

## Article

# Genetic Patterns Found in the Nuclear Localization Signals (NLSs) Associated with EBV-1 and EBV-2 Provide New Insights into Their Contribution to Different Cell-Type Specificities

Louise Zanella <sup>1,†</sup> , María Elena Reyes <sup>1,2,†</sup>, Ismael Riquelme <sup>3</sup> , Michel Abanto <sup>4</sup> , Daniela León <sup>1</sup>, Tamara Viscarra <sup>1</sup> , Carmen Ili <sup>1</sup>  and Priscilla Brebi <sup>1,\*</sup>

<sup>1</sup> Laboratory of Integrative Biology (LIBi), Scientific and Technological Bioresource Nucleus-Center for Excellence in Translational Medicine (BIOREN-CEMT), Universidad de La Frontera, Temuco 4810296, Chile; zanella.bio@gmail.com (L.Z.); m.reyes14@ufromail.cl (M.E.R.); dleon.garrido@gmail.com or d.leon02@ufromail.cl (D.L.); tamara.viscarra@ufrontera.cl (T.V.); carmen.ili@ufrontera.cl (C.I.)

<sup>2</sup> Departamento de Ciencias Básicas, Facultad de Ciencias, Universidad Santo Tomás, Santiago 8370003, Chile

<sup>3</sup> Instituto de Ciencias Biomédicas, Facultad de Ciencias de la Salud, Universidad Autónoma de Chile, Temuco 4810101, Chile; ismael.riquelme@uautonoma.cl

<sup>4</sup> Scientific and Technological Bioresource Nucleus (BIOREN), Universidad de La Frontera, Temuco 4811230, Chile; michel.abanto@ufrontera.cl

\* Correspondence: priscilla.brebi@ufrontera.cl; Tel.: +56-45-2596583

† These authors contributed equally to this work.



**Citation:** Zanella, L.; Reyes, M.E.; Riquelme, I.; Abanto, M.; León, D.; Viscarra, T.; Ili, C.; Brebi, P. Genetic Patterns Found in the Nuclear Localization Signals (NLSs) Associated with EBV-1 and EBV-2 Provide New Insights into Their Contribution to Different Cell-Type Specificities. *Cancers* **2021**, *13*, 2569. <https://doi.org/10.3390/cancers13112569>

Academic Editors: Lorenzo Leoncini, Lucia Mundo and Rosita Accardi-Gheit

Received: 8 May 2021  
Accepted: 19 May 2021  
Published: 24 May 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Simple Summary:** The Epstein–Barr virus (EBV) has been implicated in several human neoplastic diseases. The EBV-1 can transform B cells into LCL more efficiently than EBV-2, and EBV-2 preferentially infects T-cell lymphocytes. The EBNA3A oncoprotein has an essential role in B-cell transformation. The six peptide motifs called nuclear localization signals (NLSs) from EBNA3A ensure nucleocytoplasmic protein trafficking. Multiple NLSs have been suggested to enhance EBNA3 function or different specificities to different cell types; however, a comprehensive assessment of their genetic variability has not been addressed. Our objective was to study the NLSs' variability and their relationship with EBV types. Based on a comprehensive analysis of over a thousand EBNA3A sequences from different clinical manifestations and geographic locations, we found that EBNA3A from EBV-2 has two of the six NLSs altered, and genetic patterns in the NLSs are associated with EBV-1 and EBV-2.

**Abstract:** The Epstein–Barr virus (EBV) is a globally dispersed pathogen involved in several human cancers of B-cell and non-B-cell origin. EBV has been classified into EBV-1 and EBV-2, which have differences in their transformative ability. EBV-1 can transform B-cells into LCL more efficiently than EBV-2, and EBV-2 preferentially infects T-cell lymphocytes. The EBNA3A oncoprotein is a transcriptional regulator of virus and host cell genes, and is required in order to transform B-cells. EBNA3A has six peptide motifs called nuclear localization signals (NLSs) that ensure nucleocytoplasmic protein trafficking. The presence of multiple NLSs has been suggested to enhance EBNA3 function or different specificities in different cell types. However, studies about the NLS variability associated with EBV types are scarce. Based on a systematic sequence analysis considering more than a thousand EBNA3A sequences of EBV from different human clinical manifestations and geographic locations, we found differences in NLSs' nucleotide structures among EBV types. Compared with the EBNA3A EBV-1, EBNA3A EBV-2 has two of the six NLSs altered, and these mutations were possibly acquired by recombination. These genetic patterns in the NLSs associated with EBV-1 and EBV-2 provide new information about the traits of EBNA3A in EBV biology.

**Keywords:** Epstein–Barr virus (EBV); EBV nuclear antigen EBNA 3A (EBNA3A); recombination; phylogeny; EBV classification; nuclear localization signal (NLS)

## 1. Introduction

The Epstein–Barr Virus (EBV) is a human herpesvirus that belongs to the family *Herpesviridae*. EBV is related to the development of infectious mononucleosis (IM) [1], as well as potentially fatal neoplastic diseases, such as Burkitt lymphoma (BL) [2], nasopharyngeal cancer (NPC) [3], and gastric cancer (GC) [4]. The viral genome consists of about 170 kb double-stranded DNA coding over 100 proteins [5], including the nuclear antigens (EBNA)—EBNA-1, -2, -3A, -3B -3C, and leader protein (EBNA-LP)—and latent membrane proteins LMP-1 and -2 [6]. EBV is classified into EBV type 1 (EBV-1) and type 2 (EBV-2), based on specific polymorphisms of the *EBNA2* and *EBNA3A*, *3B*, and *3C* genes [7,8]. The two types have functional differences in their ability to transform human B cells. EBV-1 induces the continuous proliferation (known as “transformation”) of human B cells more efficiently than EBV-2 in vitro [8,9]. Although EBV-2 has been associated with lower efficiency in immortalizing B lymphocytes in vitro, it has been detected in BL tumors [10]. Unlike EBV-1, which has a tropism for B lymphocytes, EBV-2 preferentially infects T-cells in cell culture [11].

Among latency proteins, EBNA3 proteins are transcription regulators of virus and host cell genes [12,13]. The EBNA3 family consists of three proteins (EBNA3A, EBNA3B, and EBNA3C) that possibly originated from the triplication of a single ancestral gene [14]. The *EBNA3* genes are expressed via polycistronic transcription, and share a characteristic exon structure arranged in tandem: a short 5′ exon and a long 3′ exon, separated by a short intron [15].

EBNA3A comprises 944 amino acids, and contains multiple sequence-specific interaction regions for various proteins [16]. To regulate cellular gene expression, EBNA3A interacts with the cell’s DNA-binding protein (RBP-JK) and with the carboxyl-terminal-binding protein (CtBP) [16,17]. These proteins play an essential role in reprogramming the cellular gene transcription process to transform the B lymphocyte, contributing to the development of malignancies [18].

The EBNA3 family proteins are transported from the cytoplasm to the nucleus by the signal-mediated process. Nucleocytoplasmic trafficking occurs based on the peptide motif called nuclear localization signal (NLS) tags. The nuclear pore complexes (NPC) recognize these short, positively charged motifs and actively import EBNA3A to the nucleus [19,20]. EBNA3A has multiple (six) NLSs that ensure protein transport to the nucleus. Examining the NLSs via partial EBNA3A deletion constructs indicates that simultaneous mutations in NLS3 and NLS4 result in partial EBNA3A-mutated proteins restricted to the cell’s cytoplasm [21].

The classic EBV classification was recently revisited, updating the recombination impact on the EBV genome and demonstrating the consistent presence of sites of recombination in the *EBNA3A* gene [22]. However, an evaluation of the effect of recombination in the *EBNA3A* gene or its NLS regions has been not addressed.

Although some alterations to the peptide motif of EBNA3A NLSs have already been demonstrated, this was only in the context of partial proteins [21]. To date, these changes have not been observed in sequences recovered from natural human EBV infection, and moreover, there is a scarcity of EBV sequences from South America. In the present study, we report the first 12 EBNA3A sequences of gastric cancer bulk from Chile: 3 from EBV type 2 (the first non-Asian sequences) and 9 from EBV type 1, doubling the number of EBV EBNA3A sequences from South America. Moreover, we investigated the genetic diversity and delineated the effect of recombination in the *EBNA3A* gene structure.

Our results showed that *EBNA3A* from EBV-2 has two (NLS3 and NLS4) of the six NLSs altered (non-canonical), and that these non-synonymous mutations were possibly acquired by recombination. We found that the genetic patterns in the NLSs are associated with EBV-1 and EBV-2, and we suggested that this could be contributing to different cell-type specificities.

## 2. Materials and Methods

### 2.1. Fresh-Frozen Gastric Cancer Tissues for PCR Analysis

Gastric biopsy samples were obtained from 54 adult patients, of both sexes, after surgical resection in the Hernán Henríquez Aravena Hospital. The samples were collected with ethics approval (Universidad de La Frontera N° 20/011) between 2008 and 2012.

### 2.2. DNA Extraction and PCR Assay

DNA was isolated from the fresh-frozen GC tissues using the EZNA Tissue DNA Kit (Omega Bio-Tek, Norcross, USA, GA) according to the manufacturer's instructions. EBV screening was performed by PCR amplification of the *EBNA3A* region. We thus generated 3 overlapping fragments spanning the entire *EBNA3A* region for the 12 EBV positive samples. The sequences of the primers were as follows: 5' region (primers EBNA3A-20F-GTGCGGTGTTGGTGAGTCACACT and EBNA3A1392R-CGCCGACCCGTGACTGGTA, ~1430 pb fragment), center region (primers EBNA3A1041F-CACAKGCTTGGAATGCCGGCTT and EBNA3A2155R-GCCCATACTAGCCCKYAACG, ~1115 pb fragment), and 3' region (primers EBNA3A2136F-CGTTRMGGGCTAGTATGGGC and EBNA3A + 33R-CGGGCGTATTATCAGTGGGT, ~820 pb fragment). The PCR reactions were carried out under the following conditions: 98 °C for 3 min, followed by 35 cycles at 95 °C for 30 s, 55 °C for 30 s, and 72 °C for 1 min. The amplified *EBNA3A* fragments were sequenced using Macrogen's Sanger sequencing system (<https://dna.macrogen.com/>, accessed on 21 March 2018). The 12 *EBNA3A* sequences of gastric cancer bulk from Chile were deposited in GenBank under accession numbers MW219146–MW219157.

### 2.3. Sequence Dataset

One thousand and seventeen EBV genome sequences were obtained from Genbank (accessed on 8 April 2020). Sequence similarity searches with the Basic Local Alignment Search Tool (tBLASTn) were performed using *EBNA3A* amino acid queries P12977 and Q69138 as sequences equivalent to EBV-1 (the B95-8.Raji prototype) [23] (NC\_007605) and EBV-2 (the AG876 [24] (NC\_009334) prototype), respectively. Sequences lacking any part of the *EBNA3* gene (e.g., AP019056, MH883771, LS992262) were eliminated from the analysis. A final dataset of 1006 sequences covering the entire *EBNA3A* gene (nucleotides 79955–82877 relative to the B95-8.Raji sequence including the exons and intron) was recovered from each genome, including *EBNA3A* sequences from the 12 samples sequenced in this study. The geographic origin, clinical manifestation, year of collection, and Genbank accession numbers are available in Supplementary Table S1.

### 2.4. Recombination and Phylogenetic Analyses

The recombination analysis of *EBNA3A* was performed using two different software programs: Gubbins, and RDP4. The recombination events were detected by super alignment output using Gubbins [25]. The maximum likelihood (ML) tree reconstruction was inferred using PhyML 3.0. (<http://www.atgc-montpellier.fr/phyml/>, accessed on 25 May 2020) under an HKY +  $\Gamma$ -I substitution model determined by SMS (<http://www.atgc-montpellier.fr/sms/>, accessed on 24 May 2020) based on the Akaike information criterion (AIC), and the reliability of the topology obtained was estimated with the approximate likelihood-ratio test (aLRT). The tree was visualized using Figtree v.1.4.3 (<http://tree.bio.ed.ac.uk/software/figtree/>, accessed on 8 May 2021) and results were treated with iTOL [26]. The recombination detection based on another approach was reexamined using the RDP4 phylogenetic method as well as the Chimaera [27], GENECONV [28], MaxChi [28], SisScan [29], 3Seq [30], and nucleotide substitution methods, all implemented in the RDP4 software package with default parameters (Supplementary Tables S2 and S3). When all algorithms had the threshold *p*-value of 0.05, the potential putative recombinant events were considered significant, using the Bonferroni correction.

### 3. Results

#### 3.1. EBV-Associated Gastric Cancer from Chile

Fifty-four Chilean patients were recruited in our study. All patients with adenocarcinoma were diagnosed via surgery tumor resection and subsequent confirmation with histopathology. EBV was detected in 12 of the 54 (22%) gastric adenocarcinoma cases. The DNA from these samples was used to analyze variations in EBNA3A. Up to the present study, no EBV sequences from Chile were available. Here we reveal the first 12 EBV sequences of gastric cancer bulk from Chile (Supplementary Table S4).

#### 3.2. EBNA3A Recovery from Genomes

Since no sequences of the *EBNA3A* gene were available in the databank, it was necessary to recover them from the EBV genomes. Members of the *EBNA3* family have two exons separated by a small intron [23]; therefore, complete *EBNA3A* CDS was recovered from EBV genomes through BLAST searches. In addition to these sequences, 12 Chilean *EBNA3A*s were sequenced in this study. Altogether, the dataset contains 1006 *EBNA3A* sequences representing the 5 continents (Supplementary Table S1). The clinical status of these infected individuals was diverse, from healthy to tumor carriers. The final dataset consists of 89 samples from America (Argentina, Brazil, Chile, and the USA), 93 from Europe (France, Germany, Poland, Ukraine, and the United Kingdom), 60 from Africa (Ghana, Kenya, Nigeria, North Africa, and Uganda), 32 from Oceania (Australia and Papua New Guinea), and 732 from Asia (China, Indonesia, Japan, Singapore, South Korea, Taiwan, and Vietnam). It was evident that there is an oversampling of Asian sequences (more than 70%).

#### 3.3. Classifying EBV According to Traditional Classifications

The standard criteria (type 1 and 2) for EBV were applied to classify the *EBNA3A* sequences (including the two exons and the small intron). The ML phylogenetic reconstruction showed that the main clade containing the EBV-1 prototype sequence (B95-8.Raji) congregated most EBV strains, reflecting the considerable number of EBV-1 sequences (Figure 1, tips highlighted in pink). The EBV-1 clade also contains nine Chilean sequences from this study: CHI29, CHI33, CHI34, CHI47, CHI53, CHI55, CHI66, CHI127, and CHI130 (Figure 1, tips indicated with blue triangles).

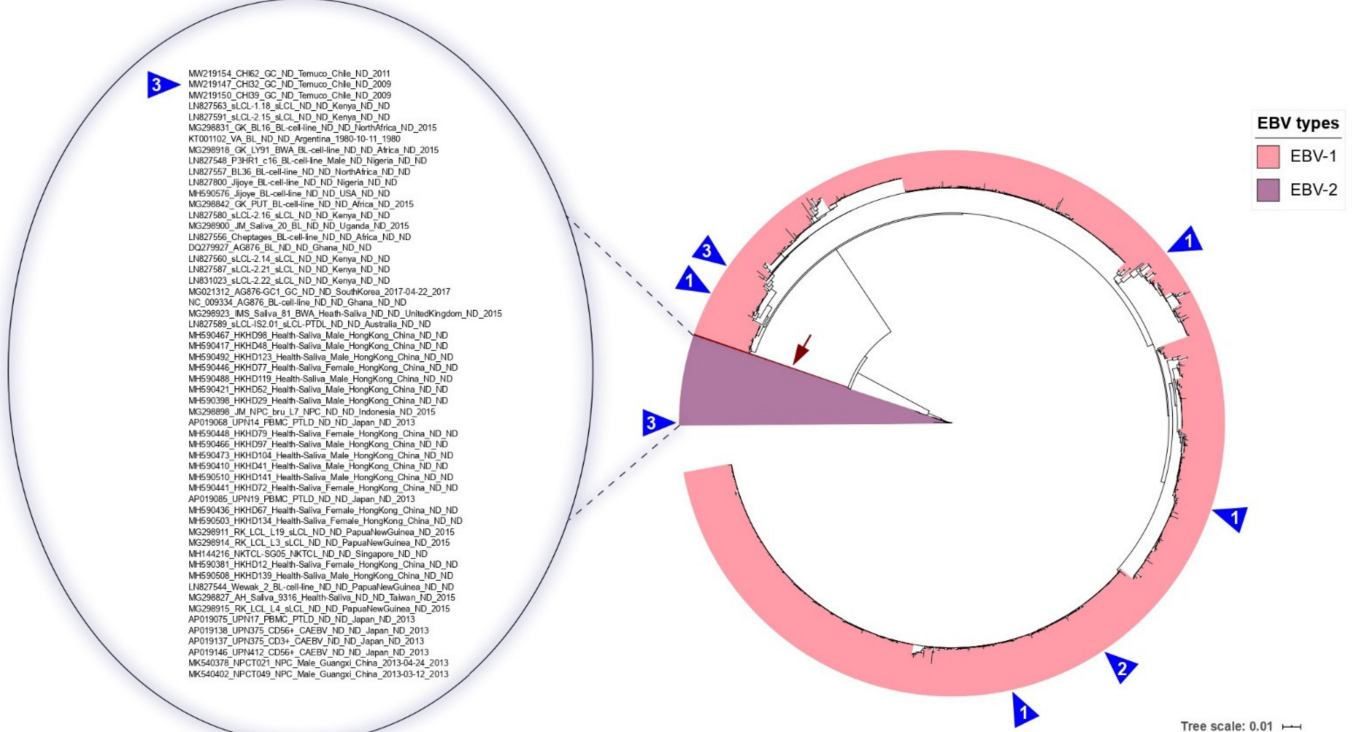
A minor clade was found to consist of ~5% of sequences ( $n = 56$ , see Figure 1, tips highlighted in lilac). This minor clade contains 13 sequences previously classified as belonging to the EBV-2 clade [22] (AG876–EBV-2 prototype, Cheptages, Jijoye, BL36, Wewak, sLCL-2.14, sLCL-2.16, sLCL1.18, sLCL2.21, sLCL-IS2.01, sLCL-2.15, sLCL-2.22, and AG876-GC1). The EBV-2 clade also contains three Chilean sequences from this study: CHI32, CHI39, and CHI62 (Figure 1, tips indicated with blue triangles), which shows the presence of EBV-2 in Chile. Two sequences (NPCT021 and NPCT049, Figure 1 tips highlighted in burgundy) were closely related to EBV-2 sequences.

The tree evidences the wide geographic distribution of EBV-1 compared with EBV-2. Half of the type 2 sequences came from the Asian continent (29/56), and the majority of the other half came from the African continent (16/56), with a meager number of samples from Oceania (5/56) and the Americas (5/56), and just one from the United Kingdom.

#### 3.4. South American EBV Characteristics

To date, there have been 12 *EBNA3A* EBV sequences from South America: 6 from Argentina and 6 from Brazil, all representing individuals with BL. Except for the Argentinian sequence VA sampled in 1980, which belongs to EBV type 2, all other South American sequences belong to the EBV type 1 clade (Table S2). Due to the scarcity of EBV sequences from South America, we have revealed the first 12 *EBNA3A* sequences of EBV-positive gastric cancer from Chile. Therefore, this study doubles the number of EBV *EBNA3A* sequences from South America (Supplementary Table S4). Contrary to expectations, three of the Chilean sequences belong to the EBV type 2 clade, indicating that both types are

circulating in Chile. In addition, this is the first non-Asian (MG021312 from South Korea) GC associated with EBV-2 sequences.



**Figure 1.** ML phylogenetic tree reconstruction of 1006 EBV *EBNA3A* sequences (2835 nt) from the 5 continents. In pink are the sequences classified as EBV-1; in lilac are the sequences classified as EBV-2. The two sequences (NPCT021 and NPCT049) closely related to the EBV-2 clade are in burgundy, and are depicted by a burgundy arrow. The zoom area shows in detail the EBV-2 sequences with their associated metadata. Blue triangles represent Chilean sequences identified in this study. The numbers inside the triangle indicate the number of sequences. The scale bar indicates nucleotide substitutions per site.

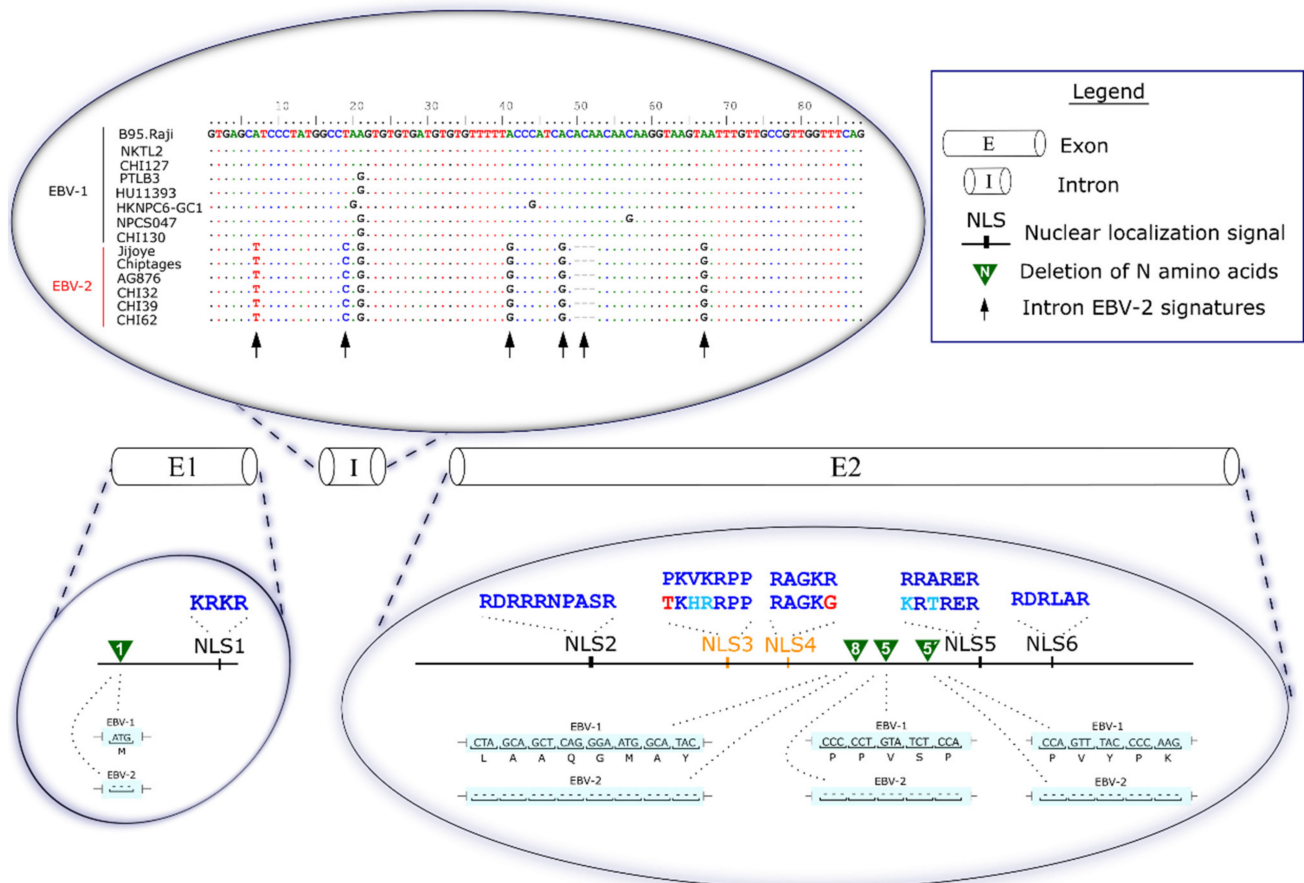
No unique signatures were found among the South American sequences. The *EBNA3A* sequence does not have sufficient signal to fine-tune discriminate subclades inside the EBV-1 group, but the phylogenetic analysis also revealed that the South American sequences belonging to the EBV-1 clade were widely dispersed in the tree, stressing the cosmopolitan nature of EBV type 1.

### 3.5. Two Non-Canonical NLSs Could Be Related to the Lower Efficiency of EBV-2 in Converting B Cells into Lymphoblastoid Cell Lines

The *EBNA3A* is essential for lymphocyte transformation, negatively affecting the transactivator EBNA-2 and the cell cycle [20]. The nucleus is the primary localization of this protein; therefore, changes in NLS residues could affect *EBNA3A* nucleocytoplasmic trafficking [20,21]. NLS3 and NLS4 have been reported previously as functional NLSs, and residue changes in these NLSs have been shown to affect *EBNA3A* partial protein transport to the nucleus [21].

Comparing the *EBNA3A* sequences made it possible to identify characteristic amino acid substitutions associated with EBV types (Figure 2). *EBNA3A* had six predicted NLSs [21], and we identified that six mutations that result in amino acid changes in the NLSs (NLS3, NLS4, and NLS5) present exclusively in EBV-2 strains (except for the NPCT021 and NPCT049 sequences from China, closely related to the EBV-2 clade). Among the six amino acid substitutions identified in the present study, two were non-conservative amino acid changes in NLS3 and NLS4 (Supplementary Table S5). P375T belongs to NLS number 3 (NLS3), and R398G belongs to NLS number 4 (NLS4) (see Figure 2, bottom). NLS3 (a

pattern 7 NLS) had the characteristic proline in the first position substituted by a glycine. NLS4 had basic residue (arginine) substituted by an aliphatic (glycine). EBV-2 sequences present both two NLS motifs altered: NLS3 and NLS4, contrasting EBV-1 sequences with the six canonical NLSs.



**Figure 2.** Schematic representation of EBNA3A amino acid architecture. Coding (exons E1 and E2) and non-coding (intron I) regions of EBNA3A are in the center of the image. The intron region's upper zoom area shows the five unique nucleotides and three-nucleotide deletions (arrows) shared only with the EBV-2 sequences. At the bottom of the image are the EBNA3A nuclear location signals (NLSs), with the respective canonical amino acids of each of the six motifs in dark blue. Non-canonical NLSs with crucial changes are highlighted in orange (NLS3 and NLS4). The amino acid changes that would not directly affect the non-canonical motifs are in light blue. The amino acid changes that would directly affect the non-canonical motifs are highlighted in red. The green triangle indicates the four EBV-2-typical deletions identified in this study, and the amino acid deletions are highlighted with a light blue background. The numbers inside the triangles (1, 8, 5, and 5') indicate the number of amino acids deleted. There were two deletions of five amino acids: the first deletion is indicated with the number 5, and the second with 5'.

### 3.6. Characteristic Deletions of EBV-2 in the EBNA3A Gene

The EBV-2 sequences show four particular regions with deletions in exonic regions (Figure 2, bottom; see Supplementary Table S6). (1) Deletion of one amino acid (Figure 2, the green triangle with the number 1 inside), M<sup>15</sup>, present in more than 96% of EBV-2 sequences (except for the NPCT021 and NPCT049 sequences from China, closely related to the EBV-2 clade). (2) Deletion of eight amino acids (Figure 2, the green triangle with the number 8 inside), <sup>451</sup>LAAQGMAY<sup>458</sup>, present in more than 92% of EBV-2 sequences (except for the UPN375\_CD3+ and UPN412\_CD56+ sequences from Japan, and the NPCT021 and NPCT049 sequences from China). (3) First deletion of five amino acids (Figure 2, the green triangle with the number 5 inside), <sup>480</sup>PPVSP<sup>484</sup>, present in more than 89% of EBV-2 sequences (except for UPN375\_CD3+, UPN375\_CD56+, and UPN412 from Japan, and

NPCT021 and NPCT049 from China). (4) Second deletion of five amino acids (Figure 2, the green triangle with the number 5' inside), <sup>548</sup>PVYPK<sup>552</sup>, present in more than 96% of EBV-2 sequences (except for the NPCT021 and NPCT049 sequences from China).

Five EBV-2 sequences present unusual InDel patterns. (1) The UPN14 sequence from Japan has a deletion of 250 nt. (2) The VA sequence from Argentina has a deletion of 115 nt. (3 and 4) The UPN17 and UPN19 sequences from Japan have a deletion of 110 nt. These four sequences have relatively extensive deletions that include the deletion of eight amino acids and the first deletion of five amino acids identified in this study. Additionally, (5) the IMS\_saliva\_81 sequence from the United Kingdom has all deletions associated with EBV-2 (deletions 1,8, 5, and 5'), and immediately after the deletion of eight amino acids has a further deletion of 8 nt. The unusual InDel patterns could be representative of natural sequence variability, or could have originated from sequencing or assembly errors.

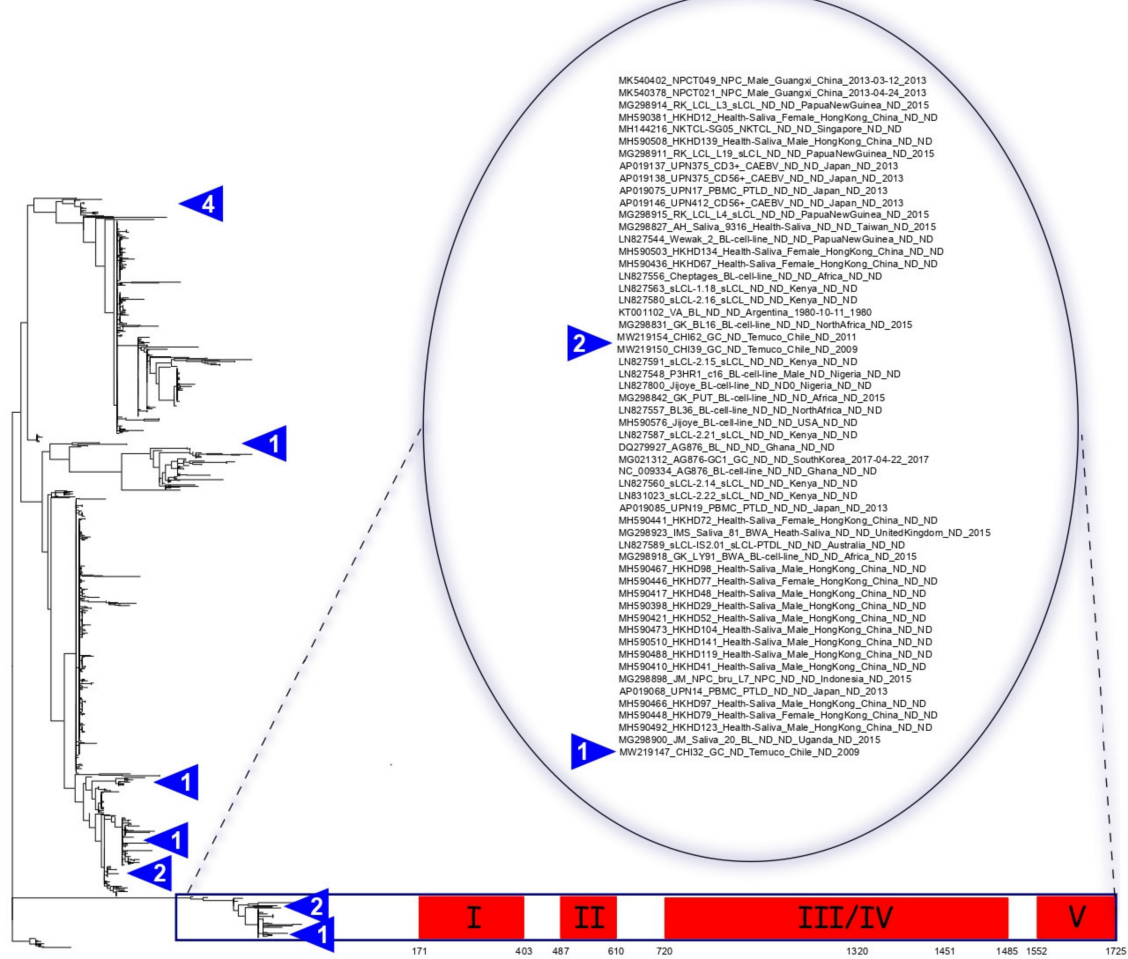
### 3.7. Unique Type 2 Signatures Retained in the Intron Region of EBNA3A

The *EBNA2* and *3A*, *3B*, and *3C* genes have a series of polymorphisms that allow the phylogenetic distinction of EBV-1 and EBV-2. We investigated such polymorphisms in the intron region of the *EBNA3A* coding region (88 nt, Figure 2 upper zoomed area). The alignment of the *EBNA3A* intron region revealed five EBV-2 signatures: A->T, T->C, A->G, A->G, and A->G, at the respective positions 80,368, 80,380, 80,402, 80,409, and 80,425. All nucleotide positions correspond to the EBV-2 prototype sequence AG876. In addition to the polymorphisms, a three-nucleotide deletion was also noted in all EBV-2 sequences. These polymorphisms and deletion patterns could be useful for further studies, or could serve as genetic markers for EBV classification.

### 3.8. EBNA3A Recombination Events Reside Exclusively in EBV-2

It was demonstrated that recombination is an important feature in the construction of the EBV genome architecture [6,22,31–36]. Based on this precedence, we identified five recombination blocks within more than a thousand sequences across the *EBNA3A* gene, based on SNP density (Figure 3 and Supplementary Table S7). Interestingly, only in EBV-2 sequences have recombination blocks been identified; conversely, no recombination blocks were identified in EBV-1 sequences. Three recombination blocks—recombinant block I (position 171–403 nt), recombinant block II (position 487–610 nt), and recombinant block V (position 1552–1725 nt)—were present in the same fifty-six sequences classified as EBV-2. One recombination block—the recombinant block III (position 720–1485 nt)—was present in all EBV-2 sequences except for the NPCT021 and NPCT049 sequences; both of these sequences, from China, are closely related to the EBV-2 clade. One recombination block—recombinant block IV (position 1320–1451 nt)—was present in only four EBV sequences from Japan (UPN375\_CD3+, UPN375\_CD56+, UPN17, and UPN412). Interestingly, both non-canonical NLS changes identified in the present study were positioned inside recombinant block III (see Figure 3, bottom).

We also reanalyzed the recombination in *EBNA3A* using different approaches via nucleotide substitution methods and phylogenetic methods to confirm these findings. Twelve possible recombination events were identified, four of which were defined as misidentifications (Supplementary Tables S2 and S3). Of the eight remaining events, only one had statistical support ( $p$ -value  $\leq 0.05$ ) for all of the evaluation methods. This statistically supported event has the potential recombinant sequences NPC021 and NPC049, the major parent of the EBV-1 sequence RK\_LCL\_L4 from Papua New Guinea, and the minor parent of the EBV-2 sequence NPC049/NPC021 from China. (Supplementary Tables S5 and S4, event 1). This result is consistent, and reinforces the different patterns presented by the NPC021 and NPC049 sequences compared to EBV type 2.



**Figure 3.** Phylogenetic tree and predicted recombination blocks in the EBNA3A gene. Phylogenetic tree reconstruction based on non-recombinant SNPs of 1006 EBV EBNA3A sequences on the Figure's left side. The red blocks (I–V) indicate recombination areas common in multiple sequences that share descent at the bottom of the image. Below the red blocks are the positions corresponding to the EBNA3A nucleotide sequence. Sequences with predicted recombination blocks are in the zoomed area with their associated metadata. Only EBV-2 sequences had recombinant blocks predicted. Chilean sequences identified in this study are represented by blue triangles. The numbers inside the triangle indicate the number of sequences.

#### 4. Discussion

EBV nuclear antigens (EBNAs) are nuclear proteins expressed during latent infection. The six different latent antigens (EBNA1, 2, 3A, 3B, 3C, and LP) influence both the virus' transcription and its host cell [37]. Several studies, using different methodological techniques, suggest the essential role of EBNA3A in B cell transformation by EBV [13,38–44]. Additionally, a recent study showed that mice infected with EBV deficient in EBNA3A and EBNA3C oncogenes had lower viral persistence levels than those infected with the wild-type virus [45].

EBV has a large genome with several repetitive regions posing a tremendous challenge to the assembly process. We found some EBV-2 sequences with unusual InDel patterns that could be representative of natural sequence variability, otherwise originating from sequence or assembly errors. Nevertheless, collecting genomic data is becoming a common issue in EBV studies. Many tumor-associated EBV genomes from different geographic areas have been sequenced, allowing access to the virus' genetic information. Further studies that include a greater representation of little-explored geographic locations and the use of long-read sequencing technologies are necessary and imperative in order to improve



and increase the confidence of the information that can be extracted from genomic analyses associated with EBV.

To dissect *EBNA3A* and characterize the gene variations, we recovered 994 *EBNA3A* sequences from EBV genomes. The scarcity of EBV sequences from South America urged us to recover and characterize the first 12 *EBNA3A* sequences of EBV-associated gastric carcinomas (EBV–GC) sequenced from Chile, generating more information about the South American sequences of these viruses of both EBV subtypes. No unique signatures were found among South American sequences. It will be necessary to recover more genome sequences from South American EBV in order to study the viral origin and evolution.

The EBNA3 family comprises EBNA3A, EBNA3B, and EBNA3C—non-redundant genes arranged in tandem and transcribed by the Cp or Wp promoters, originating along mRNA that is alternatively spliced [13,38]. These genes share a typical exon structure (a tandem array of exons: short 5' exon and longer 3' exon separated by a short intron) and a homology domain (~300 aa) N-terminal region [14]. The *EBNA3* genes possibly originated from duplication events of a single *EBNA3* ancestral gene [46]. Furthermore, EBNA3's intron has stop codons in all reading frames and suboptimal splicing sites [47]. EBNA3's intronic region, in contrast to its exons (30%), has a high sequence homology (95–98%) between EBV-1 and EBV-2 at the nucleotide level, indicating selective pressure [38,47]. We investigated *EBNA3A*'s intron nucleotide sequence for EBV-2-unique signatures. A five-nucleotide signature is a genetic marker of EBV-2 strains present in the intronic region, as is a deletion of three nucleotides. Therefore, this intron's retention and its characteristic type signatures would be related to the fine-tuning expression of EBNA3's genes [47]. *EBNA3A* phylogenetic reconstruction based on a short 5' exon and longer 3' exon of 1006 *EBNA3A* sequences could discriminate EBV types. In addition, EBV-2 is a more geographically restricted virus compared with EBV-1, reinforcing previous studies [48–51], and this is the first evidence-based sequence analysis of the presence of EBV-2 in Chile.

The presence of several nucleotide changes between EBV-1 and EBV-2 within the *EBNA3A* gene allows for a clear phylogenetic type definition. The EBNA3 genes are frequently a target of the immune system, which is a possible explanation for the notable presence of mutations in these genes [6,50,52]. Some deletions previously described exclusively in EBV-2 prototype AG876 [7] were confirmed in most EBV-2 strains.

The EBNA3 family contains several NLSs to guarantee its subcellular localization in the nucleus [20]. The NLSs of EBNA3 interact with cargo transporter proteins, such as importin alpha/beta-5, and regulate the protein's transport into the nucleus through the nuclear pore complex (NPC) [53]. Regarding the potential NLSs identified in EBNA3: EBNA3A [21] has six potential NLSs; EBNA3B [19] has four potential NLSs; and EBNA3C [54] has five potential NLSs. Therefore, EBNA3A has the largest number of NLSs compared with the other EBNA3 genes.

Computational prediction analyses combined with site-direct mutagenesis experiments using GFP-tagged partial constructs showed that mutation in NLS3 or NLS4 alone was insufficient to exclude the EBNA3A protein from the nucleus. In contrast, mutations in both NLS3 and NLS4 resulted in the EBNA3A partial protein being restricted to the cytoplasm subcellular localization [21]. We report four nucleotide substitutions in both NLS3 and NLS4, present exclusively in the EBNA3A of EBV-2 strains. The NLS3 is a monopartite NLS pattern 7, which begins with proline, and contains three or four basic residues [19]. All EBV-2 EBNA3A sequences had the NLS3 TKHRRPP motif amino acid substitutions, resulting in a motif with a lack of proline. Likewise, all EBV-2 EBNA3A sequences had the NLS4 RAGKG motif amino acid substitution, resulting in a motif with fewer basic residues (only two basic residues). There are some pieces of evidence that proline substitution in NLS affects its function. (1) Mutagenesis analyses of the transcription factor c-Myc reveal which proline amino acid is crucial to efficiently importing the protein into the nucleus [55]. (2) Another study shows alteration in protein kinase CK2 interaction with importin by altering NLSs' proline via mutagenesis [56]. Somehow, this could be affecting the EBNA3A's traffic to the nucleus.

Like previous studies, we have identified recombinant regions in the *EBNA3A* gene [22,33,57–60]. We have delineated the effect of recombination in the *EBNA3A* gene, demonstrating that only EBV-2 lineages had recombination events predicted in ancestral nodes. The NLS3 and NLS4 amino acid changes are present inside recombination block III. These common ancestor strains could have undergone a recombination event that changed the NLS3 and NLS4 from canonical to non-canonical motifs.

As previously discussed in Coleman et al. [11], why EBV-2 is less efficient than EBV-1 in transforming B lymphocytes in vitro is intriguing. EBV-2 was detected infecting the T-cells of healthy Kenyan children [61], and displayed tropism to T cells in vitro [11]. In addition, Burgess et al. [19] hypothesized that the presence of multiple NLSs could confer different functions or specificities to different cell types in the EBNA3 family proteins.

Trying to solve one piece of this puzzle, we investigated the NLSs of EBNA3A. EBV-2 sequences simultaneously presented two non-canonical NLS motifs (NLS3 and NLS4), contrasting with the EBV-1 sequences with the six canonical NLS motifs. This finding evidences the presence of genetic patterns in NLS motifs associated with EBV-1 and EBV-2. Considering that EBNA3A plays an important role in EBV transformation, our findings suggest that the NLSs from EBNA3A could be contributing to enabling EBV-1 and EBV-2 to have distinct specificity to different cell types, but further experiments using recombinant viruses are necessary in order to confirm the impact of EBNA3A NLS mutations on cell tropism.

## 5. Conclusions

In summary, our study analyzed more than one thousand EBNA3A sequences from EBV-1 and EBV-2, including the first sequences of EBV-positive gastric cancer from Chile. We identified the presence of EBV-1 and EBV-2 circulating in Chile. At a global level, we determined signatures exclusive to the EBV-2 lineage within the EBNA3A gene in both exon and intron regions, and these lineage-specific signatures could serve as EBV type markers. Therefore, EBNA3A's NLS genetic patterns associated with EBV types provide insights into their roles in different cell-type specificities, and provide new information about the traits of this protein in the EBV's biology. These findings provide the motivation for experimental validation to confirm the impact of EBNA3A NLS mutations on cell-type specificities.

**Supplementary Materials:** The following are available online at <https://www.mdpi.com/article/10.3390/cancers13112569/s1>: Supplementary Table S1: List of sequences analyzed in this study, including metadata; Supplementary Table S2: List of sequences of each putative recombinant event identified by RDP4 and breakpoint prediction; Supplementary Table S3: RDP4 methods with supporting *p*-values for putative recombinant events in the EBNA3A gene; Supplementary Table S4: Characteristics of 24 South American EBV sequences; Supplementary Table S5: EBNA3A canonical NLS amino acid (aa) positions.; Supplementary Table S6: EBNA3A amino acid (aa) deletion patterns; and Supplementary Table S7: Putative recombination regions identified by Gubbins.

**Author Contributions:** Conceptualization, validation, formal analysis, data curation, writing—original draft, writing—review and editing, and visualization: L.Z.; conceptualization, validation, writing—original draft, writing—review and editing, and visualization: M.E.R.; investigation, validation, and data curation: I.R.; validation, data curation, and writing—review and editing: M.A.; resources, validation, and data curation: D.L.; resources, validation, and data curation: T.V.; investigation, validation, and data curation: C.I.; project administration, funding acquisition, resources, supervision, and writing—review and editing: P.B. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was supported by National Funding for Scientific and Technological Development of Chile (FONDECYT) [3210687 to L.Z., 3170826 to I.R., 3210618 to D.L., 3210604 to T.V., 11150622 to CI and 1210440 to PB], Universidad de La Frontera [VRIP20P002 to L.Z.]; the National Commission for Scientific and Technological Research (CONICYT) [21201835 to M.R.]; Dirección de Investigación, Universidad de La Frontera (DIUFRO) [DIM20-0019 to LZ, DI20-0160 to CI and

DI20-0128 to PB], and Dirección de Investigación, Universidad Autónoma (UA) [DIUA 226-2021 to I.R.].

**Institutional Review Board Statement:** The study was approved by the Ethics Committee of UNIVERSIDAD DE LA FRONTERA (protocol code n°20/011 December 2011).

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Data Availability Statement:** The data presented in this study are openly available on GenBank at accession numbers [MW219146–MW219157].

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Vetsika, E.-K.; Callan, M. Infectious mononucleosis and Epstein-Barr virus. *Expert Rev. Mol. Med.* **2004**, *6*, 1–16. [[CrossRef](#)]
- Gunvén, P.; Klein, G.; Henle, G.; Henle, W.; Clifford, P. Epstein-Barr Virus in Burkitt's Lymphoma and Nasopharyngeal Carcinoma: Antibodies to EBV associated Membrane and Viral Capsid Antigens in Burkitt Lymphoma Patients. *Nat. Cell Biol.* **1970**, *228*, 1053–1056. [[CrossRef](#)] [[PubMed](#)]
- Tsao, S.W.; Yip, Y.L.; Tsang, C.M.; Pang, P.S.; Lau, V.M.Y.; Zhang, G.; Lo, K.W. Etiological factors of nasopharyngeal carcinoma. *Oral Oncol.* **2014**, *50*, 330–338. [[CrossRef](#)] [[PubMed](#)]
- Corvalan, A.; Koriyama, C.; Akiba, S.; Eizuru, Y.; Backhouse, C.; Palma, M.; Argandoña, J.; Tokunaga, M. Epstein-Barr virus in gastric carcinoma is associated with location in the cardia and with a diffuse histology: A study in one area of Chile. *Int. J. Cancer* **2001**, *94*, 527–530. [[CrossRef](#)] [[PubMed](#)]
- Baer, R.; Bankier, A.T.; Biggin, M.D.; Deininger, P.; Farrell, P.J.; Gibson, T.J.; Hatfull, G.; Hudson, G.S.; Satchwell, S.C.; Séguin, C.; et al. DNA sequence and expression of the B95-8 Epstein-Barr virus genome. *Nat. Cell Biol.* **1984**, *310*, 207–211. [[CrossRef](#)] [[PubMed](#)]
- Palser, A.L.; Grayson, N.E.; White, R.E.; Corton, C.; Correia, S.; Abdullah, M.M.B.; Watson, S.J.; Cotten, M.; Arrand, J.R.; Murray, P.G.; et al. Genome Diversity of Epstein-Barr Virus from Multiple Tumor Types and Normal Infection. *J. Virol.* **2015**, *89*, 5222–5237. [[CrossRef](#)]
- Sample, J.; Young, L.; Martin, B.; Chatman, T.; Kieff, E.; Rickinson, A. Epstein-Barr virus types 1 and 2 differ in their EBNA-3A, EBNA-3B, and EBNA-3C genes. *J. Virol.* **1990**, *64*, 4084–4092. [[CrossRef](#)] [[PubMed](#)]
- Lucchesi, W.; Brady, G.; Dittrich-Breiholz, O.; Kracht, M.; Russ, R.; Farrell, P.J. Differential Gene Regulation by Epstein-Barr Virus Type 1 and Type 2 EBNA2. *J. Virol.* **2008**, *82*, 7456–7466. [[CrossRef](#)]
- Ponnusamy, R.; Khatri, R.; Correia, P.B.; Wood, C.D.; Mancini, E.J.; Farrell, P.J.; West, M.J. Increased association between Epstein-Barr virus EBNA2 from type 2 strains and the transcriptional repressor BS69 restricts EBNA2 activity. *PLoS Pathog.* **2019**, *15*, e1007458. [[CrossRef](#)]
- Coleman, C.B.; Lang, J.; Sweet, L.A.; Smith, N.A.; Freed, B.M.; Pan, Z.; Haverkos, B.; Pelanda, R.; Rochford, R. Epstein-Barr Virus Type 2 Infects T Cells and Induces B Cell Lymphomagenesis in Humanized Mice. *J. Virol.* **2018**, *92*, 1–22. [[CrossRef](#)]
- Coleman, C.B.; Wohlford, E.M.; Smith, N.A.; King, C.A.; Ritchie, J.A.; Baresel, P.C.; Kimura, H.; Rochford, R. Epstein-Barr Virus Type 2 Latently Infects T Cells, Inducing an Atypical Activation Characterized by Expression of Lymphotactic Cytokines. *J. Virol.* **2015**, *89*, 2301–2312. [[CrossRef](#)] [[PubMed](#)]
- Cludts, I.; Farrell, P.J. Multiple Functions within the Epstein-Barr Virus EBNA-3A Protein. *J. Virol.* **1998**, *72*, 1862–1869. [[CrossRef](#)] [[PubMed](#)]
- White, R.E.; Groves, I.J.; Turro, E.; Yee, J.; Kremmer, E.; Allday, M.J. Extensive Co-Operation between the Epstein-Barr Virus EBNA3 Proteins in the Manipulation of Host Gene Expression and Epigenetic Chromatin Modification. *PLoS ONE* **2010**, *5*, e13979. [[CrossRef](#)] [[PubMed](#)]
- Ohashi, M.; Holthaus, A.M.; Calderwood, M.A.; Lai, C.-Y.; Krastins, B.; Sarracino, D.; Johannsen, E. The EBNA3 Family of Epstein-Barr Virus Nuclear Proteins Associates with the USP46/USP12 Deubiquitination Complexes to Regulate Lymphoblastoid Cell Line Growth. *PLoS Pathog.* **2015**, *11*, e1004822. [[CrossRef](#)] [[PubMed](#)]
- Allday, M.J.; Bazot, Q.; White, R.E. The EBNA3 Family: Two Oncoproteins and a Tumour Suppressor that Are Central to the Biology of EBV in B Cells. *Curr. Top. Microbiol. Immunol.* **2015**, *391*, 61–117. [[CrossRef](#)]
- Maruo, S.; Johannsen, E.; Illanes, D.; Cooper, A.; Zhao, B.; Kieff, E. Epstein-Barr Virus Nuclear Protein 3A Domains Essential for Growth of Lymphoblasts: Transcriptional Regulation through RBP-Jk/CBF1 Is Critical. *J. Virol.* **2005**, *79*, 10171–10179. [[CrossRef](#)]
- Hickabottom, M.; Parker, G.A.; Freemont, P.S.; Crook, T.; Allday, M.J. Two Nonconsensus Sites in the Epstein-Barr Virus Oncoprotein EBNA3A Cooperate to Bind the Co-repressor Carboxyl-terminal-binding Protein (CtBP). *J. Biol. Chem.* **2002**, *277*, 47197–47204. [[CrossRef](#)]
- Bazot, Q.; Paschos, K.; Allday, M.J. Epstein-Barr Virus (EBV) Latent Protein EBNA3A Directly Targets and Silences the STK39 Gene in B Cells Infected by EBV. *J. Virol.* **2018**, *92*, e01918-17. [[CrossRef](#)]
- Burgess, A.; Buck, M.; Krauer, K.; Sculley, T. Nuclear localization of the Epstein-Barr virus EBNA3B protein. *J. Gen. Virol.* **2006**, *87*, 789–793. [[CrossRef](#)]

20. Cai, M.; Liao, Z.; Chen, T.; Wang, P.; Zou, X.; Wang, Y.; Xu, Z.; Jiang, S.; Huang, J.; Chen, D.; et al. Characterization of the subcellular localization of Epstein-Barr virus encoded proteins in live cells. *Oncotarget* **2017**, *8*, 70006–70034. [[CrossRef](#)]
21. Buck, M.; Burgess, A.; Stirzaker, R.; Krauer, K.; Sculley, T. Epstein-Barr virus nuclear antigen 3A contains six nuclear-localization signals. *J. Gen. Virol.* **2006**, *87*, 2879–2884. [[CrossRef](#)]
22. Zanella, L.; Riquelme, I.; Buchegger, K.; Abanto, M.; Ili, C.; Brebi, P. A reliable Epstein-Barr Virus classification based on phylogenomic and population analyses. *Sci. Rep.* **2019**, *9*, 1–11. [[CrossRef](#)]
23. De Jesus, O.; Smith, P.R.; Spender, L.C.; Karstegl, C.E.; Niller, H.H.; Huang, D.; Farrell, P.J. Updated Epstein-Barr virus (EBV) DNA sequence and analysis of a promoter for the BART (CST, BARF0) RNAs of EBV. *J. Gen. Virol.* **2003**, *84*, 1443–1450. [[CrossRef](#)] [[PubMed](#)]
24. Dolan, A.; Addison, C.; Gatherer, D.; Davison, A.J.; McGeoch, D.J. The genome of Epstein-Barr virus type 2 strain AG. *Virology* **2006**, *350*, 164–170. [[CrossRef](#)] [[PubMed](#)]
25. Croucher, N.J.; Page, A.; Connor, T.R.; Delaney, A.J.; Keane, J.A.; Bentley, S.D.; Parkhill, J.; Harris, S.R. Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. *Nucleic Acids Res.* **2015**, *43*, e15. [[CrossRef](#)]
26. Letunic, I.; Bork, P. Interactive tree of life (iTOL) v3: An online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res.* **2016**, *44*, W242–W245. [[CrossRef](#)]
27. Posada, D.; Crandall, K.A. Evaluation of methods for detecting recombination from DNA sequences: Computer simulations. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 13757–13762. [[CrossRef](#)]
28. Padidam, M.; Sawyer, S.; Fauquet, C.M. Possible Emergence of New Geminiviruses by Frequent Recombination. *Virology* **1999**, *265*, 218–225. [[CrossRef](#)]
29. Gibbs, M.J.; Armstrong, J.S.; Gibbs, A.J. Sister-Scanning: A Monte Carlo procedure for assessing signals in recombinant sequences. *Bioinformatics* **2000**, *16*, 573–582. [[CrossRef](#)]
30. Boni, M.F.; Posada, D.; Feldman, M.W. An Exact Nonparametric Method for Inferring Mosaic Structure in Sequence Triplets. *Genetics* **2007**, *176*, 1035–1047. [[CrossRef](#)]
31. Abdullah, M.M.B.; Palermo, R.D.; Palser, A.L.; Grayson, N.E.; Kellam, P.; Correia, S.; Szymula, A.; White, R.E. Heterogeneity of the Epstein-Barr Virus (EBV) Major Internal Repeat Reveals Evolutionary Mechanisms of EBV and a Functional Defect in the Prototype EBV Strain B95. *J. Virol.* **2017**, *91*, 00920–17. [[CrossRef](#)] [[PubMed](#)]
32. Berenstein, A.J.; Lorenzetti, M.A.; Preciado, M.V.; José, B.A.; Alejandro, L.M.; Victoria, P.M. Recombination rates along the entire Epstein Barr virus genome display a highly heterogeneous landscape. *Infect. Genet. Evol.* **2018**, *65*, 96–103. [[CrossRef](#)]
33. Görzer, I.; Niesters, H.G.; Cornelissen, J.J.; Puchhammer-Stöckl, E. Characterization of Epstein-Barr virus Type I variants based on linked polymorphism among EBNA3A, -3B, and -3C genes. *Virus Res.* **2006**, *118*, 105–114. [[CrossRef](#)] [[PubMed](#)]
34. Gantuz, M.; Lorenzetti, M.A.; Chabay, P.A.; Preciado, M.V. A novel recombinant variant of latent membrane protein 1 from Epstein Barr virus in Argentina denotes phylogeographical association. *PLoS ONE* **2017**, *12*, e0174221. [[CrossRef](#)] [[PubMed](#)]
35. Kwok, H.; Wu, C.W.; Palser, A.L.; Kellam, P.; Sham, P.C.; Kwong, D.L.W.; Chiang, A.K.S. Genomic Diversity of Epstein-Barr Virus Genomes Isolated from Primary Nasopharyngeal Carcinoma Biopsy Samples. *J. Virol.* **2014**, *88*, 10662–10672. [[CrossRef](#)]
36. Santpere, G.; Darre, F.; Blanco, S.; Alcamí, A.; Villoslada, P.; Albà, M.M.; Navarro, A. Genome-Wide Analysis of Wild-Type Epstein-Barr Virus Genomes Derived from Healthy Individuals of the 1000 Genomes Project. *Genome Biol. Evol.* **2014**, *6*, 846–860. [[CrossRef](#)]
37. Young, L.S.; Rickinson, A.B. Epstein-Barr virus: 40 years on. *Nat. Rev. Cancer* **2004**, *4*, 757–768. [[CrossRef](#)]
38. Saha, A.; Robertson, E.S. Epstein-Barr Virus-Associated B-cell Lymphomas: Pathogenesis and Clinical Outcomes. *Clin. Cancer Res.* **2011**, *17*, 3056–3063. [[CrossRef](#)]
39. Tomkinson, B.; Robertson, E.; Kieff, E. Epstein-Barr virus nuclear proteins EBNA-3A and EBNA-3C are essential for B-lymphocyte growth transformation. *J. Virol.* **1993**, *67*, 2014–2025. [[CrossRef](#)]
40. Young, L.S.; Murray, P.G. Epstein-Barr virus and oncogenesis: From latent genes to tumours. *Oncogene* **2003**, *22*, 5108–5121. [[CrossRef](#)]
41. Kang, M.-S.; Kieff, E. Epstein-Barr virus latent genes. *Exp. Mol. Med.* **2015**, *47*, e131. [[CrossRef](#)]
42. Grossman, S.R.; Johannsen, E.; Tong, X.; Yalamanchili, R.; Kieff, E. The Epstein-Barr virus nuclear antigen 2 transactivator is directed to response elements by the J kappa recombination signal binding protein. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 7568–7572. [[CrossRef](#)]
43. Hertle, M.L.; Popp, C.; Petermann, S.; Maier, S.; Kremmer, E.; Lang, R.; Mages, J.; Kempkes, B. Differential Gene Expression Patterns of EBV Infected EBNA-3A Positive and Negative Human B Lymphocytes. *PLoS Pathog.* **2009**, *5*, e1000506. [[CrossRef](#)] [[PubMed](#)]
44. Maruo, S.; Zhao, B.; Johannsen, E.; Kieff, E.; Zou, J.; Takada, K. Epstein-Barr virus nuclear antigens 3C and 3A maintain lymphoblastoid cell growth by repressing p16INK4A and p14ARF expression. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 1919–1924. [[CrossRef](#)]
45. Murer, A.; McHugh, D.; Caduff, N.; Kalchschmidt, J.; Barros, M.; Zbinden, A.; Capaul, R.; Niedobitek, G.; Allday, M.; Chijioke, O.; et al. EBV persistence without its EBNA3A and 3C oncogenes in vivo. *PLoS Pathog.* **2018**, *14*, e1007039. [[CrossRef](#)] [[PubMed](#)]
46. Wang, A.; Welch, R.; Zhao, B.; Ta, T.; Keleş, S.; Johannsen, E. Epstein-Barr Virus Nuclear Antigen 3 (EBNA3) Proteins Regulate EBNA2 Binding to Distinct RBP Genomic Sites. *J. Virol.* **2016**, *90*, 2906–2919. [[CrossRef](#)] [[PubMed](#)]

47. Kienzle, N.; Young, D.B.; Liaskou, D.; Buck, M.; Greco, S.; Sculley, T.B. Intron retention may regulate expression of Epstein-Barr virus nuclear antigen 3 family genes. *J. Virol.* **1999**, *73*, 1195–1204. [[CrossRef](#)] [[PubMed](#)]
48. Zimmer, U.; Adldinger, H.K.; Lenoir, G.M.; Vuillaume, M.; Knebel-Doerberitz, M.V.; Laux, G.; Deśgranges, C.; Wittmann, P.; Freese, U.-K.; Schneider, U.; et al. Geographical prevalence of two types of Epstein-Barr virus. *Virology* **1986**, *154*, 56–66. [[CrossRef](#)]
49. Correia, S.; Bridges, R.; Wegner, F.; Venturini, C.; Palser, A.; Middeldorp, J.; Cohen, J.I.; Lorenzetti, M.A.; Bassano, I.; White, R.E.; et al. Sequence Variation of Epstein-Barr Virus: Viral Types, Geography, Codon Usage, and Diseases. *J. Virol.* **2018**, *92*, e01132-18. [[CrossRef](#)]
50. Khanna, R.; Slade, R.W.; Poulsen, L.; Moss, D.J.; Burrows, S.R.; Nicholls, J.; Burrows, J.M. Evolutionary dynamics of genetic variation in Epstein-Barr virus isolates of diverse geographical origins: Evidence for immune pressure-independent genetic drift. *J. Virol.* **1997**, *71*, 8340–8346. [[CrossRef](#)]
51. Mandage, R.; Telford, M.; Rodríguez, J.A.; Farré, X.; Layouni, H.; Marigorta, U.M.; Cundiff, C.; Heredia-Genestar, J.M.; Navarro, A.; Santpere, G. Genetic factors affecting EBV copy number in lymphoblastoid cell lines derived from the 1000 Genome Project samples. *PLoS ONE* **2017**, *12*, e0179446. [[CrossRef](#)] [[PubMed](#)]
52. Midgley, R.S.; Bell, A.I.; McGeoch, D.J.; Rickinson, A.B. Latent Gene Sequencing Reveals Familial Relationships among Chinese Epstein-Barr Virus Strains and Evidence for Positive Selection of A11 Epitope Changes. *J. Virol.* **2003**, *77*, 11517–11530. [[CrossRef](#)]
53. Bazot, Q.; Deschamps, T.; Tafforeau, L.; Siouda, M.; Leblanc, P.; Harth-Hertle, M.L.; Rabourdin-Combe, C.; Lotteau, V.; Kempkes, B.; Tommasino, M.; et al. Epstein-Barr virus nuclear antigen 3A protein regulates CDKN2B transcription via interaction with MIZ. *Nucleic Acids Res.* **2014**, *42*, 9700–9716. [[CrossRef](#)] [[PubMed](#)]
54. Krauer, K.; Buck, M.; Flanagan, J.; Belzer, D.; Sculley, T. Identification of the nuclear localization signals within the Epstein-Barr virus EBNA-6 protein. *J. Gen. Virol.* **2004**, *85*, 165–172. [[CrossRef](#)] [[PubMed](#)]
55. Makkerh, J.P.; Dingwall, C.; Laskey, R.A. Comparative mutagenesis of nuclear localization signals reveals the importance of neutral and acidic amino acids. *Curr. Biol.* **1996**, *6*, 1025–1027. [[CrossRef](#)]
56. Xiao, C.-Y.; Jans, P.; Jans, D.A. Negative charge at the protein kinase CK2 site enhances recognition of the SV40 large T-antigen NLS by importin: Effect of conformation. *FEBS Lett.* **1998**, *440*, 297–301. [[CrossRef](#)]
57. Yao, Q.Y.; Tierney, R.J.; Croom-Carter, D.; Cooper, G.M.; Ellis, C.J.; Rowe, M.; Rickinson, A.B. Isolation of intertypic recombinants of Epstein-Barr virus from T-cell-immunocompromised individuals. *J. Virol.* **1996**, *70*, 4895–4903. [[CrossRef](#)]
58. Kim, S.W.; Kim, K.P. Log-likelihood-ratio-based detection ordering in V-BLAST. *IEEE Trans. Commun.* **2006**, *54*, 302–307. [[CrossRef](#)]
59. McGeoch, D.J.; Gatherer, D. Lineage structures in the genome sequences of three Epstein-Barr virus strains. *Virology* **2007**, *359*, 1–5. [[CrossRef](#)]
60. Midgley, R.S.; Blake, N.W.; Yao, Q.Y.; Croom-Carter, D.; Cheung, S.T.; Leung, S.F.; Chan, A.T.C.; Johnson, P.J.; Huang, D.; Rickinson, A.B.; et al. Novel Intertypic Recombinants of Epstein-Barr Virus in the Chinese Population. *J. Virol.* **2000**, *74*, 1544–1548. [[CrossRef](#)]
61. Coleman, C.B.; Daud, I.I.; Ogolla, S.O.; Ritchie, J.A.; Smith, N.A.; Sumba, P.O.; Dent, A.E.; Rochford, R. Epstein-Barr Virus Type 2 Infects T Cells in Healthy Kenyan Children. *J. Infect. Dis.* **2017**, *216*, 670–677. [[CrossRef](#)] [[PubMed](#)]