





EBV genome variations enhance clinicopathological features of nasopharyngeal carcinoma in a non-endemic region

Satoru Kondo¹  | Yusuke Okuno^{2,3} | Takayuki Murata⁴ | Hirotomo Dochi¹  | Naohiro Wakisaka¹ | Harue Mizokami¹ | Makiko Moriyama-Kita¹ | Eiji Kobayashi¹ | Makoto Kano¹ | Takeshi Komori¹ | Nobuyuki Hirai¹ | Takayoshi Ueno¹ | Yosuke Nakanishi¹ | Kazuhira Endo¹ | Hisashi Sugimoto¹ | Hiroshi Kimura⁵  | Tomokazu Yoshizaki¹ 

¹Division of Otolaryngology and Head and Neck Surgery, Graduate School of Medical Science, Kanazawa University, Kanazawa, Japan

²Department of Virology, Nagoya City University Graduate School of Medical Sciences, Nagoya, Japan

³Pediatric Cancer Treatment Center, Nagoya University Hospital, Nagoya, Japan

⁴Department of Virology and Parasitology, Fujita Health University School of Medicine, Toyoake, Japan

⁵Department of Virology, Nagoya University, Graduate School of Medicine, Nagoya, Japan

Correspondence

Tomokazu Yoshizaki, Division of Otolaryngology and Head and Neck Surgery, Graduate School of Medical Science, Kanazawa University, Kanazawa, Ishikawa 920-8640, Japan.
Email: tomoy@med.kanazawa-u.ac.jp

Funding information

Ministry of Education, Culture, Sports, Science and Technology, Grant/Award Number: 16H05480 and 17H01590; Mochida Memorial Foundation for Medical and Pharmaceutical Research; Takeda Science Foundation

Abstract

Nasopharyngeal carcinoma (NPC) is caused by infection with Epstein–Barr virus (EBV) and endemic in certain geographic regions. EBV lytic gene, *BALF2*, closely associates with viral reactivation and *BALF2* gene variation, the H-H-H strain, causes NPC in endemic region, southern China. Here, we investigate whether such EBV variations also affect NPC in a non-endemic region, Japan. Viral genome sequencing with 47 EBV isolates of Japanese NPC were performed and compared with those of other EBV-associated diseases from Japan or NPC in Southern China. EBV genomes of Japanese NPC are different from those of other diseases in Japan or endemic NPC; Japanese NPC was not affected by the endemic strain (the *BALF2* H-H-H) but frequently carried the type 2 EBV or the strain with intermediate risk of endemic NPC (the *BALF2* H-H-L). Seven single nucleotide variations were specifically associated with Japanese NPC, of which six were present in both type 1 and 2 EBV genomes, suggesting the contribution of the type 2 EBV-derived haplotype. This observation was supported by a higher viral titer and stronger viral reactivation in NPC with either type 2 or H-H-L strains. Our results highlight the importance of viral strains and viral reactivation in the pathogenesis of non-endemic NPC.

KEYWORDS

BALF2, lytic infection, nasopharyngeal carcinoma, type 2 EBV, virus reactivation

Abbreviations: BART, *BamHI* A rightward transcripts; CAEBV, chronic active Epstein–Barr virus infection; EBV, Epstein–Barr virus; LMP, latent membrane protein; miRNA, microRNA; NPC, nasopharyngeal carcinoma; SNV, structural nucleotide variant; SV, structural variation; VCA, viral capsid antigen.

Satoru Kondo and Yusuke Okuno contributed equally to this work.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2022 The Authors. *Cancer Science* published by John Wiley & Sons Australia, Ltd on behalf of Japanese Cancer Association.

1 | INTRODUCTION

Epstein–Barr virus has been recorded as one of the most prevalent viruses in humans since its discovery in 1964,^{1,2} and is a causative agent of hematologic and epithelial cell neoplasms, including a subset of gastric carcinoma and NPC.^{3,4} EBV is classically divided into type 1 (T1-EBV) and type 2 (T2-EBV)^{5,6} based on the genetic differences in the Epstein–Barr nuclear antigens. The two types are reported to have different geographical distributions and biological characteristics: T1-EBV is predominant in Western and Asian countries, whereas T2-EBV is relatively frequent in Africa.⁷ T1-EBV transforms human cells more efficiently than T2-EBV. Associations between EBV types and diseases have been discussed previously; for example, T2-EBV is frequently detected in NPC in Africa,^{8–10} whereas it is rarely detected in other regions.^{8–10}

Nasopharyngeal carcinoma is an EBV-associated epithelial cell malignancy. One of its marked characteristics is the geographical difference in morbidity; the incidence rate is ~25 per 100,000 individuals in endemic areas such as southern China. In contrast, the incidence is less than 0.5–1 per 100,000 individuals in non-endemic areas such as Japan, North America, and North China.¹¹ Risk factors for NPC development include feeding habits of salty fish and genetic rearrangement of the *HLA*, *CDKN2A/B*, *TNFRSF19*, and *TERT* loci.^{12–15} In addition, a recent study revealed a strong association between an EBV strain carrying three single nucleotide variants (*BALF2* 162215A>C, 162476T>C, and 163364C>T variants) in the EBV genome (the *BALF2* H-H-H-endemic strain) and NPC in the endemic region.¹⁶ However, definitive risk factors attributable to the host, as well as the uneven geographic distribution of this disease, have not been elucidated yet, especially for NPC in non-endemic regions.

More than 500 EBV genome sequences from various EBV-associated diseases and cell lines have been identified since the whole genome sequence of B95-8,^{6,17–19} a representative EBV strain, was demonstrated in 1984.²⁰ Studies on these EBV strains have elucidated the existence of genomic variations in the geographic area that predispose individuals to certain EBV-associated diseases. In addition, our recent study revealed frequent intragenic deletions that were prevalent among hematologic malignancies.²¹ As the study was conducted in a non-endemic area of NPC in Japan, it was quite interesting to determine whether these gene modifications had also been identified in Japanese NPC patients.

In the present study, we sequenced the EBV genome of Japanese patients with NPC. We also analyzed whether viral genomic variations affected clinical and pathological features such as clinical tumor stages, serum EBV-DNA load, and viral reactivation status in tumor specimens.

2 | MATERIALS AND METHODS

2.1 | The study cohort

This study included 47 NPC patients who were diagnosed at the Department of Otolaryngology, Head and Neck Surgery, Kanazawa University Hospital between 1997 and 2020. The diagnoses of NPC

were done histologically. All patients were of Japanese ethnicity. Staging was done according to the American Joint Committee on Cancer/International Union Against Cancer Stage Classification. All the human samples used in this study were obtained from consenting patients with written informed consent. This study was conducted in accordance with the principles of all relevant ethical regulations including the Declaration of Helsinki. The institutional review boards of the Kanazawa University, Graduate School of Medical Science and the Nagoya University Graduate School of Medicine approved this study.

We also analyzed 178 EBV isolates of other EBV-associated diseases from Japan (77 patients with T/NK-type CAEBV, 23 patients with extranodal NK/T-cell lymphoma nasal type, 14 patients with EBV-positive diffuse large B-cell lymphoma, 15 patients with infectious mononucleosis, 32 patients with posttransplantation lymphoproliferative disorder, and seven patients with other EBV-associated malignancies) (European Genome-phenome Archive, accession no. EGAD00001004298), as well as 110 EBV genomes from NPC in Southern China/Hong Kong (randomly chosen from a total of 217 publicly available data in the NCBI SRA database, accessed on November 20, 2020; Table S1).

2.2 | Tissue samples

Genomic DNA was extracted from biopsy specimens embedded in paraffin or frozen tissues using the QIAamp DNA Mini kit (Qiagen) with some modifications as described previously.²² In brief, paraffin sections were incubated with 1 ml xylene at 55°C for 15 min. After centrifugation, the supernatant fluid was discarded. The pellet was washed twice with 1 ml of 1/1 xylene/100% ethanol and twice with 100% ethanol. The pellets were incubated with 180 µl ATL buffer and 20 µl proteinase K at 55°C for 48 h with intermittent vortexing, after which the manufacturer's protocol was followed.

2.3 | Resequencing analysis of the EBV genome

We used our custom SureSelect bait to capture Akata and other strains of EBV (Agilent Technologies) as described previously.²¹ Briefly, we designed 17,237 tiling probes that covered the Akata genome (RefSeq accession no. KC207813.1) with 10× coverage and added them to the original EBV bait covering six other strains.^{6,23} We performed target enrichment and library preparation according to the manufacturer's instructions. Sequencing was performed using a HiSeqX system (Illumina) with 2 × 150-bp end reads. We obtained a minimum of 5 million reads per sample.

Sequence reads were adapter trimmed using an in-house program. The processed reads were aligned to the EBV reference genome (RefSeq accession no. NC_007605.1) using NovoAlign (NovoCraft, <http://novocraft.com>) with default parameters and a '-r Random' option. Variants were detected using VarScan2 and annotated using ANNOVAR (<https://annovar.openbioinformatics>).

org/). We considered variants with P -values < 0.001 (provided by VarScan2) and VAF > 0.8 to be present. Random simulations of nucleotide variations were performed using the Mersenne Twister pseudo-random number generator to generate simulated nucleotide alterations throughout the EBV genome.

To detect structural variations, soft-clipped bases were realigned to the human (hg19) and the EBV (NC_007605.1) genomes using BLAT.²⁴ Candidate structural variations with both sides of the breakpoints in the repeat region were filtered. Visualization was performed using the Integrative Genomics Viewer.²⁵

For clustering of EBV genomes, the number of nucleotide alterations between two EBV samples was calculated using our in-house program. Hierarchical clustering and visualization were performed using the “hclust” command with the Manhattan distance calculation and the average linkage method in R (<http://www.R-project.org/>).

EBV type 1/2 classification was based on hierarchical clustering, which matched the classification using SNVs in *EBNA-2*, *EBNA-3A*, *EBNA-3B*, and *EBNA-3C*. *BALF2* subtype classification was simply based on the variant call. After these classifications, EBV genomes were classified as type 2 (regardless of the *BALF2* subtype), the H-H-H subtype (in the type 1 EBV genomes), the H-H-L subtype (in the type 1 EBV genomes), and the other subtype (H-L-L and L-L-L subtypes in the type 1 EBV genomes).

2.4 | Identification of EBV variants associated with Japanese NPC

We compared Japanese EBV genomes between NPC and other diseases. The differences in the frequency of variants were assessed individually using Fisher's exact test. The P -value of 3×10^{-7} was considered as a reasonable cutoff after considering the multiple testing correction, as the P -value was still significant (adjusted $p = 0.05$) after multiplying by the length of the viral genome (171,823 nt).

2.5 | DNA extraction from serum samples

The serum samples were stored at -80°C until further processing. DNA was extracted using the QIAamp DNA Mini kit (Qiagen) according to the manufacturer's protocol. Serum samples (200 μl) were used for DNA extraction after the addition of 10 μg of poly(A) (Roche Diagnostics K. K.) as a carrier RNA. A volume of 50 μl of the final eluate was used. Extracted DNA was stored at -20°C until further use.

2.6 | Real-time quantitative EBV-DNA PCR

A 3000P Real-Time QPCR System (Agilent Technologies) was used to perform qPCR and fluorescence measurements. Duplicate PCR mixtures of each sample were amplified using an EBV

primer probe set from the Japan Molecular Center of Excellence (Nippon Genetics) and an EagleTaq Master Mix with Rox (Roche Diagnostics K. K.).^{22,26} The PCR primers from the EBV primer probe set corresponded to the *BALF5* gene encoding the viral DNA polymerase. The upstream and downstream primer sequences were 5'-CGGAAGCCCTCTGGACTTC-3' and 5'-CCCTGTTTATCCGATGGAATG-3', respectively. The fluorogenic probe sequence was 5'-TGTACACGCACGAGAAATGCGCC-3'. qPCR was performed as follows: denaturation at 95°C for 10 min, 50 cycles at 95°C for 10 s and 62°C for 60 s, and finally, extension at 40°C for 30 s. To check whether DNA contamination had occurred, distilled water samples were also used as a negative control in every experiment.

2.7 | Cell culture and plasmids

HK1, EBV-negative nasopharyngeal cancer cell line (kind gift from Dr. George Tsao, Hong Kong University) was maintained in RPMI1640 medium (Thermo Fisher Scientific) with 10% FBS and 1% antibiotic-antimycotic (Thermo Fisher Scientific). HK1 cells were authenticated using short tandem repeats (BEX) and passaged not more than 20 times. Cells were tested for the absence of mycoplasma contamination using the CycleavePCR Mycoplasma Detection Kit (TaKaRa Bio) before further experiments. The EBV *BART3* miRNA expression vector in pLCE was a kind gift from Dr. Rebecca Skalsky (Oregon Health & Science University).²⁷ The psiCHECK-2 vector (Promega), a dual-luciferase plasmid, harbored the synthetic firefly luciferase (*Fluc*) gene and the synthetic Renilla luciferase (*hRLuc*) gene. Each gene possessed its own promoter and poly(A)-addition sites.

To generate the luciferase reporter, the 3'-UTR of the LMP1 fragment was digested from the pLSG-LMP1 3'-UTR between the *XhoI*-*NotI* restriction sites (a kind gift from Dr. Rebecca Skalsky, Oregon Health & Science University).²⁷ The fragment was ligated in the multiple cloning regions of *hRLuc* in psiCHECK-2 (wild-type: WT). Moreover, two mutant vectors were constructed by replacing two or six bases of LMP1 (mutant 1: MT1; mutant 2: MT2) using site-directed mutagenesis (Fasmac, Atsugi, Japan).²⁷ DNA sequencing was performed to confirm the nucleotide sequences of the constructed plasmids.

2.8 | Luciferase 3'-UTR reporter assay

HK1 cells (1.5×10^5) were seeded into 24-well plates and transfected with 80 ng psiCHECK-2 reporter (*LMP1* 3'-UTR WT, MT1, or MT2) and 1000 ng EBV *BART3* miRNA vector or pLCE control vector using Lipofectamine 2000 (Thermo Fisher Scientific). At 48 h post-transfection, cells were lysed in 1 \times passive lysis buffer (Promega), and luciferase expression assay was performed using the dual reporter luciferase assay system (Promega). All experiments were performed in triplicate.

2.9 | Immunofluorescence assay

Indirect and anticomplement immunofluorescence assays were carried out with Nippon Kayaku VCA slides and kits (Nippon Kayaku) for the detection of IgG and IgA antibodies to VCA according to the manufacturer's instructions. The serum samples were serially diluted (1:2), with final dilutions ranging from 1:10 to 1:10,240. The concentration of the antibodies was expressed as a titer, with the endpoint corresponding to the last dilution that clearly showed fluorescence. If serum diluted 1:10 showed no fluorescence, that titer was expressed as 0.

2.10 | Immunohistochemical analysis

The expression of BALF2 was immunohistochemically examined in NPC biopsy specimens at the first medical examination. Here, 3- μ m sections were cut from paraffin blocks of a primary lesion. The paraffin sections were deparaffinized, treated with 3% hydrogen peroxide, and incubated with a protein blocker (Agilent Technologies). Next, the sections were incubated at 4°C overnight with anti-BALF2 antibody, as described previously.²⁸ After washing with PBS, the sections were exposed to EnVision+secondary antibody (Agilent Technologies). Then, the sections were counterstained with hematoxylin. BALF2 expression was independently evaluated by two investigators (SK and HD) without information on the clinical data of the patients.

2.11 | Statistical analysis

The clinical characteristics of the patients were analyzed using Fisher's exact test and the chi-squared test. The means of the two groups were compared using Student's *t*-test. Mann-Whitney test and Kruskal-Wallis test were used for comparison of the means of EBV-specific antibodies. All *P*-values reported are two-sided and *P*-values < 0.05 were considered significant.

All statistical analyses were performed using the SPSS statistics package version 19 (IBM) and EZR (Saitama Medical Center, Jichi Medical University, Japan), which is a graphical user interface for R (The R Foundation for Statistical Computing).²⁹

3 | RESULTS

3.1 | Resequencing of EBV genomes obtained from Japanese NPC

We sequenced 47 EBV genomes isolated from Japanese NPC patients and compared them with those of other Japanese EBV (*n* = 178) and endemic NPC genomes (*n* = 110) (Figure 1A). The SNVs were similar between Japanese NPC EBV genomes and other Japanese EBV genomes with regard to effects on amino acids

(Figure 1B) and nucleotide alteration patterns (Figure 1C), and synonymous variants were more frequent than expected from random simulation ($dN/dS = 0.34$ for NPC and 0.38 for other diseases), suggesting that most of the identified variants were the result of highly conservative evolution. Nucleotide alteration patterns were biased toward transitions (cytosine-to-thymine and adenine-to-guanine alterations).

In addition to SNVs, in total five SVs were identified in four (8%) of the 47 genomes (four deletions and a tandem duplication; Figure 1D, Figure S1, and Table S2). They varied in size (231–24,787 nt) and affected genomic components such as latent membrane protein (*LMP*)-2A/2B, BART miRNA clusters 2, and several lytic cycle-associated genes; however, because of the small number of SVs and their variety, their significance is unclear. At the very least, the frequency of SVs was lower than that of hematologic malignancies (40%).²¹

Hierarchical clustering of the EBV genomes based on SNVs identified two separate clusters that contained the T1- and T2-EBV genomes (Figure 1A). In Japan, T2-EBV genomes were found more frequently in NPC than in other diseases (8/47 versus 12/178, *p* = 0.041; Figure 1E). Within the T1-EBV genomes, most EBV genomes obtained from endemic NPC patients formed a consolidated cluster. The T1-EBV genomes of Japanese NPC largely co-clustered with those of other diseases, suggesting that both Japanese EBV genomes had macroscopically similar SNVs; however, when focusing on the three SNVs associated with endemic NPC (*BALF2* 162215A>C, 162476T>C, and 163364C>T variants), 10 EBV genomes of Japanese NPC origin carried the former two SNVs (the H-H-L genotype) (Figure 1E). The frequency of the H-H-L genotype was higher in Japanese NPCs than in other diseases in Japan (10/47 versus 17/178, *P* = 0.042). All EBV isolates carrying all three SNVs (H-H-H genotype) originated from endemic NPC patients.

Taken together, the EBV genomes in Japanese NPC were significantly different from those in the endemic NPC, mainly regarding the frequency of the H-H-H genotype. In addition, Japanese EBV genomes were highly different between NPC and other diseases; the T2-EBV and H-H-L genotypes contributed to 38% of Japanese NPCs compared with 16% in other EBV-associated diseases in Japan (18/47 versus 29/178, *p* = 0.003). The association of the two types of EBV genomes with NPC suggested their specific role in the pathogenesis of this disease.

3.2 | Identification of viral SNVs associated with NPC in Japan

We compared the frequency of each SNV between Japanese NPCs and other diseases using a genome-wide approach (Figure 2A). As a result, differences with *P*-values < 3.0×10^{-7} were found in seven SNVs within two genomic regions: 107013T>G (*BBRF3* c.165T>G, p.L55L) and the other six SNVs in the 3'-untranslated region (UTR) of the latent membrane protein (*LMP*)-1 (g.167378–167470; Figure 2B,

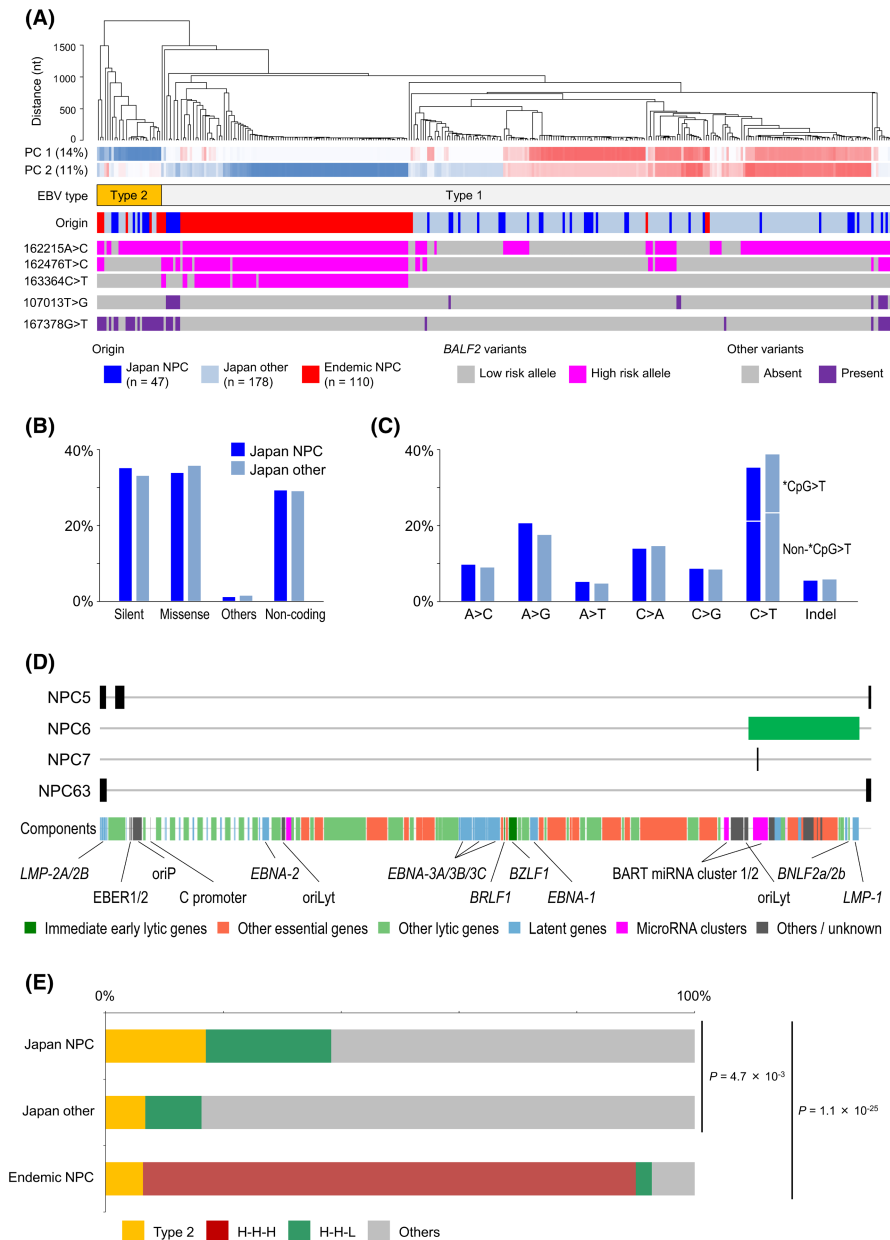


FIGURE 1 Genomic findings of the EBV genome from NPC in Japan. (A) Hierarchical clustering of the EBV genomes that originated from NPC in Japan (blue), other diseases in Japan (light blue), or NPC in the endemic region (red). Principal components (PC) of the nucleotide variations, EBV type 1/2 classification, and the BALF2 variants associated with the endemic NPC are also indicated. (B) The frequency of silent, missense, other coding (Others; nonsense and splice site variants), and noncoding SNVs in Japanese NPC and other diseases. (C) The frequency of each nucleotide alteration pattern or insertion/deletions (indels). For cytosine-to-thymine transitions, the frequency of alterations in the CpG and non-CpG contexts is also indicated. (D) Structural variations identified in Japanese NPC. Each horizontal line indicates an EBV genome, and black and green bars indicate deletions and a tandem duplication, respectively. Note that the viral genome is circular in host cells; the deletions affecting the left and right ends in NPC5 and NPC63 are continuous. The composition of the viral genome is also indicated. (E) The genotypes of EBV genome isolates. The viral genomes were classified either as type 2 (orange), type 1 H-HH (brown), type 1 H-H-L (green), or other type 1 genomes (gray). * $p < 0.01$

Table S3). The 107013T>G variant was present mainly in the H-H-L EBV genomes of Japanese NPC, suggesting that this variant was acquired in Japan (Figures 2C and 1A). In contrast, the representative SNV in the 167378–167470 region (167378G>T), as well as the other five SNVs, were present in both the T2-EBV and the H-H-L EBV genomes of Japanese NPC, suggesting that the SNVs originated from the T2-EBV. These SNVs were less frequent or absent in other diseases in Japan or endemic NPC, suggesting a specific association with Japanese non-endemic NPC (Figure 2D,E).

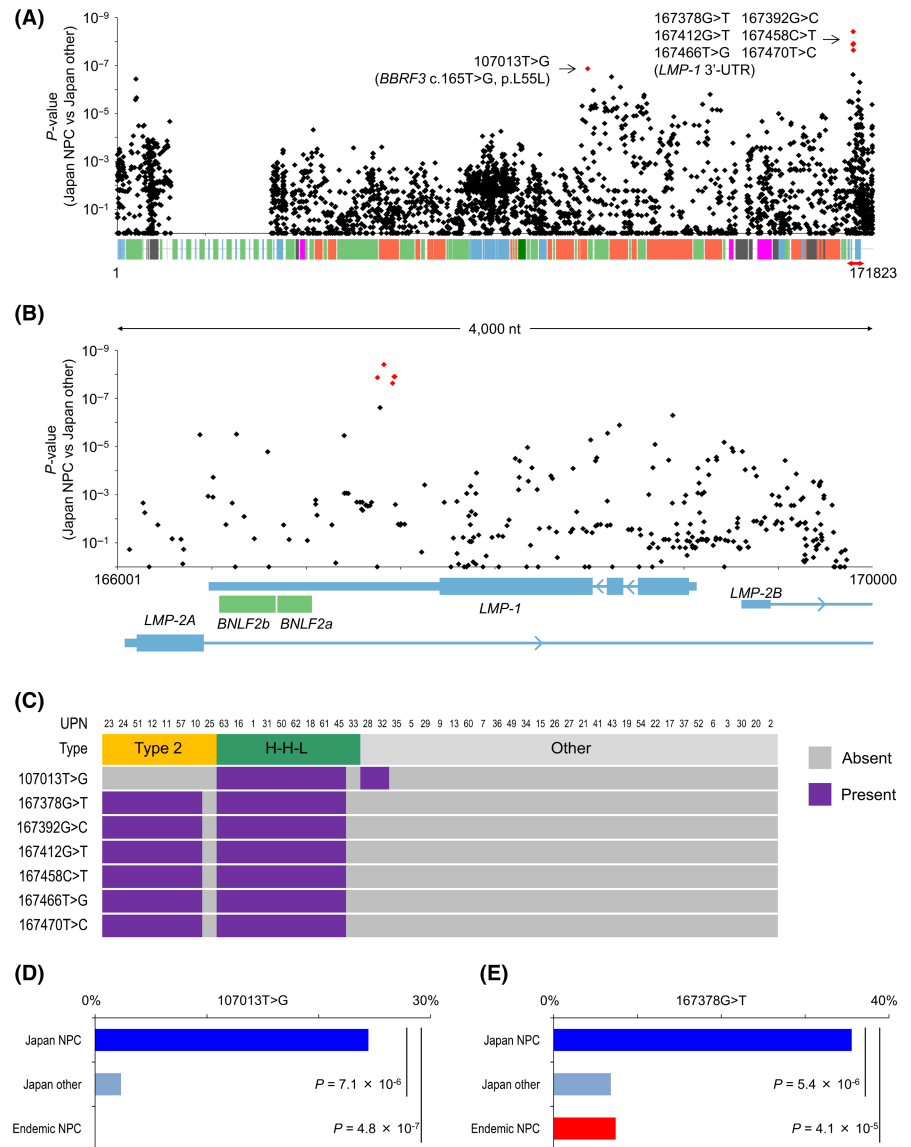
3.3 | Functional analysis of SNVs in the *LMP1* 3'-UTR

LMP1 encodes the major viral membrane protein and is transcriptionally downregulated by several viral microRNAs binding

to its 3'-UTR region.³⁰ Two of the six SNVs (167378G>T and 167392G>C) affected one of the binding sites of ebv-mir-BART3-5p (Figure 3A), suggesting the role of SNVs in transcriptional regulation, while the other four SNVs were not on the designated binding sites.

To assess the effect of the six SNVs in the *LMP1* 3'-UTR, we generated three types of the *LMP1* 3'-UTR luciferase reporter vectors (Figure 3B). The wild-type vector carried the original sequence of *LMP1* 3'-UTR sequence including the BART3-5p binding site. The *LMP1* 3'-UTR MT1 vector contained 2 SNVs (167378G>T and 167392G>C) that affected the binding site of ebv-mir-BART3-5p. In addition, *LMP1* 3'-UTR MT2 had six mutated base pairs that were all identified as SNVs in the 3'-UTR of *LMP1*. We transfected these reporters with or without the miR-BART3 expression vector into the HK1 nasopharyngeal cancer cell line. As shown in Figure 3C, miR-BART3 inhibited *LMP1* 3'-UTR WT but not *LMP1* 3'-UTR MT1 and

FIGURE 2 Discovery of Japanese NPC-specific EBV variants. (A) Manhattan plot of P -values obtained by comparing Japanese NPC and other diseases. Each dot indicates an SNV and seven red dots indicate SNVs with $P < 3.0 \times 10^{-7}$. UTR; untranslated region. (B) The focused image of the Manhattan plot. A 4-kb region (NC_007605: 166,001–170,000) is presented with the structure of genes in this region. Blue and green indicate latency- and lytic cycle-associated genes, respectively. Six SNVs with $P < 1.0 \times 10^{-7}$ were identified in the 3'-UTR of the latent membrane protein 1 (*LMP-1*) gene. (C) Correlation between EBV genotypes and SNVs in Japanese NPC. Each column indicates a genome. (D, E) The frequency of the 107013T>G variant (D) and the 167378G>T variant (E) in Japanese NPC, other diseases in Japan, and endemic NPC. * $P < 0.01$



LMP1 3'-UTR MT2. This result indicated that miR-BART3 targeted the *LMP1* 3'-UTR via the binding site of ebv-mir-BART3-5p, and that two SNVs inhibited the interaction to upregulate *LMP-1*.

3.4 | Association between EBV genotypes and clinical characteristics in Japanese NPC

As T2-EBV and H-H-L EBV were found frequently in Japanese NPC, we further characterized T2-EBV strains or the H-H-L subtype in terms of clinical presentation. As shown in Table 1, we found significant associations between these strains and the T-stage in the T-N-M classification at the time of diagnosis. Although the number of patients with T2-EBV was limited ($n = 8$), this viral strain was strongly associated with locally advanced (T3 or T4) T-stage (four and four patients with T3 and T4, respectively, $p = 0.01$). T2-EBV is also associated with relatively higher age at diagnosis compared with other subtypes (T2-EBV: 63.8 ± 12.8 years; H-H-L subtype: $58.8 \pm$

11.8 years; other subtypes: 55.0 ± 12.8 years), although the significance was marginal (T2-EBV versus other subtypes, $p = 0.09$).

3.5 | Association between EBV genotypes and serum EBV-DNA and antibody load in Japanese NPC

We measured serum EBV-DNA and EBV-VCA-IgG or EBV-VCA-IgA titers, which are predictors of disease progression in NPC,²² and investigated their association with EBV strains. Serum samples were available for eight patients with the T2-EBV strain, nine patients with the H-H-L subtype, and 27 patients with the other (H-L-L or L-L-L) subtypes. The viral DNA abundance in serum was significantly higher in NPC patients with T2-EBV (median 28,010 copies/ml; range 3550–594,975 copies/ml; $P = 0.007$) or the H-H-L subtype (median 60,075 copies/ml; range 1295–194,650 copies/ml; $P = 0.01$) than in NPC patients with the other subtypes (median 2610 copies/ml; range 0–80,300 copies/ml) (Figure 4A). In contrast, there was

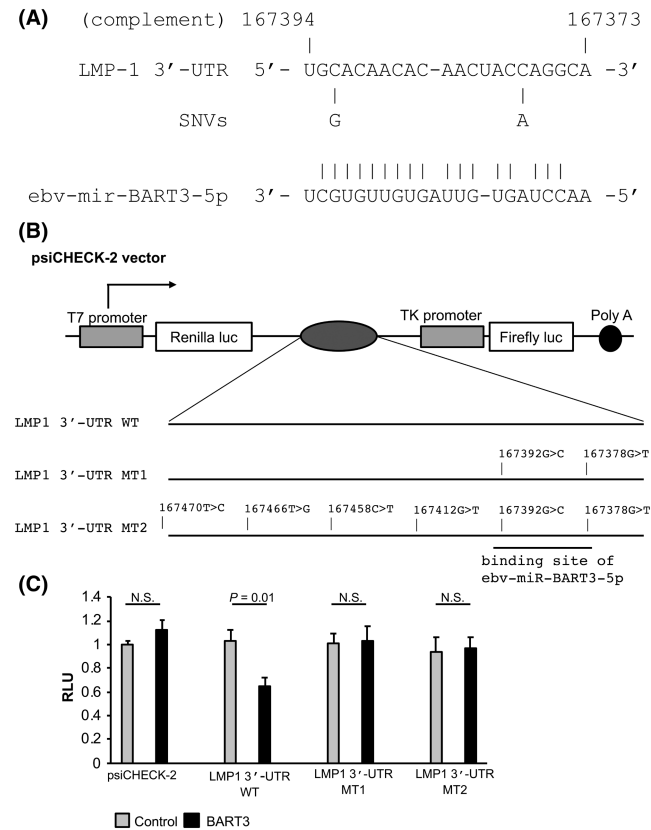


FIGURE 3 Functional analyses of the SNVs in the *LMP1* 3'-UTR region. (A) Schematic presentation of SNVs identified in the ebv-miR-BART3-5p binding sites of *LMP1*. Two SNVs (167378G>T and 167392G>C) were indicated together with the sequences of the 3'-UTR of *LMP1* (complement) and ebv-miR-BART3-5p. (B) Schematic representation of dual-luciferase reporter plasmid psiCHECK-2 vector containing the 3'-UTR of *LMP1* inserted downstream of the Renilla luciferase gene. Three types of dual-luciferase reporter plasmids were constructed, one wild-type (*LMP1* 3'-UTR WT) and two types of mutation reporters were designed. *LMP1* 3'-UTR MT1 had two replaced base pairs (167378G>T and 167392G>C) as described in (A). *LMP1* 3'-UTR MT2 replaced all six replaced base pairs identified SNVs in Japanese NPC including *LMP1* 3'-UTR MT1. (C) Either 3'-UTR of *LMP1* WT reporter or MT reporter was co-transfected into HK1 cells with or without EBV BART3 miRNA expression plasmid (black bar) or control vector (pLCE: gray bar). Lysates were assayed for luciferase activity at 48 h after transfection. All firefly luciferase units are normalized to Renilla activity (relative light units, RLU) and are reported relative to empty vector. All values are represented as mean \pm SD of three independent experiments. Statistical significance is indicated as *P*-value

no significant difference in the VCA-IgG titer among NPC patients with T2-EBV (median 1:120; range 1:40–1:2560), H-H-L subtype (median 1:160; range 1:40–1:1280), and the other subtypes (median 1:160; range 0–1:640) (Figure 4B). Similarly, there was no significant difference between the VCA-IgA titers of NPC patients with T2-EBV (median 1:25; range 0–1:160), the H-H-L subtype (median 0; range 0–1:20), and the other subtypes (median 1:10; range 0–1:80) (Figure 4C).

3.6 | Induction of the viral lytic cycle in Japanese NPC tissues

Finally, because elevated viral DNA load and the T2-EBV, which is prone to the lytic cycle compared with the T1-EBV, suggest the involvement of the viral lytic cycle, we investigated the viral latent/lytic status in tumor tissues. We immunohistochemically stained BALF2, as this protein is not expressed in the latent state and is upregulated during the lytic cycle. Biopsy specimens were available for eight NPC patients with T2-EBV strain, eight NPC patients with H-H-L subtype, and 29 NPC patients with other subtypes. The performance of the anti-BALF2 antibody was confirmed using IgG cross-linking of Akata (-) EBV-neoR cells over 48 h. As IgG cross-linking induces lytic cycle in this cell line, BALF2 expression was gradually reinforced in time-dependent manner and peaked at 24 h (Figure S2A). Similarly, the RNA expression level of *BALF2* peaked at 24 h (Figure S2B). The Akata (-) EBV-neoR cells after treatment with IgG cross-linking for 24 h were used as a positive control (Figure 4D). Akata (-) EBV-neoR cells without IgG stimulation or normal adenoid tonsil were used as negative controls (Figure 4E,F). BALF2 expression in Akata (-) EBV-neoR cells localized in the nuclei and were observed as dark brown staining (Figure 4D). In contrast, BALF2 expression in NPC tissues localized in both cytoplasm and nuclei (Figure 4G,H, $\times 1000$ enlarged image).

The expression score of BALF2 in NPC patients with the T2-EBV ($11.75 \pm 7.59\%$, $p = 0.01$) or the H-H-L subtype ($10.75 \pm 9.48\%$, $p = 0.05$) was higher than that with the other subtypes ($7.06 \pm 5.34\%$) (Figure 4G–J).

4 | DISCUSSION

In this study, we identified the frequent association of EBV in the T2-EBV and *BALF2* H-H-L subtypes in Japanese NPC. These two genotypes were associated with more frequent lytic reactivation, as shown by higher viral DNA load and more frequent *BALF2* expression, suggesting the involvement of this viral cycle in the pathogenesis of NPC. As T2-EBV is shown to be more prone to lytic reactivation than T1-EBV,³¹ this viral subtype can drive NPC pathogenesis through the lytic cycle. The H-H-L subtype is also associated with the lytic cycle and can also contribute to cancer development using the same viral program.

To date, T2-EBV is believed to be most frequently observed type in the African population.⁷ In this study, we found that T2-EBV was found frequently in Japanese NPC and had a low frequency in the Japanese cohort with other EBV-associated diseases or endemic NPC. A previous study reported that, among healthy donors and patients with tonsillitis in Japan, T1-EBV is highly dominant, while T2-EBV is quite rare.³² Therefore, T2-EBV may be specific for the EBV subtype with NPC in Japanese cohort. Although further study is mandatory to analyze a large cohort that includes Japanese healthy volunteers, we believe that T2-EBV has a frequent association with Japanese NPC.

TABLE 1 Relationship between EBV type/BALF2 haplotype and the clinicopathological features

| Variables | Type 2-EBV (n = 8) | Type 1-EBV BALF2 type H-H-L (n = 10) | Other types (n = 29) | P-value |
|---------------------|--------------------|--------------------------------------|----------------------|---------|
| Age (years ± SD) | 63.8 ± 12.8 | 58.8 ± 11.8 | 55.0 ± 12.8 | 0.09 |
| Gender | | | | |
| Female | 2 | 2 | 1 | 0.12 |
| Male | 6 | 8 | 28 | |
| WHO histologic type | | | | |
| II | 8 | 6 | 23 | 0.11 |
| III | 0 | 4 | 6 | |
| T classification | | | | |
| T1 | 0 | 2 | 10 | 0.01* |
| T2 | 0 | 6 | 8 | |
| T3 | 4 | 0 | 5 | |
| T4 | 4 | 2 | 6 | |
| N classification | | | | |
| N0 | 1 | 1 | 3 | 0.31 |
| N1 | 4 | 0 | 7 | |
| N2 | 2 | 7 | 16 | |
| N3 | 1 | 2 | 3 | |
| M classification | | | | |
| M0 | 8 | 9 | 28 | 0.55 |
| M1 | 0 | 1 | 1 | |
| Stage | | | | |
| I | 0 | 0 | 2 | 0.48 |
| II | 0 | 0 | 4 | |
| III | 3 | 6 | 13 | |
| IV | 5 | 4 | 10 | |

* $P < 0.05$.

The lytic form of EBV contributes to tumor growth under various conditions. Several lines of evidence have suggested that a small number of tumor cells with the lytic form of viral protein level may promote tumor growth in NPC.³³⁻³⁶ In addition, frequent microdeletions of the BART miR clusters in the EBV genome in hematologic malignancies suggest the contribution of the lytic cycle in neoplasms.²¹ Conversely, B cells with an EBV lacking the essential transcription factor for the lytic cycle (*BZLF1*) grew more slowly than cells transformed with the wild-type virus in a mouse xenograft model.³⁶ More recently, this lytically defective mutant was also impaired in its ability to form lymphoma in a humanized mouse model.

Serum/plasma EBV-DNA is an indicator of lytic infection, and virions are present in the circulation of NPC patients.³⁷ The majority of EBV in NPC cells is reported to be in the latent phase³⁸; however, their data also suggest the hypothesis that some EBV in NPC could enter lytic replication, leading to tumor progression. Furthermore, the expression of some lytic genes was frequently detected in NPC.³⁹ In this study, we have shown that NPC with both T2-EBV and H-H-L subtypes have higher serum EBV-DNA loads than the other subtypes. This finding suggests that T2-EBV and

the H-H-L subtype frequently enter lytic replication, which leads to tumor progression.

Although the serum EBV-DNA level was elevated in patients with NPC and the T2-EBV and H-H-L subtypes, VCA antibodies were not elevated in patients with NPC with those subtypes. Our previous study showed that the sensitivity of VCA antibodies in a Japanese NPC cohort was lower than those in other reports from endemic areas of NPC.²² This could be due to geographic differences between endemic and non-endemic areas of NPC. This suggests that EBV-DNA is a more reliable marker than VCA antibodies that enhance the progression of this malignancy. These cofactors may enhance the status of EBV-DNA load and VCA antibodies in patients with NPC and T2-EBV and H-H-L subtypes.

The ratio of the two *BALF2* high-risk subtypes (H-H-H and H-H-L) was quite high (84.35% and 8.92%, respectively) in the endemic area of southern China.¹⁶ The prevalence of the two high-risk EBV subtypes was not associated with the risk of developing other EBV-related diseases and malignancies in their study. In contrast, we found that there were no EBV isolates in Japan carrying the H-H-H subtype, including other EBV-associated diseases, which can be explained by the geographical distribution of the H-H-H subtype.

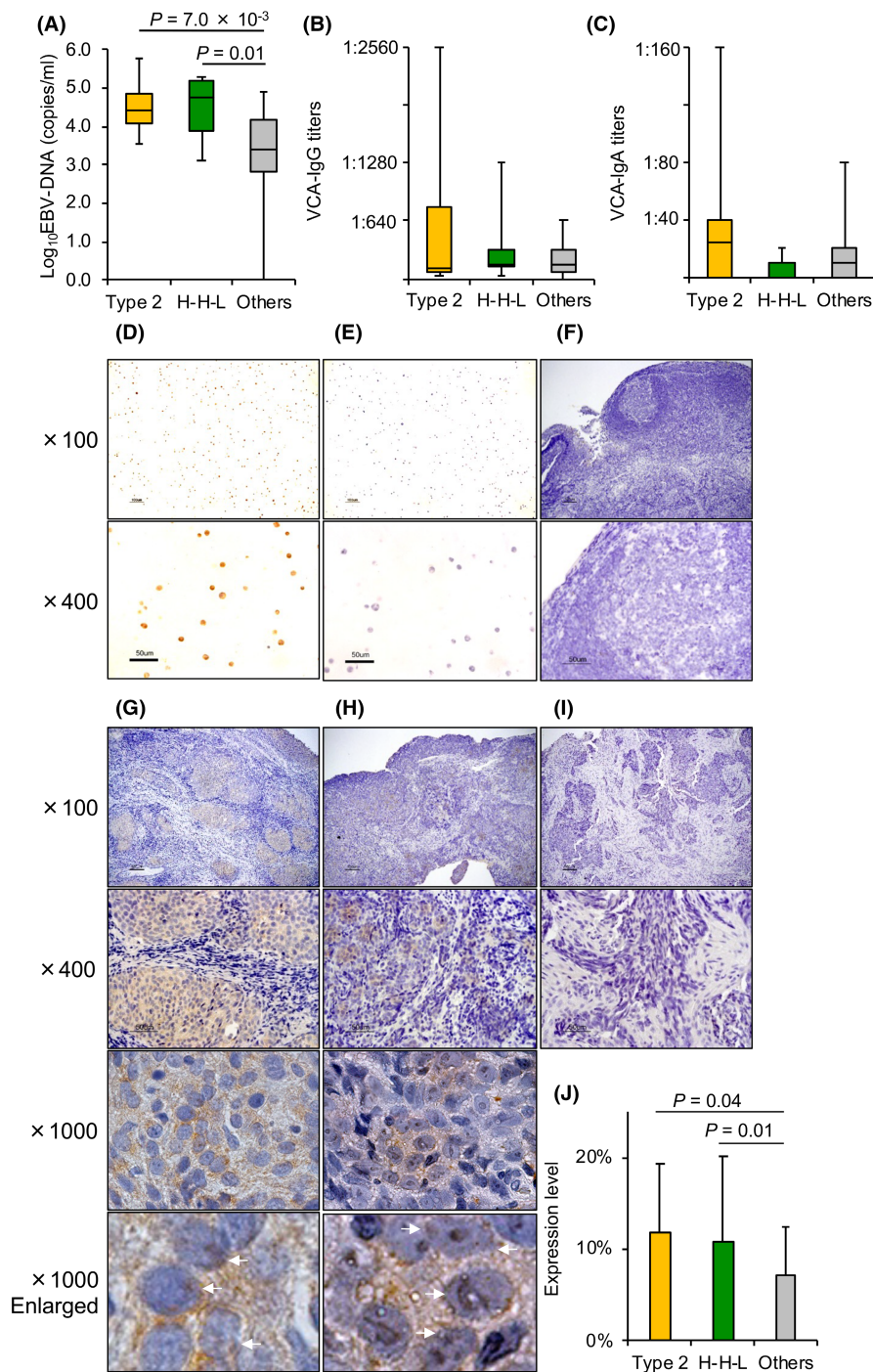


FIGURE 4 Comparison of clinical features of Japanese NPC categorized using EBV genome variations. (A–C) Comparison of serum cell-free EBV-DNA (A), VCA-IgG titers (B), VCA-IgA titers (C) in Japanese NPC among EBV genotypes. The bars and boxes indicate ranges and quantiles, respectively. (D–I) Immunohistochemical analysis of BALF2 expression in NPC tissues. Brown staining indicates positivity for BALF2 expression. Original magnification, upper figure $\times 100$, lower figure $\times 400$, respectively. For (G, H), magnified images with $\times 1000$ and enlargement $\times 1000$ are shown. White arrows indicate the representative image of nuclear expression of BALF2. Representative images with (D) positive control, paraffin block embedded Akata (-) EBV-neoR cells 24h after treatment with IgG cross-linking, (E) negative control, paraffin block embedded Akata (-) EBV-neoR cells without stimulation, (F) negative control, normal adenoid tissues. (G–I) Representative images of NPC tissues with type 2-EBV (G), H-H-L subtype (H), and others (BALF2_H-L-L and L-L-L) (I). (J) Comparison of expression scores of BALF2 in Japanese NPC. The categories are classified into type 2-EBV (yellow), BALF2 subtype H-H-L (green), and others (BALF2_H-L-L and L-L-L) (gray) indicated on the X-axis. The Y-axis indicates the expression score of BALF2 (%). Columns, mean; bars, SD

In addition, we found that the H-H-L subtype was found relatively frequently in Japanese NPCs. Although the mechanism by which *BALF2* variants contribute to NPC pathogenesis remains unclear, our findings provide new biological insights into EBV-mediated NPC oncogenesis, which is different between endemic and non-endemic areas.

BALF2 is an EBV single-stranded DNA-binding protein, which is indispensable for the lytic phase of EBV-DNA replication.^{8,10,40} Studies have shown that antibodies against EBV early lytic antigens, including *BALF2*, are highly enriched in the antibody signature for NPC risk prediction.⁴¹ In addition, *BALF2* is also a frequent target of

EBV-induced cytotoxic T-cell responses.⁴⁰ Because of the essential role of *BALF2* in EBV lytic DNA replication, these amino acid changes may influence the productive lytic cycle of EBV by altering the function of *BALF2*.

We found several SNVs in the 3'-UTR of *LMP1*, and two SNVs affected the binding site of ebv-mir-BART3-5p. BART miRs were originally identified in NPC and later identified in other EBV-associated neoplasms.⁴² BART miR expression is generally low in B lymphocytes and high in epithelial tissues.⁴³ EBV-miR-BART1-5p, miR-BART16, and miR-BART17-5p can target the *LMP1* 3'-UTR and regulate *LMP1* expression.³⁰ Another study showed that several

BART miRs, including miR-BART3, target the *LMP1* 3'-UTR.²⁷ Using luciferase reporter assays, we confirmed that miR-BART3 inhibited the *LMP1* 3'-UTR but not the *LMP1* 3'-UTR mutated in the binding site of miR-BART3. *LMP1* is believed to be important for NPC development^{44,45} and BART miRs contribute to the modulation of *LMP1* expression. Further analysis is necessary to investigate the exact mechanism by which BARTs enhance *LMP1* expression.

Finally, our results suggest a link between EBV genotypes and NPC, and a possible role of the viral lytic cycle in NPC development. More molecular and functional investigations are warranted to further clarify the association between EBV subtypes and variants, the lytic cycle, and NPC oncogenesis.

ACKNOWLEDGMENTS

We thank Dr. George Tsao (Hong Kong University) for providing HK1 cells. We also thank Dr. Rebecca Skalsky (Oregon Health & Science University) for providing the pLCE, pLCE-BART3 miRNA, and pLSG-3 *LMP1* 3'-UTR vectors. This work was supported by Grants-in-Aid from the Ministry of Education, Culture, Sports, Science, and Technology (16H05480 to SK, 17H01590 to TY) and was also supported by the Mochida Foundation to SK and the Takeda Science Foundation to YO. We would like to thank Editage (www.editage.com) for English language editing.

DISCLOSURE

The authors declare no potential conflicts of interest.

DATA AVAILABILITY STATEMENT

Nucleotide sequence data reported are available in the DDBJ Sequenced Read Archive (<https://www.ddbj.nig.ac.jp/dra/>) under the accession numbers [PRJDB12039](https://www.ddbj.nig.ac.jp/dra/) (BioProject ID) and [SAM00393851-SAM00393897](https://www.ddbj.nig.ac.jp/dra/) (BioSample ID).

ORCID

Satoru Kondo  <https://orcid.org/0000-0002-9066-0787>

Hiroto Dochi  <https://orcid.org/0000-0001-5711-7240>

Hiroshi Kimura  <https://orcid.org/0000-0001-8063-5660>

Tomokazu Yoshizaki  <https://orcid.org/0000-0002-2420-5540>

REFERENCES

- Epstein MA, Barr YM, Achong BG. A second virus-carrying tissue culture strain (Eb2) of lymphoblasts from Burkitt's lymphoma. *Pathol Biol.* 1964;12:1233-1234.
- Rickinson AB. Co-infections, inflammation and oncogenesis: future directions for EBV research. *Semin Cancer Biol.* 2014;26:99-115.
- Vockerodt M, Yap LF, Shannon-Lowe C, et al. The Epstein-Barr virus and the pathogenesis of lymphoma. *J Pathol.* 2015;235:312-322.
- Young LS, Rickinson AB. Epstein-Barr virus: 40 years on. *Nat Rev Cancer.* 2004;4:757-768.
- Young LS, Yao QY, Rooney CM, et al. New type B isolates of Epstein-Barr virus from Burkitt's lymphoma and from normal individuals in endemic areas. *J Gen Virol.* 1987;68(Pt 11):2853-2862.
- Palser AL, Grayson NE, White RE, et al. Genome diversity of Epstein-Barr virus from multiple tumor types and normal infection. *J Virol.* 2015;89:5222-5237.
- Tiwawech D, Srivatanakul P, Karalak A, Ishida T. Association between EBNA2 and *LMP1* subtypes of Epstein-Barr virus and nasopharyngeal carcinoma in Thais. *J Clin Virol.* 2008;42:1-6.
- Decaussin G, Leclerc V, Ooka T. The lytic cycle of Epstein-Barr virus in the nonproducer Raji line can be rescued by the expression of a 135-kilodalton protein encoded by the BALF2 open reading frame. *J Virol.* 1995;69:7309-7314.
- Hui KF, Chan TF, Yang W, et al. High risk Epstein-Barr virus variants characterized by distinct polymorphisms in the EBEB locus are strongly associated with nasopharyngeal carcinoma. *Int J Cancer.* 2019;144:3031-3042.
- Zeng Y, Middeldorp J, Madjar JJ, Ooka T. A major DNA binding protein encoded by BALF2 open reading frame of Epstein-Barr virus (EBV) forms a complex with other EBV DNA-binding proteins: DNAase, EA-D, and DNA polymerase. *Virology.* 1997;239:285-295.
- Kanno M, Narita N, Fujimoto Y, et al. Third epidemiological analysis of nasopharyngeal carcinoma in the central region of Japan from 2006 to 2015. *Cancers.* 2019;11(8):1180.
- Bei J-X, Li YI, Jia W-H, et al. A genome-wide association study of nasopharyngeal carcinoma identifies three new susceptibility loci. *Nat Genet.* 2010;42:599-603.
- Bei JX, Su WH, Ng CC, et al. A GWAS meta-analysis and replication study identifies a novel locus within CLPTM1L/TERT associated with nasopharyngeal carcinoma in individuals of Chinese ancestry. *Cancer Epidemiol Biomarkers Prev.* 2016;25:188-192.
- Cui Q, Feng Q-S, Mo H-Y, et al. An extended genome-wide association study identifies novel susceptibility loci for nasopharyngeal carcinoma. *Hum Mol Genet.* 2016;25:3626-3634.
- Tang MZ, Lautenberger JA, Gao XJ, et al. The principal genetic determinants for nasopharyngeal carcinoma in China involve HLA class I antigen recognition groove. *J Acquir Immune Defic Syndr.* 2012;59:88.
- Xu M, Yao Y, Chen H, et al. Genome sequencing analysis identifies Epstein-Barr virus subtypes associated with high risk of nasopharyngeal carcinoma. *Nat Genet.* 2019;51:1131.
- Dolan A, Addison C, Gatherer D, Davison AJ, McGeoch DJ. The genome of Epstein-Barr virus type 2 strain AG876. *Virology.* 2006;350:164-170.
- Lin Z, Wang X, Strong MJ, et al. Whole-genome sequencing of the Akata and Mutu Epstein-Barr virus strains. *J Virol.* 2013;87:1172-1182.
- Zeng M-S, Li D-J, Liu Q-L, et al. Genomic sequence analysis of Epstein-Barr virus strain GD1 from a nasopharyngeal carcinoma patient. *J Virol.* 2005;79:15323-15330.
- Baer R, Bankier AT, Biggin MD, et al. DNA sequence and expression of the B95-8 Epstein-Barr virus genome. *Nature.* 1984;310:207-211.
- Okuno Y, Murata T, Sato Y, et al. Defective Epstein-Barr virus in chronic active infection and haematological malignancy. *Nat Microbiol.* 2019;4:404-413.
- Kondo S, Horikawa T, Takeshita H, et al. Diagnostic value of serum EBV-DNA quantification and antibody to viral capsid antigen in nasopharyngeal carcinoma patients. *Cancer Sci.* 2004;95:508-513.
- Depledge DP, Palser AL, Watson SJ, et al. Specific capture and whole-genome sequencing of viruses from clinical samples. *PLoS One.* 2011;6:e27805.
- Kent WJ. BLAT - The BLAST-like alignment tool. *Genome Res.* 2002;12:656-664.
- Thorvaldsdottir H, Robinson JT, Mesirov JP. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief Bioinform.* 2013;14:178-192.
- Seishima N, Kondo S, Wakisaka N, et al. EBV infection is prevalent in the adenoid and palatine tonsils in adults. *J Med Virol.* 2017;89:1088-1095.

27. Skalsky RL, Kang D, Linnstaedt SD, Cullen BR. Evolutionary conservation of primate lymphocryptovirus microRNA targets. *J Virol*. 2014;88:1617-1635.
28. Daikoku T, Kudoh A, Fujita M, et al. Architecture of replication compartments formed during Epstein-Barr virus lytic replication. *J Virol*. 2005;79:3409-3418.
29. Kanda Y. Investigation of the freely available easy-to-use software 'EZ' for medical statistics. *Bone Marrow Transplant*. 2013;48:452-458.
30. Lo AKF, To KF, Lo KW, et al. Modulation of LMP1 protein expression by EBV-encoded microRNAs. *Proc Natl Acad Sci USA*. 2007;104:16164-16169.
31. Romero-Masters JC, Huebner SM, Ohashi M, et al. B cells infected with Type 2 Epstein-Barr virus (EBV) have increased NFATc1/NFATc2 activity and enhanced lytic gene expression in comparison to Type 1 EBV infection. *PLoS Pathog*. 2020;16(2):e1008365.
32. Kunimoto M, Tamura S, Tabata T, et al. One-step typing of Epstein-Barr virus by polymerase chain reaction: predominance of type 1 virus in Japan. *J Gen Virol*. 1992;73:455-461.
33. Lin X, Tsai M-H, Shumilov A, et al. The Epstein-Barr virus BART miRNA cluster of the M81 strain modulates multiple functions in primary B cells. *PLoS Pathog*. 2015;11:e1005344.
34. Hong GK, Gulley ML, Feng WH, Delecluse HJ, Holley-Guthrie E, Kenney SC. Epstein-Barr virus lytic infection contributes to lymphoproliferative disease in a SCID mouse model. *J Virol*. 2005;79:13993-14003.
35. Jones RJ, Seaman WT, Feng W-H, et al. Roles of lytic viral infection and IL-6 in early versus late passage lymphoblastoid cell lines and EBV-associated lymphoproliferative disease. *Int J Cancer*. 2007;121:1274-1281.
36. Ma S-D, Hegde S, Young KH, et al. A new model of Epstein-Barr virus infection reveals an important role for early lytic viral protein expression in the development of lymphomas. *J Virol*. 2011;85:165-177.
37. Shotelersuk K, Khorprasert C, Sakdikul S, Pornthanakasem W, Voravud N, Mutirangura A. Epstein-Barr virus DNA in serum/plasma as a tumor marker for nasopharyngeal cancer. *Clin Cancer Res*. 2000;6:1046-1051.
38. Pathmanathan R, Prasad U, Sadler R, Flynn K, Raab-Traub N. Clonal proliferations of cells infected with Epstein-Barr virus in preinvasive lesions related to nasopharyngeal carcinoma. *N Engl J Med*. 1995;333:693-698.
39. Martel-Renoir D, Grunewald V, Touitou R, Schwaab G, Joab I. Qualitative analysis of the expression of Epstein-Barr virus lytic genes in nasopharyngeal carcinoma biopsies. *J Gen Virol*. 1995;76(Pt 6):1401-1408.
40. Mumtsidu E, Makhov AM, Konarev PV, Svergun DI, Griffith JD, Tucker PA. Structural features of the single-stranded DNA-binding protein of Epstein-Barr virus. *J Struct Biol*. 2008;161:172-187.
41. Paramita DK, Fachiroh J, Artama WT, van Benthem E, Haryana SM, Middeldorp JM. Native early antigen of Epstein-Barr virus, a promising antigen for diagnosis of nasopharyngeal carcinoma. *J Med Virol*. 2007;79:1710-1721.
42. Karran L, Gao Y, Smith PR, Griffin BE. Expression of a family of complementary-strand transcripts in Epstein-Barr virus-infected cells. *Proc Natl Acad Sci USA*. 1992;89:8058-8062.
43. Cai X, Schäfer A, Lu S, et al. Epstein-Barr virus microRNAs are evolutionarily conserved and differentially expressed. *PLoS Pathog*. 2006;2:e23.
44. Yoshizaki T, Kondo S, Wakisaka N, et al. Pathogenic role of Epstein-Barr virus latent membrane protein-1 in the development of nasopharyngeal carcinoma. *Cancer Lett*. 2013;337:1-7.
45. Yoshizaki T, Kondo S, Endo K, et al. Modulation of the tumor microenvironment by Epstein-Barr virus latent membrane protein 1 in nasopharyngeal carcinoma. *Cancer Sci*. 2018;109:272-278.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Kondo S, Okuno Y, Murata T, et al. EBV genome variations enhance clinicopathological features of nasopharyngeal carcinoma in a non-endemic region. *Cancer Sci*. 2022;113:2446-2456. doi:[10.1111/cas.15381](https://doi.org/10.1111/cas.15381)