EBioMedicine 37 (2018) 101–109

Contents lists available at ScienceDirect

EBioMedicine

journal homepage: www.ebiomedicine.com

EBioMedicine
Published by THE LANCET

Research paper

Natural Variations in *BRLF1* Promoter Contribute to the Elevated Reactivation Level of Epstein-Barr Virus in Endemic Areas of Nasopharyngeal Carcinoma

Jiang-Bo Zhang [a,1], Shao-Yi Huang [a,b,1], Tong-Min Wang [a,1], Si-Qi Dong [a], Yong-Qiao He [a], Xiao-Hui Zheng [a,c], Xi-Zhao Li [a], Fang Wang [a], Mu Jianbing [d], Wei-Hua Jia [a,e,f,*]

[a] State Key Laboratory of Oncology in South China; Collaborative Innovation Center for Cancer Medicine, Key Laboratory of Nasopharyngeal Carcinoma Diagnosis and Therapy; Sun Yat-sen University Cancer Center, 651 Dongfengdong Road, Guangzhou, Guangdong 510060, China
[b] Chongqing cancer hospital, Chongqing, China
[c] Affiliated Tumor Hospital of Xinjiang Medical University, Urumqi, Xinjiang 830000, China
[d] Laboratory of Malaria and Vector Research, National Institute of Allergy and Infectious Diseases, Rockville, MD, USA
[e] Cancer Center of Guangzhou Medical University, Guangzhou, China
[f] School of Public Health, Sun Yat-Sen University, Guangzhou, China

ARTICLE INFO

Article history:
Received 5 September 2018
Received in revised form 25 October 2018
Accepted 26 October 2018
Available online 9 November 2018

Keywords:
Epstein-Barr virus
Nasopharyngeal carcinoma
BRLF1
BZLF1
Variation
Reactivation

ABSTRACT

Background: Epstein-Barr virus (EBV) infection is a crucial risk factor for nasopharyngeal carcinoma (NPC), but the mechanism for its elevated activation level in NPC endemic areas remains unclear. This study aims to identify the EBV natural variations contributed to the different reactivation potential between NPC endemic and non-endemic areas.

Methods: 1030 subjects were recruited in China, including 303 healthy individuals from two NPC non-endemic areas, 483 healthy people from three endemic areas and 244 NPC patients. Among which, saliva DNA samples from 244 participants were sequenced for the EBV immediate early (IE) genes of *BRLF1* and *BZLF1*, their promoters were included; the rest 786 subjects were used for the validation of significant variations among three different populations. Haplotype and population structure analysis were conducted. Dual-luciferase assay was used to detect the promoter activity.

Results: A total of 246 distinct variations were detected, 29 showed significant difference in the frequencies between healthy people from NPC endemic area and non-endemic area. Population structure analysis clustered EBV strains into 9 subgroups mostly in accordance with the geographical origin of samples. Interestingly, two EBV genotypes, *Rp-V1* and *Rp-V2*, were identified according to the linkage relationship of the variations in *BRLF1* promoter (Rp). *Rp-V1* has higher frequency in NPC endemic areas than in non-endemic areas (52.38% vs 18.15%, $P = 2.07 \times 10^{-14}$), and was associated with higher oral EBV DNA levels (adjusted OR = 1.64, 95% CI = 1.21–2.24, $P = .002$), suggesting a more powerful activation ability of *Rp-V1* than that of the prototype *Rp*-of the EBV strain; On the contrary, *Rp-V2* has higher frequency in NPC non-endemic areas than in endemic areas (18.48% vs 0.38%, $P = 1.17 \times 10^{-7}$), might represent a reduced activation potential of EBV. Further dual-luciferase assay showed *Rp-V1* has higher promoter activity while compared with *Rp-V2* ($P < .0001$). Notably, *Rp-V1* impaired the transcription repression effect of *YY1* while *Rp-V2* strengthened the transcription repression effect of *EBF1* on Rp. In addition, significant differences of Rta 393–407 CTL epitope which may influence the recognition of Rta by CD8+ T cells were detected between healthy people from NPC endemic area and non-endemic area.

Conclusions: This study identified natural variations in *cis*-acting elements (*YY1* and *EBF1*) of EBV Rp altering Rp transcription activities, which may contribute to the elevated EBV activation level in NPC endemic areas than non-endemic areas.

1. Introduction

Nasopharyngeal carcinoma (NPC) is an Epstein-Barr virus (EBV)-associated human epithelial malignancy of the nasopharynx with a distinct ethnic and geographical distribution, which occur commonly in southern China and Southeast Asia, but rarely in northern China

* Corresponding author at: State Key Laboratory of Oncology in South China; Collaborative Innovation Center for Cancer Medicine; Key Laboratory of Nasopharyngeal Carcinoma Diagnosis and Therapy; Sun Yat-sen University Cancer Center, 651 Dongfengdong Road, Guangzhou, Guangdong 510060, China.
E-mail address: jiawh@sysucc.org.cn (W.-H. Jia).
[1] These authors contribute equally to the work.

https://doi.org/10.1016/j.ebiom.2018.10.065

Research in context

Evidence before this study

Previous clinical and epidemiological studies have revealed that individuals with elevated antibody titers and plasma EBV DNA load have a high risk of NPC. Our previous studies also showed that serum VCA-IgA titers, sero-prevalence and oral EBV load in NPC patients and healthy people from NPC endemic area are significantly higher than those of non-endemic area, reflecting the increased reactivation levels of EBV in both NPC patients and healthy people from endemic areas. However, the mechanism of the elevated EBV reactivation level in NPC endemic areas remains unclear.

Therefore, before we undertaking this study, we searched all the existing literatures about EBV variations and their disease associations in PubMed database. The search terms we used are: Epstein-Barr virus and (variation or sequencing). We included all the relevant research from January 1, 1990 to July 31, 2018. And the reported full length of EBV sequences used in this study were before December 31, 2016. We didn't add any bias in the selection of literatures.

Added value of this study

To date, there is no relevant research about the mechanism of the elevated EBV reactivation level in NPC endemic areas. In this study, we identified EBV subtypes (*Rp-V1 and RP-V2*) which were associated with EBV lytic activation level in the study population for the first time. And we also illuminated the molecular mechanism of the newly identified EBV subtypes in the regulation of EBV activation. We believe that our findings could partially explain the elevated activation ability of EBV in NPC endemic areas.

Implications of all the available evidence

NPC is an EBV-associated human epithelial malignancy occur commonly in southern China and Southeast Asia, but rarely in northern China and most other populations around the world. Lytic activation of EBV is an early high-risk event before NPC development and endemic area showed a significant higher level of EBV activation than in non-endemic area.

Therefore, the newly identified EBV subtypes give us some choice for NPC risk prediction. In the following studies, we will investigate the interactions of EBV variations and host genetic factors in NPC development, construct the Gene-EBV-Interaction-based NPC risk prediction model, and finally improve the prediction and prevention of NPC.

and most other populations around the world [1–5]. Although the unusual geographic distributions of NPC have been linked to the host genetic (notably the variants of major histocompatibility complex (MHC) class I genes on chromosome 6p21.3) and environmental factors (mainly the consumption of dietary nitrosamines) [6–8], whether naturally occurring variations among EBV strains could affect infection or the development of NPC is still contentious.

There have been substantial studies in characterizing the natural variations, at key viral genetic regions or whole genome, among EBV strains isolated from NPC or healthy individuals. Early works focused on the targeted EBV genomic regions revealed distinct variations that associated with different phenotypic properties of EBV strains [9,10]. These genomic regions include EBV nuclear antigens (EBNAs), latent membrane proteins (LMPs) also known as latent genes, several early lytic genes and microRNAs (miR-BHRF1 and miR-BART). For example, the worldwide EBV isolates can be classified as type 1 or type 2 based on the variations on the EBNA2 gene. Although the type 1 is the common EBV subtype in NPC with efficient B cell immortalization ability in vitro, its' tumorigenic propensity in human cancer is unclear [11]. Genetic variation in c-terminus region of LMP1, an essential EBV-encoded oncogene, has classified the LMP1 into 7 main groups that showed different cell transformation properties possibly due to these sequence variations [12]. In addition, a 30-bp deletion in C-terminus of LMP1 (del-LMP1) [13] and the loss of *Xho*I restriction site in the exon 1 of LMP1 [14] were also detected in NPC patients from Southern China. Although these LMP1 variants are widely distributed in different regions, there is no substantial evidence of disease association. Variants in other EBV genes including several early lytic genes (BHRF, BARF1 and gp350) and microRNAs (miR-BHRF1 and miR-BART) [15–19] have also been reported, but again appear to show geographic restriction rather than tumor specific polymorphism. EBV genetic variations at whole genome level have also been investigated among several highly representative EBV strains, GD1 [20], GD2 [21] and HKNPC1–9 [22,23] which were isolated from the tumor tissues of NPC patients in endemic areas of southern China. However, their associations to disease have not been noted, whether a specific EBV strain involved in the pathogenesis of NPC or plays a major role in the development of NPC remains to be determined. With the improvement of sequence technology, hundreds more EBV strains are successfully sequenced recently [24–28] and the genetic variations identified in these EBV strains provide a basis for future directed association analysis of specific genetic variation with EBV biology and EBV-associated disease.

As a critical early risk event of NPC, the recurrent lytic activation of EBV has been shown to play an important role in the initiation, progression and metastasis or relapse of NPC [29,30]. Clinical and epidemiological studies have revealed that individuals with elevated antibody titers and plasma EBV DNA load have a high risk of NPC [31–33]. Chien et al. reported that the IgA antibodies against EBV capsid antigen (VCA-IgA) and neutralizing antibodies against EBV DNase, which are indicators of EBV activation, may appear in serum samples long before the development of NPC [31]. Our previous studies showed the higher serum VCA titers and sero-prevalence in NPC patient samples than those of control samples, and these elevated serological indexes were also detected in samples from healthy individuals from NPC endemic areas but not from non-endemic areas [34], reflecting the increased reactivation levels of EBV in both NPC patients and healthy people from endemic areas. Recently, a previously reported cancer-related EBV subtype, Zp-V3 [35,36], was confirmed to create a binding site for the cellular transcription factor, NFATc1, and thus enhanced Zp activity and lytic viral replication in B-cell lymphoma cell lines [37]. However, it was not linked to the EBV activation level (for example, EBV copy number or serum VCA-IgA titers) in different geographic regions or populations, and cannot explain the elevated EBV activation level in NPC endemic areas of China. Therefore, whether EBV variations contribute to the different level of EBV lytic reactivation in endemic and non-endemic region in China remain unknown.

To further identify the natural variations potentially relevant to EBV lytic reactivation, we recruited 1030 individuals from three endemic and two non-endemic regions in China. EBV genetic variations were checked extensively on the EBV encoded immediate early (IE) genes (*BZLF1* and *BRLF1*), as well as their promoters (*Zp* and *Rp*) since these genetic regions are considered as the “switch” of EBV lytic reactivation. It has been reported, during the process of EBV reactivation, IE genes *BZLF1* and *BRLF1* are initially induced, followed by transactivation of *Zp* and *Rp* by the binding of Zta (*BZLF1* encoded protein) to the *Z-responsive elements* (*ZREs*) within *Zp* and *Rp* to greatly amplify the lytic-inducing effect [38]. Therefore, any genetic variations in the IE genes or their promoters which influence the expression or function of IE genes may hold importance in the pathogenesis of EBV-related

diseases, including NPC. With the measurement of oral EBV DNA loads, the relationship between the detected natural variations in IE region and the EBV reactivation level was further explored. Identification of the natural variations within *cis*-acting elements (YY1 and EBF1) in Rp region in EBV provide valuable insight towards understanding the mechanism of the higher reactivation potential of EBV observed in NPC endemic areas.

2. Materials and methods

2.1. Study populations

The study was reviewed and approved by the Human Ethics Approval Committee at Sun Yat-sen University Cancer Center. 483 healthy people and 244 NPC case were recruited from three endemic areas in China (Guangdong province, Jiangxi province and Guangxi Zhuang Autonomous Region). All NPC cases were recruited from the Sun Yat-sen University Cancer Center in Guangzhou of Guangdong province, China (designated as GD-NPC), from October 1, 2005 to October 1, 2007, who come from whole endemic regions of Guangdong province. Healthy control recruited from endemic areas are designated as: [1] GD-H (276 from Guangdong province at same time of case recruitment [34]), [2] JX-H (94 from Ganzhou city, Jiangxi province) as well as [3] GX-H (113 from Wuzhou city, Guangxi Zhuang Autonomous Region from June 1 to September 1, 2016). We further recruited 303 healthy controls from non-endemic areas for NPC in northern China designated as: [1] BJ-H (53 healthy controls in Beijing from 2006 to 2007) and [2] SX-H (86 healthy controls in Datong from 2006 to 2007 and 164 healthy controls in Yangquan of Shanxi province from June 1 to September 1, 2014). All subjects were asked to complete an in-person interview except for the healthy controls from Beijing. Written informed consent was obtained from each study subject. Characteristics of the study populations are summarized in Supplementary Table 1.

2.2. Mouthwash sample preprocess and DNA extraction

Mouthwash samples were collected by gargling with 10 ml physiological saline (0.9% NaCl). For DNA extraction, 10 ml of each of the mouthwash sample was first digested with 1% proteinase K and then processed through standard phenol/chloroform extraction and ethanol precipitation procedures.

2.2.1. BZLF1, BRLF1 genes and their promoters (Zp and Rp) regions amplification and sequencing

The *BRLF1* gene (4056 bp) and its promoter region (~660 bp) (ranging from 89,838 to 94,560 in NC_007605.1 EBV prototype genome) were amplified by nested PCR for each sample using 8 pairs of inner and outer primers (Supplementary Table 2). *BZLF1* gene which was located at the 3`UTR region of *BRLF1* was covered by primer pair of BRLF1-F4/R4 (ranging from 89,838 to 90,943 in NC_007605.1 EBV prototype genome). DNA from the B95.8 cell line was used as a positive control, and water was used as a non-template control. PCR products were subjected to electrophoresis in 2% agarose gel and purified with a QIAEXII gel extraction kit (QIAGEN, Valencia, CA, USA). These purified PCR products were directly sequenced by using Sanger sequencing. The sequences were aligned, translated and analyzed using BIOEDIT v7.0.9.0, with the B95.8 EBV strain as a reference sequence.

2.3. Bioinformatic analysis

Linkage analysis was performed by Haploview V4.2. Detection of co-occurrence and mutual exclusivity of pairs of EBV variants was performed using R package "discover". Population structure analyses were performed using the Structure program version 2.3.4 [39]. For each cluster, we executed 10 runs with different initializations, using admixture model with a burn-in period of 10,000 and 100,000 MCMC repetitions. The most likely number of putative genetic cluster, K, was estimated using the method proposed by Evanno et al. [40]. We defined the population ID of each sample according to their geographic origin. The statistical significance was calculated using 10,000 times permutations in Arlequin followed by Bonferroni correction. We identified "pure" individuals, with low (<10%) probability of admixing, in each of the cluster.

2.4. Cell lines

The 293 T cell line purchased from American Type Culture Collection was cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum (FBS, GIBCO, Carlsbad, CA, USA); the CNE2 NPC cell line was purchased from Jennio Biotech (Jennio Biotech, Guangzhou, Guangdong, China) and cultured in RPMI-1640 medium supplemented with 10% FBS. These two cell lines were maintained in a humidified incubator containing 5% CO_2 at 37 °C.

2.5. Plasmid construction

To construct pGL3-Rp-Luc plasmids, a 672 bp genomic DNA segment, which ranged from −667~ + 5 relative to the transcription start site (TSS) of *BRLF1* containing the variations of *Rp-V1* and *Rp-V2*, were directly amplified from the samples they were detected and cloned upstream of a firefly luciferase gene in the pGL3-basic vector. To construct the *YY1*, *SP1*, *CREB*, *EBF1* and *Zta* expression vectors, complete open reading frames were amplified using pairs of primers with *Xba*I and *Eco*RI sites at their 5′ end, respectively. The PCR products were digested with these two enzymes and then inserted into the corresponding sites of the pCDH-CMV-MCS-EF1-copGFP backbone plasmid. Sequencing analysis was performed to confirm the accuracy. Primers used for plasmid constructs are listed in Supplementary Tables 3 and 4, respectively.

2.6. Luciferase reporter assay

293 T or CNE2 cells were seeded in a 12-well plate at 1×10^5 cells per well and co-transfected with PCDH-*YY1*/*EBF1*/control along with pGL3-Rp-P-Luc or pGL3-Rp-V1/V2-Luc. pRL-TK was also transfected simultaneously as an endogenous control. Luciferase activity in each well was quantified 24 h after transfection using the Dual-Luciferase Reporter Assay System (Promega, Madison, WI, USA) according to the manufacturer's protocol. The relative luciferase activity was determined by dividing the firefly luciferase activity by the pRL-TK luciferase activity.

2.7. Statistical analysis

Statistical analysis was performed using GRAPH-PAD INSTAT version 3.05 (Graphpad, La Jolla, CA, USA) and R-3.3.1. Fisher's exact test or Pearson's chi square test was used to assess the association between categorical variables. Logistic model was used for multivariable regression. All tests were two sided, and $P < .05$ was considered statistically significant.

3. Results

3.1. The variation landscape of EBV IE genes among NPC endemic and non-endemic areas of China

Saliva DNA samples of healthy individuals from the NPC non-endemic Shanxi province of China (SX-H, $n = 84$) and endemic Guangdong province (GD-H, $n = 80$), as well as NPC patients from Guangdong province (GD-NPC, n = 80) were used for the sequencing of IE genes. The EBV strain derived from the B95.8 cell line was used as a reference sequence (accession number: NC_007605.1; defined as prototype sequence). The average point variation frequencies among the three

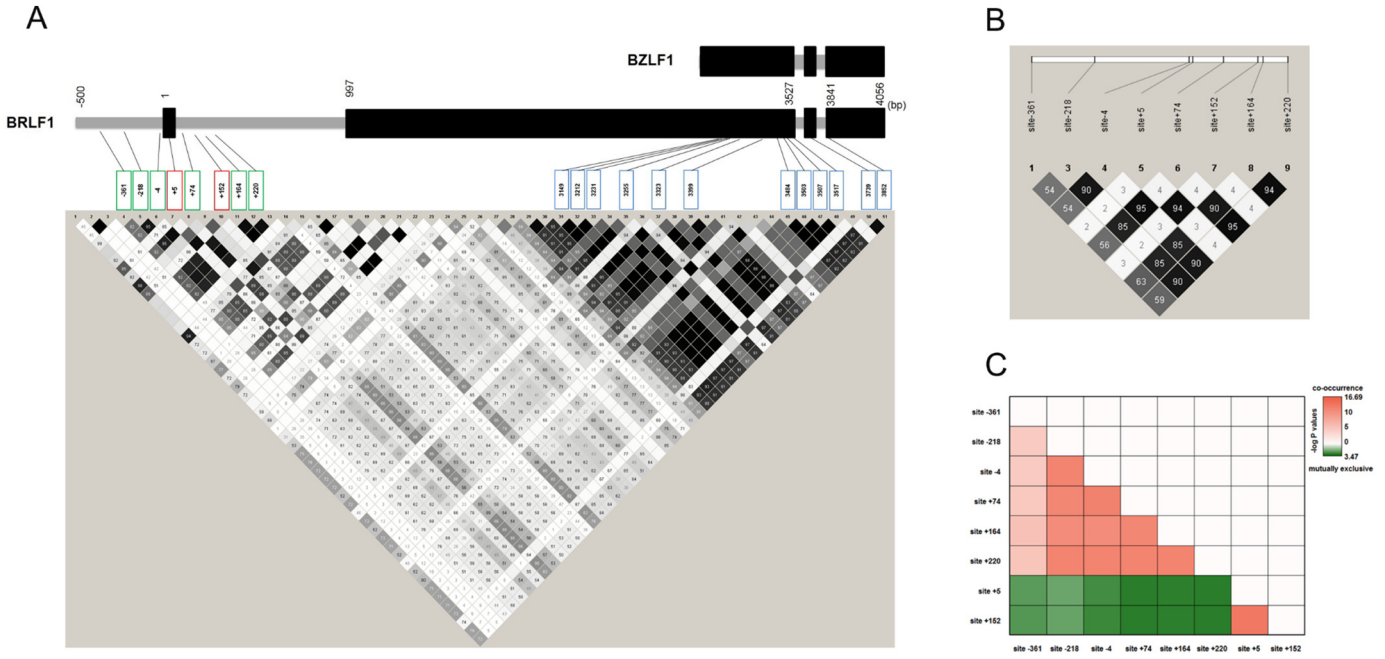

Fig. 1. Linkage analysis of significantly different variations.

populations in different gene regions of IE genes were calculated and showed in Supplementary Fig. 1. A total of 246 distinct point variations were detected within the *BRLF1* and *BZLF1* genes as well as their promoters. The variation frequencies of 29 sites showed significant differences between the SX-H and GD-H populations (Supplementary Table 5). Among these variations, we further identified three sites, site −4, +5 and + 152 relative to the TSS of *BRLF1*, with significantly different frequencies between SX-H and GD-H populations at $p < 10^{-3}$. Further linkage analysis showed that the variants at site +5 and site +152 have high linkage disequilibrium (LD) with each other ($D' = 1.0$) while site −4 was in high LD with site −361, −218, +74, +164 and + 220 ($D' = 0.92$–1.0) (Fig. 1A and B). We then defined the sequences with co-variants at sites of +5 and + 152 as *Rp-V1* and the sequences with co-variants at sites of −361, −218, −4, +74, +164 and + 220 as *Rp-V2*. We also observed that *Rp-V1* and *Rp-V2* were mutually exclusive among the samples (Fig. 1C). So, there are three distinct EBV subtypes, Rp prototype (*Rp-P*), *Rp-V1* and *Rp-V2*, defined by the polymorphisms in the promoter of *BRLF1* and they showed unbalanced distributions among populations. The frequency of *Rp-V1* is significantly higher among healthy individuals from endemic area than that from non-endemic area (43.75% vs 11.90%, $p = 1.12 \times 10^{-5}$), while the frequency of *Rp-V2* is much lower in health individuals from endemic area than those in non-endemic area (1.25% vs 16.67%, $p = .0016$).

3.2. Population structure analysis based on the variations in EBV IE genes

Structure uses multi-locus genotype data to infer population stratification, optimizing the number of sub-populations under a model whereby each sub-population is characterized by a set of allele frequencies at each locus. In this study, *Rp* (ranging from −600 to +40 relative to the TSS of *BRLF1*) and coding sequence of Rta and Zta were extracted from 99 reported EBV genomes [27] and 244 *BRLF1* sequences identified in our study, and were used to conduct population stratification by using Structure program. The results showed that, the most likely number of genetic clusters is 9, indicating that the variations within these sequences are best explained by partitioning them into 9 clusters (Supplementary Table 6). Each of the sequence was assigned to the cluster according to the highest probability (Supplementary Table 7) and the overview of the 9 clusters was shown in Fig. 2. We found that EBV from healthy people in endemic areas clustered mainly in cluster 2 and 4 which clearly separated from that of NPC patients, of which the EBV clustered mainly in cluster 3 and 7 (Table 1). The detailed geographic sources of the EBV sequences and their distribution in nine clusters are presented in Supplementary Table 9. To present the sequence features of the clusters, the consensus sequences for each of the different 9 clusters were summarized in Supplementary Table 9.

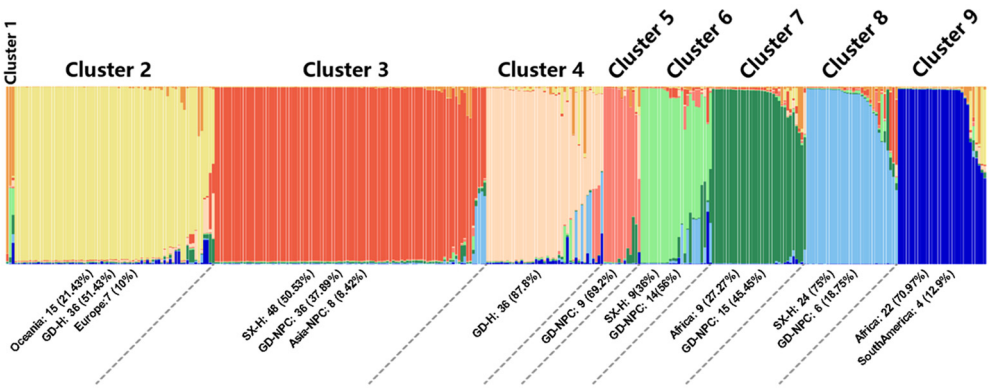

Fig. 2. Population structure analysis and the main population of each Cluster.

Table 1
Healthy people from NPC endemic area and NPC patients are gathered in distinct Clusters.

Cluster	Healthy endemic[a] (No. (%))	GD-NPC (No. (%))	P^b
Cluster 2	29 (50.0)	3 (4.2)	5.70×10^{-9}
Cluster 4	23 (39.7)	2 (2.8)	3.80×10^{-7}
Cluster 3	1 (1.7)	40 (55.6)	1.80×10^{-10}
Cluster 7	2 (3.4)	14 (19.4)	1.30×10^{-2}
Other	3 (5.2)	13 (18.1)	–

Note: Only the "pure" sequences, which are defined as those are assigned with low probabilities (<0.1) of admixing, were used in the calculation.
[a] Healthy endemic represent healthy people from NPC endemic Guangdong province.
[b] *P* values were calculated by two-sided proportional test.

3.3. Variation of CTL epitope sequences in Rta and Zta among three study populations

It has been shown previously that the recognition of immediate-early gene Rta and Zta of EBV by CD8+ T cells may have a significant impact on controlling of EBV. Therefore, to elucidate the biological importance of the variations identified in this study, the previously reported 17 CTL epitopes in Rta and 2 in Zta were further analyzed among different populations [41–43]. The results showed that Rta 393–407 epitope variation which was affected by two amino acid residues changes (P403S and L406F) was observed only in samples from endemic area (GD-H vs. SX-H, $P = .026$ by Fisher exact test). For Zta, the variation in 190–197 epitope (Q195H) showed significant difference between healthy people from endemic Guangdong province and NPC patients (GD-NPC vs. GD-H, $P = .035$ by Pearson's $\chi 2$ tests) (Table 2).

3.4. Distribution of Rp subtypes in NPC patients and healthy people from NPC endemic/non-endemic areas

Considering that the *Rp* region was found to carry a series of variations with significantly different frequencies among NPC patients and healthy people either in endemic or non-endemic areas, we enlarged the sample size of the three original populations for sequencing of the *Rp* region and included an additional healthy population BJ-H from a non-endemic area (53 healthy individuals from Beijing) as well as two healthy populations from NPC endemic areas: JX-H (94 healthy from Jiangxi province) and GX-H (113 healthy individuals from Guangxi Zhuang Autonomous Region of China). The total number of samples for sequencing of *Rp* reached to 1030.

The results showed that the frequency of *Rp-V1* was 148/276 (53.62%), 55/94 (58.51%) and 50/113 (44.25%) among three endemic healthy populations of GD-H, JX-H and GX-H, respectively. The overall frequency was 253/483 (52.38%). In contrast, the frequencies of *Rp-V1* were 13/53 (24.53%) and 42/250 (16.80%) in low-risk healthy populations of BJ-H and SX-H. The overall frequency was 55/303 (18.15%) (Fig. 3A). The prevalence of *Rp-V1* in healthy people of endemic areas is significantly higher than that of low-risk areas (52.38% vs. 18.15%, $P = 2.07 \times 10^{-14}$; Adjusted OR = 3.94, 95% CI = 2.79–5.64; Table 3).

For *Rp-V2*, the frequency was 1/276 (0.36%), 0/94 (0%) and 3/113 (2.65%) in the healthy people of endemic area GD-H, JX-H and GX-H, respectively. The overall frequency was 4/483 (0.83%); While, the frequencies were 9/53 (16.98%) and 47/250 (18.80%) in low-risk area of BJ-H and SX-H populations, with a combination *Rp-V2* frequency was 56/303 (18.48%) (Fig. 3B). The prevalence of *Rp-V2* in healthy people of endemic areas is significantly lower than that of low-risk areas (18.48% vs 0.83%, $P = 1.17 \times 10^{-7}$; Adjusted OR = 0.06, 95% CI = 0.02–0.15; Table 3).

In addition, we also observed the phenomenon of co-infections of different EBV strains within single individual. For example, the frequency of co-infection of *Rp-V2* and *Rp-P* strains in SX-H were 3.57% (3/84), co-infection of *Rp-V1* and *Rp-P* strains were 5.07% (14/276) in GD-H, 1.23% (3/244) in GD-NPC.

3.5. Association analysis of Rp subtypes with the oral EBV loads

To investigate whether the different *Rp* subtypes identified in our study can affect the level of EBV activation, EBV DNA loads of saliva (or mouth wash) samples for the 1030 subjects were analyzed (The detailed saliva EBV DNA load for each individual were listed in Supplementary Table 10). Data from different healthy populations in this study were combined and divided into two groups according to the median of the EBV copy number of saliva samples, low copy number group (<173,406 copies/ml) and high copy number group (≥173,406 copies/ml). The results showed that *Rp-V1* which was predominant in NPC endemic areas was significantly associated with higher oral EBV levels (adjusted OR = 1.64, 95% CI = 1.21–2.24; $P = .002$) (Table 4). On the contrary, *Rp-V2* which was predominant in NPC non-endemic areas was associated with lower oral EBV levels though not significant (adjusted OR = 0.76, 95% CI = 0.43–1.34; $P = .347$) (Table 4).

3.6. The variations in different Rp subtypes significantly affected the transcription activity of Rp

BRLF1 and *BZLF1* are "switch" genes of the EBV lytic cycle and reactivation, the promoter activities of them are tightly regulated by multiple cellular signaling pathways and transcription factors [44] (Fig. 4A). We have observed the association of *Rp* variants with the oral EBV loads, which may reflect different *Rp* transcriptional activities and ultimately different reactivation potentials. To investigate this hypothesis, we constructed pGL3-luc vectors containing typical sequences of three *Rp* subtypes (Fig. 4B), and tested the impact of the variations on the transcriptional activity of *Rp* by using a dual-luciferase assay in 293 T cells and an NPC cell line, CNE2. The results of the luciferase assay showed that pGL3-Rp-V1 showed a significantly increased luciferase activity compared with the B95.8 prototype, while the activity of pGL3-Rp-V2 was significantly lower than the B95.8 prototype of *Rp* (Fig. 4C). These results indicated that the dominant EBV strains from the NPC endemic areas have an enhanced *BRLF1* gene promoter activity over those from the non-endemic areas.

Based on a literature review and bioinformatic-based prediction (http://tfbind.hgc.jp), we found that the polymorphisms at sites −361, −218, −4 and + 5 in *Rp* may all affect the binding ability of *YY1* (Fig. 4A), a cellular transcriptional factor which was reported to suppress the activity of *Rp* [45]. We constructed the *YY1* expression

Table 2
Distribution of amino acid (AA) variations in IE genes` CTL epitopes.

IE genes	HLA restriction	Rta/Zta residues	No. of isolates (%)			Epitope sequence	P [a]	
			GD-NPC (n = 80)	GD-H (n = 80)	SX-H (n = 84)		GD-NPC/GD-H	GD-H/SX-H
Rta CTL epitope	HLA-Cw4	393–407	71 (88.8)	75 (93.8)	84 (100)	ERPIFPHPSKPTFLP	0.401	**0.026**
			9 (11.2)	5 (6.2)	0 (0)	ERPIFPHPSKSTFFP		
	HLA-B5801	375–383	19 (23.8)	30 (37.5)	32 (38.1)	NAAEPEQPW	0.086	0.937
			61 (76.2)	50 (62.5)	52 (61.9)	NAEEPEQPW		
Zta CTL epitope	HLA-B8	190–197	16 (20.0)	29 (36.2)	24 (28.6)	RAKFKQLL	**0.035**	0.377
			64 (80.0)	51 (63.8)	60 (71.4)	RAKFKHLL		

Note: Sequences listed are epitope sequences of B95–8 isolate. Only the variated AAs of the specimens are marked with red color. [a] Fisher exact test.

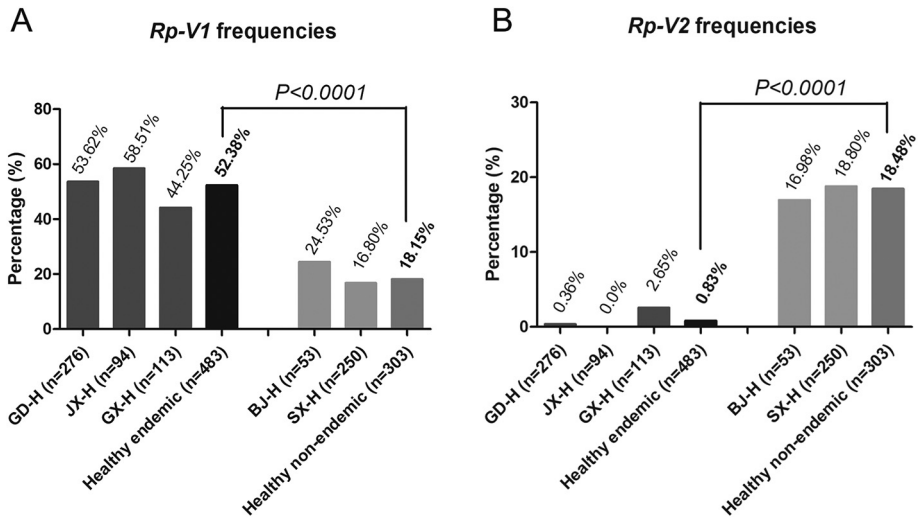

Fig. 3. Frequencies of Rp variants in each healthy population from NPC endemic and non-endemic areas.

vector and co-transfected it with pGL3-Rp-P, V1 or V2. The results of the dual-luciferase assay showed that overexpression of *YY1* significantly suppressed the activities of the *Rp* constructs. Meanwhile, the inhibition ratio of *YY1* on Rp-V1 was significantly lower than that of Rp-P ($P = .0002$ and 0.0089 for 293 T and CNE2 cells, respectively), indicating that the variation of +5 in *Rp* impaired the binding ability of *YY1*. However, no significant difference was observed between the inhibition ratio of *YY1* on *Rp-V2* and *Rp-P* (Fig. 4D).

Results of the bioinformatic analysis also suggested that the variation at −361 in *Rp* may affect the binding of the transcription repression factor *EBF1*, and the variation at −218 may affect the binding of *CREB* and *AP1*. Results of the dual-luciferase assay showed that overexpression of *EBF1* significantly suppressed the activities of the *Rp* constructs. Notably, the inhibition ratio of *EBF1* on *Rp-V2* was significantly higher than that of *Rp-P* ($P < .0001$), indicating that the variation at −361 in *Rp* could possibly strengthen the binding ability of *EBF1*. However, transcriptional activation effects of *CREB* and *AP1* showed no significant difference between *Rp-P* and *Rp-V2* (Fig. 4E).

3.7. Variation of Zta has no significant effect on Rp activation

A Zta variation subtype containing 12 point variations (defined as Zta-GD1 based on the GD1 EBV strain [20]) with close linkage showed significant differences between NPC patients and healthy people in a NPC endemic area (GD-NPC/GD-H, 78.31% vs. 62.35%, $P = .036$). Zta is a transcriptional factor regulating the expression of both Zta and Rta, as well as a subset of EBV genes, by binding the *ZREs* in their promoters (Fig. 4A). Compared with Zta-prototype (Zta-P) derived from B95.8 cells, 7 of the 12 point variations in Zta-GD1 were missense mutations. Four of the 7 missense mutations were within the transactivation domain of Zta, one was located in a DNA binding domain and another one was in the coiled coil domain [46] (Supplementary Fig. 2A), indicating that these mutations may affect the biological function of Zta. The expression vectors of Zta-P and Zta-GD1 were constructed to test the functional changes due to the variations in Zta. Results of the dual-luciferase assay showed that both Zta-GD1 and Zta-P could induce the activation of *Rp*, although with no significant difference between them (Supplementary Fig. 2B).

4. Discussion

Although EBV establishes latent infection in NPC cells with the encoded viral proteins and non-coding RNAs showing strong oncogenic properties [47,48], the recurrent lytic activation of EBV is increasingly recognized as a critical early risk event in NPC carcinogenesis [31,34,49]. In the present study, we conducted sequencing analysis on the EBV lytic activation "switch" genes (*BZLF1* and *BRLF1*) and their promoters in healthy people from NPC endemic, non-endemic areas and NPC patients in China, and identified natural variations in Rp which classified EBV into several subtypes with different EBV reactivation potentials for the first time. To date, this study is the largest investigation of EBV variations of IE genes with regard to both sample size and the number of populations from NPC endemic and non-endemic areas in China. The variations in *Rp* identified in this study provide us with valuable new clues towards understanding the predisposing factors for aberrant high reactivation potential and elevated NPC incidence in endemic areas.

Previous EBV whole genome sequencing studies indicated that the geographical factors are the main contributor to the differences in EBV genome [25,27], However, in this study, although the nine clusters each has a dominant geographical origin, e.g. Africa for cluster 1 and Asia for cluster 2, none of the geographical populations gather in just one cluster. Moreover, within Asia, the sequences are stratified into different clusters, some of which can distinguish sequences from endemic regions from those from non-endemic regions or NPC samples, indicating that there are different subtypes of IE genes in these three populations. For example, NPC patients and healthy people from low-risk areas in Asia are the main population both for cluster 6 and 7, we cannot distinguish these two clusters based on the difference of dominant populations. However, when considering only healthy people of NPC

Table 3
Comparison of the frequencies of Rp variants between healthy people from NPC endemic and non-endemic areas.

Variants	Healthy people of NPC non-endemic areas, No. (%) ($n = 303$)	Healthy people of NPC endemic areas, No. (%) ($n = 483$)	NPC patients GD-NPC, No. (%) ($n = 244$)	Adjusted OR (95%CI) Healthy endemic/Healthy non-endemic	*P*
Rp-P	192 (63.37)	225 (46.58)	146 (59.84)	*Ref*	–
Rp-V1	55 (18.15)	253 (52.38)	98 (40.16)	***3.94 (2.79, 5.64)***	2.07×10^{-14}
Rp-V2	56 (18.48)	4 (0.83)	0 (0)	***0.06 (0.02, 0.15)***	1.17×10^{-7}

Note: *P* values were calculated using logistic regression adjusting gender and age as covariates.

Table 4
The correlation of Rp subtypes with saliva EBV DNA copy numbers from NPC endemic and non-endemic areas.

Rp genotype	Low copy nuber[a] <173,406copies/ml[b]	High copy number ≥173,406copies/ml[c]	Adjusted OR (95%CI)[c]	*P*
Rp-P	222 (54.01)	189 (45.99)	1.00 (Reference)	/
Rp-V1	128 (42.38)	174 (57.62)	1.64 (1.21–2.24)	**0.002**
Rp-V2	36 (60.00)	24 (40.00)	0.76 (0.43–1.34)	0.347

[a] The oral EBV levels were divided into low and high levels according to the median of the EBV copies in saliva per milliliter.
[b] Median EBV copy number of healthy people (combination of healthy people from NPC endemic and non-endemic areas).
[c] The ORs, corresponding 95% CI and *P* values were calculated using logistic regression adjusting gender and age as covariates.

endemic area and NPC patients, these two populations gathered in completely distinct clusters: EBV from healthy donors of endemic areas gathered mainly in cluster 2 and 4 which clearly separated from that of NPC patients, of which the EBV gathered mainly in cluster 3 and 7. This result seems like that the variation pattern of IE genes may be related with NPC carcinogenesis, which needs to be further investigated.

Until now, no relationship of *Rp* variations and their potential affecting on EBV disease-related phenotypes were observed, except for a more frequent subtype of *BRLF1*, *BR1-C*, reported by Jia et al. in NPCs [50]. Our study identified two subtypes of *Rp*, *Rp-V1* and *Rp-V2*, which presented significantly disparate prevalence levels among populations with different NPC risks. The frequency of *Rp-V1* was significant higher in endemic areas than in the non-endemic area, while the frequency of *Rp-V2* was significant higher in non-endemic areas than in the endemic areas.

BRLF1 is a powerful enhancer that binds directly to the *R-responsive elements* (*RREs*) within the promoters of early lytic genes to induce the switch from latent to lytic infection of EBV [51]. The expression of *BRLF1* is tightly regulated by Zta and multiple cellular transcription factors, such as *YY1*, *ZEB1/2*, *EGR1* and *Sp1/3* [45,52,53,54]. Although the variations of *Rp-V1* and *Rp-V2* may all have affected the binding of *YY1*, the dual-luciferase assay only confirmed the impaired binding ability of *YY1* to the variation at the +5 site, indicating that variations in *Rp-V2* may not be located in the core binding sequence of *YY1*. On the other hand, *Rp-V2* was confirmed to strengthen the transcriptional repression effect of *EBF1* on *Rp*, which has not yet been reported. These results clarified the molecular mechanisms of the enhanced promoter activity of *Rp-V1* prevalent in NPC endemic areas, as well as the decreased activity of *Rp-V2* which was detected only in NPC non-endemic areas. More importantly, the significant association of *Rp-V1* with higher oral EBV DNA loads supported the critical role of this natural variation in EBV lytic reactivation. Together, these results confirm the enhanced EBV activation ability of *Rp-V1*, and partially explain the elevated reactivation of EBV in NPC endemic areas. It's important to note that, previous studies have confirmed only *BRLF1* expression (and not *BZLF1* expression) can induce lytic EBV reactivation in normal epithelial cells [55,56]. The *Rp* subtypes

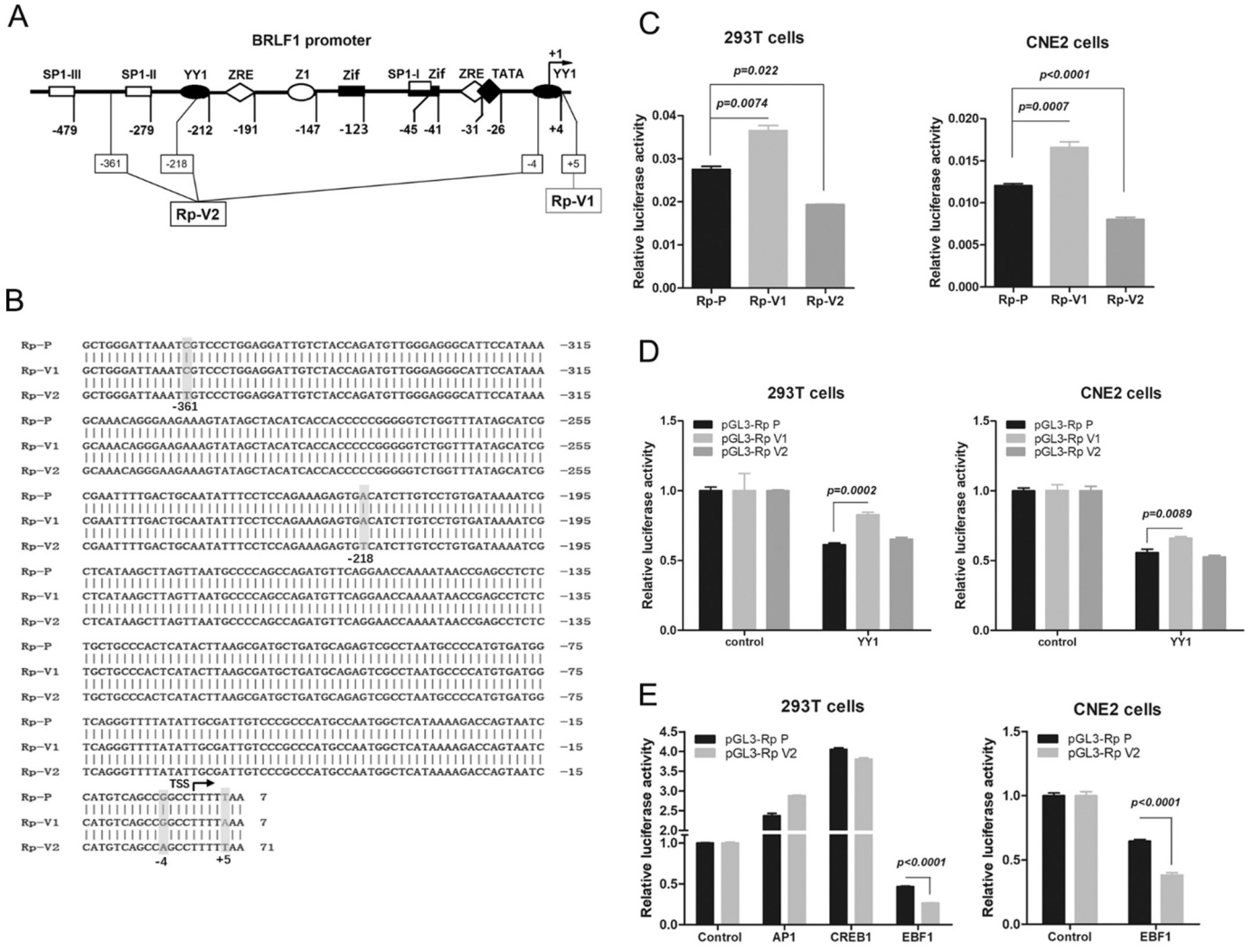

Fig. 4. Functional characteristic of Rp variants.

identified in our study which dramatically altered *Rp* activity may have critical effect on EBV lytic infection in epithelial cells.

The EBV-encoded immediate-early protein Zta (product of *BZLF1*) is one of the basic-leucine zipper (bZIP) transcription factors. It plays a critical role in both viral gene transcription and viral replication. As a lytic switch, synthesis of Zta alone is sufficient to induce the lytic reactivation of EBV. Indeed, Zta does not only regulate the expression of early lytic genes such as *BMLF1*, *BHRF1*, *BHLF1* and *BMRLF1*. By binding to *ZREs* and *TPA responsive elements* (*TREs*) in the promoters of host cellular genes or by interacting physically with host transcription factors, Zta also regulates the expression of a variety of cellular genes which are involved in the viral life cycle [57]. Therefore, elucidation of the pattern of variations in Zta is key to understanding the regional disparity of EBV reactivation levels. Previously, Ji et al. identified five *BZLF1* variants that were more prevalent in NPCs than controls based on the coding sequence changes of Zta in an NPC-endemic area of China [58]. In this study, we identified seven missense mutations with close linkage in Zta that showed significant elevated frequencies in NPC patients compared with those in healthy individuals both from endemic and non-endemic areas. However, no difference was observed between Zta-GD1 and Zta-P in the transactivation effects for *Rp*. Because *Rp* is not the only target promoter of Zta, further investigations are needed to fully understand the functional relevance of Zta variations.

Aside from the variations in Zta, other site alterations located in Zp or non-coding regions of *BZLF1* were also reported previously and analyzed by this study. Gutierrez et al. reported that the variation of *Zp-V3* (including alterations at sites −141, −106 and − 100 in *Zp* relative to the TSS of *BZLF1*) was more commonly observed in cancer cases (including non-Hodgkin's lymphomas and NPC) than in controls [35–37]. In this study, we also found the higher frequency of *Zp-V3* in NPC samples (82.4%, 70/85) than in healthy controls from the NPC endemic (61.4%, 51/83) and non-endemic (66.3%, 57/86) areas, but no significant difference was observed between healthy controls from different geographic areas.

In conclusion, our study comprehensively characterized the variation patterns of EBV IE genes in NPC endemic and non-endemic areas of China. We identified two EBV subtypes based on the natural variations in *Rp* which exerted disparate reactivation potential of EBV. Our findings partially explained the reason for the elevated reactivation level of EBV in NPC endemic areas than non-endemic areas and provided valuable insight towards understanding the mechanism of the regional disparity of NPC incidence.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ebiom.2018.10.065.

Conflict of interests

The authors declare no conflict of interest.

Author contributions

WHJ conceived, designed the study and wrote the final draft. JBZ participated in study design, performed data collection and analysis, carried out the experiments, drafted the manuscript. SYH, SQD and FW assisted the experiments. TMW performed bioinformatic analysis. YQH, XHZ and XZL assisted participant recruitment and data collection. JM assisted manuscript edition. All authors read and approved the final manuscript.

Acknowledgements

This work was supported by the National Key Research and Development Program of China [grant number: 2016YFC1302704], the Sino-Sweden joint research program [grant number: 81861138006], the Science and Technology Planning Project of Guangzhou City, China [grant number: 201804020094] and the National Natural Science Foundation of China [grant number: 81502056, 81325018 and 81220108022]. The funder had no role in the study design, participant recruitment, data collection, data analysis, data interpretation, or writing of the report. We thank Hui-Jun Li from First General Hospital of Yangquan city, Hong-Shen and Xiu-Hua Yu from General Hospital of Mishan city, Qing-hua Pan from Ganzhou cancer hospital and Cai Yong-lin from Wuzhou Red Cross hospital for their contributions in the recruitment of study population.

References

[1] Yu MC, Yuan JM. Epidemiology of nasopharyngeal carcinoma. Semin Cancer Biol 2002;12(6):421–9.
[2] Jia WH, Huang QH, Liao J, et al. Trends in incidence and mortality of nasopharyngeal carcinoma over a 20-25 year period (1978/1983-2002) in Sihui and Cangwu counties in southern China. BMC Cancer 2006;6:178.
[3] Chang CM, Yu KJ, Mbulaiteye SM, Hildesheim A, Bhatia K. The extent of genetic diversity of Epstein-Barr virus and its geographic and disease patterns: a need for reappraisal. Virus Res 2009;143(2):209–21.
[4] Torre LA, Bray F, Siegel RL, et al. Global cancer statistics, 2012. CA Cancer J Clin 2015;65(2):87–108.
[5] Lo KW, Chung GT, To KF. Deciphering the molecular genetic basis of NPC through molecular, cytogenetic, and epigenetic approaches. Semin Cancer Biol 2012;22(2):79–86.
[6] Bei JX, Li Y, Jia WH, et al. A genome-wide association study of nasopharyngeal carcinoma identifies three new susceptibility loci. Nat Genet 2010;42(7):599–603.
[7] Bei JX, Jia WH, Zeng YX. Familial and large-scale case-control studies identify genes associated with nasopharyngeal carcinoma. Semin Cancer Biol 2012;22(2):96–106.
[8] Jia WH, Qin HD. Non-viral environmental risk factors for nasopharyngeal carcinoma: a systematic review. Semin Cancer Biol 2012;22(2):117–26.
[9] Farrell PJ. Epstein-Barr Virus Strain Variation. Curr Top Microbiol Immunol 2015;390:45–69 Pt 1.
[10] Neves M, Marinho-Dias J, Ribeiro J, Sousa H. Epstein-Barr virus strains and variations: Geographic or disease-specific variants? J Med Virol 2017;89(3):373–87.
[11] Tzellos S, Farrell PJ. Epstein-barr virus sequence variation-biology and disease. Pathogens 2012;1(2):156–74.
[12] Mainou BA, Raab-Traub N. LMP1 strain variants: biological and molecular properties. J Virol 2006;80(13):6458–68.
[13] Hu LF, Zabarovsky ER, Chen F, et al. Isolation and sequencing of the Epstein-Barr virus BNLF-1 gene (LMP1) from a Chinese nasopharyngeal carcinoma. J Gen Virol 1991;72:2399–409 Pt 10.
[14] Hadhri-Guiga B, Khabir AM, Mokdad-Gargouri R, et al. Various 30 and 69 bp deletion variants of the Epstein-Barr virus LMP1 may arise by homologous recombination in nasopharyngeal carcinoma of Tunisian patients. Virus Res 2006;115(1):24–30.
[15] Li A, Zhang XS, Jiang JH, et al. Transcriptional expression of RPMS1 in nasopharyngeal carcinoma and its oncogenic potential. Cell Cycle 2005;4(2):304–9.
[16] Jing YZ, Wang Y, Jia YP, Luo B. Polymorphisms of Epstein-Barr virus BHRF1 gene, a homologue of bcl-2. Chin J Cancer 2010;29(12):1000–5.
[17] Hutajulu SH, Hoebe EK, Verkuijlen SA, et al. Conserved mutation of Epstein-Barr virus-encoded BamHI-A Rightward Frame-1 (BARF1) gene in Indonesian nasopharyngeal carcinoma. Infect Agent Cancer 2010;5:16.
[18] Kawaguchi A, Kanai K, Satoh Y, et al. The evolution of Epstein-Barr virus inferred from the conservation and mutation of the virus glycoprotein gp350/220 gene. Virus Genes 2009;38(2):215–23.
[19] Luo B, Liu M, Chao Y, et al. Characterization of Epstein-Barr virus gp350/220 gene variants in virus isolates from gastric carcinoma and nasopharyngeal carcinoma. Arch Virol 2012;157(2):207–16.
[20] Zeng MS, Li DJ, Liu QL, et al. Genomic sequence analysis of Epstein-Barr virus strain GD1 from a nasopharyngeal carcinoma patient. J Virol 2005;79(24):15323–30.
[21] Liu P, Fang X, Feng Z, et al. Direct sequencing and characterization of a clinical isolate of Epstein-Barr virus from nasopharyngeal carcinoma tissue by using next-generation sequencing technology. J Virol 2011;85(21):11291–9.
[22] Kwok H, Tong AH, Lin CH, et al. Genomic sequencing and comparative analysis of Epstein-Barr virus genome isolated from primary nasopharyngeal carcinoma biopsy. Plos One 2012;7(5):e36939.
[23] Kwok H, Wu CW, Palser AL, et al. Genomic diversity of Epstein-Barr virus genomes isolated from primary nasopharyngeal carcinoma biopsy samples. J Virol 2014;88(18):10662–72.
[24] Correia S, Palser A, Elgueta KC, et al. Natural Variation of Epstein-Barr Virus Genes, Proteins, and primary MicroRNA. J Virol 2017;91(15).
[25] Palser AL, Grayson NE, White RE, et al. Genome diversity of Epstein-Barr virus from multiple tumor types and normal infection. J Virol 2015;89(10):5222–37.
[26] Zhou L, Chen JN, Qiu XM, et al. Comparative analysis of 22 Epstein-Barr virus genomes from diseased and healthy individuals. J Gen Virol 2017;98(1):96–107.
[27] Chiara M, Manzari C, Lionetti C, et al. Geographic Population Structure in Epstein-Barr Virus Revealed by Comparative Genomics. Genome Biol Evol 2016;8(11):3284–91.
[28] Kwok H, Chiang AK. From conventional to Next Generation Sequencing of Epstein-Barr Virus Genomes. Viruses 2016;8(3):60.
[29] Tsao SW, Tsang CM, Lo KW. Epstein-Barr virus infection and nasopharyngeal carcinoma. Philos Trans R Soc Lond B Biol Sci 2017;372(1732).
[30] McKenzie J, El-Guindy A. Epstein-Barr Virus Lytic Cycle Reactivation. Curr Top Microbiol Immunol 2015;391:237–61.

[31] Chien YC, Chen JY, Liu MY, et al. Serologic markers of Epstein-Barr virus infection and nasopharyngeal carcinoma in Taiwanese men. N Engl J Med 2001;345(26):1877–82.
[32] Chan KC, Lo YM. Circulating EBV DNA as a tumor marker for nasopharyngeal carcinoma. Semin Cancer Biol 2002;12(6):489–96.
[33] Zheng XH, Lu LX, Li XZ, Jia WH. Quantification of Epstein-Barr virus DNA load in nasopharyngeal brushing samples in the diagnosis of nasopharyngeal carcinoma in southern China. Cancer Sci 2015;106(9):1196–201.
[34] Xu FH, Xiong D, Xu YF, et al. An epidemiological and molecular study of the relationship between smoking, risk of nasopharyngeal carcinoma, and Epstein-Barr virus activation. J Natl Cancer Inst 2012;104(18):1396–410.
[35] Gutierrez MI, Ibrahim MM, Dale JK, et al. Discrete alterations in the BZLF1 promoter in tumor and non-tumor-associated Epstein-Barr virus. J Natl Cancer Inst 2002;94(23):1757–63.
[36] Tong JH, Lo KW, Au FW, Huang DP, To KF. Re: Discrete alterations in the BZLF1 promoter in tumor and non-tumor-associated Epstein-Barr virus. J Natl Cancer Inst 2003;95(13):1008–1009, 1009.
[37] Bristol JA, Djavadian R, Albright ER, et al. A cancer-associated Epstein-Barr virus BZLF1 promoter variant enhances lytic infection. PLoS Pathog 2018;14(7):e1007179.
[38] Kenney SC. Reactivation and lytic replication of EBV; 2007.
[39] Hubisz MJ, Falush D, Stephens M, Pritchard JK. Inferring weak population structure with the assistance of sample group information. Mol Ecol Resour 2009;9(5):1322–32.
[40] Evanno G, Regnaut S, Goudet J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 2005;14(8):2611–20.
[41] Pepperl S, Benninger-Doring G, Modrow S, Wolf H, Jilg W. Immediate-early transactivator Rta of Epstein-Barr virus (EBV) shows multiple epitopes recognized by EBV-specific cytotoxic T lymphocytes. J Virol 1998;72(11):8644–9.
[42] Steven NM, Annels NE, Kumar A, et al. Immediate early and early lytic cycle proteins are frequent targets of the Epstein-Barr virus-induced cytotoxic T cell response. J Exp Med 1997;185(9):1605–17.
[43] Yu H, Srinivasan N, Ren E, Chan S. Identification of CD8+ T-cell epitopes specific for immediate-early transactivator Rta of Epstein-Barr virus. Hum Immunol 2005;66(5):483–93.
[44] Kenney SC, Mertz JE. Regulation of the latent-lytic switch in Epstein-Barr virus. Semin Cancer Biol 2014;26:60–8.
[45] Zalani S, Coppage A, Holley-Guthrie E, Kenney S. The cellular YY1 transcription factor binds a cis-acting, negatively regulating element in the Epstein-Barr virus BRLF1 promoter. J Virol 1997;71(4):3268–74.
[46] Petosa C, Morand P, Baudin F, et al. Structural basis of lytic cycle activation by the Epstein-Barr virus ZEBRA protein. Mol Cell 2006;21(4):565–72.
[47] Dawson CW, Port RJ, Young LS. The role of the EBV-encoded latent membrane proteins LMP1 and LMP2 in the pathogenesis of nasopharyngeal carcinoma (NPC). Semin Cancer Biol 2012;22(2):144–53.
[48] Lo AK, Dawson CW, Jin DY, Lo KW. The pathological roles of BART miRNAs in nasopharyngeal carcinoma. J Pathol 2012;227(4):392–403.
[49] Fang CY, Lee CH, Wu CC, et al. Recurrent chemical reactivations of EBV promotes genome instability and enhances tumor progression of nasopharyngeal carcinoma cells. Int J Cancer 2009;124(9):2016–25.
[50] Jia Y, Wang Y, Chao Y, et al. Sequence analysis of the Epstein-Barr virus (EBV) BRLF1 gene in nasopharyngeal and gastric carcinomas. Virol J 2010;7:341.
[51] Heilmann AM, Calderwood MA, Portal D, Lu Y, Johannsen E. Genome-wide analysis of Epstein-Barr virus Rta DNA binding. J Virol 2012;86(9):5151–64.
[52] Kraus RJ, Perrigoue JG, Mertz JE. ZEB negatively regulates the lytic-switch BZLF1 gene promoter of Epstein-Barr virus. J Virol 2003;77(1):199–207.
[53] Zalani S, Holley-Guthrie E, Kenney S. The Zif268 cellular transcription factor activates expression of the Epstein-Barr virus immediate-early BRLF1 promoter. J Virol 1995;69(6):3816–23.
[54] Ragoczy T, Miller G. Autostimulation of the Epstein-Barr virus BRLF1 promoter is mediated through consensus Sp1 and Sp3 binding sites. J Virol 2001;75(11):5240–51.
[55] Wille CK, Nawandar DM, Panfil AR, et al. Viral genome methylation differentially affects the ability of BZLF1 versus BRLF1 to activate Epstein-Barr virus lytic gene expression and viral replication. J Virol 2013;87(2):935–50.
[56] Nawandar DM, Wang A, Makielski K, et al. Differentiation-Dependent KLF4 Expression Promotes Lytic Epstein-Barr Virus Infection in Epithelial Cells. PLoS Pathog 2015;11(10):e1005195.
[57] Chen C, Li D, Guo N. Regulation of cellular and viral protein expression by the Epstein-Barr virus transcriptional regulator Zta: implications for therapy of EBV associated tumors. Cancer Biol Ther 2009;8(11):987–95.
[58] Ji KM, Li CL, Meng G, Han AD, Wu XL. New BZLF1 sequence variations in EBV-associated undifferentiated nasopharyngeal carcinoma in southern China. Arch Virol 2008;153(10):1949–53.