Research Article | **Open Access**

Genome-Wide Analysis of Epstein-Barr Virus Isolated from Extra Nodal NK/T-Cell Lymphoma, Nasal Type

Wenjing Ku[1*], Ningjing Lin[2] and Zheming Lu[1]

[1]Laboratory of Biochemistry and Molecular Biology, Peking University Cancer Hospital and Institute, Beijing, PR China

[2]Department of Medicine, Peking University Cancer Hospital and Institute, Beijing, PR China

[*]**Corresponding author:** Wenjing Ku, Laboratory of Biochemistry and Molecular Biology, Peking University Cancer Hospital and Institute, Beijing, PR China, Tel: 86-10-88196702; E-mail: luzheming@bjmu.edu.cn

Abstract

Extra-nodal natural killer (NK) cell/T cell lymphoma (NKTCL), a rare type of non-Hodgkin's lymphoma, has strongly been associated with Epstein-Barr virus (EBV) infection. However, there is no EBV genomes isolated from NKTCL, and the roles of the variations of EBV strains played in the pathogenesis of NKTCL is still unclear. In this study, whole EBV genomes from eight primary NKTCL biopsy specimens were obtained using next generation sequencing, designated NKTCL-EBV1 to 8. Compared to the six mostly referenced EBV strains, NKTCL-EBVs were closely resemble to GD1 strain, but still harbors 2072 variations, including 1938 substitutions, 58 insertions and 76 deletions. Moreover, the results from phylogenetic analysis of whole NKTCL genomes and specific genes demonstrated that all the NKTCL-EBVs were related to Asian EBV strains. Furthermore, changes in CD4+ and CD8+ T-cell epitopes of EBNA1 and LMP1 may affect the efficacy for a CTL-based therapy. In conclusion, this is the first large study to our knowledge to obtain EBV genomes isolated from NKTCL and show the diversity of EBV genomes in a whole genome level by phylogenetic analysis.

Keywords: NK/T cell lymphoma; Epstein-Barr virus; NKTCL-EBVs; Next generation sequencing

Abbreviations: EBV: Epstein-Barr Virus; EBVaGC: EBV-associated Gastric Carcinomas; EBNA: EBV-determined Nuclear Antigen; INDEL: Insertion and Deletion; IR: Internal repeats; LMP: Latent Membrane Protein; NKTCL: Natural Killer (NK) Cell/T Cell Lymphoma; NPC: Nasopharyngeal Carcinoma; TR: Terminal Repeat; SNV: Single Nucleotide Variant.

Introduction

Epstein-Barr virus infection has shown to be associated with the pathogenesis of several malignant diseases, including lymphoma, nasopharyngeal carcinoma (NPC), and a subset of gastric carcinoma [1]. Among them, extra-nodal natural killer (NK) cell/T cell lymphoma (NKTCL), a rare type of non-Hodgkin's lymphoma, is characterized by the presence of EBV in virtually all cases, irrespective of their ethnicity or geographical origin. NKTCL is an aggressive malignancy, predominantly occur in the nasal, paranasal, and oropharyngeal sites [2], and much more prevalent in East Asian [3-7] and Latin America [3,8] than in western countries. Although the association of this B lymphotropic virus with malignancies of T and NK cell origin was quite unexpected, both the presence of virus sequences in tumor cells and the virus's oncogenic potency have led to the hypothesis that whether particular EBV strains are preferentially selected in NKTCL. Moreover, understanding more about EBV sequence variations isolated from NKTCL is essential for developing cancer vaccine for immunotherapy and for monitoring dynamically in different stages of the diseases using the circulating EBV-DNA, which is derived from apoptotic and necrotic cells [9]. However, few studies have evaluated diversity in more than a single region of the EBV genome in NKTCLand the full spectrum of diversity existent within the EBV genome from NKTCL tissue has not been adequately studied. Genome-wide study of EBV provides the first opportunity to test the general validity of the EBV genetic map and explore recombination, geographic variation, and the major features of variation in this virus.

In the present study, we obtained 8 EBV genomes isolated from NKTCL biopsy samples, designated NKTCL-EBV1 to 8, and then illustrated the distribution of genetic variation in these EBV strains and further phylogenetic analysis to evaluate their genomic diversity.

Materials and Methods

NKTCL patients and DNA preparation

The tissue specimens were biopsied, and then collected after obtaining informed written consent from patients diagnosed with NKTCL at Peking University Cancer Hospital, Beijing, China, from 2014 to 2015. Collection of the tumor samples from the NKTCL patients are approved by the Institutional Review Board of the hospital for the purpose of EBV genomic sequencing. The clinical information for the NKTCL patients is summarized in Supplementary Table S1.

The DNA of the NKTCL biopsy specimens was extracted by using a Qiagen blood and tissue kit according to the manufacture's protocol (Qiagen, Hilden, Germany), and A260/A280 ratio between 1.8 and 2.0 was applied to the subsequent experiments.

Complete workflow for sequencing of EBV genomes

The workflow, including library preparation, target capture, next-generation sequencing, de novo assembly, and scaffolding and joining contigs to subsequent analysis is accounted for in detail in Supplementary Materials and Methods.

Detection of SNVs and mutation analysis

All output EBV reads were first aligned to the EBV reference genome by BWA software, removing the interferential duplicates by Sequence Alignment/Map tools (SAMtools) 3 to gain the basic classified subtype. A mass of investigates hereunto did verify that the outcomes of the NKTCL clinical cases were categorized as type II latency pattern [10]. Hence, these amino acids or epitopes changes, which may affect the corresponding immunological recognition processes mediated by CD4+ and CD8+ T-cell, should be given the special attention.

Phylogenetic analysis

The MAFFT software was used to multiple sequence alignments, comparing the NKTCL-EBV genomes with the other 19 strains [11-21] reported previously. Molecular Evolutionary Genetics Analysis software, version 7 (MEGA7) was used to perform the Phylogenetic analyses by neighbor-joining algorithm in this study. Bootstrap analysis of 1000 replicates was performed on each tree to determine the confidence.

Results

Summary of the sequencing data

In this study, we performed the EBV genome capture and deep sequencing (average depth>10) in eight NKTCL samples. A total of 1,083,561 of 75-base pair-end reads were generated, which were equivalent to 21.48 MB in the whole samples (Supplementary Table S2). Comparing with 6 mostly referenced sequences, we found that GD1 coverage percentages are higher than the rest, ranging from 86.63% in NKTCL-EBV5 to 94.94% in NKTCL-EBV8. The mean depth was 107 fold and the fraction of effective bases on EBV varied from 16.05% (31.53/418.38 Mb, NKTCL-EBV2) to 72.77% (340.12/467.39, NKTCL-EBV6).

Assembly of NTKCL-EBV 1 to 8 genomes

The number of the contigs (Supplementary Table S3) were from 20 (NKTCL-EBV5) to 32 (NKTCL-EBV7). N50 size of contigs ranged from 8,342 bp (NKTCL-EBV2) to 21,323 bp (NKTCL-EBV8). As illustrated above, the gaps would be supplemented with the tracts "N" based on the EBV reference GD1, if the PCR and Sanger sequencing did not link up the contigs effectively. Finally, eight EBV genomes were successfully obtained, named NKTCL-EBV1 to 8. The genome sizes estimated based on the reference EBV sequence were as follows: NKTCL-EBV1 (172,059 bp), NKTCL-EBV2 (171,730 bp), NKTCL-EBV3 (171,639 bp), NKTCL-EBV4 (171,613 bp), NKTCL-EBV5 (171,663 bp), NKTCL-EBV6 (171,727 bp), NKTCL-EBV7 (171,706 bp), and NKTCL-EBV8 (171,590 bp), with GC contents of approximately 55%.

Mutation analysis of the NKTCL-EBV genomes

Whole genome sequence alignments revealed extensive nucleotide variation in all of the eight NKTCL-EBV genomes. In comparison to reference GD1 sequence, the NKTCL-EBV1 to 8 harbored 2072 variations in total, including 1938 substitutions, 58 insertions and 76 deletions. Among them, 1218 substitutions, 26 deletions and 15 insertions were located in the coding regions; while the 720 substitutions, 43 insertions and 50 deletions were located in the

noncoding regions (Supplementary Table S4). Figure 1 depicts the variation of all available NKTCL-EBV genomes compared to GD1. Remarkably, the density of variations was substantially higher when compared to B95-8, the prototype originally isolated from a patient with infectious mononucleosis from America (Supplementary Figure S1).
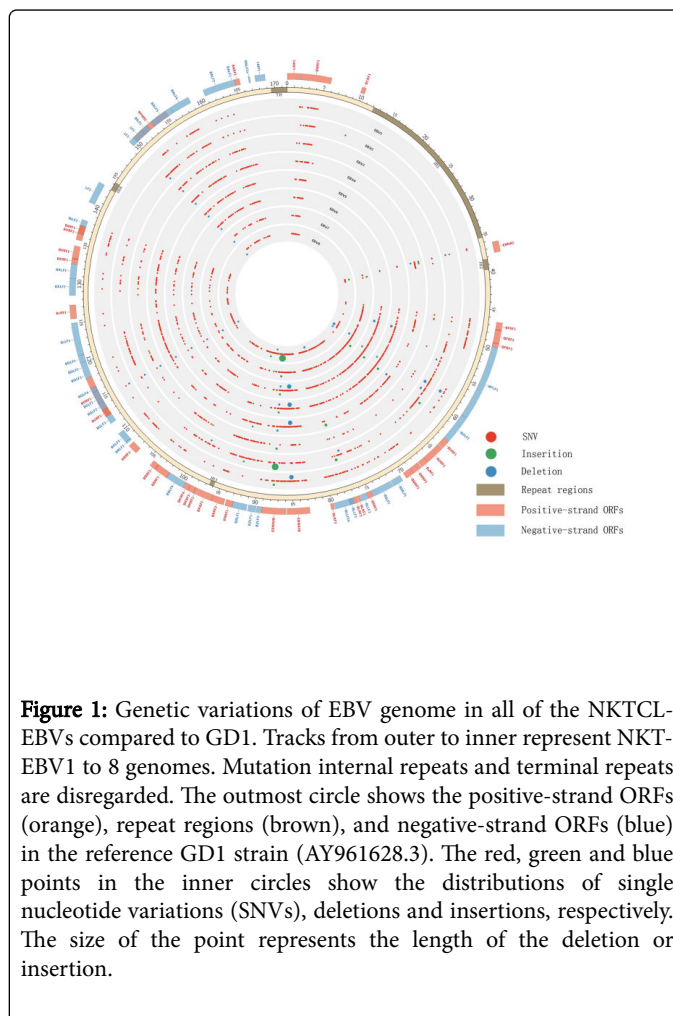

Figure 1: Genetic variations of EBV genome in all of the NKTCL-EBVs compared to GD1. Tracks from outer to inner represent NKT-EBV1 to 8 genomes. Mutation internal repeats and terminal repeats are disregarded. The outmost circle shows the positive-strand ORFs (orange), repeat regions (brown), and negative-strand ORFs (blue) in the reference GD1 strain (AY961628.3). The red, green and blue points in the inner circles show the distributions of single nucleotide variations (SNVs), deletions and insertions, respectively. The size of the point represents the length of the deletion or insertion.

The 9 categories, discriminated and defined by the different protein functions [22], were summarized in the Supplementary Table S5. In total, the mutation number of genes coding latent and tegument were higher than any other genes, averagely accounting for 56.2% of the whole nonsynonymous mutations, followed by 146 mutations (account for 10.7%) located in genes expressing membrane (glyco) protein. There were 44 to 68 nonsynonymous mutations located in the latent genes in NKTCL-EBV1, -2, -4, -5, -6, -7 and -8, whereas NKTCL-EBV3 had a minimum quantity of 20 mutations. When observing the mutations in each case, we detected the maximum quantity in NKTCL-EBV6, which harbored 194 nonsynonymous mutations, followed by the NKTCL-EBV5 (188), NKTCL-EBV4 (184), NKTCL-EBV7 (183), NKTCL-EBV8 (183), NKTCL-EBV2 (174), NKTCL-EBV1 (149) and NKTCL-EBV3 (104) (Figure 2).

Sequence variations in the EBNA1 and LMP1 genes

The total of 467 nonsynonymous mutations was located in the protein-coding genes, accounting for 38.3% of the SNVs. The EBNA1

and LMP1 genes were summarized in this study in that these two regions are the most frequent studied regions to date.
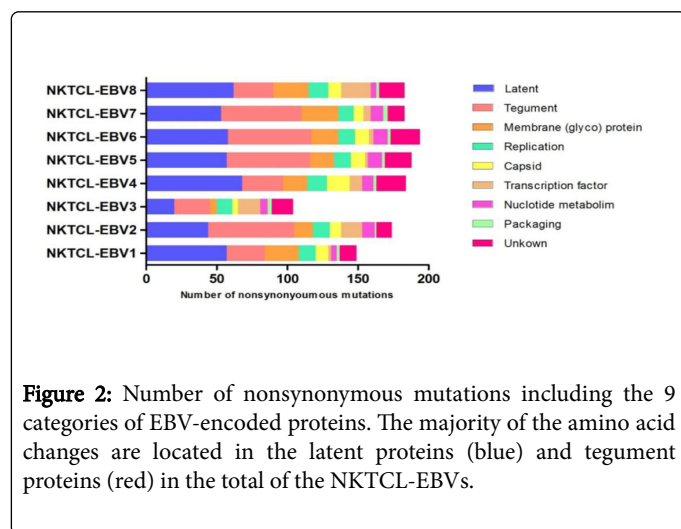

Figure 2: Number of nonsynonymous mutations including the 9 categories of EBV-encoded proteins. The majority of the amino acid changes are located in the latent proteins (blue) and tegument proteins (red) in the total of the NKTCL-EBVs.

EBNA1 is the only latent protein that is consistently expressed in all EBV-positive malignancies, and is a transcriptional activator of the EBV expression. EBNA1 have been defined P and V EBNA1, which differed from B95-8 by 15AA substitutions [23]. P and V subtypes were further classified based on the signature changes at AA residue 487 in the C-terminus of EBNA1: P-alanine (P-ala), P-threonine (P-thr), V-proline (V-pro), V-leucine (V-leu), and V-Valine (V-val) [24,25]. Compared to B95-8, 19 nonsynonymous mutations were observed in total; 6 in the N-terminal regions and 13 in the C-terminal regions. As Do et al. [26] and Wang et al. [25] had mentioned that the subvariants of V-val have AA substitution at 528 and 585, and we observed that all of the eight EBNA1 strains were identified as V-val subvariant.

LMP1 is the major transforming protein of EBV and contains a higher degree of polymorphism than most EBV genes [27]. Compare to the B95-8, 45 nonsynonymous mutation was detected in the LMP1 strains of the eight specimens, and 4 in the N-terminus (AA: 1-25), 26 in the transmembrane domains (AA:26-196), and 15 in the C-terminus (AA: 197-386). Using the Edwards classification system, based on the signature AA changes relative to the prototypic LMP-1 (B95-8) in the C-terminus (AA 189-377), six sequence variants of LMP1, termed Alaskan, China 1, China 2, Mediterranean+ (Med+), Med-, and North Carolina (NC), have been identified [28]. All eight NKTCL-EBV strains were identified as China 1 type [29]. The other two commonly classification system used to date characterized variants LMP-1 gene polymorphisms include a 30 bp deletion in the C-terminus and the loss of restriction site XhoI in the N-terminus of the gene. The LMP1 strain in NKTCL-EBV 1 to 7, but NKTCL-EBV8, harbored the 30 bp deletion, causing the loss of 10 AA (343-352).

Amino acid changes in CD4+ and CD8+ T-cell epitopes of EBNA1, LMP1 and LMP2A

Almost all of the latent proteins and lytic proteins harbored nonsynonymous mutations in epitopes induced both the CD4+ and CD8+ T cells responses. Many epitopes were defined and were mapped in EBV antigens and correlated with MHC type in previous studies. Compared to B95-8, amino acids changes were found in three CD8+ epitopes of EBNA1, eight epitopes of LMP1 and twelve in LMP2A. Eleven CD4+ epitopes of EBNA1, thirteen in LMP1 and nine in

LMP2A contained amino acids. Some of the nonsynonymous mutations were affecting multiple epitopes. For instance, a C-to-T substitution at coordinate 97121 (NC_007605) resulted in the change of residue 487 (A→V) of EBNA1 in all of the eight NKTCL specimens, where CD4+ epitopes SNP, NPK, ENI, IAE, and LRA were located. The C-to-T and C-to-G substitution at coordinate 168229 induced LMP1 epitope changes in residue 212 G→S of NKTCL-EBV1 to 6 and -8, and in residue 212 G→R in NKTCL-EBV7. Meanwhile, in the NKTCL-EBV2 strains, other two substitutions, C-to-G at coordinate 168223 and A-to-C at coordinate 168224, caused the epitope changes in residue 213 H→Q and 214 E→Q respectively, where CD4+ epitopes QAT, SSH, and SGH were located. The positions of nonsynonymous mutation changes located in the epitopes are illustrated in Figure 3 and Supplementary Table S6.
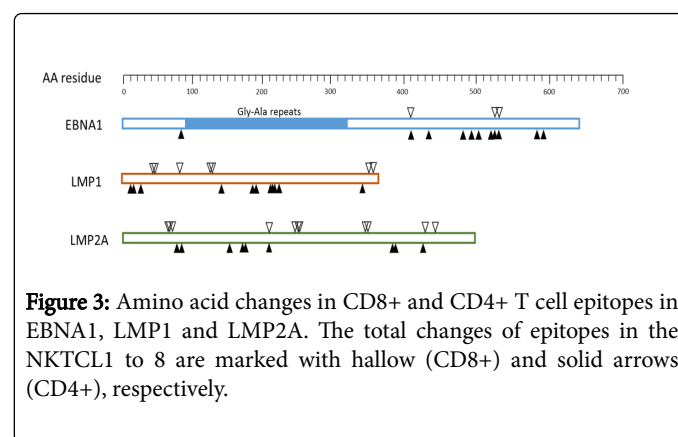

Figure 3: Amino acid changes in CD8+ and CD4+ T cell epitopes in EBNA1, LMP1 and LMP2A. The total changes of epitopes in the NKTCL1 to 8 are marked with hallow (CD8+) and solid arrows (CD4+), respectively.

Phylogenetic analysis of the NKTCL-EBV genomes and EBNA1 interstrain recombinant

The phylogenetic trees were conducted based on alignment of edited full-length of eight NKTCL-EBVs and previously published 28 strains as illustrated in Figure 4a. Of note, all NKTCL-EBVs genomes clearly sorts into type 1, based on differences in whole genome and especially EBNA2 (Supplementary Figure S2c). The phylogenetic analysis from the whole genome showed that all the NKTCL-EBVs are related to other Asian EBV strains, including EBVaGC1 to 9, HKNPC1 to 9, C666-1, GD1 and GD2 obtained from China and Akata from Japan, while none of the specimens were clustered in a branch of non-Asian strains AG876, B95-8, Mutu, K4413-Mi and K4123-Mi. Phylogenetic analysis of EBNA1 (Figure 4b) and LMP1 (Figure 4c) showed similar phylogenetic clustering as we found using full-length whole EBV genomes, albeit with minor variations. Of interest, EBNA1 of NKTCL-EBV3 sequence showed clustered away from other seven NKTCL-EBV strains. Analysis of amino acid sequences of EBNA1 supported EBNA1 of NKTCL-EBV3 may arise from recombination of GD1 and B95-8 (Figure 5 and Supplementary Table S7).

Discussion

To date, pathogenesis and genotype analysis of NKTCL have focused on genetic variations in a small fraction of EBV genes, which is limited to define the spectrum of diversity within the whole genome of EBV. The genome-wide characteristics of EBV are essential to understand the diversity of strains isolated from NKTCL. Here, we directly sequenced EBV-captured DNA from 8 primary NKTCL biopsy samples from China using Illumina HiSeq 2500 sequencer platform,

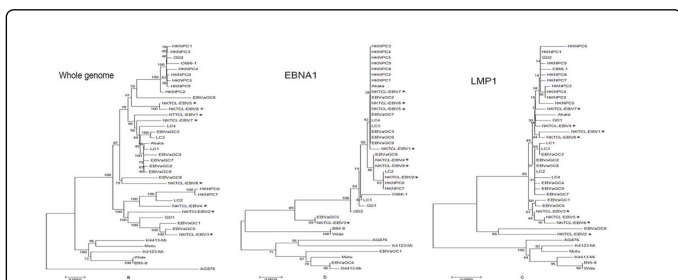and presented the eight EBV sequences obtained from primary NKTCL tumors.

Figure 4: Phylogenetic trees of the whole EBV genomes and nucleotide sequences of EBNA1 and LMP1 sequences. The evolutionary history was inferred using the Neighbor-Joining method, and evolutionary analyses were conducted in MEGA7. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) was shown next to the branches. NKTCL-EBV1 to 8 in the phylogenetic trees was marked with an asterisk.
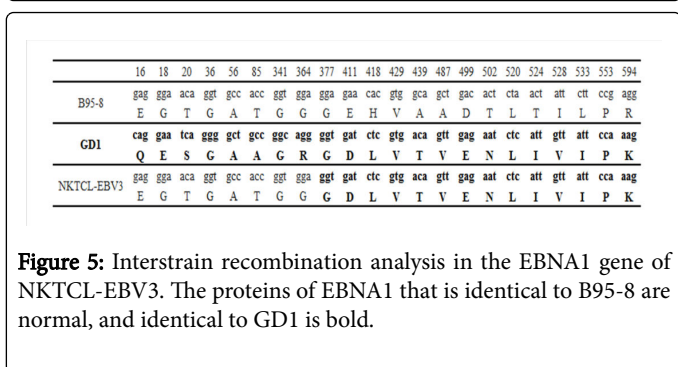

Figure 5: Interstrain recombination analysis in the EBNA1 gene of NKTCL-EBV3. The proteins of EBNA1 that is identical to B95-8 are normal, and identical to GD1 is bold.

The variability of NKTCL-EBV 1 to 8 ranged from 0.156% (NKTCL-EBV3) to 0.299% (NKTCL-EBV4), which is lower than the interstrain variability of ca. 0.5% for virus of the same type. Hence, it excludes the variability raised by coinfection of different viral strains [14,30]. In accordance with the previous reports [13], the number of the nonsynonymous mutations is highest in the gene regions encoding latent proteins in each of the NKTCL-EBV genomes, followed by genes encoding the tegument protein and membrane glycoproteins. One of the basic goals in EBV associated study is aimed at understanding EBV geographic and disease associations. The phylogenetic tree constructed based on whole genome alignment of all strains illustrated that the genomes are actually closer to the other Asian variants than non-Asian strains, indicating that a phylogenetic relationship corresponding to the geographical origin of the viral genomes instead of the type 1 and 2 dichotomy. In order to investigate how genomic variations of EBV contribute to pathogenesis of NPC, it is essential to sequence the EBV genomes of a large number of normal and diseased individuals in the same geographical region and conduct a case-control comparison.

EBNA1 and LMP1 are the most frequently studied regions to date. EBNA1 is critical in maintaining EBV in the infected cells and facilitating episomal replication, and it is also a transcriptional activator. The codon 487 is a hot spot variation in the C-terminus, classifying the P and V subtypes based on the AA present at residue 487. The EBNA1 strain in each the NKTCL-EBV case was the V-val

category. Interestingly, V-val was detected in both cases and controls almost exclusively in China according to previously reported. Chang et al. [31] had summarized the prevalence of EBNA1 variants by clinical condition and geographic region. They showed that 95.2% of the NPC and 100% of HL in China are V-Val, while in the healthy or nonmalignant people, about 48.6% is V-val subtype, and 32.4% is V-val and P-ala subtypes. LMP1 is thought to play a principal role in tumorigenesis. Based on the signature AA changes relative to the prototypic LMP-1 in the C-terminus (AA 189-377), LMP1 variants have previously been classified into seven groups named NC (North Carolina), Med, Alaskan, B95-8, China 1, China 2, and China 3. All of the NKTCL-EBV were classified China 1, consistent with results obtained from the LCs, HKNPCs and EBVaGCs from China. This showed a phylogenetic relationship corresponding to the geographical origin of the viral genomes, supporting the conclusion that LMP1 can serve as a geographical marker [29]. Moreover, the 30 bp deletion, resulting in the proteins loss between the AA 343 and AA 352, had been detected in the NKTCL-EBV1 to 7, a marked predominance of LMP1 deletion (del-LMP1) over wild-type LMP1 (wt-LMP1) variants was observed in NKTCL. Recently, Cabrera et al. [32] suggested that EBV with XhoI restriction site loss at exon 1 of the LMP1 gene was associated with the risk of NKTCL in Chile. And we found all the eight NKTCL specimens were XhoI loss; although the available data suggest the XhoI loss in NPC and HL is common in China, additional studies are needed to full evaluate geographic and disease of XhoI loss in the NKTCL.

NKTCL is associated with type II EBV latency, where only restricted EBV antigens, namely EBNA1, and LMP 1 and 2, are expressed [33,34]. These viral proteins offer the potential targets for a T cell based immunotherapy. Adoptive transfer of cytotoxic T cells (CTLs) specific for EBV antigens has proved safe and effective as prophylaxis and treatment for EBV-associated lymphoproliferative disease. Some patients with advanced stage or relapsed EBV-associated malignancies, including NKTCL, achieved complete remission (CR) after treatment with autologous LMP1/2- and EBNA 1 - specific cytotoxic T lymphocytes (CTLs) or activated by peptides derived from LMP1/2 [35,36]. Nonetheless, some cases still did not respond to LMP-CTL therapy and usually attributed this failure to immune escapes by antigen loss. It is worth to note that all these previous studies using prototype EBV sequence, B95-8, to design full-length LMP epitopes. Therefore, our work gives an alternative explanation for lacking of tumor response. As shown in Supplementary Figure S3, sequence analysis identified LMP and EBNA1 variations containing epitopes changes in comparing to the prototype B95–8 sequence in all cases of NKTCLs. Highly polymorphic LMP and EBNA1 genes lead to amino acid changes in CD4+ and CD8+ specific T cell epitopes. As an example, YFL, a HLA A2-restricted epitope of LMP1, was a variant peptide of the wild-type YLL. The changes in residue 2(L→F) and residue 5(M→I) were found in all of the NKTCL-EBV genomes, which are also found prevalently in NPC specimens [13]. The CTL recognition of YFL was abrogated compared to the wild type epitopes [37]. Thus, this LMP1 variant epitopes might contribute to the evasion of the EBV-infected cells from T cell surveillance [13]. Whether changes in such epitopes confer immune evasion of the NKTCL cells may constitute another hypothesis for future testing. Our data provide optimization proposal for selecting EBV genome for treatment from individual patient or at least predominant strains prevalent in geographical regions instead of commonly used B95-8 genome.

In summary, we reported eight EBV genomes isolated from primary NKTCL biopsy specimens based on a complete large-scale sequencing

workflow. To the best of our knowledge, this is the first report analyzing the sequence diversity on a whole-genome level of EBV obtained from primary NKTCL tumors, although their pathogenesis remains to be clarified. Furthermore, phylogenetic analysis showed that all NKTCL-EBV strains isolated in ENKTL are close to other Asian subtypes, which leading to the conclusion that EBV infection are more likely affected by different geographic regions rather than particular EBV-associated malignancies. Therefore, our data has implications for the development of effective prophylactic and therapeutic vaccine approaches targeting the personalized EBV antigens in these aggressive diseases.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgements

Availability of Data and Materials

The datasets supporting the conclusions of the article are included within the article and supplementary information.

Competing Interests

The authors declare that they have no competing interests.

Ethics Approval and Consent to Participate

This study was approved by the Peking University Cancer Hospital, Beijing, China.

References

1. Ohiole Ozoya O, Sokol L, Dalia S (2016) EBV-related malignancies, outcomes and novel prevention strategies. Infect Disord Drug Targets (Formerly Current Drug Targets-Infectious Disorders) 16: 4-21.

2. Ying Z, Zhu J (2015) Extranodal natural killer/T cell lymphoma: we should and we can do more. Chin Clin Oncol 4: 12.

3. Aozasa K, Ohsawa M, Tajima K, Sasaki R, Maeda H, et al. (1989) Nation-wide study of lethal mid-line granuloma in Japan: frequencies of wegener's granulomatosis, polymorphic reticulosis, malignant lymphoma and other related conditions. Int J Cancer 44: 63-66.

4. Harabuchi Y, Yamanaka N, Kataura A, Imai S, Kinoshita T, et al. (1990) Epstein-Barr virus in nasal T-cell lymphomas in patients with lethal midline granuloma. Lancet (London, England) 335: 128-130.

5. Harabuchi Y, Imai S, Wakashima J, Hirao M, Kataura A, et al. (1996) Nasal T-cell lymphoma causally associated with Epstein-Barr virus: clinicopathologic, phenotypic, and genotypic studies. Cancer 77: 2137-2149.

6. Jaffe ES, Chan JK, Su IJ, Frizzera G, Mori S, et al. (1996) Report of the Workshop on Nasal and Related Extranodal Angiocentric T/Natural Killer Cell Lymphomas. Definitions, differential diagnosis, and epidemiology. Am J Surg Pathol 20: 103-111.

7. Lee J, Suh C, Park YH, Ko YH, Bang SM, et al. (2006) Extranodal natural killer T-cell lymphoma, nasal-type: a prognostic model from a retrospective multicenter study. J Clin Oncol 24: 612-618.

8. Altemani A, Barbosa AC, Kulka M, Takahashi T, Endo L, et al. (2002) Characteristics of nasal T/NK-cell lymphoma among Brazilians. Neoplasma 49: 55-60.

9. Wang ZY, Liu QF, Wang H, Jin J, Wang WH, et al. (2012) Clinical implications of plasma Epstein-Barr virus DNA in early-stage extranodal nasal-type NK/T-cell lymphoma patients receiving primary radiotherapy. Blood 120: 2003-2010.

10. Harabuchi Y, Takahara M, Kishibe K, Moriai S, Nagato T, et al. (2009) Nasal natural killer (NK)/T-cell lymphoma: clinical, histological, virological, and genetic features. Int J Clin Oncol 14: 181-190.

11. Baer R, Bankier AT, Biggin MD, Deininger PL, Farrell PJ, et al. (1984) DNA sequence and expression of the B95-8 Epstein-Barr virus genome. Nature 310: 207-211.

12. Dolan A, Addison C, Gatherer D, Davison AJ, McGeoch DJ (2006) The genome of Epstein-Barr virus type 2 strain AG876. Virology 350: 164-170.

13. Kwok H, Tong AH, Lin CH, Lok S, Farrell PJ, et al. (2012) Genomic sequencing and comparative analysis of Epstein-Barr virus genome isolated from primary nasopharyngeal carcinoma biopsy. PLoS One 7: e36939.

14. Kwok H, Wu CW, Palser AL, Kellam P, Sham PC, et al. (2014) Genomic diversity of Epstein-Barr virus genomes isolated from primary nasopharyngeal carcinoma biopsy samples. J Virol 88: 10662-10672.

15. Lei H, Li T, Hung GC, Li B, Tsai S, et al. (2013) Identification and characterization of EBV genomes in spontaneously immortalized human peripheral blood B lymphocytes by NGS technology. BMC Genomics 14: 804.

16. Lin Z, Wang X, Strong MJ, Concha M, Baddoo M, et al. (2013) Whole-genome sequencing of the Akata and Mutu Epstein-Barr virus strains. J Virol 87: 1172-1182.

17. Liu P, Fang X, Feng Z, Guo YM, Peng RJ, et al. (2011) Direct sequencing and characterization of a clinical isolate of Epstein-Barr virus from nasopharyngeal carcinoma tissue by using next-generation sequencing technology. J Virol 85: 11291-11299.

18. Parker BD, Bankier A, Satchwell S, Barrell B, Farrell PJ (1990) Sequence and transcription of Raji Epstein-Barr virus DNA spanning the B95-8 deletion region. Virology 179: 339-346.

19. Pizzo PA, Magrath IT, Chattopadhyay SK, Biggar RJ, Gerber P (1978) A new tumour-derived transforming strain of Epstein-Barr virus. Nature 272: 629-631.

20. Tso KK, Yip KY, Mak CK, Chung GT, Lee SD, et al. (2013) Complete genomic sequence of Epstein-Barr virus in nasopharyngeal carcinoma cell line C666-1. Infect Agent Cancer 8: 29.

21. Zeng MS, Li DJ, Liu QL, Song LB, Li MZ, et al. (2005) Genomic sequence analysis of Epstein-Barr virus strain GD1 from a nasopharyngeal carcinoma patient. J Virol 79: 15323-15330.

22. Tarbouriech N, Buisson M, Geoui T, Daenke S, Cusack S, et al. (2006) Structural genomics of the Epstein-Barr virus. Acta Crystallogr D Biol Crystallogr 62: 1276-1285.

23. Bhatia K, Raj A, Guitierrez MI, Judde JG, Spangler G, et al. (1996) Variation in the sequence of Epstein Barr virus nuclear antigen 1 in normal peripheral blood lymphocytes and in Burkitt's lymphomas. Oncogene 13: 177-181.

24. Snudden DK, Smith PR, Lai D, Ng MH, Griffin BE (1995) Alterations in the structure of the EBV nuclear antigen, EBNA1, in epithelial cell tumours. Oncogene 10: 1545-1552.

25. Wang JT, Sheeng TS, Su IJ, Chen JY, Chen MR (2003) EBNA-1 sequence variations reflect active EBV replication and disease status or quiescent latency in lymphocytes. J Med Virol 69: 417-425.

26. Do NV, Ingemar E, Phi PT, Jenny A, Chinh TT, et al. (2008) A major EBNA1 variant from Asian EBV isolates shows enhanced transcriptional activity compared to prototype B95.8. Virus Res 132: 15-24.

27. Palser AL, Grayson NE, White RE, Corton C, Correia S, et al. (2015) Genome diversity of Epstein-Barr virus from multiple tumor types and normal infection. Journal of Virology 89: 5222-5237.

28. Mainou BA, Raab-Traub N (2006) LMP1 strain variants: biological and molecular properties. Journal of Virology 80: 6458-6468.

29. Edwards RH, Seillier-Moiseiwitsch F, Raab-Traub N (1999) Signature amino acid changes in latent membrane protein 1 distinguish Epstein-Barr virus strains. Virology 261: 79-95.

30. Liu Y, Yang W, Pan Y, Ji J, Lu Z, et al. (2016) Genome-wide analysis of Epstein-Barr virus (EBV) isolated from EBV-associated gastric carcinoma (EBVaGC). Oncotarget 7: 4903-4914.

31. Chang CM, Yu KJ, Mbulaiteye SM, Hildesheim A, Bhatia K (2009) The extent of genetic diversity of Epstein-Barr virus and its geographic and disease patterns: a need for reappraisal. Virus Res 143: 209-221.

32. Cabrera ME, Eizuru Y, Itoh T, Koriyama C, Tashiro Y, et al. (2007) Nasal natural killer/T-cell lymphoma and its association with type "i"/XhoI loss strain Epstein-Barr virus in Chile. J Clin Pathol 60: 656-660.

33. Klein E, Kis LL, Klein G (2007) Epstein-Barr virus infection in humans: from harmless to life endangering virus-lymphocyte interactions. Oncogene 26: 1297-1305.

34. Fox CP, Haigh TA, Taylor GS, Long HM, Lee SP, et al. (2010) A novel latent membrane 2 transcript expressed in Epstein-Barr virus-positive NK- and T-cell lymphoproliferative disease encodes a target for cellular immunotherapy. Blood 116: 3695-3704.

35. Bollard CM, Gottschalk S, Torrano V, Diouf O, Ku S, et al. (2014) Sustained complete responses in patients with lymphoma receiving autologous cytotoxic T lymphocytes targeting Epstein-Barr virus latent membrane proteins. Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology 32: 798-808.

36. Cho SG, Kim N, Sohn HJ, Lee SK, Oh ST, et al. (2015) Long-term Outcome of Extranodal NK/T Cell Lymphoma Patients Treated With Postremission Therapy Using EBV LMP1 and LMP2a-specific CTLs. Molecular Therapy: the Journal of the American Society of Gene Therapy 23: 1401-1409.

37. Lin HJ, Cherng JM, Hung MS, Sayion Y, Lin JC (2005) Functional assays of HLA A2-restricted epitope variant of latent membrane protein 1 (LMP-1) of Epstein-Barr virus in nasopharyngeal carcinoma of Southern China and Taiwan. J Biomed Sci 12: 925-936.