B cell. *Immunity*, *13*(4), 497–506.

Bakkalci, D., Jia, Y., Winter, J. R., Lewis, J. E. A., Taylor, G. S., & Stagg, H. R. (2020). Risk factors for Epstein Barr virus-associated cancers: a systematic review, critical appraisal, and mapping of the epidemiological evidence. *Journal of Global Health*, *10*(1).

Barth, S., Meister, G., & Grässer, F. A. (2011). EBV-encoded miRNAs. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms*, *1809*(11–12), 631–640.

Barth, S., Pfuhl, T., Mamiani, A., Ehses, C., Roemer, K., Kremmer, E., Jäker, C., Höck, J., Meister, G., & Grässer, F. A. (2008). Epstein–Barr virus-encoded microRNA miR-BART2 down-regulates the viral DNA polymerase BALF5. *Nucleic Acids Research*, *36*(2), 666–675.

Bauer, C. C., Aberle, S. W., Popow-Kraupp, T., Kapitan, M., Hofmann, H., & Puchhammer-Stöckl, E. (2005). Serum Epstein–Barr virus DNA load in primary Epstein–Barr virus infection. *Journal of Medical Virology*, *75*(1), 54–58.

Begić, V., Korać, P., Gašparov, S., Rozman, M., Simicic, P., & Zidovec-Lepej, S. (2022). Molecular Characterisation of Epstein–Barr Virus in Classical Hodgkin Lymphoma. *International Journal of Molecular Sciences*, *23*(24), 15635.

Bentz, G. L., Liu, R., Hahn, A. M., Shackelford, J., & Pagano, J. S. (2010). Epstein–Barr virus BRLF1 inhibits transcription of IRF3 and IRF7 and suppresses induction of interferon-β. *Virology*, *402*(1), 121–128.

Bergbauer, M., Kalla, M., Schmeinck, A., Göbel, C., Rothbauer, U., Eck, S., Benet-Pagès, A., Strom, T. M., & Hammerschmidt, W. (2010). CpG-methylation regulates a class of Epstein-Barr virus promoters. *PLoS Pathogens*, *6*(9), e1001114.

Blazquez, A. C., Berenstein, A. J., Torres, C., Izquierdo, A., Lezama, C., Moscatelli, G., De Matteo, E. N., Lorenzetti, M. A., & Preciado, M. V. (2021). Comprehensive evolutionary analysis of complete Epstein–Barr virus genomes from argentina and other geographies. *Viruses*, *13*(6), 1172.

Blum, K. A., Lozanski, G., & Byrd, J. C. (2004). Adult Burkitt leukemia and lymphoma. *Blood*, *104*(10), 3009–3020.

Bornkamm, G. W. (2009). Epstein-Barr virus and the pathogenesis of Burkitt's lymphoma: more questions than answers. *International Journal of Cancer*, *124*(8), 1745–1755.

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

Braz-Silva, P. H., de Rezende, N. P. M., Ortega, K. L., de Macedo Santos, R. T., & de Magalhaes, M. H. C. G. (2008). Detection of the Epstein–Barr Virus (EBV) by in situ hybridization as definitive diagnosis of hairy leukoplakia. *Head and Neck Pathology*, *2*, 19–24.

Breda, E., Catarino, R. J. F., Azevedo, I., Lobão, M., Monteiro, E., & Medeiros, R. (2010). Epstein-Barr virus detection in nasopharyngeal carcinoma: implications in a low-risk area. *Brazilian Journal of Otorhinolaryngology*, *76*, 310–315.

Brooks, J. M., Croom-Carter, D. S. G., Leese, A. M., Tierney, R. J., Habeshaw, G., & Rickinson, A. B. (2000). Cytotoxic T-lymphocyte responses to a polymorphic Epstein-Barr virus epitope identify healthy carriers with coresident viral strains. *Journal of Virology*, *74*(4), 1801–1809.

Buschle, A., & Hammerschmidt, W. (2020). Epigenetic lifestyle of Epstein-Barr virus. *Seminars in Immunopathology*, *42*(2), 131–142.

Cai, L., Ye, Y., Jiang, Q., Chen, Y., Lyu, X., Li, J., Wang, S., Liu, T., Cai, H., & Yao, K. (2015). Epstein–Barr virus-encoded microRNA BART1 induces tumour metastasis by regulating PTEN-dependent pathways in nasopharyngeal carcinoma. *Nature Communications*, *6*(1), 1–13.

Cai, Q., Chen, K., & Young, K. H. (2015). Epstein–Barr virus-positive T/NK-cell lymphoproliferative disorders. *Experimental & Molecular Medicine*, *47*(1), e133–e133.

Calderwood, M. A., Venkatesan, K., Xing, L., Chase, M. R., Vazquez, A., Holthaus, A. M., Ewence, A. E., Li, N., Hirozane-Kishikawa, T., & Hill, D. E. (2007). Epstein–Barr virus and virus human protein interaction maps. *Proceedings of the National Academy of Sciences*, *104*(18), 7606–7611.

Callan, M. F. C. (2004). The immune response to Epstein–Barr virus. *Microbes and Infection*, *6*(10), 937–945.

Callan, M. F. C., Steven, N., Krausa, P., Wilson, J. D. K., Moss, P. A. H., Gillespie, G. M., Bell, J. I., Rickinson, A. B., & McMichael, A. J. (1996). Large clonal expansions of CD8+ T cells in acute infectious mononucleosis. *Nature Medicine*, *2*(8), 906–911.

Cartwright, R. A., & Watkins, G. (2004). Epidemiology of Hodgkin's disease: a review. *Hematological Oncology*, *22*(1), 11–26.

Cesarman, E. (2013). Pathology of lymphoma in HIV. *Current Opinion in Oncology*, *25*(5), 487.

Chabay, P. A., & Preciado, M. V. (2013). EBV primary infection in childhood and its relation to B-cell lymphoma development: a mini-review from a developing region. *International Journal of Cancer*, *133*(6), 1286–1292.

Chan, C. K., Mueller, N., Evans, A., Harris, N. L., Comstock, G. W., Jellum, E., Magnus, K., Orentreich, N., Polk, B. F., & Vogelman, J. (1991). Epstein-Barr virus antibody patterns preceding the diagnosis of nasopharyngeal carcinoma. *Cancer Causes & Control*, *2*, 125–

131.

Chan, K. H., Ng, M. H., Seto, W. H., & Peiris, J. S. M. (2001). Epstein-Barr virus (EBV) DNA in sera of patients with primary EBV infection. *Journal of Clinical Microbiology*, *39*(11), 4152–4154.

Chang, C. M., Kelly, J. Y., Mbulaiteye, S. M., Hildesheim, A., & Bhatia, K. (2009). The extent of genetic diversity of Epstein-Barr virus and its geographic and disease patterns: a need for reappraisal. *Virus Research*, *143*(2), 209–221.

Chang, M. S., & Kim, W. H. (2005). Epstein-Barr virus in human malignancy: a special reference to Epstein-Barr virus associated gastric carcinoma. *Cancer Research and Treatment: Official Journal of Korean Cancer Association*, *37*(5), 257–267.

Chene, A., Donati, D., Orem, J., Björkman, A., Mbidde, E. R., Kironde, F., Wahlgren, M., & Bejarano, M. T. (2009). Endemic Burkitt's lymphoma as a polymicrobial disease: new insights on the interaction between Plasmodium falciparum and Epstein–Barr virus. *Seminars in Cancer Biology*, *19*(6), 411–420.

Chijioke, O., Azzi, T., Nadal, D., & Münz, C. (2013). Innate immune responses against Epstein Barr virus infection. *Journal of Leukocyte Biology*, *94*(6), 1185–1190.

Choi, S. J., Jung, S. W., Huh, S., Cho, H., & Kang, H. (2018). Phylogenetic comparison of Epstein-Barr virus genomes. *Journal of Microbiology*, *56*, 525–533.

Christian, M. (2014). Dendritic cells during Epstein Barr virus infection. *Frontiers in Microbiology*, *5*, 308.

Chuang, H.-C., Lay, J.-D., Hsieh, W.-C., Wang, H.-C., Chang, Y., Chuang, S.-E., & Su, I.-J. (2005). Epstein-Barr virus LMP1 inhibits the expression of SAP gene and upregulates Th1 cytokines in the pathogenesis of hemophagocytic syndrome. *Blood*, *106*(9), 3090–3096.

Cohen, J. I. (2000). Epstein–Barr virus infection. *New England Journal of Medicine*, *343*(7), 481–492.

Cohen, M., De Matteo, E., Narbaitz, M., Carreño, F. A., Preciado, M. V, & Chabay, P. A. (2013). Epstein–Barr virus presence in pediatric diffuse large B-cell lymphoma reveals a particular association and latency patterns: analysis of viral role in tumor microenvironment. *International Journal of Cancer*, *132*(7), 1572–1580.

Cohen, M., Narbaitz, M., Metrebian, F., De Matteo, E., Preciado, M. V., & Chabay, P. A. (2014). Epstein-Barr virus-positive diffuse large B-cell lymphoma association is not only restricted to elderly patients. *International Journal of Cancer*, *135*(12), 2816–2824.

Correa, R. M., Dolores Fellner, M., Durand, K., Redini, L., Alonio, V., Yampolsky, C., Colobraro, A., Sevlever, G., Teyssié, A., & Benetucci, J. (2007). Epstein Barr virus genotypes and LMP-1 variants in HIV-infected patients. *Journal of Medical Virology*, *79*(4), 401–407.

Correia, S., Bridges, R., Wegner, F., Venturini, C., Palser, A., Middeldorp, J. M., Cohen, J. I., Lorenzetti, M. A., Bassano, I., & White, R. E. (2018). Sequence variation of Epstein-Barr virus: viral types, geography, codon usage, and diseases. *Journal of Virology*, *92*(22), 10–1128.

Cui, Y., Wang, Y., Liu, X., Chao, Y., Xing, X., Zhao, C., Liu, C., & Luo, B. (2011). Genotypic analysis of Epstein-Barr virus isolates associated with nasopharyngeal carcinoma in

Northern China. *Intervirology*, *54*(3), 131–138.

De Hoon, M. J. L., Imoto, S., Nolan, J., & Miyano, S. (2004). Open source clustering software. *Bioinformatics*, *20*(9), 1453–1454.

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. *The Lancet Oncology*, *13*(6), 607–615.

De Paschale, M., & Clerici, P. (2012). Serological diagnosis of Epstein-Barr virus infection: Problems and solutions. *World Journal of Virology*, *1*(1), 31.

Dhiab, M. Ben, Ziadi, S., Mestiri, S., Gacem, R. Ben, Ksiaa, F., & Trimeche, M. (2015). DNA methylation patterns in EBV-positive and EBV-negative Hodgkin lymphomas. *Cellular Oncology*, *38*, 453–462.

Di Pietro, A. (2020). Epstein–Barr virus promotes B cell lymphomas by manipulating the host epigenetic machinery. *Cancers*, *12*(10), 3037.

Dolan, A., Addison, C., Gatherer, D., Davison, A. J., & McGeoch, D. J. (2006). The genome of Epstein–Barr virus type 2 strain AG876. *Virology*, *350*(1), 164–170.

Donzel, M., Bonjour, M., Combes, J.-D., Broussais, F., Sesques, P., Traverse-Glehen, A., & de Martel, C. (2022). Lymphomas associated with Epstein-Barr virus infection in 2020: Results from a large, unselected case series in France. *Eclinicalmedicine*, *54*, 101674.

Dowd, J. B., Palermo, T., Brite, J., McDade, T. W., & Aiello, A. (2013). Seroprevalence of Epstein-Barr virus infection in US children ages 6-19, 2003-2010. *PloS One*, *8*(5), e64921.

Ebert, J., & Brutlag, D. (2006). Development and validation of a consistency based multiple structure alignment algorithm. *Bioinformatics*, *22*(9), 1080–1087.

El-Naby, N. E. D. H., Mohamed, H. H., Goda, A. M., & Mohamed, A. E. S. (2017). Epstein-Barr virus infection and breast invasive ductal carcinoma in Egyptian women: a single center experience. *Journal of the Egyptian National Cancer Institute*, *29*(2), 77–82.

Epstein, M. A. (1964). Virus particles in cultured lymphoblasts from Burkitt's lymphoma. *Lancet*, *1*, 702–703.

Epstein, M. A., & Achong, B. G. (n.d.). YM Bart. 1964. Virus particles in cultured lymphoblasts from Burkitt's lymphoma. *Lancet*, 702–703.

Ervik, M., Lam, F., Laversanne, M., Ferlay, J., & Bray, F. (2021). Global cancer observatory: cancer over time. *Int Agency Res Cancer*.

Farrell, P. J. (2015). Epstein–Barr virus strain variation. *Epstein Barr Virus Volume 1: One Herpes Virus: Many Diseases*, 45–69.

Fellner, M. D., Durand, K., Correa, R. M., Redini, L., Yampolsky, C., Colobraro, A., Sevlever, G., Teyssié, A. R., Benetucci, J., & Picconi, M. A. (2007). Circulating Epstein–Barr virus (EBV) in HIV-infected patients and its relation with primary brain lymphoma. *International Journal of Infectious Diseases*, *11*(2), 172–178.

Fernandez, A. F., Rosales, C., Lopez-Nieva, P., Graña, O., Ballestar, E., Ropero, S., Espada, J., Melo, S. A., Lujambio, A., & Fraga, M. F. (2009). The dynamic DNA methylomes of double-stranded DNA viruses associated with human cancer. *Genome Research*, *19*(3), 438–451.

Fessahaye, G., Elhassan, A. M., Elamin, E. M., Adam, A. A. M., Ghebremedhin, A., & Ibrahim, M. E. (2017). Association of Epstein-Barr virus and breast cancer in Eritrea. *Infectious Agents and Cancer*, *12*, 1–7.

Fox, C. P., Civallero, M., Ko, Y.-H., Manni, M., Skrypets, T., Pileri, S., Kim, S. J., Cabrera, M. E., Shustov, A. R., & Chiattone, C. S. (2020). Survival outcomes of patients with extranodal natural-killer T-cell lymphoma: a prospective cohort study from the international T-cell Project. *The Lancet Haematology*, *7*(4), e284–e294.

Frappier, L. (2012). Contributions of Epstein–Barr nuclear antigen 1 (EBNA1) to cell immortalization and survival. *Viruses*, *4*(9), 1537–1547.

Fuks, F., Hurd, P. J., Wolf, D., Nan, X., Bird, A. P., & Kouzarides, T. (2003). The methyl-CpG-binding protein MeCP2 links DNA methylation to histone methylation. *Journal of Biological Chemistry*, *278*(6), 4035–4040.

Fukuda, M., Ikuta, K., Yanagihara, K., Tajima, M., Kuratsune, H., Kurata, T., & Sairenji, T. (2001). Effect of Transforming Growth Factor-β1 on the Cell Growth and Epstein–Barr Virus Reactivation in EBV-Infected Epithelial Cell Lines. *Virology*, *288*(1), 109–118.

Gan, Y., Sullivan, J. L., & Sixbey, J. W. (1994). Detection of cell-free Epstein-Barr virus DNA in serum during acute infectious mononucleosis. *Journal of Infectious Diseases*, *170*(2), 436–439.

Gärtner, B., & Preiksaitis, J. K. (2010). EBV viral load detection in clinical virology. *Journal of Clinical Virology*, *48*(2), 82–90.

Gartzonika, C., Vrioni, G., Priavali, E., Pappas, G., & Levidiotou, S. (2012). Utility of real-time PCR in the diagnosis of primary Epstein-Barr virus infection. *J Med Microb Diagn*, *2*(1), 118.

Gerle, B., Koroknai, A., Fejer, G., Bakos, A., Banati, F., Szenthe, K., Wolf, H., Niller, H. H., Minarovits, J., & Salamon, D. (2007). Acetylated histone H3 and H4 mark the upregulated LMP2A promoter of Epstein-Barr virus in lymphoid cells. *Journal of Virology*, *81*(23), 13242–13247.

Giacopelli, B., Zhao, Q., Ruppert, A. S., Agyeman, A., Weigel, C., Wu, Y.-Z., Gerber, M. M., Rabe, K. G., Larson, M. C., & Lu, J. (2019). Developmental subtypes assessed by DNA methylation-iPLEX forecast the natural history of chronic lymphocytic leukemia. *Blood, The Journal of the American Society of Hematology*, *134*(8), 688–698.

Giudice, A., D'Arena, G., Crispo, A., Tecce, M. F., Nocerino, F., Grimaldi, M., Rotondo, E., D'Ursi, A. M., Scrima, M., & Galdiero, M. (2016). Role of viral miRNAs and epigenetic modifications in epstein-barr virus-associated gastric carcinogenesis. *Oxidative Medicine and Cellular Longevity*, *2016*.

Gloghini, A., Dolcetti, R., & Carbone, A. (2013). Lymphomas occurring specifically in HIV-infected patients: from pathogenesis to pathology. *Seminars in Cancer Biology*, *23*(6), 457–467.

Gottschalk, S., & Rooney, C. M. (2015). Adoptive T-cell immunotherapy. *Epstein Barr Virus Volume 2: One Herpes Virus: Many Diseases*, 427–454.

Grömminger, S., Mautner, J., & Bornkamm, G. W. (2012). Burkitt lymphoma: the role of Epstein-Barr virus revisited. *British Journal of Haematology*, *156*(6), 719–729.

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.

Grywalska, E., & Rolinski, J. (2015). Epstein-barr virus–associated lymphomas. *Seminars in Oncology*, *42*(2), 291–303.

Gulley, M. L. (2001). Molecular diagnosis of Epstein-Barr virus-related diseases. *The Journal of Molecular Diagnostics*, *3*(1), 1–10.

Gulley, M. L., Raphael, M., Lutz, C. T., Ross, D. W., & Raab-Traub, N. (1992). Epstein-barr virus integration in human lymphomas and lymphoid cell lines. *Cancer*, *70*(1), 185–191.

Gulley, M. L., & Tang, W. (2008). Laboratory assays for Epstein-Barr virus-related disease. *The Journal of Molecular Diagnostics*, *10*(4), 279–292.

Gulley, M. L., & Tang, W. (2010). Using Epstein-Barr viral load assays to diagnose, monitor, and prevent posttransplant lymphoproliferative disorder. *Clinical Microbiology Reviews*, *23*(2), 350–366.

Habibian, A., Makvandi, M., Samarbaf-Zadeh, A., Neisi, N., Soleimani-Jelodar, R., Makvandi, K., & Izadi, S. (2018). Detection and Genotyping of Epstein-Bar Virus Among Paraffin Embedded Tissues of Hodgkin and Non-Hodgkin's Lymphoma Patients in Ahvaz, Iran. *Acta Medica Iranica*, 434–440.

Habibian, A., Makvandi, M., Samarbafzadeh, A., Neisi, N., & Ranjbari, N. (2013). Epstein-Barr Virus DNA frequency in paraffin embedded tissues of Non-Hodgkin lymphoma patients from Ahvaz, Iran. *Jentashapir J Health Res*, *4*(4), 315–320.

Harabuchi, Y., Yamanaka, N., Kataura, A., Imai, S., Kinoshita, T., & Osato, T. (1990). Epstein-Barr virus in nasal T-cell lymphomas in patients with lethal midline granuloma. *The Lancet*, *335*(8682), 128–130.

Hassan, R., White, L. R., Stefanoff, C. G., de Oliveira, D. E., Felisbino, F. E., Klumb, C. E., Bacchi, C. E., Seuánez, H. N., & Zalcberg, I. R. (2006). Epstein-Barr virus (EBV) detection and typing by PCR: a contribution to diagnostic screening of EBV-positive Burkitt's lymphoma. *Diagnostic Pathology*, *1*(1), 1–7.

Hatton, O. L., Harris-Arnold, A., Schaffert, S., Krams, S. M., & Martinez, O. M. (2014). The interplay between Epstein–Barr virus and B lymphocytes: implications for infection, immunity, and disease. *Immunologic Research*, *58*, 268–276.

Heslop, H. E., Sharma, S., & Rooney, C. M. (2021). Adoptive T-cell therapy for Epstein-Barr virus–related lymphomas. *Journal of Clinical Oncology*, *39*(5), 514.

Hess, R. D. (2004). Routine Epstein-Barr virus diagnostics from the laboratory perspective: still challenging after 35 years. *Journal of Clinical Microbiology*, *42*(8), 3381–3387.

Hislop, A. D., Taylor, G. S., Sauce, D., & Rickinson, A. B. (2007). Cellular responses to viral infection in humans: lessons from Epstein-Barr virus. *Annu. Rev. Immunol.*, *25*, 587–617.

Holman, C. J., Karger, A. B., Mullan, B. D., Brundage, R. C., & Balfour Jr, H. H. (2012). Quantitative Epstein–Barr virus shedding and its correlation with the risk of post-transplant lymphoproliferative disorder. *Clinical Transplantation*, *26*(5), 741–747.

Hong, G. K., Gulley, M. L., Feng, W.-H., Delecluse, H.-J., Holley-Guthrie, E., & Kenney, S. C. (2005). Epstein-Barr virus lytic infection contributes to lymphoproliferative disease in a SCID mouse model. *Journal of Virology*, *79*(22), 13993–14003.

Huguet, M., Navarro, J.-T., Moltó, J., Ribera, J.-M., & Tapia, G. (2023). Diffuse Large B-Cell Lymphoma in the HIV Setting. *Cancers*, *15*(12), 3191.

Ibrahim, H. A. H., Menasce, L. P., Pomplun, S., Burke, M., Bower, M., & Naresh, K. N. (2010). Epstein–Barr virus (EBV) genotypes among human immunodeficiency virus (HIV)-related B-cell Lymphomas and B-cell post-transplant lymphoproliferative disorders (B-PTLD)–late-onset lymphomas, especially in the HIV setting, are associated with type-B-EBV. *European Journal of Haematology*, *85*(3), 227–230.

Iliyasu, Y., Ayers, L. W., Liman, A. A., Waziri, G. D., & Shehu, S. M. (2014). Epstein–Barr virus association with malignant lymphoma subgroups in Zaria, Nigeria. *Nigerian Journal of Surgical Sciences: Official Journal of the Nigerian Section of International College of Surgeons*, *23*(1), 6.

Imajoh, M., Hashida, Y., Murakami, M., Maeda, A., Sato, T., Fujieda, M., Wakiguchi, H., & Daibata, M. (2012). Characterization of Epstein–Barr virus (EBV) BZLF1 gene promoter variants and comparison of cellular gene expression profiles in Japanese patients with infectious mononucleosis, chronic active EBV infection, and EBV-associated hemophagocytic lymphohistioc. *Journal of Medical Virology*, *84*(6), 940–946.

Jemal, A., Bray, F., Center, M. M., Ferlay, J., Ward, E., & Forman, D. (2011). Global cancer statistics. *CA: A Cancer Journal for Clinicians*, *61*(2), 69–90.

Jiang, L., Gu, Z.-H., Yan, Z.-X., Zhao, X., Xie, Y.-Y., Zhang, Z.-G., Pan, C.-M., Hu, Y., Cai, C.-P., & Dong, Y. (2015). Exome sequencing identifies somatic mutations of DDX3X in natural killer/T-cell lymphoma. *Nature Genetics*, *47*(9), 1061–1066.

Jin, Y., Xie, Z., Lu, G., Yang, S., & Shen, K. (2010). Characterization of variants in the promoter of BZLF1 gene of EBV in nonmalignant EBV-associated diseases in Chinese children. *Virology Journal*, *7*, 1–7.

Jochum, S., Moosmann, A., Lang, S., Hammerschmidt, W., & Zeidler, R. (2012). The EBV immunoevasins vIL-10 and BNLF2a protect newly infected B cells from immune recognition and elimination. *PLoS Pathogens*, *8*(5), e1002704.

Johannessen, I., Perera, S. M., Gallagher, A., Hopwood, P. A., Thomas, J. A., & Crawford, D. H. (2002). Expansion in scid mice of Epstein–Barr virus-associated post-transplantation lymphoproliferative disease biopsy material. *Journal of General Virology*, *83*(1), 173–178.

Kafita, D., Kaile, T., Malyangu, E., Tembo, R., Zulu, E., Chisanga, C., Kalonda, A., Samutela, M., Polepole, P., & Kwenda, G. (2018). Evidence of EBV infection in lymphomas diagnosed in Lusaka, Zambia. *Pan African Medical Journal*, *29*(1), 1–11.

Kaneda, A., Matsusaka, K., Aburatani, H., & Fukayama, M. (2012). Epstein–Barr virus infection as an epigenetic driver of tumorigenesis. *Cancer Research*, *72*(14), 3445–3450.

Kang, M.-S., & Kieff, E. (2015). Epstein–Barr virus latent genes. *Experimental & Molecular Medicine*, *47*(1), e131–e131.

Kempkes, B., & Robertson, E. S. (2015). Epstein-Barr virus latency: current and future perspectives. *Current Opinion in Virology*, *14*, 138–144.

Kenney, S. C. (2007). Reactivation and lytic replication of EBV. *Human Herpesviruses: Biology, Therapy, and Immunoprophylaxis*.

Khan, G., Fitzmaurice, C., Naghavi, M., & Ahmed, L. A. (2020). Global and regional incidence, mortality and disability-adjusted life-years for Epstein-Barr virus-attributable malignancies, 1990–2017. *BMJ Open*, *10*(8), e037505.

Khan, G., & Hashim, M. J. (2014). Global burden of deaths from Epstein-Barr virus attributable malignancies 1990-2010. *Infectious Agents and Cancer*, *9*(1), 1–11.

Kieff, E. (2001). Epstein-Barr virus and its replication. *Fields Virology*, 2511–2574.

Kimura, H., Ito, Y., Suzuki, R., & Nishiyama, Y. (2008). Measuring Epstein–Barr virus (EBV) load: the significance and application for each EBV-associated disease. *Reviews in Medical Virology*, *18*(5), 305–319.

Klein, G., & Ernberg, I. (2007). Comparative aspects of EBV, HHV-8, and previously known DNA tumor viruses in experimental systems. *Human Herpesviruses*, *29*.

Kretzmer, H., Bernhart, S. H., Wang, W., Haake, A., Weniger, M. A., Bergmann, A. K., Betts, M. J., Carrillo-de-Santa-Pau, E., Doose, G., & Gutwein, J. (2015). DNA methylome analysis in Burkitt and follicular lymphomas identifies differentially methylated regions linked to somatic mutation and transcriptional control. *Nature Genetics*, *47*(11), 1316–1325.

Kuri, A., Jacobs, B. M., Vickaryous, N., Pakpoor, J., Middeldorp, J., Giovannoni, G., & Dobson, R. (2020). Epidemiology of Epstein-Barr virus infection and infectious mononucleosis in the United Kingdom. *BMC Public Health*, *20*(1), 1–9.

Kutok, J. L., & Wang, F. (2006). Spectrum of Epstein-Barr virus–associated diseases. *Annu. Rev. Pathol. Mech. Dis.*, *1*, 375–404.

Kwok, H., Chan, K. W., Chan, K. H., & Chiang, A. K. S. (2015). Distribution, persistence and interchange of Epstein-Barr virus strains among PBMC, plasma and saliva of primary infection subjects. *PLoS One*, *10*(3), e0120710.

Kwong, Y. L., Pang, A. W. K., Leung, A. Y. H., Chim, C. S., & Tse, E. (2014). Quantification of circulating Epstein–Barr virus DNA in NK/T-cell lymphoma treated with the SMILE protocol: Diagnostic and prognostic significance. *Leukemia*, *28*(4), 865–870.

Landais, E., Saulquin, X., & Houssaint, E. (2005). The human T cell immune response to Epstein-Barr virus. *The International Journal of Developmental Biology*, *49*(2–3), 285–292.

Lenze, D., Leoncini, L., Hummel, M., Volinia, S., Liu, C. G., Amato, T., De Falco, G., Githanga, J., Horn, H., & Nyagol, J. (2011). The different epidemiologic subtypes of Burkitt lymphoma share a homogenous micro RNA profile distinct from diffuse large B-cell lymphoma. *Leukemia*, *25*(12), 1869–1876.

Leonard, S., Wei, W., Anderton, J., Vockerodt, M., Rowe, M., Murray, P. G., & Woodman, C. B. (2011). Epigenetic and transcriptional changes which follow Epstein-Barr virus infection of germinal center B cells and their relevance to the pathogenesis of Hodgkin's lymphoma. *Journal of Virology*, *85*(18), 9568–9577.

Li, H.-P., & Chang, Y.-S. (2003). Epstein-Barr virus latent membrane protein 1: structure and functions. *Journal of Biomedical Science*, *10*(5), 490–504.

Li, L., Ma, B. B. Y., Chan, A. T. C., Chan, F. K. L., Murray, P., & Tao, Q. (2018). Epstein-Barr

virus-induced epigenetic pathogenesis of viral-associated lymphoepithelioma-like carcinomas and natural killer/T-cell lymphomas. *Pathogens, 7*(3), 63.

Li, W., Duan, X., Chen, X., Zhan, M., Peng, H., Meng, Y., Li, X., Li, X.-Y., Pang, G., & Dou, X. (2023). Immunotherapeutic approaches in EBV-associated nasopharyngeal carcinoma. *Frontiers in Immunology*, *13*, 1079515.

Li, Z., Baccianti, F., Delecluse, S., Tsai, M.-H., Shumilov, A., Cheng, X., Ma, S., Hoffmann, I., Poirey, R., & Delecluse, H.-J. (2021). The Epstein–Barr virus noncoding RNA EBER2 transactivates the UCHL1 deubiquitinase to accelerate cell growth. *Proceedings of the National Academy of Sciences*, *118*(43), e2115508118.

Lieberman, P. M. (2015). Chromatin Structure of Epstein–Barr Virus Latent Episomes. *Epstein Barr Virus Volume 1: One Herpes Virus: Many Diseases*, 71–102.

Lin, J.-C., Lin, S.-C., De, B. K., Chan, W.-P., & Evatt, B. L. (1993). Precision of genotyping of Epstein-Barr virus by polymerase chain reaction using three gene loci (EBNA-2, EBNA-3C, and EBER): predominance of type A virus associated with Hodgkin's disease. *Blood, 81*(12), 3372–3381.

Ling, P. D., Vilchez, R. A., Keitel, W. A., Poston, D. G., Peng, R. S., White, Z. S., Visnegarwala, F., Lewis, D. E., & Butel, J. S. (2003). Epstein-Barr virus DNA loads in adult human immunodeficiency virus type 1-infected patients receiving highly active antiretroviral therapy. *Clinical Infectious Diseases*, *37*(9), 1244–1249.

Liu, Y., Yang, W., Pan, Y., Ji, J., Lu, Z., & Ke, Y. (2016). Genome-wide analysis of Epstein-Barr virus (EBV) isolated from EBV-associated gastric carcinoma (EBVaGC). *Oncotarget, 7*(4), 4903.

Lorenzetti, M. A., Gutiérrez, M. I., Altcheh, J., Moscatelli, G., Moroni, S., Chabay, P. A., & Preciado, M. V. (2009). Epstein–Barr virus BZLF1 gene promoter variants in pediatric patients with acute infectious mononucleosis: its comparison with pediatric lymphomas. *Journal of Medical Virology*, *81*(11), 1912–1917.

Lünemann, A., Rowe, M., & Nadal, D. (2015). Innate immune recognition of EBV. *Epstein Barr Virus Volume 2: One Herpes Virus: Many Diseases*, 265–287.

Luo, B., Wang, Y., Wang, X.-F., Liang, H., Yan, L.-P., Huang, B.-H., & Zhao, P. (2005). Expression of Epstein-Barr virus genes in EBV-associated gastric carcinomas. *World Journal of Gastroenterology: WJG*, *11*(5), 629.

Lv, K., Yin, T., Yu, M., Chen, Z., Zhou, Y., & Li, F. (2022). Treatment advances in EBV related lymphoproliferative diseases. *Frontiers in Oncology*, *12*, 838817.

Ma, S.-D., Hegde, S., Young, K. H., Sullivan, R., Rajesh, D., Zhou, Y., Jankowska-Gan, E., Burlingham, W. J., Sun, X., & Gulley, M. L. (2011). A new model of Epstein-Barr virus infection reveals an important role for early lytic viral protein expression in the development of lymphomas. *Journal of Virology*, *85*(1), 165–177.

Maeda, E., Akahane, M., Kiryu, S., Kato, N., Yoshikawa, T., Hayashi, N., Aoki, S., Minami, M., Uozaki, H., & Fukayama, M. (2009). Spectrum of Epstein-Barr virus-related diseases: a pictorial review. *Japanese Journal of Radiology*, *27*, 4–19.

Maggioncalda, A., Malik, N., Shenoy, P., Smith, M., Sinha, R., & Flowers, C. R. (2011). Clinical, molecular, and environmental risk factors for Hodgkin lymphoma. *Advances in Hematology*, *2011*.

Mainou, B. A., & Raab-Traub, N. (2006). LMP1 strain variants: biological and molecular properties. *Journal of Virology*, *80*(13), 6458–6468.

Martinez, O. M., & Krams, S. M. (2017). The immune response to Epstein Barr virus and implications for posttransplant lymphoproliferative disorder. *Transplantation*, *101*(9), 2009.

Massini, G., Siemer, D., & Hohaus, S. (2009). EBV in Hodgkin lymphoma. *Mediterranean Journal of Hematology and Infectious Diseases*, *1*(2).

Maurmann, S., Fricke, L., Wagner, H.-J., Schlenke, P., Hennig, H., Steinhoff, J., & Jabs, W. J. (2003). Molecular parameters for precise diagnosis of asymptomatic Epstein-Barr virus reactivation in healthy carriers. *Journal of Clinical Microbiology*, *41*(12), 5419–5428.

Menon, M. P., Pittaluga, S., & Jaffe, E. S. (2012). The histological and biological spectrum of diffuse large B-cell lymphoma in the WHO classification. *Cancer Journal (Sudbury, Mass.)*, *18*(5), 411.

Miller, J. L., & Grant, P. A. (2012). The role of DNA methylation and histone modifications in transcriptional regulation in humans. *Epigenetics: Development and Disease*, 289–317.

Miller, W. E., Edwards, R. H., Walling, D. M., & Raab-Traub, N. (1994). Sequence variation in the Epstein–Barr virus latent membrane protein 1. *Journal of General Virology*, *75*(10), 2729–2740.

Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., Von Haeseler, A., & Lanfear, R. (2020). IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Molecular Biology and Evolution*, *37*(5), 1530–1534.

Misganaw, A., Naghavi, M., Walker, A., Mirkuzie, A. H., Giref, A. Z., Berheto, T. M., Waktola, E. A., Kempen, J. H., Eticha, G. T., & Wolde, T. K. (2022). Progress in health among regions of Ethiopia, 1990–2019: a subnational country analysis for the Global Burden of Disease Study 2019. *The Lancet*, *399*(10332), 1322–1335.

Mitarnun, W., Pradutkanchana, J., Takao, S., Saechan, V., Suwiwat, S., & Ishida, T. (2002). Epstein-barr virus-associated non-Hodgkin's lymphoma of B-cell origin, Hodgkin's disease, acute leukemia, and systemic lupus erythematosus: a serologic and molecular analysis. *Journal of the Medical Association of Thailand= Chotmaihet Thangphaet*, *85*(5), 552–559.

Monteiro, T. A. F., Costa, I. B., Costa, I. B., Corrêa, T. L. dos S., Coelho, B. M. R., Silva, A. E. S., Ramos, F. L. de P., Monteiro, J. L. F., Siqueira, J. A. M., & Gabbay, Y. B. (2020). Genotypes of Epstein–Barr virus (EBV1/EBV2) in individuals with infectious mononucleosis in the metropolitan area of Belém, Brazil, between 2005 and 2016. *Brazilian Journal of Infectious Diseases*, *24*, 322–329.

Monteiro, T. A. F., Costa, I. B., Costa, I. B., dos Santos Corrêa, T. L., Coelho, B. M. R., Polaro, A. A., da Silva, A. E. S., de Paula Ramos, F. L., Martins Filho, A. J., & Monteiro, J. L. F. (2019). *Gene EBNA3C: Type of Infection by EBV (EBV1 and EBV2) Correlation With Clinical and Biochemical Parameters (AST, ALT and GGT) in Individuals With Infectious Mononucleosis in the Metropolitan Area of Belém, Pará, 2005-2016*.

Moore, M. D., Cannon, M. J., Sewall, A., Finlayson, M., Okimoto, M., & Nemerow, G. R. (1991). Inhibition of Epstein-Barr virus infection in vitro and in vivo by soluble CR2 (CD21) containing two short consensus repeats. *Journal of Virology*, *65*(7), 3559–3565.

Münz, C. (2015). *Epstein Barr virus volume 2: one herpes virus: many diseases* (Vol. 391). Springer.

Münz, C., Bickham, K. L., Subklewe, M., Tsang, M. L., Chahroudi, A., Kurilla, M. G., Zhang, D., O'Donnell, M., & Steinman, R. M. (2000). Human CD4+ T lymphocytes consistently respond to the latent Epstein-Barr virus nuclear antigen EBNA1. *The Journal of Experimental Medicine*, *191*(10), 1649–1660.

Murata, T., Kondo, Y., Sugimoto, A., Kawashima, D., Saito, S., Isomura, H., Kanda, T., & Tsurumi, T. (2012). Epigenetic histone modification of Epstein-Barr virus BZLF1 promoter during latency and reactivation in Raji cells. *Journal of Virology*, *86*(9), 4752–4761.

Murata, T., Narita, Y., Sugimoto, A., Kawashima, D., Kanda, T., & Tsurumi, T. (2013). Contribution of myocyte enhancer factor 2 family transcription factors to BZLF1 expression in Epstein-Barr virus reactivation from latency. *Journal of Virology*, *87*(18), 10148–10162.

Murata, T., Sugimoto, A., Inagaki, T., Yanagi, Y., Watanabe, T., Sato, Y., & Kimura, H. (2021). Molecular basis of epstein–barr virus latency establishment and lytic reactivation. *Viruses*, *13*(12), 2344.

Mushtaq, S., Akhtar, N., Jamal, S., Mamoon, N., Khadim, T., Sarfaraz, T., & Waqar, A. (2008). Malignant lymphomas in Pakistan according to the WHO classification of lymphoid neoplasms. *Asian Pac J Cancer Prev*, *9*(2), 229–232.

Nanbo, A., Inoue, K., Adachi-Takasawa, K., & Takada, K. (2002). Epstein–Barr virus RNA confers resistance to interferon-α-induced apoptosis in Burkitt's lymphoma. *The EMBO Journal*, *21*(5), 954–965.

Neparidze, N., & Lacy, J. (2014). Malignancies associated with epstein-barr virus: pathobiology, clinical features, and evolving treatments. *Clin Adv Hematol Oncol*, *12*(6), 358–371.

Neves, M., Marinho-Dias, J., Ribeiro, J., & Sousa, H. (2017). Epstein–Barr virus strains and variations: Geographic or disease-specific variants? *Journal of Medical Virology*, *89*(3), 373–387.

Niedobitek, G., Meru, N., & Delecluse, H. (2001). Epstein-Barr virus infection and human malignancies. *International Journal of Experimental Pathology*, *82*(3), 149–170.

Niller, H. H., Banati, F., Salamon, D., & Minarovits, J. (2016). Epigenetic alterations in Epstein-Barr virus-associated diseases. *Patho-Epigenetics of Infectious Disease*, 39–69.

Niller, H. H., Tarnai, Z., Decsi, G., Zsedényi, Á., Bánáti, F., & Minarovits, J. (2014). Role of epigenetics in EBV regulation and pathogenesis. *Future Microbiology*, *9*(6), 747–756.

Niller, J. M., Salamon, D., Banati, F., Szenthe, K., Wolf, H., Helmut, H., Gerle, B., Koroknai, A., Fejer, G., & Bakos, A. (2007). Acetylated Histone H3 and H4 Mark the. *J. Virol*, *81*(23), 13242.

Nystad, T. W., & Myrmel, H. (2007). Prevalence of primary versus reactivated Epstein-Barr virus infection in patients with VCA IgG-, VCA IgM-and EBNA-1-antibodies and suspected infectious mononucleosis. *Journal of Clinical Virology*, *38*(4), 292–297.

Ocheni, S., Olusina, D. B., Oyekunle, A. A., Ibegbulam, O. G., Kröger, N., Bacher, U., & Zander, A. R. (2010). EBV-associated malignancies. *Open Infectious Diseases Journal*, *4*(1), 101–112.

Odumade, O. A., Hogquist, K. A., & Balfour Jr, H. H. (2011). Progress and problems in understanding and managing primary Epstein-Barr virus infections. *Clinical Microbiology Reviews*, *24*(1), 193–209.

Oh, J.-K., & Weiderpass, E. (2014). Infection and cancer: global distribution and burden of diseases. *Annals of Global Health*, *80*(5), 384–392.

Ojeniyi, O. O. (2017). *INVESTIGATING THE MECHANISM OF CELLULAR GENE ACTIVATION AND REPRESSION BY THE EBV TRANSCRIPTION FACTOR EBNA 2*.

Olu-Eddo, A. N., & Omoti, C. E. (2011). Hodgkin lymphoma: Clinicopathologic features in Benin City, Nigeria and update on its biology and classification. *Nigerian Journal of Clinical Practice*, *14*(4), 440–444.

Orem, J., Mbidde, E. K., Lambert, B., De Sanjose, S., & Weiderpass, E. (2007). Burkitt\'s lymphoma in Africa, a review of the epidemiology and etiology. *African Health Sciences*, *7*(3).

Palma, I., Sánchez, A. E., Jiménez-Hernández, E., Alvarez-Rodríguez, F., Nava-Frias, M., Valencia-Mayoral, P., Salinas-Lara, C., Velazquez-Guadarrama, N., Portilla-Aguilar, J., & Pena, R. Y. (2013). Detection of Epstein-Barr virus and genotyping based on EBNA2 protein in Mexican patients with Hodgkin lymphoma: a comparative study in children and adults. *Clinical Lymphoma Myeloma and Leukemia*, *13*(3), 266–272.

Palser, A. L., Grayson, N. E., White, R. E., Corton, C., Correia, S., Ba abdullah, M. M., Watson, S. J., Cotten, M., Arrand, J. R., & Murray, P. G. (2015). Genome diversity of Epstein-Barr virus from multiple tumor types and normal infection. *Journal of Virology*, *89*(10), 5222–5237.

Papesch, M., & Watkins, R. (2001). Epstein-Barr virus infectious mononucleosis. *Clinical Otolaryngology & Allied Sciences*, *26*(1), 3–8.

Parkin, D. M. (2006). The global health burden of infection-associated cancers in the year 2002. *International Journal of Cancer*, *118*(12), 3030–3044.

Paulson, E. J., Fingeroth, J. D., Yates, J. L., & Speck, S. H. (2002). Methylation of the EBV genome and establishment of restricted latency in low-passage EBV-infected 293 epithelial cells. *Virology*, *299*(1), 109–121.

Paya, C. V, Fung, J. J., Nalesnik, M. A., Kieff, E., Green, M., Gores, G., Habermann, T. M., Wiesner, R. H., Swinnen, L. J., & Woodle, E. S. (1999). Epstein-Barr virus-induced posttransplant lymphoproliferative disorders. *Transplantation*, *68*(10), 1517–1525.

Pei, Y., Wong, J. H. Y., & Robertson, E. S. (2020). Targeted therapies for Epstein-Barr virus-associated lymphomas. *Cancers*, *12*(9), 2565.

Peng, Q., Wang, L., Qin, Z., Wang, J., Zheng, X., Wei, L., Zhang, X., Zhang, X., Liu, C., & Li, Z. (2020). Phase separation of Epstein-Barr virus EBNA2 and its coactivator EBNALP controls gene expression. *Journal of Virology*, *94*(7), 10–1128.

Peng, R.-J., Han, B.-W., Cai, Q.-Q., Zuo, X.-Y., Xia, T., Chen, J.-R., Feng, L.-N., Lim, J. Q., Chen, S.-W., & Zeng, M.-S. (2019). Genomic and transcriptomic landscapes of Epstein-Barr virus in extranodal natural killer T-cell lymphoma. *Leukemia*, *33*(6), 1451–1462.

Pereira, L. M. S., França, E. dos S., Costa, I. B., Lima, I. T., Freire, A. B. C., Ramos, F. L. de P., Monteiro, T. A. F., Macedo, O., Sousa, R. C. M., & Freitas, F. B. (2022). Epstein–Barr Virus (EBV) Genotypes Associated with the Immunopathological Profile of People Living with HIV-1: Immunological Aspects of Primary EBV Infection. *Viruses*, *14*(2), 168.

Petrara, M. R., Giunco, S., Serraino, D., Dolcetti, R., & De Rossi, A. (2015). Post-transplant

lymphoproliferative disorders: from epidemiology to pathogenesis-driven treatment. *Cancer Letters*, *369*(1), 37–44.

Pike, B. L., Greiner, T. C., Wang, X., Weisenburger, D. D., Hsu, Y.-H., Renaud, G., Wolfsberg, T. G., Kim, M., Weisenberger, D. J., & Siegmund, K. D. (2008). DNA methylation profiles in diffuse large B-cell lymphoma and their relationship to gene expression status. *Leukemia*, *22*(5), 1035–1043.

Pileri, S. A., Ascani, S., Leoncini, L., Sabattini, E., Zinzani, P. L., Piccaluga, P. P., Pileri, A., Giunti, M., Falini, B., & Bolis, G. B. (2002). Hodgkin's lymphoma: the pathologist's viewpoint. *Journal of Clinical Pathology*, *55*(3), 162–176.

Pontejo, S. M., Murphy, P. M., & Pease, J. E. (2018). Chemokine subversion by human herpesviruses. *Journal of Innate Immunity*, *10*(5–6), 465–478.

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.

Raab-Traub, N. (2002). Epstein–Barr virus in the pathogenesis of NPC. *Seminars in Cancer Biology*, *12*(6), 431–441.

Ravanfar, P., Mendoza, N., Satyaprakash, A., Creed, R., & Tyring, S. (2016). 17 Infections of the Skin and Mucosa. *Lennette's Laboratory Diagnosis of Viral Infections*, 285.

Razzouk, B. I., Srinivas, S., Sample, C. E., Singh, V., & Sixbey, J. W. (1996). Epstein-Barr virus DNA recombination and loss in sporadic Burkitt's lymphoma. *Journal of Infectious Diseases*, *173*(3), 529–535.

Renzette, N., Somasundaran, M., Brewster, F., Coderre, J., Weiss, E. R., McManus, M., Greenough, T., Tabak, B., Garber, M., & Kowalik, T. F. (2014). Epstein-Barr virus latent membrane protein 1 genetic variability in peripheral blood B cells and oropharyngeal fluids. *Journal of Virology*, *88*(7), 3744–3755.

Ressing, M. E., van Gent, M., Gram, A. M., Hooykaas, M. J. G., Piersma, S. J., & Wiertz, E. J. H. J. (2015). Immune evasion by Epstein-Barr virus. *Epstein Barr Virus Volume 2: One Herpes Virus: Many Diseases*, 355–381.

Rickinson, A. B. (2007). Epstein-barr virus. *Fields Virology*, *5*, 2655–2700.

Robertson, K. D., & Wolffe, A. P. (2000). DNA methylation in health and disease. *Nature Reviews Genetics*, *1*(1), 11–19.

Roschewski, M., & Wilson, W. H. (2012). EBV-associated lymphomas in adults. *Best Practice & Research Clinical Haematology*, *25*(1), 75–89.

Rosemarie, Q., & Sugden, B. (2020). Epstein–Barr virus: how its lytic phase contributes to oncogenesis. *Microorganisms*, *8*(11), 1824.

Roughan, J. E., & Thorley-Lawson, D. A. (2009). The intersection of Epstein-Barr virus with the germinal center. *Journal of Virology*, *83*(8), 3968–3976.

Rowe, M., & Zuo, J. (2010). Immune responses to Epstein–Barr virus: molecular interactions in the virus evasion of CD8+ T cell immunity. *Microbes and Infection*, *12*(3), 173–181.

Rüeger, S., Hammer, C., Loetscher, A., McLaren, P. J., Lawless, D., Naret, O., Khanna, N., Bernasconi, E., Cavassini, M., & Günthard, H. F. (2021). The influence of human genetic variation on Epstein–Barr virus sequence diversity. *Scientific Reports*, *11*(1), 4586.

Ryan, J. L., Fan, H., Swinnen, L. J., Schichman, S. A., Raab-Traub, N., Covington, M., Elmore, S., & Gulley, M. L. (2004). Epstein-Barr Virus (EBV) DNA in plasma is not encapsidated in patients with EBV-related malignancies. *Diagnostic Molecular Pathology*, *13*(2), 61–68.

Sachithanandham, J., Kannangai, R., Pulimood, S. A., Desai, A., Abraham, A. M., Abraham, O. C., Ravi, V., Samuel, P., & Sridharan, G. (2014). Significance of Epstein-Barr virus (HHV-4) and CMV (HHV-5) infection among subtype-C human immunodeficiency virus-infected individuals. *Indian Journal of Medical Microbiology*, *32*(3), 261–269.

Saha, A., & Robertson, E. S. (2013). Impact of EBV essential nuclear protein EBNA-3C on B-cell proliferation and apoptosis. *Future Microbiology*, *8*(3), 323–352.

Salahuddin, S., Khan, J., Azhar, J., Whitehurst, C. B., Qadri, I., Shackelford, J., Pagano, J. S., Muhammad, D., & Richards, K. L. (2018). Prevalence of Epstein–Barr virus genotypes in Pakistani lymphoma patients. *Asian Pacific Journal of Cancer Prevention: APJCP*, *19*(11), 3153.

Saldanha, A. J. (2004). Java Treeview—extensible visualization of microarray data. *Bioinformatics*, *20*(17), 3246–3248.

Sample, J., Young, L., Martin, B., Chatman, T., Kieff, E., Rickinson, A., & Kieff, E. (1990). Epstein-Barr virus types 1 and 2 differ in their EBNA-3A, EBNA-3B, and EBNA-3C genes. *Journal of Virology*, *64*(9), 4084–4092.

Santpere, G., Darre, F., Blanco, S., Alcami, A., Villoslada, P., Mar Albà, M., & Navarro, A. (2014). Genome-wide analysis of wild-type Epstein–Barr virus genomes derived from healthy individuals of the 1000 Genomes Project. *Genome Biology and Evolution*, *6*(4), 846–860.

Scott, R. S. (2017). Epstein–Barr virus: a master epigenetic manipulator. *Current Opinion in Virology*, *26*, 74–80.

Shaklawoon, K., Altagazi, N., Altorjman, F., Alturki, A., Eltaweel, M., & Alqawi, O. (2020). Molecular detection of Epstein–Barr virus in different types of lymphoma. *Molecular Biology Reports*, *47*(3), 1803–1807.

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

Shannon-Lowe, C., Rickinson, A. B., & Bell, A. I. (2017). Epstein–Barr virus-associated lymphomas. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1732), 20160271.

Shannon-Lowe, C., & Rowe, M. (2014). Epstein Barr virus entry; kissing and conjugation. *Current Opinion in Virology*, *4*, 78–84.

Shareena, G., & Kumar, D. (2023). Epigenetics of Epstein Barr virus—A review. *Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease*, *1869*(8), 166838.

Shibata, D., & Weiss, L. M. (1992). Epstein-Barr virus-associated gastric adenocarcinoma. *The American Journal of Pathology*, *140*(4), 769.

Shinozaki-Ushiku, A., Kunita, A., & Fukayama, M. (2015). Update on Epstein-Barr virus and gastric cancer. *International Journal of Oncology*, *46*(4), 1421–1434.

Sinclair, A. J. (2013). Epigenetic control of Epstein–Barr virus transcription–relevance to viral

life cycle? *Frontiers in Genetics*, *4*, 161.

Sixbey, J., Chesney, P. J., Shirley, P., Buntin, D., & Resnick, L. (1989). Detection of a second widespread strain of Epstein-Barr virus. *The Lancet*, *334*(8666), 761–765.

Sixbey, J. W. (2000). Epstein-Barr virus DNA loss from tumour cells and the geography of Burkitt's lymphoma. *EPSTEIN BARR VIRUS REPORT*, *7*(2).

Skalsky, R. L., & Cullen, B. R. (2015). EBV noncoding RNAs. *Epstein Barr Virus Volume 2: One Herpes Virus: Many Diseases*, 181–217.

Smatti, M. K., Al-Sadeq, D. W., Ali, N. H., Pintus, G., Abou-Saleh, H., & Nasrallah, G. K. (2018). Epstein–Barr virus epidemiology, serology, and genetic variability of LMP-1 oncogene among healthy population: an update. *Frontiers in Oncology*, *8*, 211.

Smatti, M. K., Yassine, H. M., AbuOdeh, R., AlMarawani, A., Taleb, S. A., Althani, A. A., & Nasrallah, G. K. (2017). Prevalence and molecular profiling of Epstein Barr virus (EBV) among healthy blood donors from different nationalities in Qatar. *PLoS One*, *12*(12), e0189033.

Smith, Z. D., & Meissner, A. (2013). DNA methylation: roles in mammalian development. *Nature Reviews Genetics*, *14*(3), 204–220.

Sohn, D.-H., Sohn, H.-J., Lee, H.-J., Lee, S.-D., Kim, S., Hyun, S.-J., Cho, H.-I., Cho, S.-G., Lee, S.-K., & Kim, T.-G. (2015). Measurement of CD8+ and CD4+ T cell frequencies specific for EBV LMP1 and LMP2a using mRNA-transfected DCs. *Plos One*, *10*(5), e0127899.

Speck, S. H., & Ganem, D. (2010). Viral latency and its regulation: lessons from the γ-herpesviruses. *Cell Host & Microbe*, *8*(1), 100–115.

Spitzer, M., Wildenhain, J., Rappsilber, J., & Tyers, M. (2014). BoxPlotR: a web tool for generation of box plots. *Nature Methods*, *11*(2), 121–122.

Stanland, L. J., & Luftig, M. A. (2020). The role of EBV-induced hypermethylation in gastric cancer tumorigenesis. *Viruses*, *12*(11), 1222.

Stephenson, H. N., Herzig, A., & Zychlinsky, A. (2016). Beyond the grave: When is cell death critical for immunity to infection? *Current Opinion in Immunology*, *38*, 59–66.

Stern-Ginossar, N., Elefant, N., Zimmermann, A., Wolf, D. G., Saleh, N., Biton, M., Horwitz, E., Prokocimer, Z., Prichard, M., & Hahn, G. (2007). Host immune system gene targeting by a viral miRNA. *Science*, *317*(5836), 376–381.

Strockbine, L. D., Cohen, J. I., Farrah, T., Lyman, S. D., Wagener, F., DuBose, R. F., Armitage, R. J., & Spriggs, M. K. (1998). The Epstein-Barr virus BARF1 gene encodes a novel, soluble colony-stimulating factor-1 receptor. *Journal of Virology*, *72*(5), 4015–4021.

Strowig, T., Brilot, F., Arrey, F., Bougras, G., Thomas, D., Muller, W. A., & Münz, C. (2008). Tonsilar NK cells restrict B cell transformation by the Epstein-Barr virus via IFN-γ. *PLoS Pathogens*, *4*(2), e27.

Su, Z. Y., Siak, P. Y., Leong, C.-O., & Cheah, S.-C. (2023). The role of Epstein–Barr virus in nasopharyngeal carcinoma. *Frontiers in Microbiology*, *14*, 1116143.

Sugiura, M., Imai, S., Tokunaga, M., Koizumi, S., Uchizawa, M., Okamoto, K., & Osato, T. (1996). Transcriptional analysis of Epstein-Barr virus gene expression in EBV-positive gastric carcinoma: unique viral latency in the tumour cells. *British Journal of Cancer*, *74*(4),

625–631.

Sun, K., Jia, K., Lv, H., Wang, S.-Q., Wu, Y., Lei, H., & Chen, X. (2020). EBV-positive gastric cancer: current knowledge and future perspectives. *Frontiers in Oncology*, *10*, 583463.

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: A Cancer Journal for Clinicians*, *71*(3), 209–249.

Swaminathan, S., & Kenney, S. (2009). The Epstein–Barr virus lytic life cycle. *DNA Tumor Viruses*, 285–315.

Swerdlow, S. H., Campo, E., Harris, N. L., Jaffe, E. S., Pileri, S. A., Stein, H., Thiele, J., & Vardiman, J. W. (2008). *WHO classification of tumours of haematopoietic and lymphoid tissues* (Vol. 2). International agency for research on cancer Lyon, France.

Swerdlow, S. H., Campo, E., Pileri, S. A., Lee Harris, N., Stein, H., Siebert, R., Advani, R., Ghielmini, M., Salles, G. A., Zelenetz, A. D., & Jaffe, E. S. (2016). The 2016 revision of the World Health Organization classification of lymphoid neoplasms. *Blood*, *127*(20), 2375–2390. https://doi.org/10.1182/blood-2016-01-643569

Tabibzadeh, A., Niya, M. H. K., Esghaei, M., Bokharaei-Salim, F., Ataei-Pirkooh, A., Kiani, S. J., & Monavari, S. H. R. (2020). Molecular epidemiology of epstein-barr virus (ebv) in patients with hematologic malignancies. *Asian Pacific Journal of Cancer Prevention: APJCP*, *21*(3), 693.

Tabtieng, T., & Gaglia, M. M. (2018). Emerging proviral roles of caspases during lytic replication of gammaherpesviruses. *Journal of Virology*, *92*(19), 10–1128.

Takacs, M., Segesdi, J., Banati, F., Koroknai, A., Wolf, H., Niller, H. H., & Minarovits, J. (2009). The importance of epigenetic alterations in the development of epstein-barr virus-related lymphomas. *Mediterranean Journal of Hematology and Infectious Diseases*, *1*(2).

Tang, C.-L., Li, X.-Z., Zhou, T., Deng, C.-M., Jiang, C.-T., Zhang, Y.-M., Liao, Y., Wang, T.-M., He, Y.-Q., & Xue, W.-Q. (2024). EBV DNA methylation profiles and its application in distinguishing nasopharyngeal carcinoma and nasal NK/T-cell lymphoma. *Clinical Epigenetics*, *16*(1), 1–13.

Tao, Q., & Robertson, K. D. (2003). Stealth technology: how Epstein–Barr virus utilizes DNA methylation to cloak itself from immune detection. *Clinical Immunology*, *109*(1), 53–63.

Tao, Q., Young, L. S., Woodman, C. B. J., & Murray, P. G. (2006). Epstein-Barr virus (EBV) and its associated human cancers--genetics, epigenetics, pathobiology and novel therapeutics. *Frontiers in Bioscience-Landmark*, *11*(3), 2672–2713.

Tefera, B., Assefa, M., Abebe, B., & Rauch, D. (2016). Patterns of cancer in University of Gondar Hospital: north-west Ethiopia. *J Oncol Med Pract*, *1*(106), 2.

Telenti, A., Uehlinger, D. E., Marchesi, F., Germann, D., Malinverni, R., & Matter, L. (1993). Epstein-Barr virus infection in HIV-positive patients. *European Journal of Clinical Microbiology and Infectious Diseases*, *12*, 601–609.

Tempera, I., & Lieberman, P. M. (2010). Chromatin organization of gammaherpesvirus latent genomes. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms*, *1799*(3–4), 236–245.

Tempera, I., & Lieberman, P. M. (2014). Epigenetic regulation of EBV persistence and oncogenesis. *Seminars in Cancer Biology*, *26*, 22–29.

Teras, L. R., DeSantis, C. E., Cerhan, J. R., Morton, L. M., Jemal, A., & Flowers, C. R. (2016). 2016 US lymphoid malignancy statistics by World Health Organization subtypes. *CA: A Cancer Journal for Clinicians*, *66*(6), 443–459.

Thompson, M. P., & Kurzrock, R. (2004). Epstein-Barr virus and cancer. *Clinical Cancer Research*, *10*(3), 803–821.

Thorley-Lawson, D. A. (2015). EBV persistence—introducing the virus. *Epstein Barr Virus Volume 1: One Herpes Virus: Many Diseases*, 151–209.

Thorley-Lawson, D. A., & Gross, A. (2004). Persistence of the Epstein–Barr virus and the origins of associated lymphomas. *New England Journal of Medicine*, *350*(13), 1328–1337.

Thorley-Lawson, D., Deitsch, K. W., Duca, K. A., & Torgbor, C. (2016). The link between Plasmodium falciparum malaria and endemic Burkitt's lymphoma—new insight into a 50-year-old enigma. *PLoS Pathogens*, *12*(1), e1005331.

Timotewos, G., Solomon, A., Mathewos, A., Addissie, A., Bogale, S., Wondemagegnehu, T., Aynalem, A., Ayalnesh, B., Dagnechew, H., & Bireda, W. (2018). First data from a population based cancer registry in Ethiopia. *Cancer Epidemiology*, *53*, 93–98.

Tinguely, M., Brundler, M. A., Gogus, S., Kerl, K., & Borisch, B. (2000). Epstein-Barr virus association in pediatric abdominal non-Hodgkin-lymphomas from Turkey. *ARCHIVUM IMMUNOLOGIAE ET THERAPIAE EXPERIMENTALIS-ENGLISH EDITION-*, *48*(4), 317–322.

Tiwawech, D., Srivatanakul, P., Karalak, A., & Ishida, T. (2008). Association between EBNA2 and LMP1 subtypes of Epstein-Barr virus and nasopharyngeal carcinoma in Thais. *Journal of Clinical Virology*, *42*(1), 1–6.

Trimèche, M., Bonnet, C., Korbi, S., Boniver, J., & Leval, L. de. (2007). Association between Epstein-Barr virus and Hodgkin's lymphoma in Belgium: a pathological and virological study. *Leukemia & Lymphoma*, *48*(7), 1323–1331.

Trottier, H., Buteau, C., Robitaille, N., Duval, M., Tucci, M., Lacroix, J., & Alfieri, C. (2012). Transfusion-related Epstein-Barr virus infection among stem cell transplant recipients: a retrospective cohort study in children. *Transfusion*, *52*(12), 2653–2663.

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., Tsang, C. M., Pang, P. S., Zhang, G., Chen, H., & Lo, K. W. (2012). The biology of EBV infection in human epithelial cells. *Seminars in Cancer Biology*, *22*(2), 137–143.

Tzellos, S., & Farrell, P. J. (2012). Epstein-Barr virus sequence variation—biology and disease. *Pathogens*, *1*(2), 156–175.

Umakanthan, S., & Bukelo, M. M. (2021). Molecular Genetics in Epstein–Barr Virus-Associated Malignancies. *Life*, *11*(7), 593.

Uner, A., Akyurek, N., Saglam, A., Abdullazade, S., Uzum, N., Onder, S., Barista, I., & Benekli, M. (2011). The presence of Epstein–Barr virus (EBV) in diffuse large B-cell lymphomas

(DLBCLs) in Turkey: special emphasis on 'EBV-positive DLBCL of the elderly.' *Apmis*, *119*(4-5), 309–316.

Van den Bosch, C. A. (2004). Is endemic Burkitt's lymphoma an alliance between three infections and a tumour promoter? *The Lancet Oncology*, *5*(12), 738–746.

Vetsika, E.-K., & Callan, M. (2004). Infectious mononucleosis and Epstein-Barr virus. *Expert Reviews in Molecular Medicine*, *6*(23), 1–16.

Vockerodt, M., Cader, F. Z., Shannon-Lowe, C., & Murray, P. (2014). Epstein-Barr virus and the origin of Hodgkin lymphoma. *Chinese Journal of Cancer*, *33*(12), 591.

Wachsman, J. T. (1997). DNA methylation and the association between genetic and epigenetic changes: relation to carcinogenesis. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, *375*(1), 1–8.

Walling, D. M., Etienne, W., Ray, A. J., Flaitz, C. M., & Nichols, C. M. (2004). Persistence and transition of Epstein-Barr virus genotypes in the pathogenesis of oral hairy leukoplakia. *The Journal of Infectious Diseases*, *190*(2), 387–395.

Wan, Z., Chen, Y., Hui, J., Guo, Y., Peng, X., Wang, M., Hu, C., Xie, Y., Su, J., & Huang, Y. (2023). Epstein-Barr virus variation in people living with human immunodeficiency virus in southeastern China. *Virology Journal*, *20*(1), 1–11.

Wang, M., Yu, F., Wu, W., Wang, Y., Ding, H., & Qian, L. (2018). Epstein-Barr virus-encoded microRNAs as regulators in host immune responses. *International Journal of Biological Sciences*, *14*(5), 565.

Wang, S., Medeiros, L. J., Xu-Monette, Z. Y., Zhang, S., O'Malley, D. P., Orazi, A., Zuo, Z., Bueso-Ramos, C. E., Yin, C. C., & Liu, Z. (2014). Epstein-Barr virus–positive nodular lymphocyte predominant Hodgkin lymphoma. *Annals of Diagnostic Pathology*, *18*(4), 203–209.

Wang, Y.-J., Ba, Y., Chen, Q.-Y., & Han, Y.-Q. (2021). Clinical Significance of Peripheral Blood EBV-DNA Determination and Genotyping in Lymphoma Patients. *Zhongguo Shi Yan Xue Ye Xue Za Zhi*, *29*(6), 1802–1806.

Wegner, F., Lassalle, F., Depledge, D. P., Balloux, F., & Breuer, J. (2019). Coevolution of sites under immune selection shapes Epstein–Barr virus population structure. *Molecular Biology and Evolution*, *36*(11), 2512–2521.

Wen, W., Iwakiri, D., Yamamoto, K., Maruo, S., Kanda, T., & Takada, K. (2007). Epstein-Barr virus BZLF1 gene, a switch from latency to lytic infection, is expressed as an immediate-early gene after primary infection of B lymphocytes. *Journal of Virology*, *81*(2), 1037–1042.

Williams, H., & Crawford, D. H. (2006). Epstein-Barr virus: the impact of scientific advances on clinical practice. *Blood*, *107*(3), 862–869.

Williams, H., McAulay, K., Macsween, K. F., Gallacher, N. J., Higgins, C. D., Harrison, N., Swerdlow, A. J., & Crawford, D. H. (2005). The immune response to primary EBV infection: a role for natural killer cells. *British Journal of Haematology*, *129*(2), 266–274.

Winter, J. R., Taylor, G. S., Thomas, O. G., Jackson, C., Lewis, J. E. A., & Stagg, H. R. (2019). Predictors of Epstein-Barr virus serostatus in young people in England. *BMC Infectious Diseases*, *19*, 1–9.

Woellmer, A., Arteaga-Salas, J. M., & Hammerschmidt, W. (2012). *BZLF1 governs CpG-methylated chromatin of Epstein-Barr Virus reversing epigenetic repression*.

Wong, Y., Meehan, M. T., Burrows, S. R., Doolan, D. L., & Miles, J. J. (2022). Estimating the global burden of Epstein–Barr virus-related cancers. *Journal of Cancer Research and Clinical Oncology*, 1–16.

Wu, L., Borza, C. M., & Hutt-Fletcher, L. M. (2005). Mutations of Epstein-Barr virus gH that are differentially able to support fusion with B cells or epithelial cells. *Journal of Virology*, *79*(17), 10923–10930.

Wu, L., & Hutt-Fletcher, L. M. (2007). Point mutations in EBV gH that abrogate or differentially affect B cell and epithelial cell fusion. *Virology*, *363*(1), 148–155.

Xiao, K., Yu, Z., Li, X., Li, X., Tang, K., Tu, C., Qi, P., Liao, Q., Chen, P., & Zeng, Z. (2016). Genome-wide analysis of Epstein-Barr virus (EBV) integration and strain in C666-1 and Raji cells. *Journal of Cancer*, *7*(2), 214.

Yakovleva, L. S., Senyuta, N. B., Goncharova, E. V, Scherback, L. N., Smirnova, R. V, Pavlish, O. A., & Gurtsevitch, V. E. (2015). Epstein–Barr Virus LMP1 oncogene variants in cell lines of different origin. *Molecular Biology*, *49*, 714–722.

Yang, H.-J., Huang, T.-J., Yang, C.-F., Peng, L.-X., Liu, R.-Y., Yang, G.-D., Chu, Q.-Q., Huang, J.-L., Liu, N., & Huang, H.-B. (2013). Comprehensive profiling of Epstein-Barr virus-encoded miRNA species associated with specific latency types in tumor cells. *Virology Journal*, *10*(1), 1–13.

Yao, Q. Y., Tierney, R. J., Croom-Carter, D., Dukers, D., Cooper, G. M., Ellis, C. J., Rowe, M., & Rickinson, A. B. (1996). Frequency of multiple Epstein-Barr virus infections in T-cell-immunocompromised individuals. *Journal of Virology*, *70*(8), 4884–4894.

Yeole, B. B. (2008). Trends in the incidence of Non-Hodgkin's lymphoma in India. *Asian Pac J Cancer Prev*, *9*(3), 433–436.

Yin, H., Qu, J., Peng, Q., & Gan, R. (2019). Molecular mechanisms of EBV-driven cell cycle progression and oncogenesis. *Medical Microbiology and Immunology*, *208*, 573–583.

Young, L. S., & Rickinson, A. B. (2004). Epstein–Barr virus: 40 years on. *Nature Reviews Cancer*, *4*(10), 757–768.

Zalani, S., Holley-Guthrie, E., & Kenney, S. (1996). Epstein-Barr viral latency is disrupted by the immediate-early BRLF1 protein through a cell-specific mechanism. *Proceedings of the National Academy of Sciences*, *93*(17), 9194–9199.

Zealiyas, K., Teshome, S., Haile, A. F., Weigel, C., Alemu, A., Amogne, W., Yimer, G., Abebe, T., Berhe, N., & Ahmed, E. H. (2023). Genotype characterization of Epstein–Barr virus among adults living with human immunodeficiency virus in Ethiopia. *Frontiers in Microbiology*, *14*, 1270824.

Zhang, N. (2018). Role of methionine on epigenetic modification of DNA methylation and gene expression in animals. *Animal Nutrition*, *4*(1), 11–16.

Zhang, W., Wang, M.-Y., Wei, X.-L., Lin, Y., Su, F.-X., Xie, X.-M., Tang, L.-Y., & Ren, Z.-F. (2017). Associations of Epstein-Barr virus DNA in PBMCs and the subtypes with breast cancer risk. *Journal of Cancer*, *8*(15), 2944.

Zhou, L., Chen, J., Qiu, X., Pan, Y., Zhang, Z., & Shao, C. (2017). Comparative analysis of 22 Epstein–Barr virus genomes from diseased and healthy individuals. *Journal of General Virology*, *98*(1), 96–107.

ANNEXES

Annex 1: Isolation of Peripheral Blood Mononuclear Cells (PBMC)

1. Draw blood from donor into a heparinized (or any other anti-coagulant) blood tube
2. Aliquot up to 10 ml of blood into 50 ml conical tube
3. Dilute blood 1:2 with PBS at room temperature (20 ml PBS)
4. Underlay diluted blood with 10 ml Ficoll Hypaque lymphocyte separation medium (Or can overlay diluted blood slowly on top of 10mL of Ficoll)
5. Centrifuge at 1500rpm without brake at room temperature for 30 minutes
6. Remove buffy coat and transfer into a new 50 ml conical tube
7. Raise volume to 50 ml with PBS and spin at 1500 rpm at room temperature for 10 minutes
8. Pour off supernatant
9. Add ~10 ml RBC lysis buffer to lyse red blood cells, you may skip RBC lysis (steps 9, 10 and 11) if RBC lysis is not necessary or the lysis buffer might affect downstream analysis.
10. Incubate at room temperature for 10 min,
11. Wash cells by adding up to 50 ml PBS and spinning at 1500 at room temperature for 10 minutes.
12. Resuspend washed cells in 10 ml of complete RPMI medium.
13. Count live cells using a hemocytometer
14. Adjust the volume as needed to obtain the desired cell concentration

Annex 2: Lymph node tissue Processing Procedure

It is important to note that ALL procedures should be carried out as sterilely as possible. These cells will be used in experiments and need to be properly collected and preserved.

Supplies:

- Freezing Media at room temperature (90% FBS, 10% DMSO)
- Sterile 1x PBS
- Sterile 50 mL conical tubes
- Sterile 70 uM cell strainer
- Sterile 1-5mL syringe
- Petri dish
- Ice Bucket (filled with ice)
- Cryovials
- Mr. Freezie or Styrofoam container for freezing

If we would like to use RBC lysis

- *Sterile filtered 1x RBC lysis at room temperature*

Procedure:

1) Sterilely harvest the lymph node and put it in sterile conical tube with PBS.
2) In the hood, place a 70uM cell strainer in a petri dish. Pour the lymph node and PBS into the strainer, allowing the fluid to flow into the petri dish. Use the backside of a 1-5mL syringe plunger to mash the lymph node and pass it through the strainer.
3) Once lymph node is fully dissociated, place the strainer back on the 50mL conical tube. Use a pipet to filter the liquid through 70uM cell strainer into the 50 mL conical tube (more than one 70uM cell strainer might be needed depending on the lymph node size).
4) Rinse the petri dish and syringe plunger again with PBS and pass this through the strainer for maximum recovery. Once filtered, fill the tube to 50mL with PBS. Skip to step 8 if you are not doing RBC Lysis

USE THESE STEPS IF YOU ARE DOING RBC LYSIS

5) *Return the tube to the culture hood and remove supernatant. Add 3mL (scale up as necessary for larger pellets) of the RBC lysis buffer and gently resuspend the pellet by pipetting. For the next 3 minutes, gently rock the conical tube.*

6) *Add 20mL of PBS to the tube to stop the RBC lysis reaction. Filter the liquid into another conical tube using the 70uM filter if there are any solids. Fill up the conical tube to 50mL with PBS. Centrifuge at 1200 RPM for 5 minutes.*
7) *Then, take the tube to the hood and discard the supernatant. Add 50mL of PBS. Gently disrupt pellet so that the cells are suspended.*

8) Put 10uL of the 50mL solution into tube for counting. Place the 50mL conical tube with LMNCs on ice. Count the cells and determine cell number per mL and viability.

9) Place X mL of freezing media into room temperature water bath for X cryovials needed. Example: 10mL freezing media and 10 cryovials.

10) Centrifuge the tube at 1200 RPM for 10 minutes.

11) During the centrifuge, print out X labels for the cryovials.
    a. Concentrations are typically 10e6 upto 50e6.
    b. Put labels onto the cryovials before adding freezing media to the cells. Remember, DMSO is toxic to the cells so freezing media should be added to the cells once everything is ready and then immediately stored inside the Mr. Freezie and transferred to the -80 deg C freezer.

12) Remove supernatant and add the freezing media. Gently suspend the pellet in the freezing media by pipetting up and down for proper re-suspension. Once evenly distributed, pipette 1mL of the solution into each cryovial. Place the cryovials into a Mr. Freezie.

13) Place the Mr. Freezie into a -80 freezer for at least twenty-four hours before storing them in liquid nitrogen.

Annex 3: Isolation of Genomic DNA from FFPE

Deparaffinizing using Deparaffinization Solution

1. Add Deparaffinization Solution: add **600ul** Deparaffinization Solution to the FFPE sample. For more sample material, add up to **1ml** Deparaffinization Solution.

2. Vortex vigorously for 10 sec, sit at room temperature (15–25°C) for 10 min and then vortex briefly.
3. Incubate at 56°C for 10 min, then vortex briefly and allow to cool at room temperature (15–25°C) and then centrifuge at maximum speed for 10 min.
4. Carefully remove the Deparaffinization Solution by **repeat pipetting and centrifuging** without disturbing the pellet.
5. Remove any residual Deparaffinization Solution using a fine pipet tip. Resuspend the pellet by adding 300 µl Buffer PKD and loosen the pellet by pipetting and vortexing. Add 10 µl proteinase K and vortex briefly.
6. Incubate at 56°C for 15 min. After incubation, the sample may not be completely lysed. This does not affect the procedure.
7. Incubate on ice for 3 min (complete cooling is important for efficient precipitation in next step). Meanwhile, adjust water bath (or heat block) temperature to 80°C.
8. Centrifuge for 15 min at maximum speed at room temperature.
9. Carefully transfer the clear supernatant, **without disturbing the pellet**, to a 1.5 ml tube for RNA extraction by **repeat pipetting and centrifuging**.
10. **Note**: Depending on the amount and nature of the FFPE sample, the pellet may be very small or difficult to see. Can keep the pellet with little residual supernatant (but as less as possible).

Keep the pellet on ice for DNA extraction in the later steps

11. The DNA-containing pellet can be stored for 2 h at room temperature, for up to 1 day at 2–8°C or for longer periods at –15 to –30°C.
12. Resuspend the pellet in 180µl Buffer ATL. Add 40 µl proteinase K, and mix by vortexing.
13. Incubate the solution overnight in a heat block (water bath) at 56°C for 12 hours, then 80°C for 1 hours for proteinase K digestion and de-crosslinking.
14. Briefly centrifuge the tube to remove condensation from the inside of the lid.
15. Add 220µl Buffer AL and vortex vigorously for 10 sec. There is precipitate in the solution. Stay in room temperature for 10 min.
16. Add 220 µl ethanol (96–100%) to the sample and vortex vigorously for 10 sec to obtain a homogenous solution.

17. Briefly centrifuge the 1.5 ml tube to remove drops from the inside of the lid.
18. Carefully transfer the entire lysate to the center of QIAamp MinElute column (in a 2ml collection tube).
19. Close the lid and centrifuge at 6000 x *g* (8000 rpm) for 1min.
20. **Note**: If the lysate has not completely passed through the membrane after centrifugation, centrifuge again at a higher speed until the column is empty.
21. Place the column in a new 2 ml collection tube, and discard the collection tube containing the flow-through.
22. Carefully open the column and add 500 µl Buffer AW1.
23. Close the lid and centrifuge at 6000 x *g* (8000 rpm) for 1 min.
24. Place the column in a new 2ml collection tube, and discard the collection tube containing the flow-through.
25. Carefully open the column and add 500 µl Buffer AW2.
26. Close the lid and centrifuge at 6000 x *g* (8000 rpm) for 1 min.
27. Place the column in a new 2ml collection tube, and discard the collection tube containing the flow-through.
28. Open and close the cap to release the pressure. Centrifuge at maximum speed for 3 min to dry the membrane completely.
29. Place the column in a clean 1.5 ml microcentrifuge tube (not provided), and discard the collection tube containing the flow-through. Apply 20–100µl Buffer ATE to the center of the membrane.
30. Close the lid and incubate at room temperature for 5 min. Centrifuge at maximum speed for 1 min

Annex 4: Isolation of Genomic DNA from PBMCs, LMNCs and plasma

1. Pipet 20 µl QIAGEN Protease (or proteinase K) into the bottom of a 1.5 ml microcentrifuge tube.

2. Add 200 µl sample to the microcentrifuge tube. Use up to 200 µl whole blood, plasma, serum, buffy coat, or body fluids, or up to 5 x 106 lymphocytes in 200 µl PBS.

- If the sample volume is less than 200 µl, add the appropriate volume of PBS. QIAamp Mini spin columns copurify RNA and DNA when both are present in the sample. RNA may inhibit some downstream enzymatic reactions, but not PCR. If

RNA-free genomic DNA is required, 4 µl of an RNase A stock solution (100 mg/ml) should be added to the sample before addition of Buffer AL.

3. Add 200 µl Buffer AL to the sample. Mix by pulse-vortexing for 15 s.

- To ensure efficient lysis, it is essential that the sample and Buffer AL are mixed thoroughly to yield a homogeneous solution.
- If the sample volume is larger than 200 µl, increase the amount of QIAGEN Protease (or proteinase K) and Buffer AL proportionally; for example, a 400 µl sample will require 40 µl QIAGEN Protease (or proteinase K) and 400 µl Buffer AL. If sample volumes larger than 400 µl are required, use of QIAamp DNA Blood Midi or Maxi Kits is recommended; these can process up to 2 ml or up to 10 ml of sample, respectively.

Note: Do not add QIAGEN Protease or proteinase K directly to Buffer AL.

4. Incubate at 56°C for 10 min.

- DNA yield reaches a maximum after lysis for 10 min at 56°C. Longer incubation times have no effect on yield or quality of the purified DNA.

5. Briefly centrifuge the 1.5 ml microcentrifuge tube to remove drops from the inside of the lid.

6. Add 200 µl ethanol (96–100%) to the sample, and mix again by pulse-vortexing for 15 s. After mixing, briefly centrifuge the 1.5 ml microcentrifuge tube to remove drops from the inside of the lid.

- If the sample volume is greater than 200 µl, increase the amount of ethanol proportionally; for example, a 400 µl sample will require 400 µl of ethanol.

7. Carefully apply the mixture from step 6 to the QIAamp Mini spin column (in a 2 ml collection tube) without wetting the rim. Close the cap, and centrifuge at 6000 x *g* (8000 rpm) for 1 min. Place the QIAamp Mini spin column in a clean 2 ml collection tube (provided), and discard the tube containing the filtrate.

- Close each spin column to avoid aerosol formation during centrifugation. Centrifugation is performed at 6000 x *g* (8000 rpm) to reduce noise. Centrifugation at full speed will not affect the yield or purity of the DNA. If the lysate has not completely passed through the column after centrifugation, centrifuge again at higher speed until the QIAamp Mini spin column is empty.

- Note: When preparing DNA from buffy coat or lymphocytes, centrifugation at full speed is recommended to avoid clogging.

8. Carefully open the QIAamp Mini spin column and add 500 µl Buffer AW1 without wetting the rim. Close the cap and centrifuge at 6000 x *g* (8000 rpm) for 1 min.

- Place the QIAamp Mini spin column in a clean 2 ml collection tube (provided), and discard the collection tube containing the filtrate.
- It is not necessary to increase the volume of Buffer AW1 if the original sample volume is larger than 200 µl. Carefully open the QIAamp Mini spin column and add 500 µl Buffer AW2 without wetting the rim. Close the cap and centrifuge at full speed (20,000 x *g*; 14,000 rpm) for 3 min.

10. Recommended: Place the QIAamp Mini spin column in a new 2 ml collection tube (Not provided) and discard the old collection tube with the filtrate. Centrifuge at full speed for 1 min.

11. Place the QIAamp Mini spin column in a clean 1.5 ml microcentrifuge tube (not provided), and discard the collection tube containing the filtrate. Carefully open the QIAamp Mini spin column and add 200 µl Buffer AE or distilled water.

12. Incubate at room temperature (15–25°C) for 1 min, and then centrifuge at 6000 x *g* (8000 rpm) for 1 min.

Annex 5: Agarose Gel Preparation

Agarose Gel preparation:

1. Add 500ml of 1X TAE buffer into a 1000ml bottle (leave excess space to boil, do not use 500ml bottles for 500ml total volume)
2. Add 10 g of Agarose to the buffer.
3. Heat in a microwave for 4-5min until boiling. Briefly swirl every 2min to help dissolve the Agarose.
4. Take out the bottle, add a magnetic stir bar, and put on a magnetic stirrer. Make sure the agarose solution is no longer boiling when handling it.
5. Wait for about 15min until the agarose solution has cooled to approximately 60 degrees.
6. Add Ethidium bromide (10 µl of a 1% (m/v) dye solution per 500ml agarose solution n),

7. Stir for another 10 sec until the dye is mixed thoroughly.
8. Pour into agarose gel rack and let sit until Agarose is solid, about 10min

* The agarose gel can be wrapped in a plastic foil and stored at 4 degrees for later use. Use within two months. Make sure agarose gel does not break when handling (especially when preparing low-concentration gels)

Annex 6: qPCR from genomic DNA for EBV load/copy number determination

qPCR is carried out with Fast SYBR green master mix (Applied Biosystems) and run on a Viia7 qPCR machine (Applied Biosystems). EBV DNA was quantified with primers specific to the EBV EBNA1 locus (forward: TCATCATCATCCGGGTCTCC, reverse: CCTACAGGGTGGAAAAATGGC), and signals were normalized to host genome DNA using primers specific for ACTB (forward: CAGGCAGCTCGTAGCTCTTC, reverse: TTGTGGCTCAGGGAAAATGT). The relative gene expression was calculated as 2-Δct method with Δct = ct(EBNA1) - ct(ACTB).

Protocol

1. Thaw DNA samples, mix, and in a separate tube prepare a 5ng/ul dilution of the DNA
2. Set up qPCR master mix for samples (run technical triplicates for each sample):

1 reaction:

2X Fast SYBR green master mix (Applied Biosystems)	5.0ul
ddH2O	2.5ul
primer mix (10uM forward+10uM reverse primer)	0.5ul
Total volume per reaction:	8.0ul

3. Add master mix to 384 well-plates
4. Seal plate with adhesive foil, make sure it's covered tightly (apply pressure to streak foil)
5. spin down at 2000g for 1min
6. add 2ul of DNA dilution (from step 1.) per well to plate

7. Seal plate with qPCR-compatible adhesive foil, make sure it's covered tightly (apply pressure to streak foil)
8. spin down at 2000g for 1min
9. insert plate in ViiA7 qPCR machine: In software, start new experiment
10. under "experiment properties",

    select ViiA7 as device

    select "comparative CT (ΔΔCT)" as readout

    select "SYBR green" and "Fast" as assay

    check "include melt curve" to have melt curve added after run (recommended)
11. under "assign", assign sample AND assay/amplicon names to each well that holds a reaction (CAUTION: Wells that do not have a sample and an assay/amplicon assigned will NOT be read by the detector, and will not show up in your results)
12. under "run method" make sure PCR program is correct:

    SYBR green fast PCR:

95 deg.	20 sec	
95 deg	1 sec	
60 deg.	20 sec	x40 cycles

13. run plate by clicking on "Run" on the left side, then click green "start run" button
14. once run is complete, export data: click "export" button on the left, then check "results" tab only and export to a file location of your choice to get one single file that has ct values and melting temperature in it.

Annex 7: EBNA 3C detection using conventional PCR

A. Run PCR product on 2% agarose gel, 120 Volt for 50min.

B. Load 5 uμ of PCR product per lane.

C. Visualize gel in a Gel imager to ensure PCR worked and the product is of the correct size. (a failed PCR with primer dimers is approximately 100bp band)

Annex 8: Bisulfate conversion of DNA

Annex 9: Bisulfite iPLEX Assay for Methylation Analysis

1. Thaw bisulfite DNA (btDNA), keep on ice
2. Prepare PCR master mix, **add DNA polymerase last**

   1 reaction:

ddH2O	2.40 ul
10X PCR buffer (Agena)	0.70 ul
magnesium chloride (Agena)	0.56 ul
dNTP mix (Agena)	0.14 ul
capture primer mixes	1.00 ul
Taq polymerase	0.20 ul
Total master mix:	5.00 ul

3. Add master mix to a 384 well-plate,
4. Seal with adhesive cover, spin down at 2000g for 1min
5. Add 2.0 ul of btDNA to 384-well plate.
6. Seal with adhesive cover, make sure it's covered tightly (apply pressure to streak cover), spin down at 2000g for 1min.
7. Run in 384-well PCR cycler (annealing temperature of primer multiplex reaction usually 54 deg.)

95 deg.	2min	
95 deg	30sec	x45-50 cycles
54 deg.	30sec	
72 deg.	1min	
72 deg.	5min	
4 deg.	hold	

8. Cycler can be left on 4deg. overnight, store PCR product at 4 deg. or -20 deg. afterwards
9. Prepare shrimp alkaline phosphatase (SAP) master mix:

   1 reaction:

ddH2O	1.53 ul
Shrimp Alkaline phosphatase buffer (Agena)	0.17 ul
Shrimp Alkaline Phosphatase (Agena)	0.30 ul
Total master mix:	2.00 ul

10. Spin down PCR plate, add 2.0ul SAP master mix to the top of each PCR well
11. Seal plate with fresh adhesive cover, make sure it sits tightly, spin down

12. Incubate in a thermocycler at 37 deg. for 30min, followed by inactivation of the enzyme at 85 deg. for 5min
13. Prepare extension reaction master mix, keep on ice after preparation:

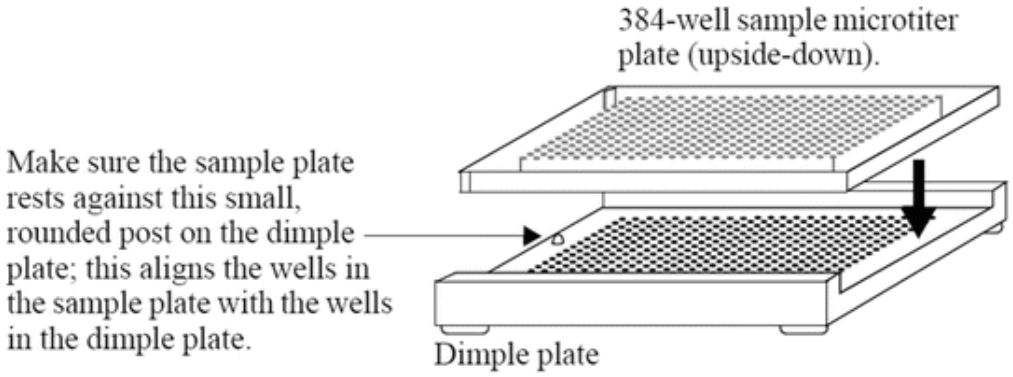
1 reaction:

ddH2O	0.74 ul
extension primer mixes	0.94 ul
iPLEX extension buffer	0.20 ul
iPLEX termination reaction mix	0.10 ul
iPLEX enzyme	0.02 ul
Total master mix:	2.00 ul

14. Spin down SAPed PCR plate, add 2ul of extension master mix to the top of each PCR well, seal plate spin down again.
15. Incubate in a thermocycler using "iPLEX extension program", keep on hold at 4 deg. after that (may be left overnight at 4 deg.):

94 deg.	30sec	
94 deg	5sec	
52 deg.	5sec	
80 deg.	5sec	
52 deg.	5sec	
80 deg.	5sec	
52 deg.	5sec	40 cycles
80 deg.	5sec	
52 deg.	5sec	
80 deg.	5sec	
52 deg.	5sec	
80 deg.	5sec	
72 deg.	3min	
4 deg.	hold	

16. Add 18ul of ddH2O to each well, seal plate spin down.
17. Desalting: Spread out desalting resin (Agena) on the 6mg/well 384-well Dimple plate. Make sure no excess desalting resin is left on the plate
18. Let Dimple plate sit at room temp for approx. 2 min to let resin dry slightly.
19. Load Resin: Take place from step 18, remove seal, turn it upside down and put it on top of Dimple plate from step 19 (see image below). While holding the plate and Dimple plate together firmly, Flip both of them together upside down. Have resin fall into place by lightly tapping the top of the Dimple plate frame with a pen or empty 50ml tube.



20. Seal plate with adhesive cover. Briefly spin down at 500g and put in a rotating mixer. Rotate at room temperature for 30min.
21. Chip spotting onto Agena SpectroChIP:
    a) Let Agena Nanodispenser RS1000 chip spotter complete one cleanup cycle for 24-well pin setup (tools -> daily clean)
    b) Spin down plate at 2000g, then transfer to Chip spotter and place in MTP1 spot.
    c) Place Agena Spectrochip on chip holder spot.
    d) Pipette 80ul of 3-point calibrant (Agena) into white chip spotter reservoir. DO NOT USE 4-POINT CALIBRANT for iPLEX assays
    e) Drain sonicator bath and refill with 96-100% Ethanol
22. Set up plate on attached PC by first defining a plate transfer map (Mapping -> New). Make sure “tuning” is enabled when assigning samples. Save plate map. Then define method (Method) for the map you just created. Make sure “auto tuning” is checked. Save methods file. Start Spectrochip transfer with the methods file (Transfer)
23. Once completed, remove Spectrochip and insert into mass spectrometer chip holder. Make sure not to touch the chip surface. Insert a second (empty) chip to fill out space in chip holder.
24. Insert chip holder into mass spectrometer. Close chamber and draw vacuum using button on the device.
25. Set up plate map at PC attached to mass spectrometer:
    a) Open TYPER software -> plate editor -> new
    b) Type in sample names and multiplex assay info on 384-well platemap using available sample/amplicon names to the left.
    c) Save plate
    d) Go to chip linker software, select iPLEX plate from the list to the left, type in chip barcode
    e) Press “add”, then “create” (plate file will be created and can now be loaded in SpectroACQUIRE)
    f) Go to SpectroACQUIRE software, load plate in microchip spot A by going to “chip autorun setup” in lower right screen panel
    g) Load parameter file (Tools -> load parameters -> IPLEX)

h) Click “start autorun” in upper field of SpectroACQUIRE software. Chip will be run one well at a time in automatic program (384 wells take approx. 2h). Note: “Start Autorun” button will remain inactive until vacuum is drawn completely in chip chamber. Check Mass spectrometer parameters window to monitor pressure.

26. Once completed, release vacuum, remove Spectrochip and store at room temperature in plastic housing for future analysis.
27. Export data to text file:
    a) TYPER software -> Analyzer
    b) select and load plate data from list to the left
    c) View -> plate data pane -> save as .csv file
28. Open and analyze file with excel or comparable software. Methylation values for each analyzed CpG can be calculated from C-extended vs. T-extended signal ratio. Make sure to apply cutoff for technical data quality (usually <0.85 unextended primer, and >10.0 peak area)

Annex 10: EBV Library Construction protocol

1. End Repair and A-tailing

1.1 Assemble each end repair and A-tailing reaction in a tube or well of a PCR plate as follows:

Component	Volume
Fragmented, double-stranded DNA	50 µL
End Repair & A-Tailing Buffer*	7 µL
End Repair & A-Tailing Enzyme Mix*	3 µL
Total volume:	60 µL

*The buffer and enzyme mix should preferably be pre-mixed and added in a single pipetting step. Premixes are stable for ≤24 hrs at room temperature, for ≤3 days at 2°C to 8°C, and for ≤4 weeks at -15°C to -25°C.

1.2 Vortex gently and spin down briefly. Return the plate/tube(s) to ice. Proceed immediately to the next step.

1.3 Incubate in a thermocycler programmed as outlined below:

Step	Temp	Time
End repair and A-tailing	20°C	30 min
	65°C*	30 min
HOLD	4°C**	∞

*A heated lid is required for this incubation. If possible, set the temperature of the lid at 85°C, instead of the usual ~105°C.

**If proceeding to the adapter ligation reaction setup without any delay, the

reaction may be cooled to 20°C instead of 4°C.

1.4 Proceed immediately to Adapter Ligation (step 2).

2. Adapter Ligation

2.1 Dilute adapter stocks to the appropriate concentration, as outlined in Table 2 (p. 5).

2.2 In the same plate/tube(s) in which end repair and A-tailing was performed, assemble each adapter ligation reaction as follows:

Component	Volume
End repair and A-tailing reaction product	60 µL
Adapter stock (concentration as required)	5 µL
PCR-grade water*	5 µL
Ligation Buffer*	30 µL
DNA Ligase*	10 µL
Total volume:	110 µL

*The water, buffer and ligase enzyme should preferably be premixed and added in a single pipetting step. Premixes are stable for ≤24 hrs at room temperature, for ≤3 days at 4°C, and for ≤4 weeks at -20°C.

2.3 Mix thoroughly and centrifuge briefly.

2.4 Incubate at 20°C for 15 min.

Note: to achieve higher conversion rates and library yields, particularly for low-input samples, consider increasing the ligation time—to a maximum of 4 hrs at 20°C, or overnight at 2°C to 8°C. Please note that longer ligation times may lead to increased levels of adapter-dimer. Adapter concentrations may have to be optimized if ligation times are extended significantly.

2.5 Proceed immediately to the next step.

3. Post-ligation Cleanup

3.1 In the same plate/tube(s), perform a 0.8X bead- based cleanup by combining the following:

Component	Volume
Adapter ligation reaction product	110 µL
KAPA cleanup beads	88 µL
Total volume:	198 µL

3.2 Mix thoroughly by vertexing and/or pipetting up and down multiple times.

3.3 Incubate the plate/tube(s) at room temperature for 5 – 15 min to bind DNA to the beads.

3.4 Place the plate/tube(s) on a magnet to capture the beads. Incubate until the liquid is clear.

3.5 Carefully remove and discard the supernatant.

3.6 Keeping the plate/tube(s) on the magnet, add 200 µL of 80% ethanol.

3.7 Incubate the plate/tube(s) on the magnet at room temperature for ≥30 sec.

3.8 Carefully remove and discard the ethanol.

3.9 Keeping the plate/tube(s) on the magnet, add 200 µL of 80% ethanol.

3.10 Incubate the plate/tube(s) on the magnet at room temperature for ≥30 sec.

3.11 Carefully remove and discard the ethanol. Try to remove all residual ethanol without disturbing the beads.

3.12 Dry the beads at room temperature for 3 – 5 min, or until all of the ethanol has evaporated. *Caution: over-drying the beads may result in reduced yield.*

3.13 Remove the plate/tube(s) from the magnet.

3.14 Thoroughly resuspend the beads:

- in 25 µL of elution buffer (10 mM Tris-HCl, pH 8.0 – 8.5) to proceed with Library Amplification (step 4), or in 55 µL of elution buffer (10 mM Tris-HCl, pH 8.0 – 8.5) to proceed with double-sided size selection

3.15 Incubate the plate/tube(s) at room temperature for 2 min to elute DNA off the beads.

3.16 Place the plate/tube(s) on a magnet to capture the beads. Incubate until the liquid is clear.

3.17 Transfer the clear supernatant to a new plate/tube(s):

- to proceed with Library Amplification (step 4), transfer 20 µL of supernatant, or
- to proceed with double-sided size selection (Appendix 1), transfer 50 µL of supernatant.

4. Library Amplification

Note: Please refer to Important Parameters: Library Amplification (pp. 6 – 7) and the KAPA NGS Library Preparation Technical Guide (available on request from Technical Support at sequencing.roche.com/support) for more information on optimizing library amplification.

4.1 Assemble each library amplification reaction as follows:

Component	Volume
KAPA HiFi HotStart ReadyMix (2X)	25 µL
KAPA Library Amplification Primer Mix (10X) *	5 µL
Adapter-ligated library	20 µL
Total volume:	50 µL

*Or another, suitable 10X library amplification primer mix. The recommended final concentration of each primer in the library amplification reaction is 0.5 – 4 µM. Also refer to Important Parameters: Library Amplification (p. 6).

4.2 Mix thoroughly and centrifuge briefly.

4.3 Amplify using the following cycling protocol:

Step	Temp	Duration	Cycles
Initial denaturation	98°C	45 sec	1
Denaturation	98°C	15 sec	Minimum number required for optimal amplification (Table 3 or 4)
Annealing*	60°C	30 sec	
Extension	72°C	30 sec	
Final extension	72°C	1 min	1
HOLD	4°C	∞	1

*Optimization of the annealing temperature may be required for non- standard (i.e., other than Illumina TruSeq) adapter/primer combinations.

4.4 Proceed directly to Post-amplification Cleanup (step 5).

5. Post-amplification Cleanup

5.1 In the library amplification plate/tube(s) perform a 1X bead-based cleanup by combining the following:

Component	Volume
Library amplification reaction product	50 µL
KAPA cleanup beads	50 µL
Total volume:	100 µL

5.2 Mix thoroughly by vortexing and/or pipetting up and down multiple times.

5.3 Incubate the plate/tube(s) at room temperature for 5 – 15 min to bind DNA to the beads.

5.4 Place the plate/tube(s) on a magnet to capture the beads. Incubate until the

liquid is clear.

5.5 Carefully remove and discard the supernatant.

5.6 Keeping the plate/tube(s) on the magnet, add 200 µL of 80% ethanol.

5.7 Incubate the plate/tube(s) on the magnet at room temperature for ≥30 sec.

5.8 Carefully remove and discard the ethanol.

5.9 Keeping the plate/tube(s) on the magnet, add 200 µL of 80% ethanol.

5.10 Incubate the plate/tube(s) on the magnet at room temperature for ≥30 sec.

5.11 Carefully remove and discard the ethanol. Try to remove all residual ethanol without disturbing the beads.

5.12 Dry the beads at room temperature for 3 – 5 min, or until all of the ethanol has evaporated. *Caution: over-drying the beads may result in reduced yield.*

5.13 Remove the plate/tube(s) from the magnet.

5.14 Thoroughly resuspend the beads in an appropriate volume of elution buffer (10 mM Tris-HCl, pH 8.0 – 8.5) or PCR-grade water.

Note: If proceeding with a second post- amplification cleanup, or double-sided size selection (Appendix 1), resuspend the beads in 55 µL of elution buffer.

5.15 Incubate the plate/tube(s) at room temperature for 2 min to elute DNA off the beads.

5.16 Place the plate/tube(s) on a magnet to capture the beads. Incubate until the liquid is clear.

5.17 Transfer the clear supernatant to a new plate/tube(s) and proceed with size selection (refer to Appendix 1), library QC, target capture or sequencing, as appropriate. Store purified, amplified libraries at +2°C to +8°C for 1 – 2 weeks, or at -15°C to -25°C for up to 3 months.

Annex 11. EBV capture sequencing using Illumina tube protocol

This protocol has been developed for a maximum of 6 capture reactions using individual tubes.

Guidelines

During the 4 hr incubation, the tube needs to be sealed properly to avoid evaporation.Excessive evaporation during the hybridization can lead to capture failure.

The duration of hybridization should be kept consistent for all

samples within a project. For GC-rich or small panels (<1000 probes), longer hybridization times (up to 16 hr) may improve performance.

Before you start

Two cycling programs, set at different incubation temperatures, are used for hybridization capturein this protocol.

1. Create the following PCR programs:

HYB program (lid set at 100°C)	
95°C	30 sec
65°C	4 hr
65°C	Hold
WASH program (lid set at 70°C*)	
65°C	Hold

* It is critical to reduce the lid temperature to 70°C for the WASH program.

Perform hybridization reaction

1. In a 1.7 mL MAXYMum Recovery microtube (low-bind), add the following components:

Blocker component	Volume per reaction (µL)
Human Cot DNA	5
xGen Blocking Oligos based on your library adapters	2

2. Add 500 ng of library to each tube containing Blocker components.If multiplexing samples, use 500 ng of each library.

3. Dry down the mixture in a SpeedVac system.

Safe Stop: Be sure to seal the sample tube. Store the sample at RT overnight,or –20°C for longer.

4. Thaw all contents of the xGen Hybridization and Wash Kit to room temperature.

Note: Inspect the tube of 2X Hybridization Buffer for crystallization of salts. If crystals are present, heat the tube at 65°C, shaking intermittently, until the buffer is completely solubililzed; this may require heating for several hours.

5. Create the Hybridization Master Mix by adding the following components to the tube from step 2 (above).

Hybridization Master Mix Component	Volume per reaction (µL)
xGen 2X Hybridization Buffer	8.5
xGen Hybridization Buffer Enhancer	2.7
xGen Lockdown Panel or custom probes	4
Nuclease-Free Water*	1.8

* **Do not use water** if using an xGen spike-in panel. See Appendix B for more information.

6. Pipet the mix, then incubate at room temperature for 5–10 min.
7. Vortex, then briefly centrifuge.
8. Transfer 17 µL of the capture to a low-bind 0.2 mL PCR tube, then briefly centrifuge.

   .
9. Place the sample tube in the thermal cycler and start the HYB program.

Prepare buffers

Note: Before preparing the buffers, remove the Dynabeads M-270 Streptavidin beads from storage at 4°C. The beads need to be at room temperature for a minimum of 30 min before performing the washes.

1. Dilute the following xGen buffers to create 1X working solutions:

Component	Nuclease-Free Water (µL)	Buffer (µL)	Total (µL)	Storage
xGen 2X Bead Wash Buffer	160	160	320	Keep at room temperature.
xGen 10X Wash Buffer 1	252	28	280	Aliquot 110 µL of the 1X Buffer into a separate tube and heat to 65°C. The remaining solution should be kept at room temperature.
xGen 10X Wash Buffer 2	144	16	160	Keep at room temperature.
xGen 10X Wash Buffer 3	144	16	160	Keep at room temperature.
xGen 10X Stringent Wash Buffer	288	32	320	Aliquot into two tubes (160 µL each). Heat tubes to 65°C in a water bath or heating block.

Note: If Wash Buffer 1 is cloudy, heat the bottle in a 65°C water bath to allow resuspension.

Tip: The 1X working solutions are stable at room temperature (15–25°C) for up to 4 weeks.

2. Prepare the following Bead Resuspension Mix in a low-bind tube:

Bead Resuspension Mix component	Volume per reaction (µL)
xGen 2X Hybridization Buffer	8.5
xGen Hybridization Buffer Enhancer	2.7
Nuclease-Free Water	5.8

Wash Streptavidin beads

Important! Only perform bead washes with beads that have equilibrated to room temperature.

1. Mix the beads thoroughly by vortexing for 15 sec.
2. Aliquot 50 µL of streptavidin beads per capture into a single 1.7 mL low-bind tube. For

example, for 1 capture, prepare 50 µL of beads and for 2 captures, prepare 100 µL of beads.

3. Add 100 µL of Bead Wash Buffer per capture. Gently pipet the mix 10 times.

4. Place the tube on a magnetic rack, allowing the beads to fully separate from thesupernatant (approximately 1 min).

5. Remove and discard the clear supernatant, ensuring that the beads remain in the tube.

6. Remove the tube from the magnet.

7. Perform the following wash:

    a. Add 100 µL of Bead Wash Buffer per capture, then pipet the mix 10 times.

    b. Place the tube on a magnetic rack for approximately 1 min, allowing the beads to fullyseparate from the supernatant.

    c. Carefully remove and discard the clear supernatant.

8. Perform an additional wash by repeating step 7 (above) for a total of 2 washes.

9. Resuspend the beads in 17 µL per capture of Bead Resuspension Mix from

    Prepare buffers, step 2.

10. Mix thoroughly to ensure that the beads are not left to dry in the tube.

    If needed, briefly centrifuge the tube at 25 x *g*.

11. Aliquot 17 µL of resuspended beads into a new low-bind 0.2 mL tube for eachcapture reaction.

Perform bead capture

1. Place the 1X Wash Buffer 1 (110 µL aliquot) and the 1X Stringent Wash Buffer (both aliquots)in a 65°C water bath for at least 15 min.

    Tip: The buffers will be used during the Heated washes, but we recommend startingthis incubation at the same time as the

bead capture, so that the buffers will be at thecorrect temperature when needed.

2. After the 4 hr incubation, take the tube out of the thermal cycler.
3. Once removed, stop the HYB program.
4. Immediately after the HYB program completes, start the WASH program.
5. Transfer 17 µL of resuspended streptavidin beads to the 0.2 mL tube containingthe sample.
6. Vortex to ensure that sample is fully resuspended. Gently and briefly centrifuge, if needed (10 sec at 25 x *g*).
7. Place the sample tube in the thermal cycler and set a timer for 45 min.

Note: It is safe to place the sample tubes in the thermal cycler before the lidtemperature has fully cooled to 70°C when starting the incubation.

8. Every 10–12 min, remove the tube from the thermal cycler and gently vortex to ensure thesample is fully resuspended.
9. At the end of the 45 min, take the sample off the thermal cycler. Proceed immediately to Heated washes.

Perform washes

Important! It is critical to ensure that the buffers have reached 65°C in a water bath beforestarting the Heated washes.

Heated washes

1. Transfer 100 µL of heated Wash Buffer 1 to the sample, pipet the mix 10 times, being carefulto minimize bubble formation.
2. Place the tube on a magnetic rack for 1 min. Remove the supernatant.

Note: Due to the supernatant having a high concentration of hybridization bufferand enhancer, use appropriate disposal methods.

Tip: If you do not have a magnetic rack that holds 0.2 mL tubes, transfer the entirereaction to a 1.7 mL tube.

3. Remove the tube from the magnet and add 150 µL of heated Stringent Wash Buffer to thesample.
4. Pipet the mix 10 times, being careful to not introduce bubbles.
5. Incubate in the water bath at 65°C for 5 min.
6. Place the sample on the magnet for 1 min. Remove the supernatant.
7. Remove the tube from the magnet and add 150 µL of heated Stringent Wash Buffer to thesample.
8. Pipet the mix 10 times, being careful to not introduce bubbles.
9. Incubate in the water bath at 65°C for 5 min.
10. Place the tube on a magnet for 1 min.

Room temperature washes

Important! To ensure the beads remain fully resuspended, vigorously mix the samplesduring the room temperature washes.

1. Remove and discard supernatant. Add 150 µL of Wash Buffer 1 equilibrated to roomtemperature.
2. Vortex thoroughly until fully resuspended.
3. Incubate for 2 min while alternating between vortexing for 30 sec and resting for 30 sec,to ensure the mixture remains homogenous.
4. At the end of the incubation, briefly centrifuge the tube.
5. Place on the magnet for 1 min.
6. Remove the supernatant. Add 150 µL of Wash Buffer 2.
7. Vortex thoroughly until fully resuspended.
8. Incubate for 2 min while alternating between vortexing for 30 sec and resting for 30 sec,to ensure the mixture remains homogenous.

9. At the end of the incubation, briefly centrifuge the tube.
10. Place on the magnet for 1 min.
11. Remove the supernatant. Add 150 µL of Wash Buffer 3.
12. Vortex thoroughly until fully resuspended.
13. Incubate for 2 min while alternating between vertexing for 30 sec and resting for 30 sec,to ensure the mixture remains homogenous.
14. At the end of the incubation, briefly centrifuge the tube.
15. Place the sample tube on the magnet for 1 min.
16. Remove and discard the supernatant.
17. With the sample tube still on the magnet, use a fresh pipette tip to remove residual WashBuffer 3 from the tube, then remove the tube from the magnet.
18. Add 20 µL of Nuclease-Free Water to each capture.
19. Pipet the mix 10 times to resuspend any beads stuck to the side of the tube.

Important! Do not discard the beads. Use the entire 20 µL of resuspended beads with captured DNA in Perform post-capture PCR.

Perform post-capture PCR

1. If a 1.7 mL tube was used for the washes, transfer the sample to a low-bind
   0.2 mL PCR tube.
2. Add the following components to create the Amplification Reaction Mix:

Amplification Reaction Mix component	Volume per reaction (µL)
2X KAPA HiFi HotStart ReadyMix*	25
xGen Library Amplification Primer	1.25
Nuclease-Free Water	3.75

* If using a PCR master mix other than Kapa HiFi, the magnesium concentration may need to be optimized for on-bead PCR.

3. Place the tube in a thermal cycler, and run the following program with the heated lid set at105°C:

Step	Number of cycles	Temperature (°C)	Time
Polymerase activation	1	98	45 sec
Amplification Denaturation Annealing Extension	Variable—refer to thePanel table below.	98 60 72	15 sec 30 sec 30 sec
Final extension	1	72	1 min
Hold	1	4	Hold

Note: The number of PCR cycles should be optimized per panel size and thenumber of pooled libraries per capture, to ensure there is enough yield for sequencing.

We recommend starting optimization with the following:

Panel size	1-plex	4-plex	8-plex	12-plex
>100,000 probes (xGen Exome)	10 cycles	8 cycles	7 cycles	6 cycles
10,000–100,000 probes	12 cycles	10 cycles	9 cycles	8 cycles
500–10,000 probes	13 cycles	11 cycles	10 cycles	10 cycles
1–500 probes	14 cycles	12 cycles	11 cycles	11 cycles

Optional stopping point. Amplified captures may be stored at 4°C overnight.

Purify post-capture PCR fragments

Important! Ensure Agencourt AMPure XP beads have been equilibrated to roomtemperature before proceeding.

1. Prepare 250 µL of fresh 80% ethanol per sample, multiplied by the number of samples with
   a 10% overfill.
2. Add 75 µL (1.5X volume) of Agencourt AMPure XP beads to each amplified capture (transfer

to a larger 1.7 mL tube, if needed).

3. After adding the beads, mix thoroughly and incubate for 5–10 min.
4. Place the sample tube on a magnet until the supernatant is clear (2–5 min).
5. Remove supernatant without disturbing the beads.
6. While keeping the tube on the magnet, add 125 µL of 80% ethanol, then incubate for1 min.
7. Remove the ethanol, then repeat another ethanol wash.
8. Allow the beads to air dry for 1–3 min. Do not overdry the beads.
9. Remove the sample tube from the magnet and elute in 22 µL of Buffer EB, or equivalent(10mM Tris-Cl, pH 8.5). Mix thoroughly. Alternatively, TE can be used.
10. Incubate for 5 min at room temperature.
11. Place the tube on a magnet until the supernatant is clear (1–2 min).
12. Transfer 20 µL of eluate to a fresh tube making sure that no beads are carried over.

Optional stopping point. Purified PCR fragments may be stored at –20°C for up to

1 week.

Validate and quantify library

1. Measure the concentration of the captured library using a fluorescence-based method for
   DNA quantitation (such as Qubit dsDNA HS Assay kit) or qPCR.

   Note: Knowing the concentration of your captured library helps ensure your library iswithin the detection limits of the chip or tape (as measured in step 2, below) for use on your digital electrophoresis system.

2. Measure the average fragment length of the captured library on a digital electrophoresis system (e.g., the BioRad Experion™ System using a DNA 1K chip, the Agilent 2100 Bioanalyzer using

a high sensitivity DNA chip, or the Agilent 2200 TapeStation system usinga DNA tape or other similar system).

Perform sequencing

Perform sequencing according to the instructions for your Illumina instrument.

Annex 12: IRB approval letters and MTA

የኢትዮጵያ ፌደራላዊ ዲሞክራሲያዊ ሪፐብሊክ
ትምህርት ሚኒስቴር

FEDERAL DEMOCRATIC REPUBLIC OF ETHIOPIA
MINISTRY OF EDUCATION

ስ: +251 11 155 3133
ፓ.ሳ.ቁ 1367፣ አራዳ ክፍለ-ከተማ,
አዲስ አበባ፣ ኢትዮጵያ
T: +251 11 155 3133
P.O. BOX 1367, ARADA SUB-CITY,
ADDIS ABABA, ETHIOPIA
www.moe.gov.et
www.facebook.com/fdremoe
@fdremoe
info@moe.gov.et

ቀን/DATE: 23 SEP 2022
ማጣ/REF NO: 311521552722

Addis Ababa University College of Health Sciences
Addis Ababa

Subject*: *Letter of Approval

The Ministry of Education (MoE) via its National Research Ethics Review Committee has reviewed **"Molecular Epidemiology and Genetic Diversity of Epstein - Barr virus among EBV associated lymphoproliferative in Ethiopia"** project protocol in an expedited manner. We are writing to advise you that MoE has granted full approval to the above named project, for a period of **one year (September 23, 2022- September 22, 2023)**.

All your most recently submitted documents have been approved for use in this study. The study should comply with the international and national scientific and ethical standard guidelines. Any change to the approved protocol or consent material must be reviewed and approved through the amendment process prior to its implementation. In addition, any adverse or unanticipated events should be reported within 24-48 hours to MoE. Please ensure that you submit biannual progress report to MoE once in six months and annual renewal application 30 days prior to the expiry date.

We are confident that your esteemed organization and the PI (copied) of the project should monitor the ethical implication of the project as stipulated in the latest approved document.

Sincerely

Solomon Behor Belay (PhD)
Chief Executive Officer
Research and Community Engagement

Cc

- State Minister for Higher Education Development Sector
- CEO, Research and Community Engagement Affairs
- Research Ethics Desk
  Ministry of Education

- Mr. Seifegebriel Teshome (PI)
  AAUCHS

ADDIS ABABA UNIVERSITY, COLLEGE OF HEALTH SCIENCES (IRB)
አዲስ አበባ ዩኒቨርስቲ ጤና ሳይንስ ኮሌጅ
Institutional Review Board

ANNEX 3
Form AAUMF 03-008

IRB's Decision

Meeting No: 08/2022 Meeting Date: August 15, 2022
Protocol number: 056/21/DMIP

Protocol Title: Molecular Epidemiology and Genetic Diversity of Epstein - Barr virus among EBV associated lymphoproliferative disorders in Ethiopia		
Principal Investigator:	Seifegebriel Teshome	
Institute:	College of Health Sciences, AAU	
Elements Reviewed (AAUMF 01-008)	☑ Attached	☐ Not attached
Review of Revised Application ☐ Yes ☐ No	Date of Previous review:	
Decision of the meeting:	☑Approved ☐ Approved with Recommendation ☐Resubmission ☐ Disapproved	

I. Elements approved- 1. Protocol Version No: 3
2. Protocol Version Date:
3. Informed consent Version No. 3
4. Informed Consent Version Date:

II. Obligations of the PI-
1. Should comply with the standard international & national scientific and ethical guidelines
2. All amendments and changes made in protocol and consent form need IRB approval
3. The PI should report SAE within 10 days of the event
4. End of the study, including manuscripts and thesis works should be reported to the IRB
5. The PI should report non-compliance and unanticipated events

III. TO NERC ☑

Institution Review Board (IRB) Approval: Period from October 13, 2021, to October 12, 2022
Follow up report expected in 3 Months ___ 6 months ___ 9 months ✔ one year ___

Chairperson, IRB
Dr. Adamu Addissie
Signature _____
Date: August 23, 2022

Institutional Review Board Office
IRB
Faculty of Medicine
ADDIS ABABA UNIVERSITY

Director of Research & Technology Transfer, CHS
Dr. Wondwossen Amogne
Signature _____
Date 8/4/2022

College of Health Science
Addis Ababa University

Material Transfer Agreement

This Material Transfer Agreement (MTA) has been prepared for use by Addis Ababa University, Department of Microbiology, Immunology and Parasitology, Ethiopia and The Ohio State University, Technology Commercialization Office, 1524 N. High Street, Columbus, OH 43201 USA in all transfer of research material (samples, derivatives, and specimens) related to the protocol ***"Molecular Epidemiology and Genetic Diversity of Epstein - Barr virus among patients with lymphoproliferative disorders in Ethiopia".***

Provider: Mr. Seifegebriel Teshome Addis Ababa University, Department of Microbiology, Immunology and Parasitology, Ethiopia

Recipient: Robert Baiocchi MD, PhD, Department of Internal Medicine, The Ohio State University, USA. Provider agrees to transfer to recipient's designated (provider) the following research materials /specimen. Samples: DNA, RNA, PBMCs, Plasma and Serum for sequencing and microbiome analysis under high-tech laboratory condition, specifically isolated human DNA (390 samples, 20-100µleach), Human mononuclear cells (578samples, 1mL each), Human plasma (380samples, 1mL each) and Serum (424samples 0.5mL each).

The research material will only be used for research purposes as described in the protocol by recipient's investigator in designated laboratory for the research project described below, under suitable containment conditions. This research material will not be used for commercial purposes such as screening, production or sale for which a commercialization license may be required. Recipient agrees to comply with all National and International guidelines rules and regulations applicable to the Research Project and the handling of the Research Material.

a) Are the research materials of human origin?

Yes [√] No [ ]

b) If yes, are they collected according to the details in the protocol and in adherence to National Research Ethics Review Committee (NRERC) and Addis Ababa University College of Health Sciences (AAU-CHS) Ethics Review Committee recommendations and their approval?

Yes [√] No [ ]

2. This research material and its derivatives will be used by recipient's investigator solely in connection with the following research project ("Research Project") described with specificity as follows ***"Molecular Epidemiology and Genetic Diversity of Epstein - Barr virus among patients with lymphoproliferative disorders in Ethiopia".***
3. In all presentations or written publications concerning the research project, recipient will acknowledge provider's contribution of this research material unless requested otherwise.

4. This research material represents a significant contribution on the part of provider and is considered proprietary to provider. Recipient therefore agrees to retain control over this research material and further agrees not to transfer the research material to other people not under her/his direct supervision without advance written approval of provider. The research material will be disposed of as agreed upon per protocol at the end of completion of the project.
5. The provider does not take any responsibility for loss, damage, wastage or spoilage of the research material during or after shipment to the address provided by the recipient under conditions agreed to in the protocol on shipment of the samples. This research material is provided as a service to the research community. IT IS BEING SUPPLIED TO RECIPIENT WITH NO WARRANTIES, EXPRESS OR IMPLIED, INCLUDING ANY WARRANTY OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Provider makes no representations that the use of the research material will not infringe any patent or proprietary right of third parties.
6. The recipient shall notify the provider in witting of any intention, improvement, modification discovery or development to the material or the information made by recipient or parties, collaborating with recipient, here in after referred to as "invention". Nothing in this agreement shall, however, be construed as conveying to the provider any rights under any patents or other intellectual property to such invention, and other than as explicitly provided herein. At its option the provider shall be entitled to receive sample of any materials derived from the Materials for its own research and evaluation purposes only.
7. The under- signed provider and recipient expressly certify and affirm that the contents of any statements made herein are truthful and accurate.
8. Any additional terms (use an attached page if necessary):
9. The provider maintains, ownership right of the research material and its unmodified derivatives unless stated otherwise.

The provider will retain a copy (aliquot) of every sample sent abroad as much as possible for local research needs.

Material Transfer Agreement

Signature page

For The Ohio State University:

Read and Acknowledged by:

Recipient's Investigator

Dr./Prof. Robert Baiocchi

Senior scientist
Signature
DocuSigned by:
Robert Baiocchi
E7CEA78B9AF941B...
Date 06/04/2022

Mailing Address for Material:
400 west 12th avenue, Room 481 Wiseman Hall
C/O Jessica Weist, Columbus OH 43210

Tel: … 614-370-9942 (J. Weist)

Fax: ….

Duly Authorized
DocuSigned by:
Kristin Cress
9F6382D1EE744C2...
Signature/ Stamp
Kristin Cress
Manager, Contracts
Date 06/02/2022

Mailing Address for Notices:

Technology Commercialization Office,
1524 N. High Street,
Columbus, OH 43201
Tel: + 614.247.6633
Email:contracts@osu.edu

For Provider:

Provider's Investigator
Mr. Seifegebriel Teshome
Signature

Mailing Address:
Addis Ababa University, Department of Microbiology, Immunology and Parasitology, College of Health Sciences, School of Medicine, Tikur Anbessa Specialized Hospital, 6th floor room number 610,
P.o.Box 9086, Addis Ababa, Ethiopia
Tel: +251913059887

Duly Authorized

Prof. Daniel Asrat

Signature/ Stamp

Date: 13-06-22

Mailing Address for Notices:

Addis Ababa University, Department of Microbiology, Immunology and Parasitology,
Tikur Anbessa Specialized Hospital
P.O.Box. 9086, Addis Ababa University
Addis Ababa, Ethiopia
Office 2nd floor room no.69 (Old building)
Tel: +251-911-22-30-19

Molecular Epidemiology and Genetic Diversity of Epstein-Barr virus among lymphoma patients in Ethiopia

ORIGINALITY REPORT

13%	10%	12%	5%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	cn.bio-protocol.org Internet Source	2%
2	Xue Sun, Xingyu Liu, Ying Zhao, Geng Tian, Weiwei Wang. "Chapter 3 Detection of Circulating Tumor DNA in Plasma Using Targeted Sequencing", Springer Science and Business Media LLC, 2023 Publication	1%
3	Submitted to The Energy and Resources Institute Student Paper	1%
4	Kidist Zealiyas, Seifegebriel Teshome, Aklilu Feleke Haile, Christoph Weigel et al. "Genotype characterization of Epstein–Barr virus among adults living with human immunodeficiency virus in Ethiopia", Frontiers in Microbiology, 2023 Publication	1%
5	www.n-genetics.com Internet Source	1%