

Article

# Intensive Distribution of G<sub>2</sub>-Quadruplexes in the Pseudorabies Virus Genome and Their Sensitivity to Cations and G-Quadruplex Ligands

Hui Deng <sup>1,†</sup>, Bowen Gong <sup>1,†</sup>, Zhiqian Yang <sup>2</sup>, Zhen Li <sup>1</sup>, Huan Zhou <sup>1</sup>, Yashu Zhang <sup>3</sup>, Xiaohui Niu <sup>2</sup>, Sisi Liu <sup>1,\*</sup>  and Dengguo Wei <sup>1,\*</sup> 

<sup>1</sup> Laboratory of Medicinal Biophysical Chemistry, College of Science, Huazhong Agricultural University, Wuhan 430070, China; denghui0923@163.com (H.D.); night@webmail.hzau.edu.cn (B.G.); lizhen5a@163.com (Z.L.); zhouhuan@webmail.hzau.edu.cn (H.Z.)

<sup>2</sup> Hubei Key Laboratory of Agricultural Bioinformatics, College of Informatics, Huazhong Agricultural University, Wuhan 430070, China; yang\_zq@foxmail.com (Z.Y.); niuxiaoh@126.com (X.N.)

<sup>3</sup> Laboratory of Medicinal Biophysical Chemistry, College of Plant Science and Technology, Huazhong Agricultural University, Wuhan 430070, China; yashu@webmail.hzau.edu.cn

\* Correspondence: liusisi@mail.hzau.edu.cn (S.L.); dgwei@mail.hzau.edu.cn (D.W.)

† These authors contributed equally to this work.

Academic Editor: Sara N. Richter

Received: 17 January 2019; Accepted: 15 February 2019; Published: 21 February 2019



**Abstract:** Guanine-rich sequences in the genomes of herpesviruses can fold into G-quadruplexes. Compared with the widely-studied G<sub>3</sub>-quadruplexes, the dynamic G<sub>2</sub>-quadruplexes are more sensitive to the cell microenvironment, but they attract less attention. Pseudorabies virus (PRV) is the model species for the study of the latency and reactivation of herpesvirus in the nervous system. A total of 1722 G<sub>2</sub>-PQs and 205 G<sub>3</sub>-PQs without overlap were identified in the PRV genome. Twelve G<sub>2</sub>-PQs from the CDS region exhibited high conservation in the genomes of the *Varicellovirus* genus. Eleven G<sub>2</sub>-PQs were 100% conserved in the repeated region of the annotated PRV genomes. There were 212 non-redundant G<sub>2</sub>-PQs in the 3' UTR and 19 non-redundant G<sub>2</sub>-PQs in the 5' UTR, which would mediate gene expression in the post-transcription and translation processes. The majority of examined G<sub>2</sub>-PQs formed parallel structures and exhibited different sensitivities to cations and small molecules in vitro. Two G<sub>2</sub>-PQs, respectively, from 3' UTR of *UL5* (encoding helicase motif) and *UL9* (encoding sequence-specific ori-binding protein) exhibited diverse regulatory activities with/without specific ligands in vivo. The G-quadruplex ligand, NMM, exhibited a potential for reducing the virulence of the PRV Ea strain. The systematic analysis of the distribution of G<sub>2</sub>-PQs in the PRV genomes could guide further studies of the G-quadruplexes' functions in the life cycle of herpesviruses.

**Keywords:** alphaherpesviruses; pseudorabies virus; genome; G-quadruplex; G-quadruplex ligand; nucleic acids conformation; regulatory element

## 1. Introduction

The life-long latency of herpesviruses poses potential threats to the host at any time, and the reason for the wide existence of the GC-rich sequences in herpesvirus genomes remains unknown [1]. Guanine-rich sequences have been discovered to form special DNA or RNA secondary structures called G-quadruplexes. They are composed of  $\pi$ -stacked G-quartets via hydrogen-bonded structure in DNA or RNA [2]. Each G-quartet contains four guanines held by eight hydrogen bonds, and coordinated with a central monovalent cation, such as K<sup>+</sup> and Na<sup>+</sup>. The G<sub>3</sub>-putative quadruplex sequences (G<sub>3</sub>-PQs) in the form of (G<sub>3+</sub>N<sub>1-7</sub>)<sub>3</sub>G<sub>3+</sub> form three or more G-quartets in the structures (Figure 1), and they exhibited



the virulence of the PRV Ea strain. This study will further investigate G<sub>2</sub>-quadruplexes working as sensors in response to small molecules, proteins, and physiological cation conditions in the specific microenvironment, to reveal the latency-reactivation mechanism of herpesviruses.

## 2. Results

### 2.1. Bioinformatic Analysis

#### 2.1.1. Genome-Wide Analysis of G<sub>3</sub>-PQSs and G<sub>2</sub>-PQSs Distributions in the PRV Genome

The putative G<sub>3</sub>-quadruplex sequences (G<sub>3</sub>-PQSs) in the PRV genome were predicted with the Quadparser program [3] in the form of the G<sub>n</sub>N<sub>m</sub> sequence (where  $n \geq 3$  and  $1 \leq m \leq 7$ ). The putative G<sub>2</sub>-quadruplex sequences (G<sub>2</sub>-PQSs) were predicted with the same program, but in the modified sequence form ( $n \geq 2$ ). The genome size of PRV was 143,461 bp, and it encoded 69 proteins. The analysis of the PRV reference genome (NC\_006151.1) indicated that 1722 G<sub>2</sub>-PQSs and 205 G<sub>3</sub>-PQSs were distributed in the PRV genome (Figure 2; File S1), with the density of the G<sub>2</sub>-PQSs being 12 PQS/kb.

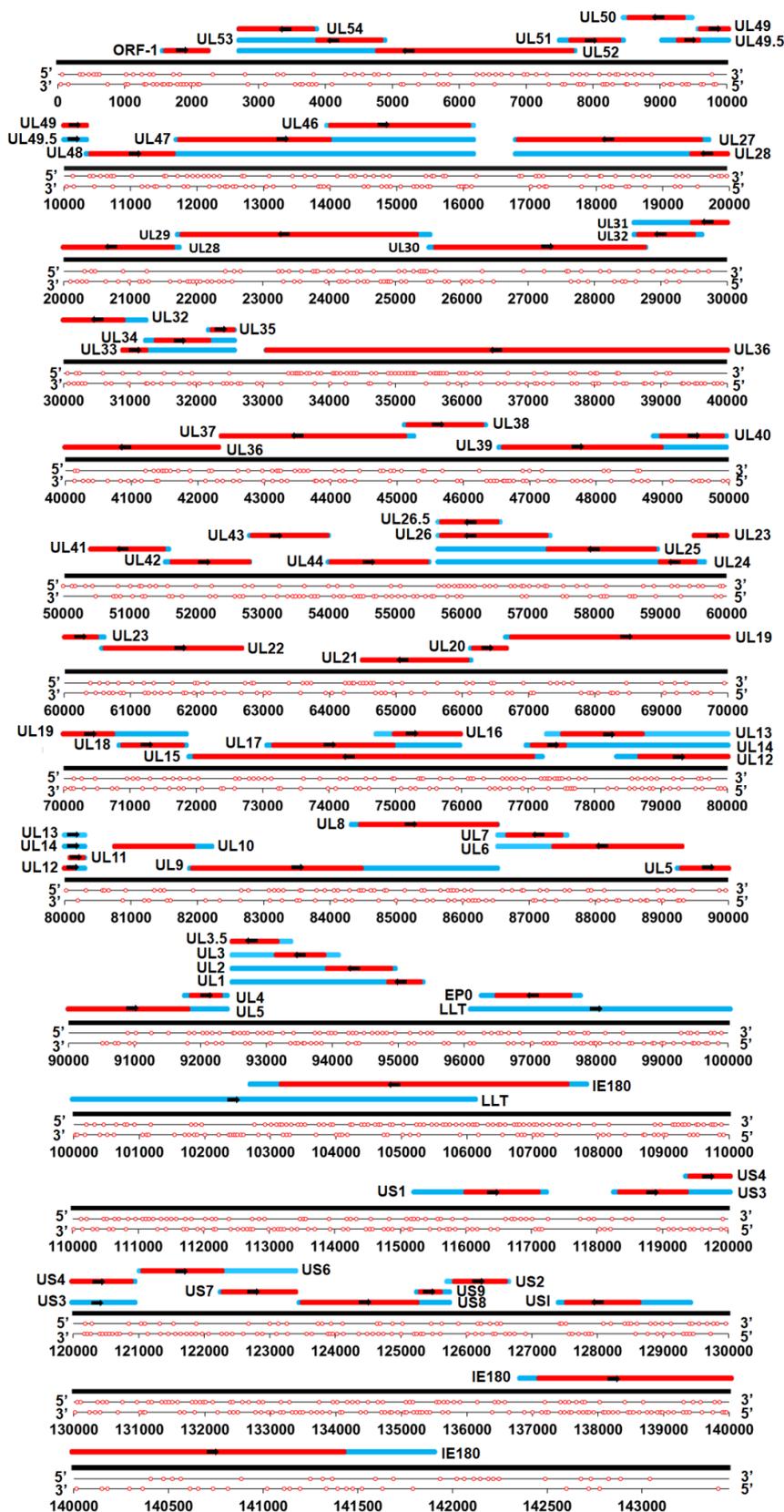
As the formation of the G-quadruplex could affect either transcription or translation, our analysis of PQSs was based on double strands with the well-annotated regions, and the predicted regulatory regions in the PRV reference genome (File S2). The 3' end untranslated regions (3' UTRs) from 63 genes, and the 5' end untranslated regions (5' UTRs) from 61 genes were annotated in the PRV reference genome in The National Center for Biotechnology Information (NCBI) Genome database [24]. The promoters of the annotated PRV genes were predicted to be 1 kb upstream of the annotated transcription start site of each gene. G<sub>2</sub>-PQSs were higher in density in the CDS region and in the large latency transcript (LLT) than in the repeat region, while the G<sub>3</sub>-PQSs were densely distributed in the repeat region (Table 1). The density of the G<sub>2</sub>-PQSs in the 3' UTR was higher than that in the 5' UTR (Table 1). In the promoter regions, the G<sub>2</sub>-PQSs density was 5.34 PQS/kb, and it was more than eight-fold that of the G<sub>3</sub>-PQSs density (Table 1).

The PQS monomer was named the single G-quadruplex-forming sequence, and the sequences forming more than two possible simultaneous G-quadruplexes were defined as the PQS cluster. The G<sub>3</sub>-quadruplex cluster, forming a highly stable structure, was identified in the repetitive region of the Herpes simplex virus type 1 (HSV-1) genome [25]. This study found that 86.8% of G<sub>3</sub>-PQSs and 77% of G<sub>2</sub>-PQSs were monomers in the genome of PRV (File S1). 27 non-redundant G<sub>3</sub>-PQSs clusters were located in the regulatory regions, rather than in the coding region in the PRV genome. Sixty-nine G<sub>2</sub>-PQS clusters were distributed in the repeat regions, and 122 G<sub>2</sub>-PQS clusters were located in the coding region in the PRV genome (Table 1).

**Table 1.** Number of putative G-quadruplex sequences in three herpesvirus genomes.

Herpesviruses	Region †	Length (bp)	Putative G <sub>2</sub> -quadruplex Sequences				Putative G <sub>3</sub> -quadruplex Sequences			
			Number	Monomer	Cluster	Density (PQS/kb)	Number	Monomer	Cluster	Density (PQS/kb)
Pseudorabies virus (PRV, NC_006151.1)	coding sequence (CDS)	105,363	1126	1004	122	10.69	51	51	0	0.48
	3' end untranslated region (3' UTR)	35,688	166	149	17	4.65	6	6	0	0.17
	5' end untranslated region (5' UTR)	6760	19	18	1	2.81	5	3	2	0.74
	Promoter	56,324	301	276	25	5.34	36	32	4	0.64
	large latency transcript (LLT)	13,040	139	109	30	10.66	27	25	2	2.07
	Repeat region	36,234	338	269	69	9.33	131	109	22	3.62
Human alphaherpesvirus 1 (Herpes simplex virus type 1, HSV-1, NC_001806.2)	CDS	121,089	1351	1248	103	11.16	109	104	5	0.90
	3' UTR	49,322	235	223	12	4.76	27	27	0	0.55
	5' UTR	3434	18	18	0	5.24	3	1	2	0.87
	Promoter	64,966	346	317	29	5.33	64	56	8	0.99
	latency-associated transcript (LAT)	15,944	158	146	12	9.91	54	44	10	3.39
	Repeat region	31,875	338	307	31	10.60	120	87	33	3.76
Human alphaherpesvirus 3 (Varicella-zoster virus, VZV, NC_001348.1)	CDS	111,496	296	282	14	2.65	10	10	0	0.09
	3' UTR	29,758	37	36	1	1.24	1	1	0	0.03
	5' UTR *	496	0	0	0	0	0	0	0	0
	Promoter	60,900	71	69	2	1.17	5	3	2	0.08
	VZV latency-associated transcript (VLT)	2417	13	12	1	5.38	0	0	0	0
	Repeat region	15,514	119	111	8	7.67	11	9	2	0.71

†: The CDS, 3' UTR, 5' UTR and repeat regions were analyzed according to the annotation information in the reference genomes. The promoter regions were predicted as 1kb upstream of the transcription start site of each annotated gene. LLT: large latency transcript; LAT: latency-associated transcript; VLT: VZV latency-associated transcript [26]. \*: The annotated 5' UTR of the genes from the Human alphaherpesvirus 3 were analyzed for the PQSs.

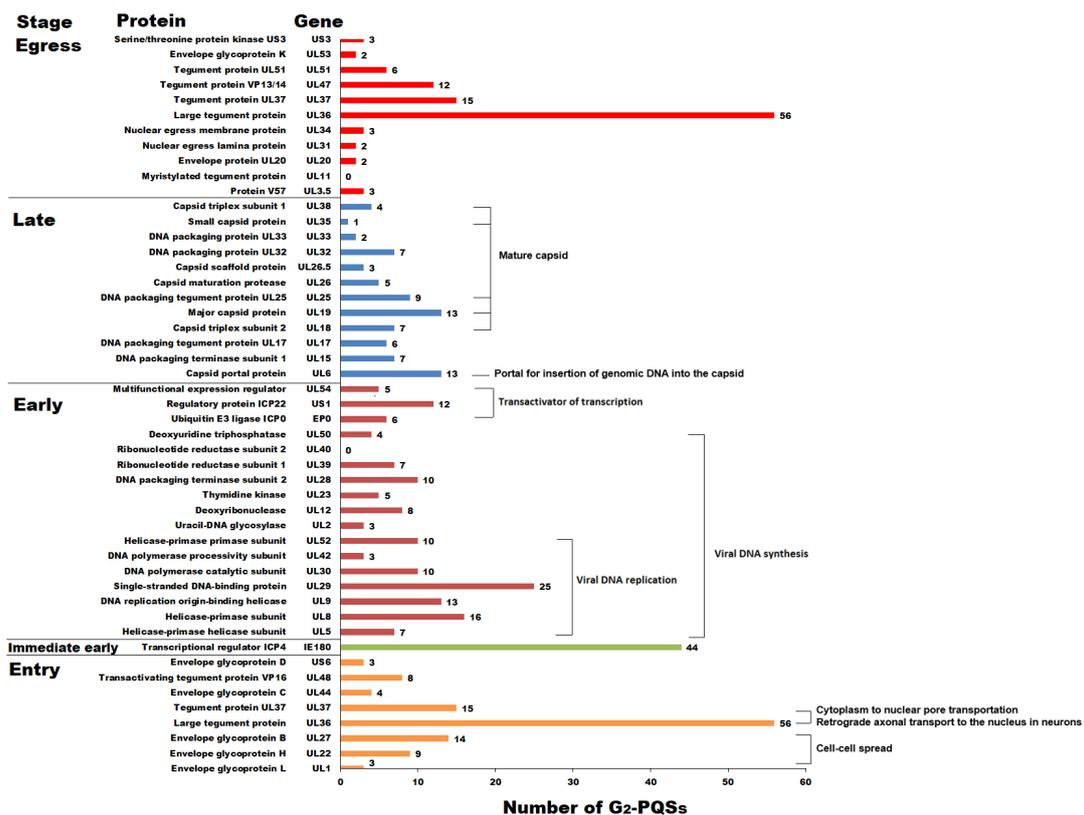


**Figure 2.** Distribution of  $G_2$ -putative quadruplex sequences ( $G_2$ -PQSs) in the double-strand genome of pseudorabies virus (PRV). One red circle indicates one  $G_2$ -PQS. The red bars indicate the coding sequence (CDS) region. The blue bars indicate the untranslated region.

### 2.1.2. Conserved G<sub>2</sub>-PQSs in the Coding Sequences of PRV Genes

#### G<sub>2</sub>-PQS Distribution in the Coding Regions of Genes Involved in the Replication Cycle of PRV

The G<sub>2</sub>-quadruplex formation in the open reading frame (ORF) of Epstein–Barr virus-encoded nuclear antigen 1 (EBNA1) led to decreased mRNA translation in the Epstein–Barr virus (EBV), which suggested that the G-quadruplex in viral transcripts acts as a specific regulatory element to regulate translation level and immune evasion [7]. The PRV replication cycle contains five main processes, including entry, immediate early stage, early stage, late stage, and egress [10]. There were 481 G<sub>2</sub>-PQSs in the ORFs of the genes in above five stages (Figure 3). 112 G<sub>2</sub>-PQSs were found in the entry stage, and 30 of them were involved in important envelope glycoproteins recognizing host cells. After the entry stage, PRV is in the immediate early stage. It transcribes and expresses only one immediate early gene, *IE180*; this was different from human herpesviruses, which transcribe three to five immediate early genes. *IE180* is required for the effective transcription of early viral genes [10]. There were two copies of *IE180* in the PRV genome, and 22 G<sub>2</sub>-PQSs were located in the coding sequence of each copy. The genes in the early stage contained 144 G<sub>2</sub>-PQSs, which was twice as many as the total number of G<sub>2</sub>-PQSs in the late stage ( $n = 77$ ). The early-stage proteins had functions in transactivation and viral DNA synthesis, while the late-stage proteins were mainly responsible for DNA packaging and capsid maturation. A total of 104 G<sub>2</sub>-PQSs were identified in the egress stage, and most of them were present in the tegument proteins, which are important for virion formation before cell-to-cell movement (Figure 3).



**Figure 3.** Number of G<sub>2</sub>-PQSs in the coding sequences of the genes that are involved in the PRV replication cycle.

#### G<sub>2</sub>-PQSs Involved in the Processes of DNA Replication and Encapsidation

Eighty-four G<sub>2</sub>-PQSs were identified in the genes involved in the viral DNA replication process (Figure 3), including *UL5*, *UL8*, *UL9*, *UL29*, *UL30*, *UL42*, and *UL52*. In total, 39% of these G<sub>2</sub>-PQSs were

in the coding region of the helicase–primerase complex formed by *UL5*, *UL8*, and *UL52*. Twenty-five G<sub>2</sub>-PQSs were found in the gene *UL29*, encoding the single-stranded DNA-binding protein ICP8. The DNA replication origin-binding helicase *UL9* contained 13 G<sub>2</sub>-PQSs in its coding sequence, and the DNA polymerase complex *UL30/UL42* contained 13 G<sub>2</sub>-PQSs in total.

In the late stage, *UL38*, *UL35*, *UL25*, *UL19*, *UL18*, and *UL6* encoded the proteins of the mature capsid constituents; these are required for the capsid assembly in the nucleus of the cell, and two scaffolding proteins (*UL26* and *UL26.5*) have been reported to have participated in capsid formation [10]. The above late proteins accounted for 61% of the G<sub>2</sub>-PQSs in the late stage (Figure 3).

#### G<sub>2</sub>-PQSs Involved in the Processes of Virus Entry and Egress

Both *UL36* and *UL37* were the tegument protein genes required in both the entry and egress stages [10]. The VP1/2 protein (*UL36*) is the important component of the inner layer of tegument proteins, and it was reported to be associated with the capsid during PRV transport across cytoplasm into the nuclear pore [27]. The production of VP1/2 protein is cut off by truncating the translation of *UL36* gene, resulting in the failed transportation of the virus particle. Fifty-six G<sub>2</sub>-PQSs were observed in the CDS region of *UL36*, accounting for 50% of the total number of G<sub>2</sub>-PQSs in the stages it involved (Figure 3). The viral glycoproteins gC (*UL44*), gD (*UL6*), gB (*UL27*), gH (*UL22*), and gL (*UL1*) mediated a cascade required by PRV virions to enter specific cells, and totally there were 29 G<sub>2</sub>-PQSs in the ORFs of glycoproteins (Figure 3).

#### Conservation and Potential Function of G<sub>2</sub>-PQSs in CDS Region in the Varicellovirus Genus

Since coding sequences had higher conservation than regulatory regions among different viruses in the same genus, the inter-species analysis of the conservation of the G<sub>2</sub>-PQSs in the coding region was performed within 11 *Varicellovirus* species including PRV. The conservation of the 494 G<sub>2</sub>-PQSs in the CDS region was analyzed. The analysis indicated that 55.3% of G<sub>2</sub>-PQSs exhibited the conservation score higher than 0.2 (Figure S1). Twelve of these 494 G<sub>2</sub>-PQSs were found with the conservation score higher than 0.7 (Table 2). Five G<sub>2</sub>-PQSs of them were derived from the immediate early gene *IE180*, and two G<sub>2</sub>-PQSs were derived from the gene *UL30* encoding the DNA polymerase catalytic subunit. One G<sub>2</sub>-PQS derived from the gene *UL13* encoding protein-serine/threonine kinase was 100% conserved in the *Varicellovirus* genus (Table 2). The other G<sub>2</sub>-PQSs with their conservation scored as 1.0 were from the gene *UL33* associated with *UL28* and *UL15* to contribute to DNA cleavage and PRV package. The DNA cleavage and encapsidation related gene *UL17* and major capsid protein gene *UL19* contained one conserved G<sub>2</sub>-PQS in their coding sequences. Two tegument protein genes *UL16* and *UL47* also contained one conserved G<sub>2</sub>-PQS, respectively (Table 2).

**Table 2.** G<sub>2</sub>-PQS with high conservation score in the coding sequence of *Varicellovirus* genes.

Gene Common Name	Function	Putative G <sub>2</sub> -quadruplex Sequence (5'-3')	Score	Nucleotide Identity
<i>IE180</i>	Transactivator, homolog of ICP4	GGGCCGGAACTGGACCGGG	1.000	0.579
		GGACTGGCCCCGCGGACGGCCCCGCGTGGGGG	0.900	0.579
		GGCTCGGCGCGGCGCGGCGCCCG	0.800	0.579
		GGCCAACGTGGCCCGCGGCCCGG	0.700	0.579
		GGGCCCCGGTCCCCGGGCGGCTCCGGGCCCCCGG	0.700	0.579
<i>UL30</i>	DNA replication	GGACGACGGCGGCGGCTACCAGGGCGCCAAGG	0.727	0.678
		GGTGTACGGGTTACGGGCGTGCCCAACGGG	0.727	0.678
<i>UL33</i>	DNA cleavage and packaging	GGGGAGGCGCTGCGGGCGCGG	1.000	0.664
<i>UL17</i>	DNA cleavage and encapsidation	GGGGCGGCGGCGGCGGGCCCCCGG	0.778	0.589
<i>UL16</i>	Tegument protein	GGTCTGGCCCCCGGCGCGTGGTGGGCGCGCGG	0.800	0.580
<i>UL47</i>	Tegument protein	GGGGACGAGGAGGAGGAGGAGGAGGAGGAGGAGG AGAGCGAGGGGGCGCGTGGTCCGACGGGG	0.700	0.486
<i>UL13</i>	Protein-serine/threonine kinase	GGCCGTCGGGGCCGGATCGTACGG	1.000	0.525

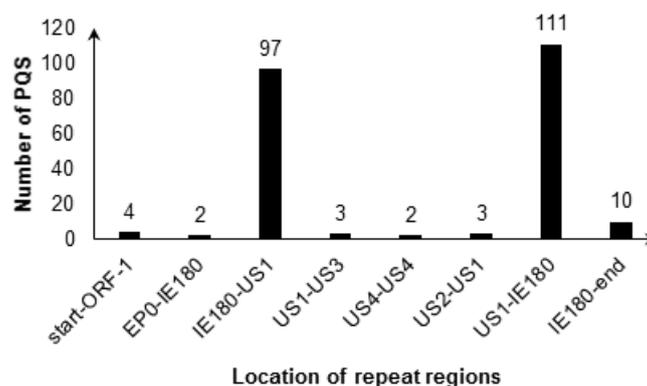
### 2.1.3. Distribution Analysis of G<sub>2</sub>-PQSs in Regulatory Regions in PRV Genomes

Promoter, untranslated region, and intergenic region, which were important regulatory regions, had multiple functions in gene regulation. G<sub>3</sub>-PQSs were reported to be mainly located in the regulatory regions in the genomes of human herpesviruses [5], while less report on G<sub>2</sub>-PQSs was available. Systematic analysis of G<sub>2</sub>-PQS could be conducive to the exploration of their regulatory function. In this study, PRV was found to have higher density of G<sub>2</sub>-PQS than G<sub>3</sub>-PQS in the repeat regions (Table 1).

#### Dozens of Conserved G<sub>2</sub>-PQSs in the Repeat Regions Related to Genome Recombination

Terminal repeat (TR) region is important for genome replication in some herpesviruses. The G-quadruplex in the TR region of gammaherpesvirus Kaposi sarcoma-associated herpesvirus (KSHV) altered the latent DNA replication and episomal persistence [28]. Furthermore, the stabilization of HSV-1 G-quadruplexes in the repeat region inhibited DNA polymerase processing and viral DNA replication [25]. The PRV genome was similar to the HSV-1 genome which was characterized by two unique regions (UL and US), and the US region was flanked by the internal and terminal repeat sequences (IRS and TRS, respectively). During PRV infection, the recombination between the inverted repeats produced two possible isomers of the genome with the U<sub>S</sub> region in opposite orientation. Both isomers were infectious and were in equimolar amount after infection, and the PRV genome was circularized upon entry into the host nucleus through blunt end ligation independent of any viral protein synthesis (reviewed in [10]).

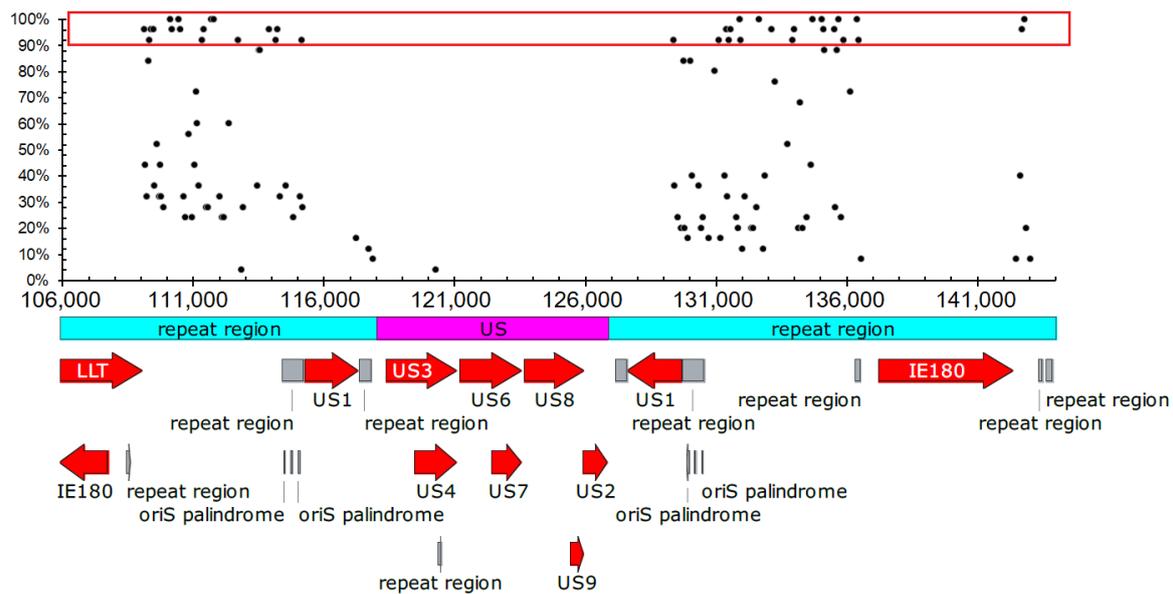
4 G<sub>2</sub>-PQSs were located in 0–656 bp region, and 208 G<sub>2</sub>-PQSs were located in the two regions between genes *IE180* and *US1* (Figure 4). The conservation percentage of 117 G<sub>2</sub>-PQSs on the sense strand of PRV genome ranged from 4% to 100% (Figure 5). There were 38 G<sub>2</sub>-PQSs in the repeated region between *IE180* and *US1* with conservation percentage more than 90% (Figure 5), including 11 G<sub>2</sub>-PQSs with 100% conservation rate (Table S1). These conserved GC-rich sequences might form G-quadruplexes in the inverted repeat region, mediating the genome recombination after infection. 14 G<sub>2</sub>-PQSs existing between *IE180* and *ORF1* may be switches of PRV replication in circular genome.



**Figure 4.** Distribution of G<sub>2</sub>-PQS in repeat regions in PRV genome.

#### G<sub>2</sub>-PQSs in the Untranslated Regions of PRV Genes

The untranslated regions (UTR) are important regulatory region for gene expression. G-quadruplex-forming sequences in the 5' UTR acting as translational repressor have been reported in several human genes [29,30]. The 3' UTR is related to mRNA stability, alternative splicing, polyadenylation, and localization. G-quadruplexes in the 3' UTR of CaMKIIa (Ca<sup>2+</sup>/calmodulin-dependent protein kinase II) and PDS-95 (post-synaptic density protein 95) mRNAs are responsible for the transport of those mRNA in neurites in vivo [31].



**Figure 5.** Conservation percentage of putative G-quadruplex sequences in the repeat region. Red block indicated the G<sub>2</sub>-PQSs with conservation percentage more than 90%.

Nineteen non-redundant G<sub>2</sub>-PQSs and five G<sub>3</sub>-PQSs were found in the 5' UTR of the annotated PRV genes (Figure S2). Among all the genes in PRV genome, the gene *UL13* contained four G<sub>2</sub>-PQSs in its 212-bp 5' UTR with the highest density of G<sub>2</sub>-PQS in the 5' UTR regions, and the gene *UL28* contained one G<sub>3</sub>-PQS in its 72-bp 5' UTR with the highest density of G<sub>3</sub>-PQS in the 5' UTR (Figure 6). The transactivator gene *US1* contained four G<sub>2</sub>-PQSs in its 5' UTR, accounting for 17% of G<sub>2</sub>-PQS monomers found in 5' UTR region in the whole genome (Figure 7A). *UL13* had three G<sub>2</sub>-PQS monomers in its 5' UTR, accounting for 13% of G<sub>2</sub>-PQS monomers in 5' UTR in the PRV genome (Figure 7A). In the early protein genes, the deoxyribonuclease gene *UL12* and dUTPase gene *UL50* had two G<sub>2</sub>-PQS monomers in their 5' UTR, respectively (Figure 7A). One G<sub>2</sub>-PQS cluster was found in the gene *UL13*, which was the unique G<sub>2</sub>-PQS cluster in the all annotated 5' UTR of PRV genes (File S2).

The G<sub>2</sub>-PQSs in the 3' UTR of annotated PRV genes were more than ten folds of those found in the 5' UTR of the annotated PRV genes. There existed 212 non-redundant G<sub>2</sub>-PQSs and 8 G<sub>3</sub>-PQSs in the 3' UTR of PRV genes (Figure S2). These results suggested that G<sub>2</sub>-quadruplexes might regulate gene expression frequently by effecting 3' UTR secondary structure. The genes *UL51* and *UL29* had the highest density of G<sub>3</sub>-PQSs in their 3' UTR regions (Figure 6).

More than 70 G<sub>2</sub>-PQS monomers were found in the 3' UTR of the genes related to viral genome replication, envelopment, and packaging processes (Figure S2). *UL48* encoded the VP16 protein exerting multiple functions in the PRV, including transactivation in gene regulation and secondary envelopment in viral egress [10]. 21 G<sub>2</sub>-PQS monomers were found in the 3' UTR of gene *UL48*, accounting for 11% of all the G<sub>2</sub>-PQS monomers in the 3' UTR in the PRV genome (Figure 7B). The late stage gene *UL15* encoding DNA packaging terminase subunit 1 had 15 G<sub>2</sub>-PQS monomers in its 3' UTR. The *UL9* encoded the sequence-specific ori-binding protein OBP which formed the ATP-dependent helicase motif. OBP was essential for the replication of the viral DNA [10]. *UL9* had 13 G<sub>2</sub>-PQS monomers in its 3' UTR. The tegument protein coding gene *UL14* and type III membrane protein coding gene *UL24* had the same number of G<sub>2</sub>-PQSs in their 3' UTR (Figure S2). The G<sub>2</sub>-PQS monomers from *UL9*, *UL14* and *UL24* accounted for 21% of the entire number of G<sub>2</sub>-PQS monomers in the 3' UTR (Figure 7B).

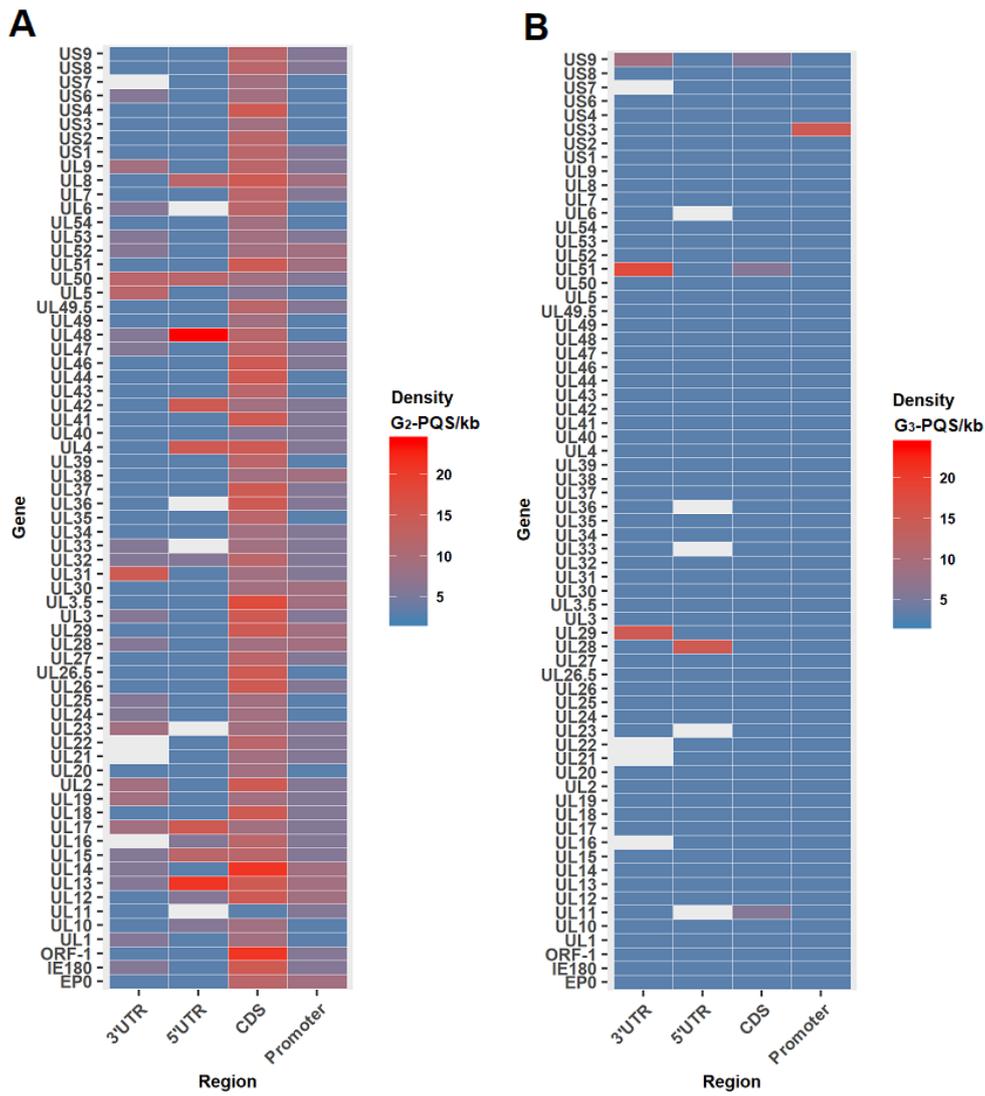


Figure 6. Heat maps of PQS density in coding sequence and regulatory regions of each PRV gene. (A) Heat map of G<sub>2</sub>-PQS; (B) Heat map of G<sub>3</sub>-PQS.

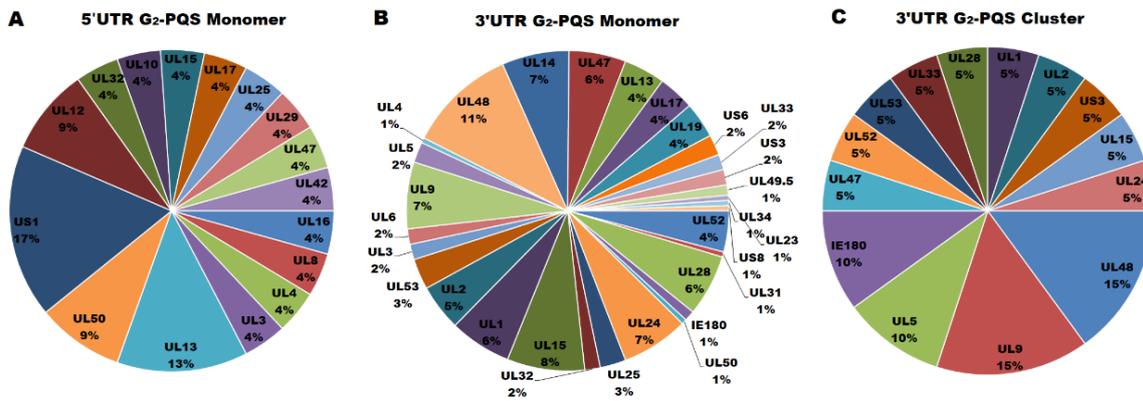
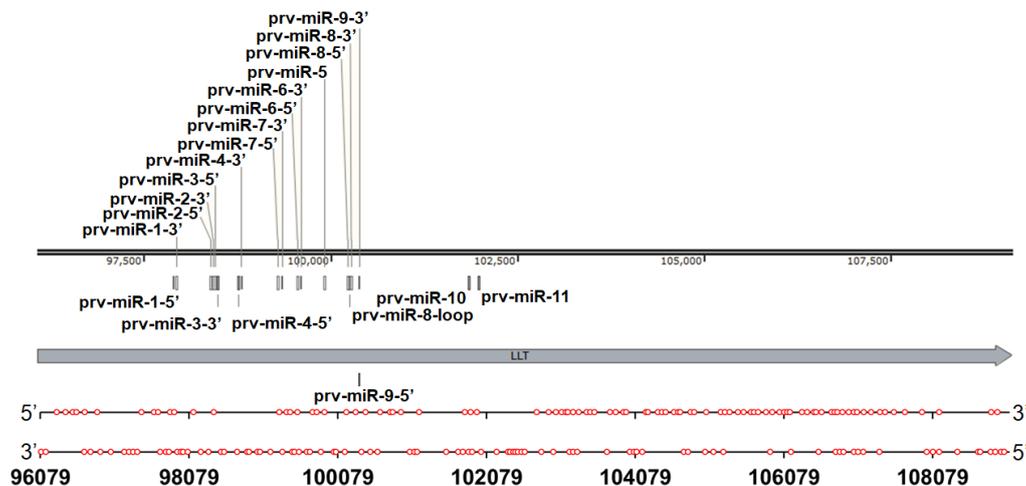


Figure 7. Quantitative distribution of G<sub>2</sub>-PQS in the untranslated regions of annotated genes in PRV genome. (A) G<sub>2</sub>-PQS monomers in the 5' UTR of the PRV genes. (B) G<sub>2</sub>-PQS monomers in the 3' UTR of the PRV genes. (C) G<sub>2</sub>-PQS clusters in the 3' UTR of the PRV genes.

Another important characteristic was that G<sub>2</sub>-PQS clusters existed in the 3' UTR of PRV gene densely. The gene *UL48* had 3 G<sub>2</sub>-PQS clusters in its 3' UTR. The DNA replication related gene group had far more G<sub>2</sub>-PQS clusters than other functional gene groups. The *UL9/UL5/UL52* group had 6 G<sub>2</sub>-PQS clusters in the 3' UTR region (Figure 7C). The proteins coded by above three genes included the sequence-specific ori-binding protein, helicase motif, and primase subunit. The G<sub>2</sub>-PQSs from 3' UTR of *UL5* and *UL9* were validated in further experiments. The identification of the functional G<sub>2</sub>-PQSs in the untranslated regions of genes in transactivation, and viral DNA synthesis processes will provide the elements for post-transcriptional regulation of virus genes.

#### A Wide Distribution of G<sub>2</sub>-PQSs in the Large Latency Transcript

The large latency transcript was the only gene reported to be transcribed during PRV latency [10]. It overlapped with the oppositely transcribed *IE180* gene and *EPO* gene, and it was one of the spliced transcripts in the PRV genome [9]. Sixty-seven G<sub>2</sub>-PQSs and 72 G<sub>2</sub>-PQSs were predicted, respectively, in the positive strand and in the negative strand of the LLT (Figure 8). Forty G<sub>2</sub>-PQSs were found to be 100% conserved among the annotated PRV genomes (Table S2). Thirty G<sub>2</sub>-PQSs in the negative strand were overlapped with the G<sub>2</sub>-PQSs in the *EPO* and *IE180* transcripts, suggesting that these overlapped G<sub>2</sub>-PQSs could be potential dual-functional regulatory elements. Twenty-seven G<sub>3</sub>-PQSs were found in the LLT transcription region (Figure S3), and only one G<sub>3</sub>-PQS was overlapped with *EPO* transcript, while the *IE180* transcript did not have any overlapped G<sub>3</sub>-PQS. The LLT intron had a microRNA cluster that was involved in PRV replication and affecting virulence [32,33]. Fifteen G<sub>3</sub>-PQSs and 35 G<sub>2</sub>-PQSs were found to be located in the intron of LLT. Thus, it could be speculated that these PQSs in the intron might be involved in the transcription regulation of microRNAs during PRV latency.



**Figure 8.** Distribution of G<sub>2</sub>-PQSs in the double-strands of the PRV large latency transcript region. One red circle indicates one G<sub>2</sub>-PQS.

#### Density of the G<sub>2</sub>-PQSs in the Promoter Regions

The G<sub>3</sub>-quadruplexes in the promoter regions of Human immunodeficiency virus 1 (HIV-1) and human herpesviruses were reported to negatively regulate gene expression [5,34–36]. Since the promoters of PRV genes were not annotated in the reference genome, we predicted the promoters to be 1 kb upstream of the transcription start site of each gene. The gene *UL14* coding tegument protein and the gene *UL28*, coding DNA packaging terminase subunit 2, were found to contain nine G<sub>2</sub>-PQSs in their promoters, respectively, which was the largest number among all of the PRV genes. Viral DNA replication-related genes *UL29/UL30/UL8* had eight G<sub>2</sub>-PQSs in each promoter (Figure 6).

#### 2.1.4. Comparison of PQS Distribution among Three Herpesvirus Genomes

Inter-species comparison was performed among PRV, HHV-1 (Herpes simplex virus type 1, HSV-1) and HHV-3 (Varicella-zoster virus, VZV). These three viruses were neurotropic alphaherpesviruses. PRV and HHV-3 belonged to the same genus called *Varicellovirus*. These herpesviruses always caused serious diseases after reactivation from latency. Trigeminal ganglia (TG) was the common latency site for these three herpesviruses, and the dorsal root ganglia (DRG) was another latency site for HHV-3 [10,25,37]. It was difficult to control the switch between latency and reactivation of these herpesviruses. An analysis of the common features and differences in the PQS distribution among these herpesviruses will provide an insight into latency modulation.

G<sub>3</sub>-PQSs with the highest density among three herpesviruses were located in the repeat regions (Table 1) related to genome recombination during infection. Though they were in different genres, PRV and HHV-1 shared similar features in the distribution and high density of G<sub>2</sub>-PQSs, which were different from HHV-3. G<sub>2</sub>-PQSs with the highest density in PRV and HHV-1 were located in the coding regions. The density of G<sub>2</sub>-PQSs in the coding regions of PRV and HHV-1 genes was found to be five-fold as much as that of HHV-3 (Table 1). In the PRV and HHV-1 genomes, the number of the G<sub>2</sub>-PQS monomers in the 3' UTR regions was 8–12 fold of that of the 5' UTR region. The latency-associated transcripts in HHV-1 and PRV were much longer than that in HHV-3. In PRV, the densities of G<sub>2</sub>-PQSs and G<sub>3</sub>-PQSs were 10.66 G<sub>2</sub>-PQS/kb and 2.07 G<sub>3</sub>-PQS/kb in its LLT, respectively. In HHV-1, it was 9.91 G<sub>2</sub>-PQS/kb and 3.39 G<sub>3</sub>-PQS/kb in its LAT, while in HHV-3, no G<sub>3</sub>-PQS was observed in its VLT, and the density of G<sub>2</sub>-PQS was half that of PRV (Table 1). Compared with HHV-1, the PRV had more G<sub>3</sub>-PQS monomers distributed in the repeat region, but less G<sub>3</sub>-PQS monomers in the CDS, UTR, promoter, and LLT regions (Table 1).

#### 2.1.5. Summary of Bioinformatics Analysis

In herpesvirus genomes, G<sub>3</sub>-PQSs were especially abundant in the repeat regions, while G<sub>2</sub>-PQSs were distributed genome-wide. The non-coding region had many more G<sub>2</sub>-PQSs than G<sub>3</sub>-PQSs, and G<sub>2</sub>-PQSs showed a higher density in the CDS region and LLT than in the repeat region. The genes involved in transactivation, genomic DNA replication, and the virus maturation processes had rich and conserved G<sub>2</sub>-PQSs in their mRNA sequences. In the *Varicellovirus* genus, some G<sub>2</sub>-PQSs in the coding sequences of the immediate early protein ICP4, DNA cleavage and packaging protein (*UL33*), and serine/threonine kinase (*UL13*) exhibited conservations of 100%. The highly conserved G<sub>2</sub>-PQSs provide universal target sites to control *Varicellovirus* by disturbing the translation of the above proteins. The density of G<sub>2</sub>-PQS in 3' UTR was higher than that in the 5' UTR, and more G<sub>2</sub>-PQS clusters were found in the 3' UTR of PRV genes than those in the 5' UTR. The 38 G<sub>2</sub>-PQSs in the repeated region between *IE180* and *US1* exhibited conservations of higher than 90% among all the annotated PRV genomes, indicating that the conserved elements might be involved in PRV genome recombination. The LLT overlapped with the oppositely transcribed *IE180* gene and the *EPO* gene. The PQSs in the overlapped regions might be potential dual-functional regulatory elements.

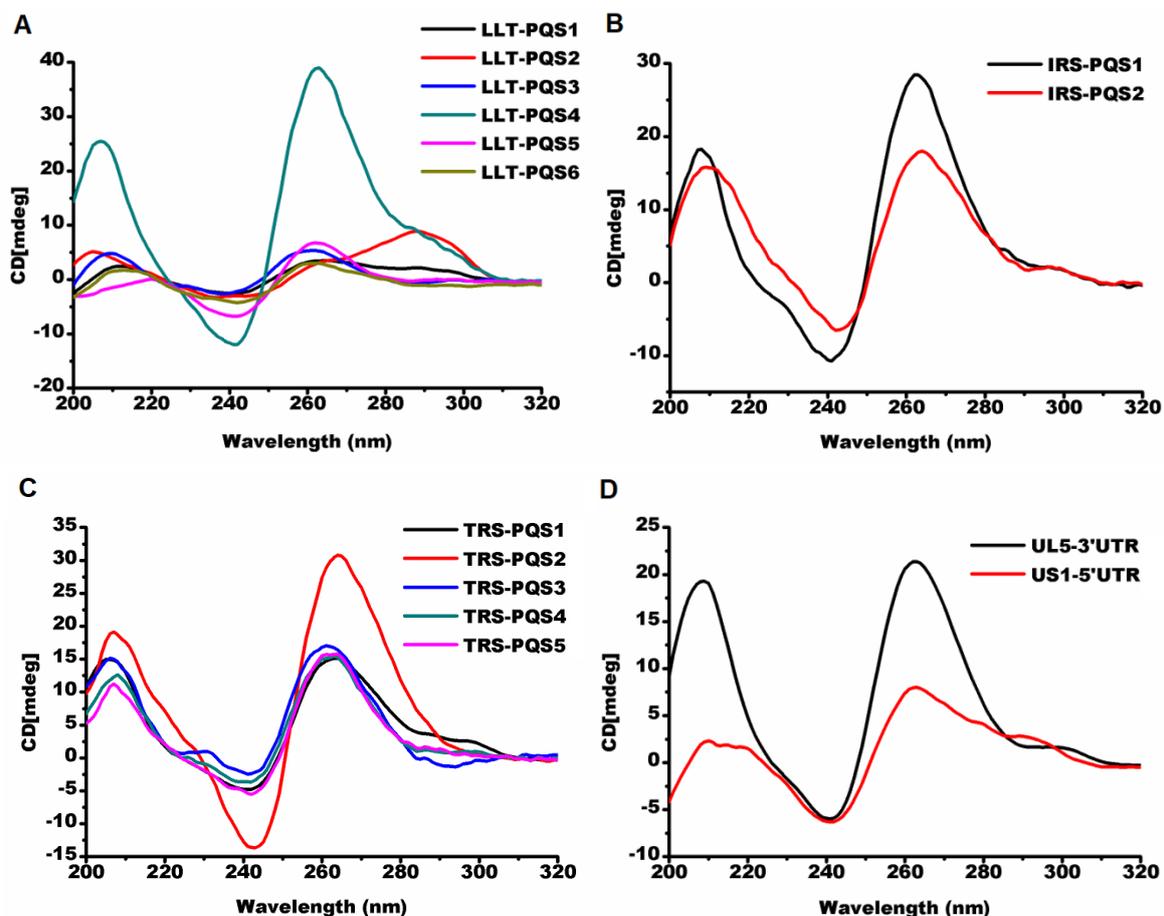
### 2.2. Experimental Validation

The genome-wide analysis of the G<sub>2</sub>-PQS distribution indicated that one-third of the PRV G<sub>2</sub>-PQSs were distributed in the non-coding regions of PRV genome, such as the regions of UTR, IRS, TRS, and LLT (File S1). Further study is needed to determine whether these PQSs could serve as *cis*-regulatory elements in gene expression.

#### 2.2.1. Parallel G-Quadruplexes formed by G<sub>2</sub>-PQSs In Vitro

Circular dichroism (CD) spectroscopy is widely used to distinguish parallel, anti-parallel or hybrid G-quadruplex structures [38]. A parallel G-quadruplex showed a negative peak near 240 nm, and a positive peak near 260 nm. An anti-parallel G-quadruplex exhibited a negative peak near 260 nm

and a positive peak around 290 nm. The hybrid G-quadruplex always displayed one negative peak at 240 nm and two positive peaks at 260 nm and 290 nm. Of the 15 G<sub>2</sub>-PQS oligonucleotides (details in Table S3) selected from the regulatory regions of the PRV genome, 13 G<sub>2</sub>-PQS folded into parallel G-quadruplexes in a buffer containing 100 mM potassium. LLT-PQS1 and LLT-PQS2, between the start of LLT and the Prv-miR-1-5', formed hybrid G-quadruplexes, while other LLT-PQSs between Prv-miR-11-1 and the end of LLT formed parallel G-quadruplexes (Figure 9A). The three G<sub>2</sub>-PQSs were located in the repeat region between *IE180* and *US1*, and the five G<sub>2</sub>-PQSs were located in the repeat region complementary to *US1* CDS, and all of these eight G<sub>2</sub>-PQSs mentioned above formed parallel G-quadruplexes (Figure 9B,C). One G<sub>2</sub>-PQS from the 3' UTR of *UL5*, and another from the 5' UTR of *US1* exhibited typical parallel G-quadruplex peaks (Figure 9D).



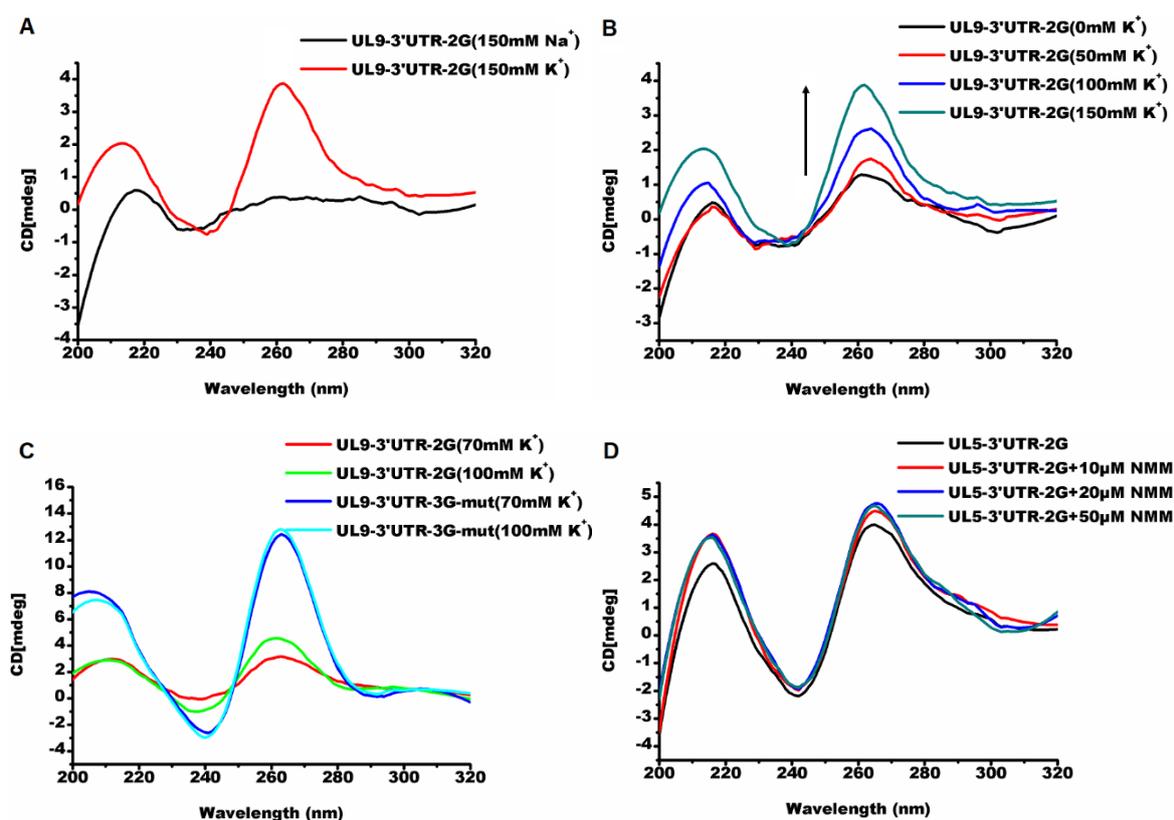
**Figure 9.** Circular dichroism (CD) spectra of G-quadruplex structures formed by selected oligonucleotides. (A) G<sub>2</sub>-PQS from the large latency-associated transcript (LLT) region. The LLT-PQS1 and LLT-PQS2 are located between the start site of LLT and the Prv-miR-1-5'; the LLT-PQS3, LLT-PQS4, LLT-PQS5, and LLT-PQS6 are located between Prv-miR-11-1 and the end of the LLT. (B) G<sub>2</sub>-PQS from the internal repeat sequence (IRS) region. The IRS-PQS1 and IRS-PQS2 are located between *IE180* and *US1*. (C) G<sub>2</sub>-PQS from the terminal repeat sequence (TRS) region. TRS-PQS1, TRS-PQS2, TRS-PQS3, and TRS-PQS4 are located in the sequence complementary to *US1* CDS; TRS-PQS5 is located between *IE180* and *US1*. (D) G<sub>2</sub>-PQS from the untranslated region (UTR). The UL5-3UTR is from the 3' UTR of the gene *UL5*; US1-5UTR is from the 5' UTR of the gene *US1*.

### 2.2.2. G<sub>2</sub>-quadruplexes in the 3' UTR Affects Gene Expression In Vivo with Varying Sensitivities

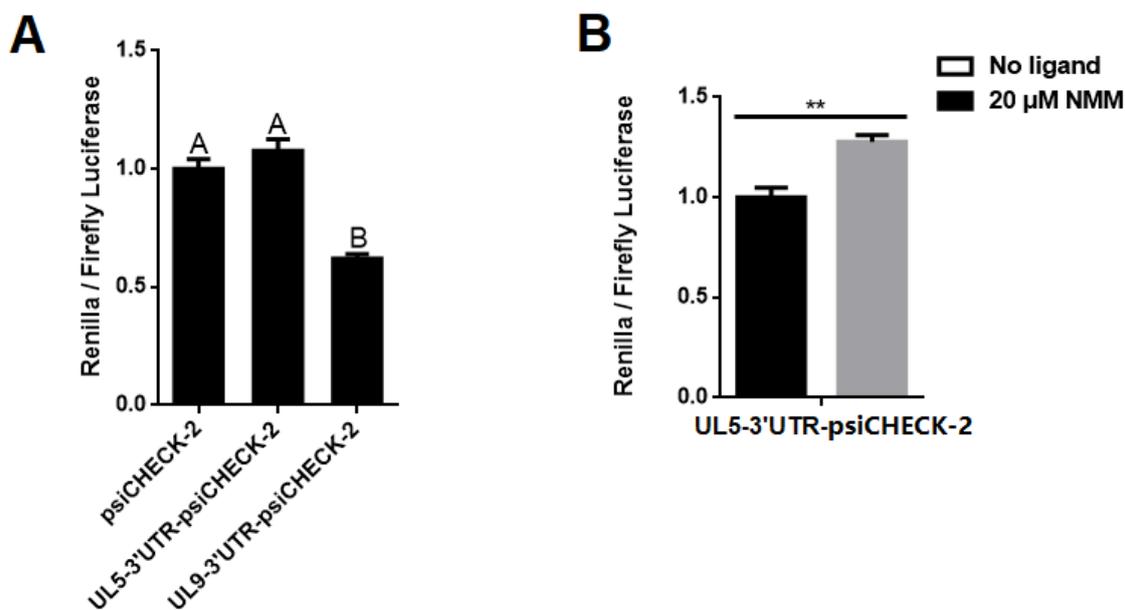
The formation of a G-quadruplex in the 3' UTR was reported to result in gene expression decrease, alternative polyadenylation [39], and retrotransposition [40]. This study found that there were more

G<sub>2</sub>-PQSs in the 3' UTR than in the 5' UTR, in PRV (Table 1). In this study, G<sub>2</sub>-PQSs from the 3' UTR were inserted into the dual luciferase vector, to check their functions.

In the Circular dichroism (CD) experiment, the PQS UL9-3'UTR-2G exhibited no G-quadruplex signal in 100 mM sodium solution, while it folded into a parallel G-quadruplex in the buffer containing 100 mM potassium (Figure 10A). The stability of the G-quadruplex increased with the addition of the potassium cation. (Figure 10B). UL9-3'UTR-2G showed a higher sensitivity to the increase of the potassium than the 3G mutant oligonucleotide obtained by replacing "GG" with "GGG" (Figure 10C). The insertion of PQS UL9-3'UTR-2G into the 3' UTR region of the renilla luciferase gene in the psiCHECK-2 vector resulted in a 38% decrease in relative luciferase activity (Figure 11A). The PQS UL5-3'UTR-2G folded into a parallel G-quadruplex in a buffer with 100 mM K<sup>+</sup> (Figure 10D), and the addition of the G-quadruplex ligand NMM stabilized the structure. The insertion of PQS UL5-3'UTR into the 3' UTR region of renilla luciferase gene in the psiCHECK-2 vector made no difference in the protein expression (Figure 11A). The addition of NMM at 20 μM increased the relative luciferase activity ( $p < 0.01$ ) (Figure 11B, Figures S5). The addition of other typical G-quadruplex ligands such as PDS and BRACO-19 made no significant difference in structural stability (Figures S5, S6A, S6B) and relative luciferase activity (Figures S5, S7A, S7B). The above results suggested that the G-quadruplex formed from G<sub>2</sub>-PQSs in 3' UTR affected the protein expression, and that G<sub>2</sub>-PQSs were more sensitive to different cations and ligands than G<sub>3</sub>-PQSs. These features made the G<sub>2</sub>-quadruplex suitable as sensitive switches for gene expression regulation in response to different environmental factors, such as various proteins, small molecules, and signal cations.



**Figure 10.** G<sub>2</sub>-PQSs from UL5-3'UTR and UL9-3'UTR are sensitive to cations and ligands. (A) CD spectroscopy of the G<sub>2</sub>-PQS (UL9-3'UTR-2G) in the presence of 150 mM NaCl or KCl. (B) CD spectroscopy of the G<sub>2</sub>-PQS (UL9-3'UTR-2G) with increasing KCl concentration (0–150 mM). (C) CD spectroscopy of G<sub>2</sub>-PQS (UL9-3'UTR-2G) and the G<sub>3</sub>-PQS mutant (UL9-3'UTR-3G-mut) with increasing KCl concentration (70–100 mM). (D) CD spectroscopy of the G<sub>2</sub>-PQS (UL5-3'UTR-2G) with increasing NMM concentration (10–50 μM).



**Figure 11.** G<sub>2</sub>-PQSs from UL5-3'UTR- and UL9-3'UTR-regulated gene expression. (A) Dual-luciferase assay of G<sub>2</sub>-PQS from *UL5* and *UL9* in gene expression regulation. Different uppercase letters indicate the significant differences among the different constructs ( $p < 0.05$ , Tukey's honestly significant difference (HSD) test) (B) Dual-luciferase assay of G<sub>2</sub>-PQS from *UL5* in gene expression regulation with the ligand NMM. \*\*  $p < 0.01$ , Student *t*-test.

### 2.2.3. The G-Quadruplex Ligand Decreases the Virulence of PRV

G-quadruplex ligands were reported to interfere with biological processes related to tumor growth, by binding, stabilizing, converting, or unwinding G-quadruplex structures [41–43]. NMM at a series of concentrations (150 nM, 100 nM, 50 nM) was employed to treat the PRV-infected PK15 cells for 24 h. The plaque assay indicated that the NMM had the potential for reducing the virus titer at 24 h post-treatment (Figure S8).

## 3. Discussion

A high density of G<sub>2</sub>-PQSs was distributed widely in the genomes of PRV, HHV-1, and HHV-3, and they could provide sensitive regulatory switches in response to environmental factors in gene expression regulation. Compared to the G<sub>3</sub>-quadruplexes, the G<sub>2</sub>-quadruplexes exhibited less thermal stability and higher sensitivity to loop size and compositions, and they presented multiple conformations in different solutions [44]. As they were sensitive to the microenvironment in the cells, the G<sub>2</sub>-quadruplexes could act as the receptors of specific proteins or metabolites in the cells, to identify the cells that are suitable for PRV latency. In this study, PQS UL9-3'UTR-2G exhibited a parallel structure in potassium solution, while it could not fold into a G-quadruplex in a sodium cation solution (Figure 9A). The insertion of PQS UL9-3'UTR-2G into the psiCHECK-2 vector led to the decrease in reporter gene expression, while the insertion of PQS UL5-3'UTR-2G resulted in an increase in gene expression only with the addition of the ligand NMM (Figure 11, Figure S4). The G-quadruplex structural sensitivity enabled the virus to be easily mediated by physiological cations, different small molecules, or proteins. These findings are conducive to revealing the latency–reactivation mechanism of PRV in specific tissues under certain conditions.

Highly conserved G<sub>2</sub>-PQSs were discovered in the coding regions of genes related to virus replication and maturation, in the *Varicellovirus* genus. The formation of the RNA G<sub>2</sub>-quadruplex in the open reading frame of EBNA1 of EBV resulted in a downregulation of the expression level of the maintenance protein, facilitating virus escape from immune recognition [7]. The conserved G<sub>2</sub>-PQSs in the *Varicellovirus* genus were located in the unique immediate early protein (*IE180*), viral DNA

replication protein (*UL30*), DNA cleavage and package proteins (*UL33/UL17*), and tegument proteins (*UL16/UL47*). *IE180* had five conserved G<sub>2</sub>-PQSs among all the *Varicellovirus* (Table 2). *IE180* was found to be a unique immediate early gene of PRV and the first viral gene transcribed during PRV infection. *IE180* is reported to mediate latency and reactivation by regulating early genes, and it activates *US4*, *UL12* (alkaline exonuclease), *UL22* (type I membrane protein), *UL23* (thymidine kinase), and *UL41* (RNase) [10]. Based on previous reports, we speculated that the formation of the G-quadruplex in the CDS region of *IE180* would result in a reduction of immediate early protein levels, disturbing downstream gene expression in the life-cycle of PRV. Therefore, it could be further inferred that the G<sub>2</sub>-quadruplex might have a significant regulatory effect on the *Varicellovirus* genus, especially for the immediate early protein ICP4, the homolog of immediate early protein from PRV.

A high density of G<sub>2</sub>-PQSs in the untranslated regions of the PRV genes could provide cis-elements to regulate the post-transcription and translation processes. Though the G<sub>3</sub>-quadruplexes in 3' UTR and 5' UTR were widely reported to have regulated human gene translation [5,29–31,45], few reports on the G-quadruplexes in the untranslated regions of human herpesvirus genes are available. In this study, many G<sub>2</sub>-PQSs were found in the untranslated regions of herpesvirus genes (Figure 7). Various G-quadruplexes formed by those G<sub>2</sub>-PQSs regulated gene expression diversely, either under physiological conditions or with a stabilizing ligand (Figure 11). The data suggested that the formation of these G<sub>2</sub>-quadruplexes might result in the truncated 3' UTR, and in turn, disturb 3'-end polyadenylation, finally leading to unstable virus mRNA. The G-quadruplex in the 3' UTR was also reported to have played a role in regulating microRNA binding in humans [46,47], suggesting that the G-quadruplex might be involved in the interaction between virus/host microRNAs and virus genes.

A large number of G<sub>2</sub>-PQSs and G<sub>3</sub>-PQSs in the repeat regions of the herpesvirus genomes might play an important role in genomic integration or recombination during herpesvirus latency. Some conserved G<sub>3</sub>-PQSs in herpesviruses were reported to have played important role in virus integration, latent DNA replication, and episomal persistence. Several herpesviruses, like Marek's Disease virus (MDV) [48], gallid herpesvirus 2 (GaHV-2) [49], and human herpesvirus 6 (HHV6) [50], were found to establish latent infections, with viral genomes integrated into telomere repeat tracts in host chromosomes through homologous recombination. In this study, the pseudorabies virus had 205 G<sub>3</sub>-PQSs in the genome, with 43.9% located in the repeat region (File S1). A high density of conserved G<sub>2</sub>-PQSs existed in the repeat regions between the two diverging transcripts *IE180* and *US1*, which were close to the origin of replication (OriS) (Figure 5). Telomeres were reported to form G-quadruplex clusters with a repeated sequence of (GGGTTA)<sub>n</sub> [51]. This study found that the PRV genome had imperfect telomeric repeats, with the unit sequence 5'-GGGGTGGAGACGGTGGAGGGAGAGGGGAGTGGG-3' repeated 12 times. Thus, these G<sub>2</sub>-PQSs were speculated to be related to genomic recombination and integration during latency.

## 4. Methods and Materials

### 4.1. Virus Sequences

The complete genome sequence of the Suid herpesvirus 1 (Pseudorabies virus, PRV) (NC\_006151.1) was retrieved from NCBI Genome database [24]. It was annotated into 12 features, including 69 CDS regions, 70 genes, four introns, 119 misc. feature regions, one misc. RNA, 20 polyA sites, 14 protein binding sites, 117 regulatory region, 28 repeat regions, seven stem loop regions, one sequence tagged sites (STS), and one variation. The UTR was determined through alignment between the CDS sequence and the gene sequence. The sequences included in the gene sequences and located upstream of the CDS sequence were designated as 5' UTRs, and the sequences downstream of the CDS were defined as 3' UTRs. The putative G-quadruplex sequences (PQS) were searched in the standard genome sequence of PRV. The number of PQSs in the CDS, 3' UTR, 5' UTR and repeat regions was counted, respectively.

The standard genome sequences of the *Varicellovirus* genus retrieved from the NCBI Genome database [24] for conservation analysis were Human herpesvirus 3 (NC\_001348.1), Equid herpesvirus 1 (NC\_001491.2), Equid herpesvirus 3 strain AR/2007/C3A (NC\_024771.1), Equid herpesvirus 4 (NC\_001844.1), Equid herpesvirus 8 (NC\_017826.1), Equid herpesvirus 9 (NC\_011644.1), Bovine herpesvirus 1 (NC\_001847.1), Bovine herpesvirus 5 (NC\_005261.2), Suid herpesvirus 1 (NC\_006151.1), Cercopithecine herpesvirus 9 (NC\_002686.2), and Felid herpesvirus 1 (NC\_013590.2).

#### 4.2. Identification of Putative G-Quadruplex Sequences in the PRV Genome

The putative G<sub>3</sub>-quadruplex sequences were searched by Quadparser software [3]. The formula of the putative G<sub>2</sub>-quadruplex sequence was defined as G<sub>2+</sub>N<sub>1-7</sub>G<sub>2+</sub>N<sub>1-7</sub>G<sub>2+</sub>N<sub>1-7</sub>G<sub>2+</sub>, where G is guanine, and N is any nucleotide including G. With the same software, the putative I-motif sequences were searched with the formula C<sub>2+</sub>N<sub>1-7</sub>C<sub>2+</sub>N<sub>1-7</sub>C<sub>2+</sub>N<sub>1-7</sub>C<sub>2+</sub>, where C is cytosine, and N is any nucleotide including C. The Quadparser coded the sequence in the format *x:y:z*, where *x* stands for the number of guanine tracts or C-runs, *y* stands for the number of locations of putative G<sub>4</sub> or I-motif formations, and *z* stands for the number of possible simultaneous G<sub>4</sub>- or I-motif structures.

The number of I-motif sequences in the sense strand was the same as the number of G<sub>2</sub>-PQSs in its complementary strand. The sum of the G<sub>2</sub>-PQSs and I-motif sequences in the sense strand was the total number of G<sub>2</sub>-PQSs in the double-stranded genome of PRV. The density of G<sub>2</sub>-PQS in the PRV standard genome was calculated by dividing the total number of G<sub>2</sub>-PQSs in the double-stranded genome of PRV by the genome size. If *z* = 1, the G<sub>2</sub>-PQS was counted as a PQS monomer. If *z* ≥ 2, the G<sub>2</sub>-PQS was counted as a PQS cluster.

#### 4.3. Comparison of the Distribution of G<sub>2</sub>-PQS and G<sub>3</sub>-PQS between PRV and Two Herpesviruses

The G<sub>3</sub>-PQS preferred to be located in the repeat regions in human herpesvirus genomes and some repeated G<sub>3</sub>-PQS clusters among the analyzed genomes were reported to be conserved [5]. In order to determine the distribution features of PQSs in PRV genome, we compared PRV, Human herpesvirus 1 (HHV-1), also known as Herpes simplex virus type 1 (HSV-1), and Human herpesvirus 3 (HHV-3), also known as Varicella-zoster virus (VZV). The reference genome sequences of PRV (NC\_006151.1), HHV-1 (NC\_001806.2) and HHV-3 (NC\_001348.1) were downloaded from NCBI [24] and saved as FASTA files. The genome features of viruses above were analyzed with the software BEDTools (<https://github.com/ark5x/bedtools2>) [52] and the length of CDS region, 3' UTR region, 5' UTR region, repeat region, and latency associate transcript was recorded. The promoter region of each gene was predicted as 1kb upstream of the transcription start site of each gene. The PQSs in CDS, latency associate transcript, and repeat regions were counted in both positive strand and negative strand, and the PQSs in the untranslated regions and promoters were counted in terms of the transcription direction of each PRV gene. The number of G<sub>3</sub>-PQS monomer, G<sub>3</sub>-PQS cluster, G<sub>2</sub>-PQS monomer and G<sub>2</sub>-PQS cluster was counted with the Quadparser software. Quadparser was modified as described in last section. The density of PQSs was calculated by dividing the total number of PQSs located in each region or gene by the length of the corresponding region or gene.

#### 4.4. Conservation of Putative G-Quadruplex Sequences and I-Motif Sequences in the PRV CDS Region in the *Varicellovirus* Genus

The reference genome sequences of 11 virus species from the *Varicellovirus* genus were downloaded from NCBI. All the PQS and I-motif sequences in the genome of viruses from *Varicellovirus* genus were searched with Quadparser software and output into file *g4\_cds.txt* (File S3). The protein sequences in each genome were listed in file *common\_protein.txt* (File S4). Following data preparation, CDS and amino acid sequences of all proteins in the eleven virus species were used for multiple sequence alignment with MAFFT software (<https://www.ebi.ac.uk/Tools/msa/mafft/>). The identity of the above nucleotide sequences and amino acid sequences of each protein was calculated with infoalign program of the EMBOSS package (<http://emboss.sourceforge.net/apps/release/6>).

[6/emboss/apps/infoalign.html](#)). The PQS sequences (GG\*\*GG\*\*GG\*\*GG) and I-motif sequences (CC\*\*CC\*\*CC\*\*CC) from the multiple alignment result of nucleotide sequences were identified and counted, and then the ratio of the common PQSs and I-motif sequences to the nucleotide sequences of all the proteins was calculated and output as conservation score of each PQS and I-motif sequence. The conservation of PQS in PRV CDS in the *Varicellovirus* genus was analyzed.

#### 4.5. Percentage Conservation of Putative G-Quadruplex Sequences in the Repeat Regions and LLTs of the PRV Genomes

Twenty-five PRV complete genome sequences from different strains were downloaded from the NCBI Genome database [24]. These sequences were Suid herpesvirus 1 (NC\_006151.1, KU056477.1, BK001744.1), Suid alphaherpesvirus 1 isolate Ea (Hubei) (KX423960.1), Suid alphaherpesvirus 1 isolate LA (KU552118.1), Suid herpesvirus 1 strain NIA3 (KU900059.1), Suid herpesvirus 1 isolate DL14/08 (KU360259.1), Suid herpesvirus 1 strain Ea (KU315430.1), Suid herpesvirus 1 strain ADV32751/Italy2014 (KU198433.1), Suid herpesvirus 1 strain Kolchis (KT983811.1), Suid herpesvirus 1 isolate HLJ8 (KT824771.1), Suid herpesvirus 1 strain HN1201 (KP722022.1), Suid herpesvirus 1 strain HNB (KM189914.3), Suid herpesvirus 1 strain Fa (KM189913.1), Suid herpesvirus 1 strain Kaplan (KJ717942.1, JF797218.1, JQ809328.1), Suid herpesvirus 1 strain HNX (KM189912.1), Suid herpesvirus 1 isolate SC (KT809429.1), Suid herpesvirus 1 strain JS-2012 (KP257591.1), Suid herpesvirus 1 strain HeN1 (KP098534.1), Suid herpesvirus 1 isolate ZJ01 (KM061380.1), Suid herpesvirus 1 strain TJ (KJ789182.1), Suid herpesvirus 1 strain Becker (JF797219.1), and Suid herpesvirus 1 strain Bartha (JF797217.1). All of the putative G-quadruplex sequences located in the repeat regions or LLT of the reference genome sequence (Accession Number: NC\_006151.1) were searched from the other 24 genome sequences, and the percentage conservation was calculated through dividing the number of genome sequences containing the same putative G-quadruplex sequences by the total number of genome sequences.

#### 4.6. Oligonucleotide Folding Conditions

All oligonucleotides purchased from Sangon Biotech, Shanghai, China were salt-free, purified, and dissolved in ddH<sub>2</sub>O to a concentration of 100  $\mu$ M. Oligonucleotide sequences of PQS in regulatory region were selected from the PQSs predicted by Quadparser (Table S3). G<sub>2</sub>-PQSs from the *UL5* 3' UTR, *UL9* 3' UTR, and mutant sequences (Table S4) were folded under the same conditions, as follows. Oligonucleotides were diluted to 10  $\mu$ M in 10 mM phosphate buffer at pH 7.0 supplemented with 100 mM KCl, then they were heated to 95 °C for 5 min in a 1.5 mL Eppendorf tube in water bath, and subsequently slowly cooled for ~8 h to room temperature, then used for spectra or stored at 4 °C.

Under the induced folding conditions for G<sub>2</sub>-PQS from *UL9* 3' UTR, the G<sub>2</sub>-PQS was diluted to 10  $\mu$ M in 10mM Tris-HCl buffer at pH 7.0, supplemented with an increasing concentration of KCl (50 mM, 100 mM, and 150 mM). The G<sub>2</sub>-PQS in 10mM sodium phosphate buffer at pH 7.0 supplemented with 100 mM NaCl was tested. These samples were placed at room temperature for 30 min, and then applied to CD spectra.

Under the induced folding condition for G<sub>2</sub>-PQS from *UL5* 3' UTR, the G<sub>2</sub>-PQS was diluted to 10  $\mu$ M in 10 mM phosphate buffer at pH 7.0, supplemented with 100 mM KCl, and then incubated with different ligands at room temperature for 30 min before CD spectroscopy.

#### 4.7. Circular Dichroism Spectroscopy

CD spectra of the 10  $\mu$ M folded oligonucleotide samples were collected at 25 °C on a JASCO 1500 CD spectrometer by using a quartz cuvette with 1 mm optical path. Data within a 200–320 nm range were collected using two scans at 100 nm/min with 1 s settling time and 1 nm bandwidth. The buffer baseline was recorded with the same parameters, and it was subtracted from the sample spectra before plotting.

#### 4.8. Plasmid Construction

The plasmids UL5-3'UTR-psiCHECK-2 and UL9-3'UTR-psiCHECK-2 were constructed by using specific oligonucleotides (Table S5), with the ClonExpress II One Step Cloning Kit (Vazyme, Nanjing, China) according to the manufacturer's instructions.

#### 4.9. Cell Culture

Cell lines, human embryonic kidney (HEK) 293T, porcine kidney cell (PK-15), and bovine kidney cells (MDBK) were provided by the State Key Laboratory of Agricultural Microbiology, College of Veterinary Medicine, Huazhong Agricultural University in China. The above cell lines were cultured in Dulbecco's modified Eagle medium (DMEM) containing 10% fetal bovine serum (FBS), 0.044 M NaHCO<sub>3</sub>, and 0.025 M HEPES. Cells were grown at 37 °C in a humidified atmosphere with 5% CO<sub>2</sub>.

#### 4.10. Cytotoxicity Assay

The cytotoxicity of the G-quadruplex ligands, *N*-methyl mesoporphyrin IX (NMM) [53] (J&K Scientific, Beijing, China), BRACO-19 [54] (Sigma-Aldrich, Saint Louis, MO, USA), and pyridostatin (PDS) [55] (J&K Scientific, Beijing, China) to HEK293T was determined by the MTT assay, which is dependent on the measurement of mitochondrial dehydrogenase enzyme activity of viable cells. MTT, 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide, is a yellow tetrazole, and it can be reduced to a purple formazan in living cells. The HEK293T cells, at a density of  $1 \times 10^4$  cells per well were seeded into a 96-well microplate. When the cell confluence reached 90%, the growth medium was replaced with 100  $\mu$ L of DMEM containing 2% FBS and G-quadruplex ligands at a series of final concentrations (Figure S5), and the cells were incubated at 37 °C for 24 h. Then, 50  $\mu$ L of MTT solution was added into each well, and the cells were incubated at 37 °C for another 4 h. After incubation, the supernatant was removed and 150  $\mu$ L of dimethyl sulfoxide (DMSO) was added into each well to dissolve the formazan. The relative cell viability was analyzed by measuring the absorbance of formazan at 570 nm on the Synergy<sup>TM</sup> HTX microplate reader (BioTek, Winooski, VT, USA). The cytostatic concentration, which will be applied in the subsequent dual luciferase assays, was required to maintain a cell viability of more than 90%.

#### 4.11. Transfection and Dual Luciferase Assays

The HEK293T cells were seeded in the 24-well plates at the concentration of  $1 \times 10^5$  cells/well. The cells were transfected with 0.8  $\mu$ g of psiCHECK-2 reporter plasmids and Lipofectamine 2000 (Thermo Fisher Scientific, Carlsbad, CA, USA) according to the manufacturer's instructions. NMM at 20  $\mu$ M, BRACO-19 at 10  $\mu$ M, or PDS at 10  $\mu$ M was respectively added to the cells, with the medium being replaced at 6 h after transfection. The activities of firefly and renilla luciferase were measured 24 h after addition of the above G-quadruplex ligands, using the Dual-Luciferase Reporter Assay Kit (Promega, Madison, WI, USA) on a GloMax 20/20 luminometer (Promega, Madison, WI, USA).

#### 4.12. Plaque Assay

The antiviral activity of NMM to PRV Ea strain was examined through plaque assay. The PK15 cells at the density of  $1.2 \times 10^6$  cells per well were seeded in the 6-well plate. When cell confluence reached 80–90%, the cells were placed in 4 °C for 1 h, and then the PRV Ea strain virus solution (MOI = 5) was added into each well. Then, the cells were placed back to 4 °C for 2 h, and the plate with cells was shaken once every 15 min during the incubation. After the incubation at 4 °C, the virus solution was removed and replaced with 2 mL of DMEM supplemented with NMM at different final concentrations (150 nM, 100 nM, 50 nM). Afterwards, the cells were incubated at 37 °C for 24 h. After incubation, the infectious viral particles were isolated from each well and prepared for plaque assay.

MDBK cells at the density of  $2 \times 10^5$  cells per well were seeded in the 12-well plate. When the cell confluence reached 90%, the cells were used for the plaque assay. Exactly 200  $\mu$ L of virus solution was added into each well. Afterwards, the cells were incubated at 37 °C for 2 h. Then, the infected cells were covered by 4% sodium carboxymethylcellulose (CMC-Na) supplemented with 3% FBS and 1% penicillin–streptomycin solution. After incubation at 37 °C for another 48 h, the infected cells were fixed and stained with the crystal violet solution (0.35%, *w/v* in ethanol) at room temperature. After 15 min, the crystal violet solution was removed, and the plate was washed with tap water. Then, the plate was put in the dry oven for a few minutes. Finally, the viral titer was determined by the plaque assay.

#### 4.13. Statistical Analysis

A Student's *t*-test was used to determine the significant differences in luciferase activity between the ligand treatment and control of each construct. One-way analysis of variance (ANOVA) with Tukey's multiple comparison was applied to determining the significant difference among various treatments dual luciferase assay.

## 5. Conclusions

In summary, the systematic analysis of the distribution of G<sub>2</sub>-PQS in the PRV genomes provides a clear guide for elaborated studies of their functions, related to the establishment of latency and its reactivation in herpesviruses. We analyzed the putative G<sub>2</sub>-quadruplex and G<sub>3</sub>-quadruplex sequences in the PRV genome systematically, and then compared it with typical human herpesviruses in the same subfamily, then evaluated the structures and functions of G<sub>2</sub>-quadruplexes. G<sub>2</sub>-quadruplex sequences in the form of both monomers and clusters were found to be distributed in the entire PRV genome, especially in the CDS, LLT, and repeat regions. Extremely conserved G<sub>2</sub>-quadruplex sequences existed in the CDS of the genes related to viral genome replication and maturation processes. G<sub>2</sub>-quadruplex sequences tended to be located in the repeat regions close to the origin of replication site, which may contribute to genome replication and recombination. Most G<sub>2</sub>-quadruplexes from the regulatory regions formed parallel-type G-quadruplex. There were more G<sub>2</sub>-quadruplex sequences in the 3' UTR regions than in the 5' UTR regions. These G<sub>2</sub>-quadruplexes showed different sensitivities to physiological cations and small molecules. Thus, it could be inferred that the G<sub>2</sub>-quadruplex could act as a switch to control the expression of genes involved in virus latency establishment, viral genome replication cascade, and virus cell-to-cell movement. The G-quadruplex ligand, NMM, exhibited the potential for inhibition of the proliferation of PRV in its host cells. These massive and sensitive G<sub>2</sub>-quadruplexes could serve as a class of receptors in response to intracellular environments, guiding herpesviruses to choose specific cells or conditions for latency or reactivation.

**Supplementary Materials:** The following are available online.

**Author Contributions:** Conceptualization, S.L., X.N. and D.W.; methodology, S.L., X.N. and D.W.; software, B.G., Z.L., Z.Y. and X.N.; validation, H.Z. and Y.Z.; formal analysis, S.L., H.D. and B.G.; investigation, H.D., B.G., Z.L., and Z.Y.; resources, D.W.; data curation, Z.L.; writing—original draft preparation, H.D. and B.G.; writing—review and editing, S.L. and D.W.; visualization, S.L., H.D., B.G. and Z.L.; supervision, D.W.; project administration, S.L. and D.W.; funding acquisition, S.L. and D.W.

**Funding:** This research was funded by the National Natural Science Foundation of China awarded to D.W. (grant number 31672558, 21502060 and 21732002) and S.L. (grant number 31701791); Huazhong Agricultural University Scientific & Technological Self-innovation Foundation (grant number 2015RC013, 2662017PY113 and 2662015PY208), and the Open fund of The State Key Laboratory of Bio-organic and Natural Products Chemistry, Chinese Academy of Sciences (grant number SKLBNPC16343) to D.W.; Hubei Provincial Natural Science Foundation of China (grant number 2017CFB233) and Fundamental Research Funds for the Central Universities (grant number 2662015QD046) to S.L.

**Acknowledgments:** The authors thank Shaobo Xiao and Zhengfei Liu in Huazhong Agricultural University for providing equipment for the cell culture and luciferase assay. The authors thank Huazhong Agricultural University and the State Key Laboratory of Agricultural Microbiology for providing access to the JASCO 1500 CD spectrometer.

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

## References

1. Lee, K.; Kolb, A.W.; Sverchkov, Y.; Cuellar, J.A.; Craven, M.; Brandt, C.R. Recombination Analysis of Herpes Simplex Virus 1 Reveals a Bias toward GC Content and the Inverted Repeat Regions. *J. Virol.* **2015**, *89*, 7214–7223. [[CrossRef](#)]
2. Balasubramanian, S.; Neidle, S. G-quadruplex nucleic acids as therapeutic targets. *Curr. Opin. Chem. Biol.* **2009**, *13*, 345–353. [[CrossRef](#)]
3. Huppert, J.L.; Balasubramanian, S. Prevalence of quadruplexes in the human genome. *Nucleic Acids Res.* **2005**, *33*, 2908–2916. [[CrossRef](#)]
4. Andorf, C.M.; Kopylov, M.; Dobbs, D.; Koch, K.E.; Stroupe, M.E.; Lawrence, C.J.; Bass, H.W. G-Quadruplex (G4) Motifs in the Maize (*Zea mays* L.) Genome Are Enriched at Specific Locations in Thousands of Genes Coupled to Energy Status, Hypoxia, Low Sugar, and Nutrient Deprivation. *J. Genet. Genomics* **2014**, *41*, 627–647. [[CrossRef](#)]
5. Biswas, B.; Kandpal, M.; Jauhari, U.K.; Vivekanandan, P. Genome-wide analysis of G-quadruplexes in herpesvirus genomes. *BMC Genomics* **2016**, *17*. [[CrossRef](#)]
6. Lavezzo, E.; Berselli, M.; Frasson, I.; Perrone, R.; Palu, G.; Brazzale, A.R.; Richter, S.N.; Toppo, S. G-quadruplex forming sequences in the genome of all known human viruses: A comprehensive guide. *PLoS Comput. Biol.* **2018**, *14*, e1006675. [[CrossRef](#)]
7. Murat, P.; Zhong, J.; Lekieffre, L.; Cowieson, N.P.; Clancy, J.L.; Preiss, T.; Balasubramanian, S.; Khanna, R.; Tellam, J. G-quadruplexes regulate Epstein-Barr virus-encoded nuclear antigen 1 mRNA translation. *Nat. Chem. Biol.* **2014**, *10*, 358–364. [[CrossRef](#)]
8. Lightfoot, H.L.; Hagen, T.; Clery, A.; Allain, F.H.; Hall, J. Control of the polyamine biosynthesis pathway by G2-quadruplexes. *eLife* **2018**, *7*. [[CrossRef](#)]
9. Adler, B.; Sattler, C.; Adler, H. Herpesviruses and Their Host Cells: A Successful Liaison. *Trends Microbiol.* **2017**, *25*, 229–241. [[CrossRef](#)]
10. Pomeranz, L.E.; Reynolds, A.E.; Hengartner, C.J. Molecular biology of pseudorabies virus: Impact on neurovirology and veterinary medicine. *Microbiol. Mol. Biol. Rev.* **2005**, *69*, 462–500. [[CrossRef](#)]
11. Muller, T.; Hahn, E.C.; Tottewitz, F.; Kramer, M.; Klupp, B.G.; Mettenleiter, T.C.; Freuling, C. Pseudorabies virus in wild swine: A global perspective. *Arch. Virol.* **2011**, *156*, 1691–1705. [[CrossRef](#)]
12. Masot, A.J.; Gil, M.; Risco, D.; Jimenez, O.M.; Nunez, J.I.; Redondo, E. Pseudorabies virus infection (Aujeszky's disease) in an Iberian lynx (*Lynx pardinus*) in Spain: A case report. *BMC Vet. Res.* **2017**, *13*, 6. [[CrossRef](#)]
13. Musante, A.R.; Pedersen, K.; Hall, P. First reports of pseudorabies and winter ticks (*Dermacentor albipictus*) associated with an emerging feral swine (*Sus scrofa*) population in New Hampshire. *J. Wildlife Dis.* **2014**, *50*, 121–124. [[CrossRef](#)]
14. Pedersen, K.; Bevins, S.N.; Baroch, J.A.; Cumbee, J.C., Jr.; Chandler, S.C.; Woodruff, B.S.; Bigelow, T.T.; DeLiberto, T.J. Pseudorabies in feral swine in the United States, 2009–2012. *J. Wildlife Dis.* **2013**, *49*, 709–713. [[CrossRef](#)]
15. An, T.Q.; Peng, J.M.; Tian, Z.J.; Zhao, H.Y.; Li, N.; Liu, Y.M.; Chen, J.Z.; Leng, C.L.; Sun, Y.; Chang, D.; et al. Pseudorabies virus variant in Bartha-K61-vaccinated pigs, China, 2012. *Emerg. Infect. Dis.* **2013**, *19*, 1749–1755. [[CrossRef](#)]
16. Hu, H.; Yu, T.; Fan, J.; Cheng, S.; Liu, C.; Deng, F.; Guo, N.; Wu, B.; He, Q.G. Genetic Properties of the New Pseudorabies Virus Isolates in China in 2012. *Pak. Vet. J.* **2016**, *36*, 264–269.
17. Liu, H.; Li, X.T.; Hu, B.; Deng, X.Y.; Zhang, L.; Lian, S.Z.; Zhang, H.L.; Lv, S.; Xue, X.H.; Lu, R.G.; et al. Outbreak of severe pseudorabies virus infection in pig-offal-fed farmed mink in Liaoning Province, China. *Arch. Virol.* **2017**, *162*, 863–866. [[CrossRef](#)]
18. Wang, X.; Wu, C.X.; Song, X.R.; Chen, H.C.; Liu, Z.F. Comparison of pseudorabies virus China reference strain with emerging variants reveals independent virus evolution within specific geographic regions. *Virology* **2017**, *506*, 92–98. [[CrossRef](#)]
19. Wu, R.; Bai, C.; Sun, J.; Chang, S.; Zhang, X. Emergence of virulent pseudorabies virus infection in northern China. *J. Vet. Sci.* **2013**, *14*, 363–365. [[CrossRef](#)]

20. Yu, T.; Chen, F.; Ku, X.; Fan, J.; Zhu, Y.; Ma, H.; Li, S.; Wu, B.; He, Q. Growth characteristics and complete genomic sequence analysis of a novel pseudorabies virus in China. *Virus Genes* **2016**, *52*, 474–483. [CrossRef]
21. Yu, T.; Chen, F.; Ku, X.; Zhu, Y.; Ma, H.; Li, S.; He, Q. Complete Genome Sequence of Novel Pseudorabies Virus Strain HNB Isolated in China. *Genome. Announc.* **2016**, *4*. [CrossRef]
22. Yu, X.; Zhou, Z.; Hu, D.; Zhang, Q.; Han, T.; Li, X.; Gu, X.; Yuan, L.; Zhang, S.; Wang, B.; et al. Pathogenic pseudorabies virus, China, 2012. *Emerg. Infect. Dis.* **2014**, *20*, 102–104. [CrossRef]
23. Dong, B.; Zarlenga, D.S.; Ren, X. An overview of live attenuated recombinant pseudorabies viruses for use as novel vaccines. *J. Immunol. Res.* **2014**, *2014*, 824630. [CrossRef]
24. The National Center for Biotechnology Information Genome Home Page. Available online: <https://www.ncbi.nlm.nih.gov/genome> (accessed on 13 August 2018).
25. Artusi, S.; Nadai, M.; Perrone, R.; Biasolo, M.A.; Palu, G.; Flamand, L.; Calistri, A.; Richter, S.N. The Herpes Simplex Virus-1 genome contains multiple clusters of repeated G-quadruplex: Implications for the antiviral activity of a G-quadruplex ligand. *Antiviral. Res.* **2015**, *118*, 123–131. [CrossRef]
26. Depledge, D.P.; Ouwendijk, W.J.D.; Sadaoka, T.; Braspenning, S.E.; Mori, Y.; Cohrs, R.J.; Verjans, G.; Breuer, J. A spliced latency-associated VZV transcript maps antisense to the viral transactivator gene 61. *Nat. Commun.* **2018**, *9*, 1167. [CrossRef]
27. Granzow, H.; Klupp, B.G.; Mettenleiter, T.C. Entry of pseudorabies virus: An immunogold-labeling study. *J. Virol.* **2005**, *79*, 3200–3205. [CrossRef]
28. Madireddy, A.; Purushothaman, P.; Loosbrock, C.P.; Robertson, E.S.; Schildkraut, C.L.; Verma, S.C. G-quadruplex-interacting compounds alter latent DNA replication and episomal persistence of KSHV. *Nucleic Acids Res.* **2016**, *44*, 3675–3694. [CrossRef]
29. Bugaut, A.; Balasubramanian, S. 5'-UTR RNA G-quadruplexes: Translation regulation and targeting. *Nucleic Acids Res.* **2012**, *40*, 4727–4741. [CrossRef]
30. Kumari, S.; Bugaut, A.; Huppert, J.L.; Balasubramanian, S. An RNA G-quadruplex in the 5' UTR of the NRAS proto-oncogene modulates translation. *Nat. Chem. Biol.* **2007**, *3*, 218–221. [CrossRef]
31. Subramanian, M.; Rage, F.; Tabet, R.; Flatter, E.; Mandel, J.L.; Moine, H. G-quadruplex RNA structure as a signal for neurite mRNA targeting. *EMBO Rep.* **2011**, *12*, 697–704. [CrossRef]
32. Wang, X.; Zhang, M.M.; Yan, K.; Tang, Q.; Wu, Y.Q.; He, W.B.; Chen, H.C.; Liu, Z.F. The full-length microRNA cluster in the intron of large latency transcript is associated with the virulence of pseudorabies virus. *Virology* **2018**, *520*, 59–66. [CrossRef]
33. Mahjoub, N.; Dhorne-Pollet, S.; Fuchs, W.; Endale Ahanda, M.L.; Lange, E.; Klupp, B.; Arya, A.; Loveland, J.E.; Lefevre, F.; Mettenleiter, T.C.; et al. A 2.5-kilobase deletion containing a cluster of nine microRNAs in the latency-associated-transcript locus of the pseudorabies virus affects the host response of porcine trigeminal ganglia during established latency. *J. Virol.* **2015**, *89*, 428–442. [CrossRef]
34. Amrane, S.; Kerkour, A.; Bedrat, A.; Vialet, B.; Andreola, M.L.; Mergny, J.L. Topology of a DNA G-quadruplex structure formed in the HIV-1 promoter: A potential target for anti-HIV drug development. *J. Am. Chem. Soc.* **2014**, *136*, 5249–5252. [CrossRef]
35. Perrone, R.; Nadai, M.; Frasson, I.; Poe, J.A.; Butovskaya, E.; Smithgall, T.E.; Palumbo, M.; Palu, G.; Richter, S.N. A dynamic G-quadruplex region regulates the HIV-1 long terminal repeat promoter. *J. Med. Chem.* **2013**, *56*, 6521–6530. [CrossRef]
36. Kong, J.N.; Zhang, C.; Zhu, Y.C.; Zhong, K.; Wang, J.; Chu, B.B.; Yang, G.Y. Identification and characterization of G-quadruplex formation within the EP0 promoter of pseudorabies virus. *Sci. Rep.* **2018**, *8*, 14029. [CrossRef]
37. Depledge, D.P.; Sadaoka, T.; Ouwendijk, W.J.D. Molecular aspects of Varicella-Zoster Virus latency. *Viruses* **2018**, *10*. [CrossRef]
38. Kypr, J.; Kejnovska, I.; Renciuik, D.; Vorlickova, M. Circular dichroism and conformational polymorphism of DNA. *Nucleic Acids Res.* **2009**, *37*, 1713–1725. [CrossRef]
39. Beaudoin, J.D.; Perreault, J.P. Exploring mRNA 3'-UTR G-quadruplexes: Evidence of roles in both alternative polyadenylation and mRNA shortening. *Nucleic Acids Res.* **2013**, *41*, 5898–5911. [CrossRef]
40. Sahakyan, A.B.; Murat, P.; Mayer, C.; Balasubramanian, S. G-quadruplex structures within the 3' UTR of LINE-1 elements stimulate retrotransposition. *Nat. Struct. Mol. Biol.* **2017**, *24*, 243–247. [CrossRef]

41. Kim, M.Y.; Gleason-Guzman, M.; Izbicka, E.; Nishioka, D.; Hurley, L.H. The different biological effects of telomestatin and TMPyP4 can be attributed to their selectivity for interaction with intramolecular or intermolecular G-quadruplex structures. *Cancer Res.* **2003**, *63*, 3247–3256.
42. Grand, C.L.; Han, H.; Munoz, R.M.; Weitman, S.; Von Hoff, D.D.; Hurley, L.H.; Bearss, D.J. The cationic porphyrin TMPyP4 down-regulates c-MYC and human telomerase reverse transcriptase expression and inhibits tumor growth in vivo. *Mol. Cancer Ther.* **2002**, *1*, 565–573.
43. Mikami-Terao, Y.; Akiyama, M.; Yuza, Y.; Yanagisawa, T.; Yamada, O.; Yamada, H. Antitumor activity of G-quadruplex-interactive agent TMPyP4 in K562 leukemic cells. *Cancer Lett.* **2008**, *261*, 226–234. [[CrossRef](#)]
44. Varizhuk, A.M.; Protopopova, A.D.; Tsvetkov, V.B.; Barinov, N.A.; Podgorsky, V.V.; Tankevich, M.V.; Vlasenok, M.A.; Severov, V.V.; Smirnov, I.P.; Dubrovin, E.V.; et al. Polymorphism of G4 associates: From stacks to wires via interlocks. *Nucleic Acids Res.* **2018**, *46*, 8978–8992. [[CrossRef](#)]
45. Beaudoin, J.D.; Perreault, J.P. 5'-UTR G-quadruplex structures acting as translational repressors. *Nucleic Acids Res.* **2010**, *38*, 7022–7036. [[CrossRef](#)]
46. Rouleau, S.; Glouzon, J.S.; Brumwell, A.; Bisailon, M.; Perreault, J.P. 3' UTR G-quadruplexes regulate miRNA binding. *RNA* **2017**, *23*, 1172–1179. [[CrossRef](#)]
47. Bolduc, F.; Garant, J.M.; Allard, F.; Perreault, J.P. Irregular G-quadruplexes Found in the Untranslated Regions of Human mRNAs Influence Translation. *J. Biol. Chem.* **2016**, *291*, 21751–21760. [[CrossRef](#)]
48. Greco, A.; Fester, N.; Engel, A.T.; Kaufer, B.B. Role of the short telomeric repeat region in Marek's disease virus replication, genomic integration, and lymphomagenesis. *J. Virol.* **2014**, *88*, 14138–14147. [[CrossRef](#)]
49. Lieberman, P.M. Epigenetics and Genetics of Viral Latency. *Cell Host Microbe* **2016**, *19*, 619–628. [[CrossRef](#)]
50. Gilbert-Girard, S.; Gravel, A.; Artusi, S.; Richter, S.N.; Wallaschek, N.; Kaufer, B.B.; Flamand, L. Stabilization of Telomere G-Quadruplexes Interferes with Human Herpesvirus 6A Chromosomal Integration. *J. Virol.* **2017**, *91*. [[CrossRef](#)]
51. Sun, D.; Thompson, B.; Cathers, B.E.; Salazar, M.; Kerwin, S.M.; Trent, J.O.; Jenkins, T.C.; Neidle, S.; Hurley, L.H. Inhibition of human telomerase by a G-quadruplex-interactive compound. *J. Med. Chem.* **1997**, *40*, 2113–2116. [[CrossRef](#)]
52. Quinlan, A.R.; Hall, I.M. BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* **2010**, *26*, 841–842. [[CrossRef](#)]
53. Nicoludis, J.M.; Miller, S.T.; Jeffrey, P.D.; Barrett, S.P.; Rablen, P.R.; Lawton, T.J.; Yatsunyk, L.A. Optimized end-stacking provides specificity of N-methyl mesoporphyrin IX for human telomeric G-quadruplex DNA. *J. Am. Chem. Soc.* **2012**, *134*, 20446–20456. [[CrossRef](#)]
54. Campbell, N.H.; Parkinson, G.N.; Reszka, A.P.; Neidle, S. Structural basis of DNA quadruplex recognition by an acridine drug. *J. Am. Chem. Soc.* **2008**, *130*, 6722–6724. [[CrossRef](#)]
55. Koirala, D.; Dhakal, S.; Ashbridge, B.; Sannohe, Y.; Rodriguez, R.; Sugiyama, H.; Balasubramanian, S.; Mao, H. A single-molecule platform for investigation of interactions between G-quadruplexes and small-molecule ligands. *Nat. Chem.* **2011**, *3*, 782–787. [[CrossRef](#)]

**Sample Availability:** Samples of the compounds are not available from the authors.



© 2019 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 (<http://creativecommons.org/licenses/by/4.0/>).