Introduction
1. Epstein-Barr virus

Epstein-Barr virus (EBV) also known as human gammaherpesvirus 4, is a virus that infects more than 90% of the world’s population. It causes asymptomatic, persistent life-long infection in most people while a small number of infected individuals develop infectious mononucleosis in adolescence or early adulthood.

EBV can infect both B cells and epithelial cells. It was thought to have definite tropism for B cells as epithelial cells lack expression of CD21, which functions as a receptor for EBV glycoprotein gp350/220. Studies eventually ascertained that the binding of gp350/220 and CD21 only aids the attachment of EBV to its target cells but is not necessary for EBV entry. The essential proteins for EBV fusion and entry into target cells are glycoproteins known as gH, gL and gB. The gH/gL heterodimer either forms a complex with gp42 which binds to human leukocyte antigen (HLA) class II molecules on B cells or directly binds to integrins and NMHC-IIA followed by interaction with NRP1 and EphA2 on epithelial cells. Lastly, gB, a fusion protein, facilitates fusion and EBV into the host cells.

Although the first contact in the human host is the epithelium of the oropharynx, EBV does not infect many epithelial cells and mostly passes through the epithelium to infect naïve B cells. The initial lytic infection invokes a cytotoxic T cell response, which EBV can escape by switching to a latent type of viral gene expression. Four different latency EBV infection patterns are recognized in B cells, and these are referred to as latency type 0, I, II and III. Initial infection of naïve B cells usually leads to latency type III infection, which is characterized by expression of all EBV nuclear antigens (EBNA-1, -2, -3A, 3B, -3C, -LP) and latent membrane proteins (LMP-1, -2A, -2B). This infection induces proliferation and entry of the B cells into the germinal center. Within the germinal center, the latency type changes to type II, which is characterized by expression of only EBNA-1 and all three LMP proteins. The LMP proteins provide survival signals, allowing differentiation of the EBV infected B cell within the germinal center. Upon exit of the germinat center, EBV persists in memory B cells in a latency I stage, characterized by expression of only EBNA-1 and/or a latency 0 stage with no expression of any of the EBV proteins at all. EBNA-1 has an essential role in the replication and maintenance of the EBV genome. Reactivation of the lytic phase occurs infrequently in a small proportion of the EBV infected memory B cells. Additional infection of epithelial cells is likely to occur at this point where EBV from B cells can infect epithelial cells with higher efficiency, resulting in productive replication and shedding of EBV from differentiated epithelial cells in the oral cavity. The host immune system keeps the persistence and lytic activities of EBV at a low and benign level. When the virus-host equilibrium is out of balance, extensive proliferation of EBV-infected cells can lead to diseases.
2. **EBV-associated malignancies and diseases**

Although EBV infection is usually asymptomatic, some of the infected individuals develop infectious mononucleosis (IM) caused by a strong T-cell mediated immune response against lytic infection. IM usually causes fever, sore throat and enlarged lymph nodes in the neck. In general, these symptoms will recover within a month. EBV infection increases the risk of developing malignancies such as Hodgkin lymphoma (HL), diffuse large B-cell lymphoma (DLBCL), Burkitt lymphoma (BL), nasopharyngeal carcinoma (NPC) and gastric carcinoma (GC). Individuals who are immunosuppressed following organ transplant are at risk to develop post-transplant lymphoproliferative disorder (PTLD) due to uncontrolled growth of EBV-infected B cells. These malignancies exhibit one of the three types of latency programs with variable expression of several EBV proteins. The percentage of cases associated with EBV varies in different geographic location for each of the cancers (Table 1). In this thesis, the main focus is on HL, monomorphic PTLD and NPC.

HL tumor cells, known as Hodgkin Reed-Sternberg cells, arise from germinal center B cells. Four histological HL subtypes are recognized, with nodular sclerosis being the most common subtype. In the Caucasian population, about 30% of all HL cases are EBV positive and presence of EBV is most common in mixed cellularity HL (80-90%). In nodular sclerosis HL, the presence of EBV is much lower, accounting for 15-20% of the cases. HL is more common in Caucasian than in African and Asian, with estimated age-standardized incidence rates per 100,000 of >2, 0.92 and 0.59 respectively. HL patients are usually treated with a combination of chemotherapy and radiotherapy. Treatment does not differ between EBV-positive and EBV-negative HL cases. In stage I and II disease the treatment generally leads to complete remission, while in stage III and IV the 5-year disease free survival is about 70%. EBV-positive adults aged between 50 to 74 years have inferior survival.

PTLD develops in patients receiving immunosuppressive treatment following either solid organ transplantation or hematopoietic stem cell transplantation. Varying percentages of PTLDs are EBV-positive, with decreasing percentage of EBV-positive cases over the years. EBV-positive PTLDs accounted for more than 90% of the cases prior to 1995 and this percentage decreased to 52% in the period of 2008-2013. One of the risk factors for PTLD is being sero-negative for EBV prior to transplant. PTLDs are categorized into four subtypes, namely non-destructive, polymorphic, monomorphic and classic Hodgkin lymphoma. Treatment of PTLD includes reduction of immunosuppression or immunotherapy targeting the B cells.
<table>
<thead>
<tr>
<th>Associated malignancies</th>
<th>Target cell</th>
<th>Latency type</th>
<th>Percentage of EBV positive cases</th>
<th>Geographic location</th>
</tr>
</thead>
<tbody>
<tr>
<td>Burkitt lymphoma</td>
<td>B cell</td>
<td>Latency I</td>
<td>&gt; 95%</td>
<td>Endemic (Sub-Saharan Africa)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Sporadic (Western Europe, Northern America)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>85%</td>
<td>Southern America</td>
</tr>
<tr>
<td>Hodgkin lymphoma</td>
<td>B cell</td>
<td>Latency II</td>
<td>~30%</td>
<td>Europe, Northern America</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>&gt; 60%</td>
<td>Asia, Africa</td>
</tr>
<tr>
<td>Diffuse large B-cell lymphoma</td>
<td>B cell</td>
<td>Latency II or III</td>
<td>5-10%</td>
<td>Asia</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>&lt; 5%</td>
<td>Western</td>
</tr>
<tr>
<td>Post-transplant lymphoproliferative disorders</td>
<td>B cell</td>
<td>Latency III</td>
<td>60-80%</td>
<td>Worldwide</td>
</tr>
<tr>
<td>Plasmablastic lymphoma</td>
<td>B cell</td>
<td>Latency I or II</td>
<td>50-70%</td>
<td>Worldwide</td>
</tr>
<tr>
<td>Nasopharyngeal carcinoma</td>
<td>Epithelial cell</td>
<td>Latency II (or I/II)</td>
<td>&gt; 90%</td>
<td>Endemic (South China, Southeast Asia)</td>
</tr>
<tr>
<td>Gastric carcinoma</td>
<td>Epithelial cell</td>
<td>Latency I</td>
<td>~10%</td>
<td>Worldwide</td>
</tr>
<tr>
<td>Extranodal NK/T-cell lymphoma</td>
<td>Natural killer</td>
<td>Latency I/II</td>
<td>90%</td>
<td>Worldwide</td>
</tr>
</tbody>
</table>

NPC is the most common EBV-associated epithelial malignancy with over 90% of cases being EBV positive. The incidence of NPC is high especially in endemic regions, i.e., Southern China and Southeast Asia. The estimated age-standardized incidence rate in Asia is 2.1 per 100,000 per year, while it is 1.0 per 100,000 per year or less in other continents. Bidayuh, a minority population in Malaysia has one of the world’s highest incidence rate of 31 per 100,000. NPC cases in endemic regions are predominantly non-keratinizing.
subtype and EBV-positive. NPC has a higher incidence rate in males which is two- to three-fold higher than in females. The mainstay of treatment for NPC is radiotherapy or concurrent chemoradiotherapy in late stage disease. NPC is fairly sensitive to ionizing radiation, with 5-year overall survival rate ranging from 78% to 100% for early stage and decreases remarkably to about 26% for late stage and recurrent cases\textsuperscript{15–17}. Treatment of metastatic NPC primarily involves palliative systemic chemotherapy and patients usually had median survival ranging from about seven to 34 months\textsuperscript{18}.

3. EBV-specific HLA associations

HLA genes encode for membrane proteins that present antigenic peptides to T cells and thereby are crucial factors for the regulation of the immune response. HLA proteins are divided into HLA class I molecules, which are expressed in almost all nucleated cells, and HLA class II molecules, for which expression is restricted to professional antigen presenting cells like dendritic cells, macrophages and B cells. HLA class I is responsible for the presentation of viral peptides to CD8+ T cells and can elicit cytotoxic responses. HLA class II presents antigens to CD4+ T cells which helps to orchestrate both humoral and cytotoxic immune responses.

There are nine classical HLA genes, including three class I (HLA-A, -B, and –C) and six class II (HLA-DRA, -DRB1, -DQA1, -DQB1, -DPA1 and -DPB1) molecules. HLA genes are highly polymorphic and allele frequencies vary between different ethnicities. Early studies demonstrated that certain alleles are risk alleles for HL patients\textsuperscript{19,20}. Similarly, increased frequencies of certain HLA alleles have also been demonstrated in NPC patients\textsuperscript{21}. More recently performed genome-wide association studies have confirmed the strong susceptibility of variants in the HLA region for EBV-positive HL\textsuperscript{22} and NPC\textsuperscript{23–25}. Although no genome-wide association studies have been performed for PTLDs, several studies identified HLA associations with PTLD\textsuperscript{26–28}.

The underlying mechanisms for the observed increased susceptibility to EBV-associated diseases has been associated to efficiency of triggering EBV latency II-specific cytotoxic immune responses. HLA-A*01, HLA-B*37 and HLA-DRB1*10 are risk alleles for developing EBV-positive HL, while HLA-A*02:01 is a protective allele\textsuperscript{22,29}. To date, HLA-A*01-restricted EBV epitopes, as well as EBV-specific CTL response have not been elucidated\textsuperscript{30,31}. On the other hand, a high affinity has been shown for the protective HLA-A*02 allele that upon binding, resulted in an effective anti-EBV immune response\textsuperscript{31}. Interestingly, the HLA-A*02:07 allele has been shown to be a risk allele in both EBV-positive HL and NPC. Although this risk allele differs at only one amino acid position from the protective HLA-A*02:01 allele, it is less efficient in inducing LMP2-restricted CTL responses\textsuperscript{22}. Individuals carrying HLA-B*46:01, along with HLA-A*02:07 have an increased risk of developing NPC\textsuperscript{33}. These two alleles are in strong linkage disequilibrium in the Southern Chinese population.
However, studies on the susceptibility of the combination of HLA-A*02:07 and HLA-B*46:01 showed contradictory results. One study reported that HLA-A*02:07 is a risk allele even in the absence of HLA-B*46:01\(^\text{34}\), while, another study demonstrated no effect for individuals that have HLA-A*02:07 alone without HLA-B*46:01\(^\text{35}\). Both studies reported that having HLA-B*46:01 without HLA-A*02:07 results in a mild protective effect and individuals with the HLA-A*02:07-HLA-B*46:01 haplotype have an increased susceptibility to NPC\(^\text{35}\).

4. **EBV-based circulating biomarkers**

Testing circulating biomarkers is especially valuable for detecting early stage cancers because of its non-invasive nature. For EBV positive cancers, testing presence of EBV-derived biomarkers may result in high sensitivity as most infected cells have high EBV copy numbers. An overview of the sensitivity and specificity of commonly used EBV biomarkers to detect EBV-associated cancers is shown in Table 2.

**Table 2. Commonly studied circulating EBV biomarkers in EBV-associated diseases.**

<table>
<thead>
<tr>
<th>Biomarker</th>
<th>Disease</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>VCA-IgA</td>
<td>NPC</td>
<td>79.8 % – 91.9 %</td>
<td>70.9 % – 89.5%</td>
</tr>
<tr>
<td>EA-IgA</td>
<td>NPC</td>
<td>75 % – 92.6 %</td>
<td>90.8 % – 94.7 %</td>
</tr>
<tr>
<td>EA-IgG</td>
<td>NPC</td>
<td>85.2%</td>
<td>82.3%</td>
</tr>
<tr>
<td>EBNA1-IgA</td>
<td>NPC</td>
<td>80.0 % – 90.7 %</td>
<td>80.0 % – 92.0 %</td>
</tr>
<tr>
<td>EBV DNA</td>
<td>NPC, HL, PTLD</td>
<td>58.7 % – 96.5 %</td>
<td>66.7 % – 100 %</td>
</tr>
</tbody>
</table>

4.1. **EBV antibodies**

EBV serology tests to determine the presence of antibodies against viral components showed that EBV-associated cancer patients have higher EBV antibody titers as compared to controls\(^\text{36,37}\). The antibodies that are normally tested include those directed against viral capsid antigen (VCA), EBV-nuclear antigen (EBNA) and early antigen (EA). These proteins are expressed at different stages of the EBV life cycle. VCA and EA are expressed during lytic cycle, while EBNA1, which is essential for the persistence of the EBV genome in the host cell, is expressed during the latent stage. In NPC patients, the EBV antibody titers were positively correlated with tumor stage\(^\text{38}\). Probability of relapse was higher in patients with increased IgG and IgA antibodies to EA in blood samples obtained one year after complete remission\(^\text{39}\).
4.2. EBV DNA

EBV DNA was detected in plasma of patients with EBV-associated diseases such as HL, PTLD and NPC\textsuperscript{40–43}. Presence of plasma EBV DNA in NPC patients had prognostic value, with high plasma EBV DNA loads in pre-treatment, during-treatment and in post-treatment being associated with poor survival\textsuperscript{44}. In HL, a decrease in circulating EBV DNA load was observed in patients who responded to treatment, while an increase of the EBV DNA load was shown in patients who did not respond to treatment\textsuperscript{42}. In another HL study, presence of EBV DNA in pre-treatment plasma was shown to be associated with inferior failure-free survival in multivariate analysis\textsuperscript{40}. In view of high EBV positivity in PTLD cases, EBV DNA is frequently used to monitor high-risk post-transplant patients. High EBV DNA load is often seen before presentation of clinical symptoms of PTLD\textsuperscript{45}, hence preemptive treatment could be applied in these patients to halt the development of PTLD\textsuperscript{46,47}.

4.3. EBV-derived microRNAs

MicroRNA (miRNAs) are short non-coding RNAs that alter gene expression at the post-transcriptional level by either inducing mRNA degradation or by repression of translation. Target recognition is dependent on sequence homology, which in general is high in the seed region of the miRNA, i.e., nucleotide 2-7. The expression of miRNAs is dysregulated in many diseases including cancers. In accordance with this, tumor suppressors or oncogenic roles have been shown for many miRNAs.

EBV-derived miRNAs were first described in 2004\textsuperscript{48}. These miRNAs were shown to target both viral and human mRNAs and regulate expression in the same way as human miRNAs. In EBV-associated malignancies, EBV-derived miRNAs were detected in primary tumor tissue of NPC\textsuperscript{49}, GC\textsuperscript{50}, PTLD\textsuperscript{51} and HL\textsuperscript{52} as well as in the circulation, e.g., blood plasma, serum and in extracellular vesicles. Thus, EBV-derived miRNAs might potentially be used as biomarkers for EBV-associated cancers.

5. Tumor-cell derived circulating biomarkers

5.1 Cell-free DNA

Cell-free (cf)DNA is a term that is used for DNA fragments released into the blood circulation from cells undergoing apoptosis, necrosis or via active secretion\textsuperscript{53}. In healthy individuals, cfDNA is considered to be derived from apoptotic cells, cleaved by endonucleases in between the nucleosomes that leads to fragments that are approximately 167 bp in size. On the other hand, DNA released from tumor cells originates from cell death that occurs as a result of necrosis, autophagy or mitotic catastrophe. These circulating tumor (ct)DNA differ in size and can range from 70 to 200 bp or even be up to 30 kb.
Cancer patients were shown to have higher levels of cfDNAs as compared to controls already in 1977\textsuperscript{54}. It was suggested that ctDNA constitutes a proportion of the total cfDNA amount. More recent studies indeed confirmed presence of ctDNA in cancer patients based on detection of copy number losses and gains\textsuperscript{55,56} and somatic mutations\textsuperscript{57} in plasma samples that were similar to those observed in tumor tissue. In recent years, next generation sequencing technology greatly increased knowledge on ctDNA and its value as a clinical biomarker. A few studies showed presence of genomic aberrations in cfDNA of HL patients and showed that these aberrations correspond to those found in paired primary tissue samples\textsuperscript{58,59}.

As the majority of NPC cases are EBV positive, most studies are focused on circulating EBV DNA as a biomarker. The utilization of non-EBV cfDNA as biomarker in NPC is less explored with exception of studies defining hypermethylation profile. Hypermethylation of tumor suppressor genes is frequently detected in cancers including NPC. The cfDNA methylation profile could be used as a potential biomarker in addition to EBV DNA for early detection of NPC. The sensitivity to detect early stage NPC was improved from 65\% to 80\% with the addition of EBV DNA to the methylation panel\textsuperscript{60}. Similar to NPC, the presence and potential clinical value of cfDNA in PTLD is largely unexplored.

5.2 Other biomarkers

Thymus and Activation Regulated Chemokine (TARC) is a chemokine that is produced in high amounts by the tumor cells of HL and can be detected at high levels in blood samples of the vast majority of HL patients. TARC levels remain high in HL patients that were unresponsive to treatment and were re-elevated in patients with relapsed disease\textsuperscript{61,62}. High TARC levels at diagnosis have been correlated with negative prognosis\textsuperscript{63}.

A study demonstrated that a model with three long non-coding (Inc)RNAs MALAT1, AFAP1-AS1 and AL359062 showed high sensitivity in discriminating NPC from controls. Higher levels of these IncRNAs were observed in EBV-positive NPC as compared with EBV-negative NPC. The levels of these IncRNAs were elevated with tumor progression and reduced tremendously after treatment\textsuperscript{64}.
6. Scope of the thesis

This thesis focuses on two EBV-associated malignancies that both display a latency type II infection program, i.e., HL and NPC. In addition, PTLD which is commonly associated with EBV is included. The main aims of this thesis are to explore i) risk factors of HL and NPC, and ii) biomarkers in NPC, HL and PTLD.

Unlike most malignancies, HL and NPC both have an extremely abundant infiltration of immune cells in the tumor microenvironment. The presence and relevance of these cells in the tumor microenvironment has been reviewed in Chapter 2 of this thesis. Next to HL and NPC, GC, another common EBV-associated malignancy, has been included in this chapter.

In Chapter 3 and Chapter 4, we studied the association between expression of HLA on the tumor cells in relation to the phenotype of known susceptibility HLA alleles in HL (Chapter 3) and NPC (Chapter 4) patients. The hypotheses of these studies were that risk allele carriers have selective pressure to retain HLA expression while protective allele carriers have pressure to lose expression of HLA to avoid recognition by cells of the immune system.

In Chapter 5, the value of new and existing circulating biomarkers for NPC have been compared systematically. Plasma biomarkers including EBV antibodies, EBV DNA and miRNAs were evaluated to compare their diagnostic and prognostic value.

In Chapter 6, we explored the feasibility of using nasal washings as a novel and non-invasive method to detect NPC biomarkers. Nasal washings were used to detect both EBV DNA and NPC-associated miRNAs. We hypothesized that due to the proximity of the tumor site to the sampling site, tumor-derived molecules (i.e., EBV DNA and miRNAs) might be detectable. This method might be particularly useful in countries with high NPC incidence but low socio-economic status where accessibility to healthcare facilities is limited.

In Chapter 7 and Chapter 8, we explored the feasibility of using cfDNA as a disease biomarker in HL and PTLD patients. The aim was to detect copy number variations and EBV in retrospectively collected plasma samples with next-generation sequencing.

In Chapter 9, we summarize our studies and present general conclusions based on our findings. In addition, we discuss our findings and present future perspectives.
References


25. Bei, J. X., Zuo, X. Y., Liu, W. S., Guo, Y. M. & Zeng, Y. X. Genetic susceptibility to the endemic form of
Introduction


64. He, B. et al. Serum long non-coding RNAs MALAT1, AFAP1-AS1 and AL359062 as diagnostic and prognostic biomarkers for nasopharyngeal carcinoma. Oncotarget 8, 41166–41177 (2017).