



DNA Zip Codes in Herpesvirus Genomes

Jay C Brown*

Abstract

The genomes of 39 herpesviruses were examined for the presence of DNA zip codes, sequence elements found in yeast to be involved in gene localization to the nuclear periphery and transcriptional memory. Tests were carried out with three different zip codes and with viruses in all three subfamilies of the Herpesviridae. All genomes were found to have at least three zip codes with 30 the most observed. Zip codes in many alphaherpesviruses and in human cytomegalovirus were found to be symmetrically arranged within the S genome segment, a region that inverts actively during virus replication. This suggests zip codes may be involved in the segment inversion process. Zip codes located in 16 alphaherpesvirus promoters were examined separately from those in other locations. The genes most likely to have a zip code containing promoter were found to be those encoding glycoprotein B and transcription factor ICP4. Of the two, zip codes associated with ICP4 are of particular interest as their ability to mediate transcriptional memory is well-suited for a role in reactivation of alphaherpesviruses from the latent state.

Keywords

Herpesvirus genomes; DNA zip codes; Segment inversion; Virus latency; Transcriptional memory

Introduction

More than 200 viruses are now classified in the Herpesviridae family. Three subfamilies are recognized, the familiar alpha-, beta- and gammaherpesviruses [1,2] and evolutionary relationships among them have been defined [3,4]. All Herpesviridae infect vertebrate animals with a high degree of specificity for the host species, as observed in other DNA viruses. Humans are the hosts for a total of eight Herpesviridae including herpes simplex virus, varicella-zoster virus, human cytomegalovirus and Epstein-Barr virus. Following an initial productive infection, viruses in the Herpesviridae family are able to enter a latent state that lasts for the lifetime of the host. Reactivation from latency occurs after an appropriate stimulus that can differ depending on the virus species.

Herpes family viruses all have the same basic structure, an icosahedral capsid containing the virus DNA surrounded by a membrane envelope. A layer of protein, the tegument, is found between the capsid and membrane [5,6]. In all Herpesviridae, the genome is a single molecule of dsDNA whose length is in the range of ~110kb-240kb depending on the virus species [1,2]. In the intact virion the DNA is present inside the capsid, but it is not bound to

protein [7]. Herpes genomes encode ~60-169 genes depending on the virus species and genome length. Genomes are said to be "gene dense" because compared to host cell genomes there is little space between the genes and because many herpes genes lack introns. Although the genome of most herpes viruses is a single, unique sequence, the genomes of alpha- and some betaherpesviruses consist of two distinct segments (long and short [L and S]) that invert during virus replication creating four genome isomers [8,9].

Although herpes family viruses differ in the species and cell types they infect, there are clear similarities in their pathway of replication [2]. Infection is initiated by a fusion event between the virus and host cell membranes, a process that can occur at the cell surface or in an endocytic vesicle. Fusion results in deposition of the virus DNA-containing capsid into the peripheral cytoplasm of the host cell. The capsid then traffics to the nucleus where it binds to a nuclear pore and injects its DNA through the pore into the nucleoplasm. There the virus DNA is replicated and its genes are expressed with transcription carried out by the host DNA-dependent RNA polymerase. Progeny capsids are formed in the infected cell nucleus and packaged there with the virus DNA. The filled capsid is then enveloped in the cytoplasm to create the infectious virion that exits the cell to spread the virus infection. Compared to productive virus replication, latent herpesvirus infection differs to a greater extent depending on the virus species [10-12]. In all cases, however, there is little or no virus replication or virus gene expression in latently-infected cells. Both resume normally upon reactivation resulting in an infection that is indistinguishable from productive replication including the release of progeny virus.

DNA zip codes are short sequence elements that have been particularly well characterized in yeast. Four are known and these can affect the location of a gene within the nucleus, the extent of gene expression and the rate at which a repressed gene is reactivated [13-16]. Two zip codes have been particularly well-characterized, a gene recruitment sequence (GRS1; GGGTTGGA) and the memory recruitment sequence (MRS; TCCTTCTTCC). Both are most effective when they are located in a gene promoter. Zip codes are able to relocate an active gene from the nucleoplasm to a nuclear pore complex or to nuclear pore proteins in the nucleoplasm. MRS sequences have the additional function of promoting transcriptional memory of genes that have been activated and later repressed [15]. Zip codes are found to be present in yeast species separated evolutionarily by 10⁹ years or more suggesting they are ancient and possibly widespread among living organisms [13]. Genes have been found to bind nuclear pore proteins in *Drosophila* and in higher eukaryotes, but zip code sequences have not been reported in either case [17,18].

The present study was designed to test the idea that DNA zip codes may be present in herpesvirus genomes. In particular, it was attractive to think that since MRS zip codes are able to mediate reactivation of repressed yeast genes, they may have a role in reactivation of latent herpesviruses. Virus genomes were computationally scanned for the presence of zip codes, and the results were interpreted to evaluate the potential role of zip codes in aspects of herpesvirus replication including latency.

*Corresponding authors: Jay C Brown, Department of Microbiology, Immunology and Cancer Biology, University of Virginia School of Medicine Box 800734, Charlottesville, Virginia 22908, USA, Tel: 01-434-924-1814; E-mail: jcb2g@virginia.edu

Received: June 29, 2016 Accepted: July 14, 2016 Published: July 25, 2016

Material and Methods

Herpesvirus DNA sequences were retrieved from the NCBI database using the accession numbers shown in Table 1 and examined with Genome Workbench (www.ncbi.nlm.nih.gov/tools/gbench/). Accession numbers for Individual strains of HHV1 were: 17: NC_001806.2; CR38: HM585508; E06: HM585496; F: GU734771; H129: GU734772; KOS: JQ673480. For HHV4: AG876: NC_009344; GD1: AY961628; WT: NC_007605.1. All genomes were scanned for three zip codes, MRS' (TCCTCCTT; [15,19], GRS1 (GGGTTGGA; [13]) and Put3B (CGGGGTTA; [16,20]). Scans were done on both virus DNA strands. A randomized version of each virus sequence was produced using the shuffleseq (default parameters) program in

EMBOSS (<http://cys.genomics.purdue.edu/emboss/>). One cycle of randomization was employed in each case. Counts and locations of zip codes were determined with locally-written Python scripts based on data.count and re.finditer, respectively. Data were plotted with Sigma Plot and rendered graphically with Adobe Illustrator CS3

Results

Zip code abundance

Scans were carried out with 39 herpesvirus species including viruses from all three subfamilies of the Herpesviridae. Studies included two other related viruses, channel catfish virus (ItalHV1) and Ostreid herpesvirus (OstHV1) from the related taxonomic orders

Table 1: Herpesvirus genomes examined.

	Abbreviation	Accession No.	Genome Length (bp)
Alpha herpesviruses			
Papiine herpesvirus 2	PapHV2	NC_007653.1	156487
Cercopithecine herpesvirus 2	CercoHV2	NC_006560.1	150715
Bovine herpesvirus 5	BHV5	NC_005261.2	137821
Macacine herpesvirus 1	MacHV1	NC_004812.1	156789
Suid herpesvirus 1	SHV1	NC_006151.1	143461
Bovine herpesvirus 1	BHV1	NC_001847.1	135301
Human herpesvirus 2	HHV2	NC_001798.1	154746
Human herpesvirus 1	HHV1	NC_001806.2	152261
Psittacid herpesvirus 1	PsitHV1	NC_005264.1	163025
Equine herpesvirus 1	EHV1	NC_001491.2	150224
Gallid herpesvirus 3	MDV2	AB049735	164270
Equine herpesvirus 4	EHV4	NC_001844.1	145597
Meleagrid herpesvirus 1	MelHV1	NC_002641.1	159160
Human herpesvirus 3	HHV3	NC_001348.1	124884
Gallid herpesvirus 2	MDV1	AF147806.2	174077
Cercopithecine herpesvirus 9	CercoHV9	NC_002686.2	124784
Beta herpesviruses			
Tupauid herpesvirus 1	TupHV1	NC_002794.1	195859
Murid herpesvirus 2	MHV2	NC_002512.2	230138
Panine herpesvirus 2	PanHV2	NC_003521.1	241087
Murid herpesvirus 1	MHV1	NC_004065.1	230278
Human herpesvirus 5	HHV5	NC_006273.2	235646
Macacine herpesvirus 3	MacHV3	NC_006150.1	221454
Human herpesvirus 6A	HHV6A	NC_001664.2	159322
Human herpesvirus 6B	HHV6B	NC_000898.1	162114
Human herpesvirus 7	HHV7	NC_001716.2	153080
Gamma herpesviruses			
Macacine herpesvirus 4	MacHV4	NC_006146.1	171096
Human herpesvirus 4	HHV4	NC_007605.1	171823
Equine herpesvirus 2	EHV2	NC_001650.1	184427
Human herpesvirus 8	HHV8	NC_009333.1	137969
Ovine herpesvirus 2	OHV2	NC_007646.1	135135
Macacine herpesvirus 5	MacHV5	NC_003401.1	133719
Callitrichine herpesvirus 3	CallHV3	NC_004367.1	149696
Murid herpesvirus 4	MHV4	NC_001826.2	119451
Alcelaphine herpesvirus 1	AlcHV1	NC_002531.1	130608
Bovine herpesvirus 4	BHV4	NC_002665.1	108873
Ateline herpesvirus 3	AtelHV3	NC_001987.1	108409
Saimiriine herpesvirus 2	SamHV2	NC_001350.1	112930
Allo herpesvirus; Malaco herpesvirus			
Ictalurid herpesvirus 1	ItalHV1	NC_001493.2	134226
Ostreid herpesvirus 1	OstHV1	NC_005881.1	207439

Table 2: DNA zip codes in 39 herpesviruses.

Virus	Zip Code			Total-
	MRS*	GRS1	Put3B	
PapHV2	5 (0; 1)a	1 (1; 4)	0 (1; 4)	6
CercoHV2	2 (2; 1)	1 (5; 4)	0 (2; 4)	3
BHV5	8 (1; 2)	2 (5; 4)	0 (5; 4)	10
MachV1	5 (1; 2)	1 (8; 5)	2 (6; 5)	8
SHV1	16 (1; 2)	5 (4; 5)	0 (5; 5)	21
BHV1	7 (2; 2)	1 (4; 5)	0 (4; 5)	8
HHV2	12 (1; 2)	4 (5; 6)	1 (7; 6)	17
HHV1	6 (5; 3)	9 (8; 6)	9 (3; 6)	24
PsitHV1	4 (7; 4)	3 (7; 6)	4 (7; 6)	11
EHV1	4 (5; 4)	8 (4; 5)	3 (11; 5)	15
MDV2	9 (2; 5)	5 (5; 6)	10 (3; 6)	24
EHV4	1 (5; 4)	8 (3; 4)	5 (3; 4)	14
MelHV1	3 (6; 5)	3 (3; 4)	5 (3; 4)	11
HHV3	3 (0; 4)	6 (4; 3)	7 (1; 3)	16
MDV1	8 (7; 5)	5 (4; 4)	1 (5; 4)	14
CercoHV9	3 (2; 3)	5 (2; 2)	2 (2; 2)	10
Sub-total	96	67	49	212
TupHV1	7 (6; 4)	4 (9; 8)	2 (8; 8)	13
MHV2	14 (7; 6)	6 (11; 9)	0 (9; 9)	20
PanHV2	20 (3; 6)	5 (2; 9)	0 (9; 9)	25
MHV1	18 (4; 6)	8 (9; 9)	4 (10; 9)	30
HHV5	14 (8; 7)	1 (9; 9)	4 (7; 9)	19
MachV3	5 (5; 7)	6 (3; 7)	3 (4; 7)	14
HHV6A	6 (5; 4)	2 (1; 3)	3 (4; 3)	11
HHV6B	3 (2; 5)	4 (1; 3)	4 (3; 3)	11
HHV7	4 (7; 3)	0 (3; 2)	5 (2; 2)	9
Sub-total	91	36	25	152
MachV4	10 (4; 4)	5 (5; 7)	2 (4; 7)	17
HHV4	16 (3; 5)	6 (10; 7)	3 (11; 7)	25
EHV2	18 (3; 5)	3 (4; 7)	2 (8; 7)	23
HHV8	9 (2; 4)	4 (7; 5)	1 (2; 5)	14
OHV2	12 (4; 4)	3 (5; 4)	1 (1; 4)	16
MachV5	1 (3; 4)	4 (5; 4)	6 (2; 6)	11
CallHV3	6 (3; 4)	9 (5; 4)	1 (6; 4)	16
MHV4	1 (5; 4)	7 (0; 3)	1 (4; 3)	9
AlcHV1	7 (4; 4)	1 (3; 3)	2 (1; 3)	10
BHV4	5 (3; 3)	7 (4; 2)	1 (5; 2)	13
AtelHV3	6 (1; 2)	3 (3; 1)	1 (3; 1)	10
SamHV2	5 (3; 2)	1 (3; 1)	1 (1; 1)	7
Sub-total	96	53	22	171
ItalHV1	6 (4; 4)	4 (5; 5)	6 (4; 5)	16
OstHV1	6 (6; 5)	10 (5; 3)	3 (2; 3)	19
Total	295	170	105	570

^a Entries show the virus zip code count followed by zip code count in a randomized virus genome and the statistically expected number in the virus genome. Entries in yellow indicate virus zip code counts higher than those in both controls.

Alloherpesviruses and Malacoherpesviruses, respectively (Table 2). Genomes were scanned for the presence of three zip codes: (1) a core form of MRS found to activate yeast genes and identified as MRS⁺ (TCCTCCTT; [15,19]); (2) GRS1 (GGGTTGGA; [13]); and a binding site for the transcription factor Put3 (CGGGGTTA; called Put 3B; [16,20]) also able to recognize GRS1. Put3B was included because although Put3 recognizes GRS1, it has additional binding specificities

including Put3B that could also have zip code function due to their ability to bind Put3 [20].

A total of 570 zip codes were observed among the viruses tested (see Supplementary Table 1). In each virus, the zip code count in the wild type (wt) virus genome was compared to the count obtained with a control, randomized version of the virus sequence and with the statistically expected number of zip codes in the virus sequence

Table 3: Zip code count in herpesvirus genomic features.

Herpesvirus Group	Zip Code	Count				Total count
		Gene	Promoter	Between genes	Gene > end	
Alpha	MRS*	29	44	23	0	96 (45%)
	GRS1	43	21	3	0	67 (32%)
	Put3B	20	13	13	3	49 (23%)
		92 (43%)	78 (37%)	39 (18%)	3 (1%)	212 (100%)
Beta	MRS*	40	40	10	0	91 (60%)
	GRS1	21	12	4	0	36 (24%)
	Put3B	16	5	3	1	25 (16%)
		77 (51%)	57 (38%)	17 (11%)	1 (1%)	152 (100%)
Gamma	MRS*	36	50	6	4	96 (56%)
	GRS1	33	17	3	0	53 (31%)
	Put3B	8	10	4	0	22 (13%)
		76 (44%)	78 (46%)	13 (8%)	4 (2%)	171 (100%)
Allo/Malaco	MRS*	7	3	2	0	12 (34%)
	GRS1	8	6	0	0	14 (40%)
	Put3B	5	2	2	0	9 (26%)
		20 (57%)	11 (31%)	4 (11%)	0 (0%)	35 (100%)
	All	265 (46%)	224 (39%)	73 (13%)	8 (1%)	570 (100%)

^a Viruses are those identified in Table 1.

^b A zip code was considered to be in a gene promoter if it is upstream of an annotated gene start site with no intervening feature. Promoter zip code distances to the relevant gene start are given in Supplementary Table 1. 8bp was the shortest distance of a promoter zip code to the gene start and 1277bp was the longest. The mean was 384bp.

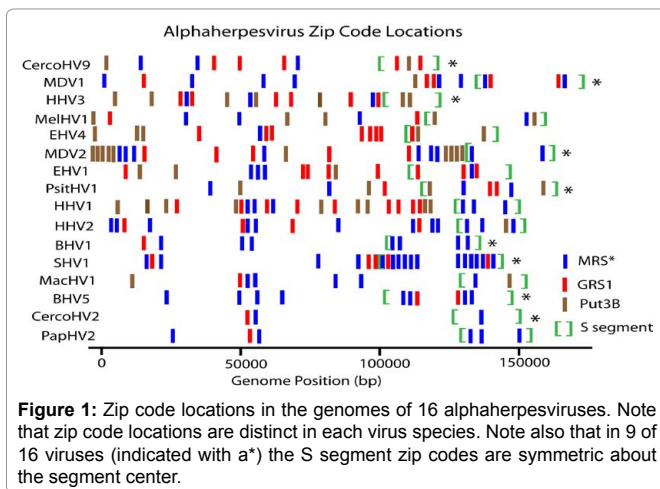


Figure 1: Zip code locations in the genomes of 16 alphaherpesviruses. Note that zip code locations are distinct in each virus species. Note also that in 9 of 16 viruses (indicated with a*) the S segment zip codes are symmetric about the segment center.

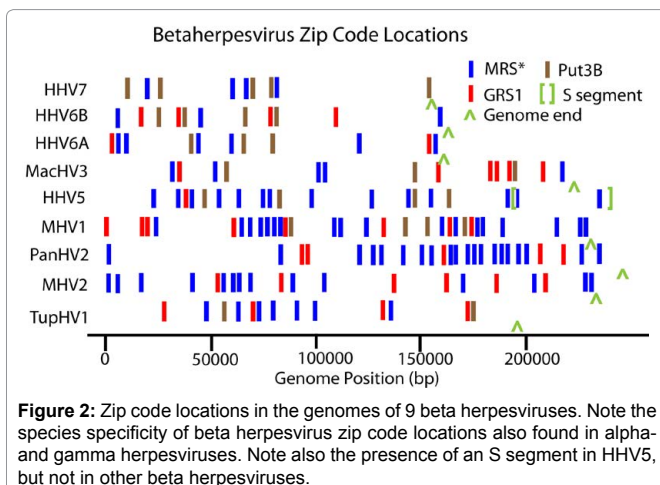


Figure 2: Zip code locations in the genomes of 9 beta herpesviruses. Note the species specificity of beta herpesvirus zip code locations also found in alpha and gamma herpesviruses. Note also the presence of an S segment in HHV5, but not in other beta herpesviruses.

(Table 2). All but four of the viruses tested were found to have counts above both control counts in at least one zip code. In one virus (HHV1) the count is above background in all three zip codes tested (Table 2). Three is the lowest number of zip codes found in any of the 39 viruses (CercoHV2) and 30 the highest (MHV1). Among the three zip codes tested, the highest count was observed for MRS* (295; 52%) followed by GRS1 (170; 30%) and Put3B (105; 18%). The overall abundance of GRS1 zip codes was found to be similar in the herpesvirus genomes examined here and the value reported for the *Saccharomyces cerevisiae* genome [13]. Values are 1 GRS1/36.5kbp and 1/49.3kbp for herpesviruses and *S. cerevisiae*, respectively.

Zip code locations

The genomic locations were determined for all 570 herpesvirus zip codes identified and the results are summarized in Table 3. Most zip codes (46%) were located in gene coding regions followed by promoters (39%), regions between genes (13%) and regions between a genome end and the nearest gene (1%). The same order of abundance was observed in the alpha- and beta herpesviruses. In gamma herpesviruses the abundance was slightly higher in promoters (46%) compared to gene coding regions (44%). MRS* was the most abundant zip code in the alpha-, beta- and gamma herpesviruses accounting for 45%-60% of the total number of zip codes. This was followed by GRS1 and Put3B in all three virus subfamilies (Table 3).

Zip code locations in the alpha-, beta- and gamma herpesviruses are shown graphically in Figures 1-3, respectively. In all viruses the gene order is the one shown in GenBank (see Table 1) with the left ends aligned. If an S segment is present, it is indicated with green square brackets ([]). If not, the right end is shown with an arrowhead (^). Note that S segments are found only in alphaherpesviruses and in one beta herpesvirus (HHV5).

The results show that zip codes are distributed across the entire genome in most of the viruses examined. There was little evidence for large scale clustering in a particular region. Exceptions to this

generalization involve closely-spaced groups of a few zip codes found, for instance, in SHV1 and MDV2 among the alphaherpesviruses, PanHV2 and MHV1 among the betaherpesviruses and HHV4 and AtelHV3 among the gammas. In some viruses there are large regions that lack zip codes entirely, although this is rare. Examples are found in the genomes of PapHV2 (alpha), HHV7 (beta) and MacHV4 (gamma).

Divergent zip code locations in different herpesvirus species

In most viruses the distribution of zip codes is distinctive. Viruses with the same or similar zip code location patterns were not observed. An exception to this rule occurs in the case of three alphaherpesviruses,

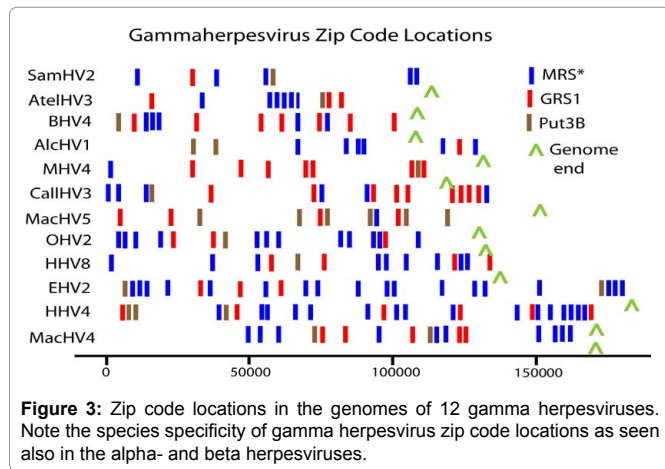


Figure 3: Zip code locations in the genomes of 12 gamma herpesviruses. Note the species specificity of gamma herpesvirus zip code locations as seen also in the alpha- and beta herpesviruses.

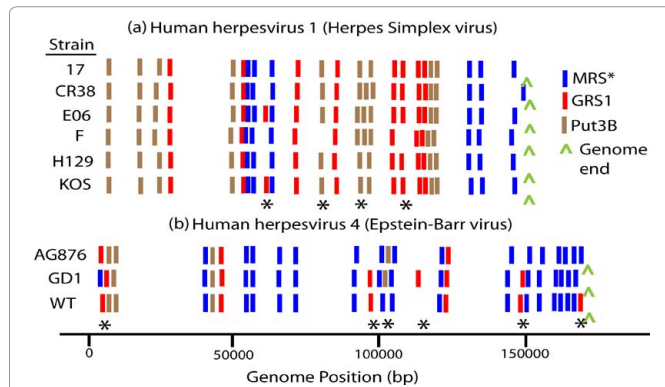


Figure 4: Zip code locations in the genomes of six HHV1 strains (a) and three strains of HHV4 (b). An asterisk marks zip codes present in less than all the strains examined. Note that most zip code locations are conserved in HHV1 strains. Similar conservation is observed with HHV4 zip codes.

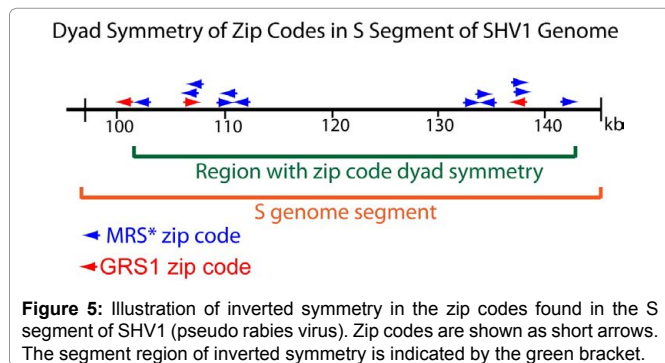


Figure 5: Illustration of inverted symmetry in the zip codes found in the S segment of SHV1 (pseudo rabies virus). Zip codes are shown as short arrows. The segment region of inverted symmetry is indicated by the green bracket.

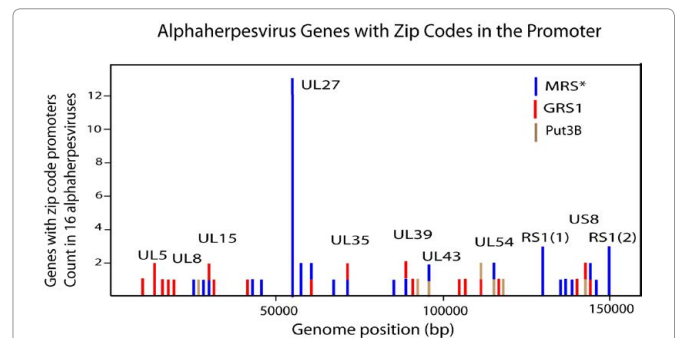


Figure 6: Genes with zip code-containing promoters in 16 alphaherpesviruses. Analysis was carried out with the genomes of the 16 alphaherpesviruses shown in Table 1. An entry is made for a gene if it has a zip code in its promoter and a homologous gene in the HHV1 genome. Note that genes that meet the above criteria are found throughout the alphaherpesvirus genome. The genes most likely to have a zip code-containing promoter encode glycoprotein B (UL27 gene) and transcription factor ICP4 (RS1 gene).

PapHV2, CercoHV2 and BHV1, where a recognizable pattern of similarity was observed in 5 of 8 zip codes (Figure 1).

Conservation of zip code locations within individual herpesvirus species

In contrast to the distinct patterns of zip code locations observed in different herpesvirus species, clear similarities were observed when strains of a single species were compared. Figure 4 shows the results of relevant studies carried out with six strains of HHV1 (Figure 4a) and with three strains of HHV4 (Figure 4b). In both cases a clear pattern of similarity was observed. In HHV1, for instance, the locations of 20 zip codes were conserved among the 24 present. Likewise in the HHV4 genome, 19 of 25 zip code locations were conserved.

L and S genome segments

A noteworthy feature of zip code locations has to do with viruses that have distinct L and S genome segments, regions that invert during virus DNA replication [21-23]. Among the viruses examined here, this feature is observed in all alphaherpesviruses and in human cytomegalovirus [24,25]. In some cases S segment zip codes are found to be arranged with inverted repeat (dyad) symmetry about the center of the segment. This feature is found in 9 of 16 alphaviruses (in Figure 1) and in human cytomegalovirus (HHV5; Figure 2). Individual zip codes related by dyad symmetry were found to be transcribed from opposite DNA strands.

This feature is illustrated in Figure 5 for SHV1. A similar symmetric arrangement of L segment zip codes is observed in MDV2, but not in any other virus tested. The symmetric distribution of zip codes in the S and L segments is of interest because the same symmetry is not observed in viruses lacking S and L segments (i.e. 22 of the 39 viruses examined). Symmetry is also not observed in other major features of the segments such as the gene order or nucleotide sequence. The presence of symmetry specifically in inverting segments raises the possibility that zip codes may be functionally involved in the segment inversion process [24,26,27].

Zip codes in alphaherpesvirus gene promoters

A study was carried out to focus specifically on zip codes present in gene promoters. The goal was to determine whether there are specific genes most likely to have a zip code-containing promoter. Analysis was restricted to the 16 alphaherpesviruses examined here

(Table 1) as clear homologies are found in many of the genes in these species. The study was further restricted to virus genes with a homolog in human herpesvirus 1 (HHV1). For each HHV1 gene, the number of alphaherpesvirus homologs with a zip code-containing promoter was counted and the results are plotted in Figure 6.

A total of 65 zip code-containing promoters were identified among the genes of the 16 alphaherpesviruses examined (Figure 6). The 65 were distributed among 38 different genes with all three types of zip codes represented. The count was highest for MRS' zip codes (38), with 20 and 7 found for GRS1 and Put3B zip codes, respectively. For each gene with a zip code-containing promoter, this feature was most often found in only one or two of the 16 viruses tested as shown in Figure 6. As a group these genes were distributed throughout the HHV1 genome. The highest counts were observed in UL27 (13 of 16 viruses) and both RS1 alleles (3 of 16 viruses for each). UL27 encodes a glycoprotein present on the virus surface and involved in entry of the virus into the host cell. RS1 encodes ICP4, a virus transcriptional transactivator. All the zip codes in both UL27 and RS1 promoters are MRS'. UL27 genes with zip code-containing promoters stand out because of the high number (13) of UL27 zipcode-containing promoters and also because 7 of the 13 viruses also have a GRS1 zip code in the UL27 coding region (see Supplementary Table 1). The high number of UL27 zip codecontaining promoters may relate to the close proximity of the UL28 coding region to the UL27 promoter in many alphaherpesviruses. In such cases the UL27 promoter zip code would be present in the UL28 gene.

Discussion

Gene relocation to the nuclear periphery

The ability to relocate an active gene from the nucleoplasm to the nuclear periphery is a defining function of DNA zip codes as established in studies with yeast. Relocation of yeast genes has been observed with multiple genes and with multiple zip codes [13,14,16]. The presence of zip codes in herpesvirus genomes as described here raises the possibility that zip code sequences may have an analogous function during virus replication. For instance, shortly after infection of fibroblasts with HHV1 in cell culture, the virus genome is found to be localized to the inner surface of the nuclear envelope [28,29]. It is suggested that HHV1 zip codes may cooperate with other factors such as nuclear lamins and virus-encoded VP16 protein to cause the virus genome to remain at or relocate to the nuclear rim.

Zip code locations

Further clues about the functions of herpesvirus zip codes were obtained from an analysis of zip code locations. Locations were found to be highly specific for the virus species. No two species were alike among the 39 virus genomes examined; even regions of local similarity were difficult to identify (see Figures 1-3). In contrast to the divergence of herpesvirus species in zip code location, a high level of conservation of was observed in different strains of the same virus. The latter conclusion is documented in Figure 4 which shows conservation of zip code locations in six distinct strains of human herpesvirus 1 (HHV1); similar conservation of zip code locations was observed in three Epstein-Barr virus strains.

Herpesvirus latency and transcriptional memory

A part of the reason the present study was initiated has to do with the properties of the memory recruitment sequence zip code (MRS; [13,15]). Studies with yeast have demonstrated that this sequence has the ability

to mediate transcriptional memory, an acceleration of the rate at which expression of a gene is reactivated if it has previously been expressed and later down regulated. Acceleration is most pronounced if an MRS is present in the promoter of a reactivated gene.

Involvement in transcriptional memory makes MRS well-suited for a role in reactivation of a herpesvirus from the latency. This possibility is illustrated by information available about herpes simplex virus latency. HHV1 causes latent infections specifically in neurons with infection beginning when the virus DNA enters the neuronal cell nucleus. Thereafter, virus genes are expressed for a short time (2-3 days) before their expression is almost entirely extinguished creating the latent state [30,31]. Latency can persist for a period of weeks or months. Expression of lytic virus genes can be reactivated in response to specific signals initiating a sequence of events that results in virus replication. It is easy to see how the ability of an MRS' sequence to accelerate expression of a target virus gene might potentiate the reactivation of virus genes, particularly if the MRS is located in a gene promoter.

Results obtained in the present study are consistent with the proposed role of MRS' zip codes to favor reactivation of latent herpesviruses. Of the three zip codes examined, MRS' was found to be the most abundant in all three subfamilies of the Herpesviridae. A substantial proportion of MRS' zip codes are located in gene promoters. For instance, among the alphaherpesviruses, 44 of 96 MRS' zip codes are found in promoters (Table 3).

Finally, it is intriguing to note that in three viruses, an MRS' zip code is located in the promoter of a gene encoding a transcription factor or regulatory protein able to affect initiation of virus lytic replication. This applies to the HHV1 RS1 gene encoding transcription factor ICP4, HHV5 genes encoding IRS and TRS regulatory proteins and the ZTA transcription factor encoded by the HHV4 BZLF1 gene (see Supplementary Table 1 and Figure 6). By affecting the expression of other virus genes, it is expected that accelerated expression of the above transcription factors will have an enhanced effect on the overall reactivation process.

Role of zip codes in herpesvirus segment inversion

In view of the irregular zip code locations present in most genome regions, it was unusual to note the inverted repeat symmetry of zip codes in the S segment. Such symmetry was observed in the S segment of 9 of 16 alphaherpesviruses and in human cytomegalovirus (HHV5; see Figures 1 and 2), but not in any of the other viruses tested. The presence of features with inverted repeat symmetry resembles results obtained with herpes simplex virus (HHV1) demonstrating that segment inversion is dependent on the presence of inverted repeat sequences at the segment ends [24,26,27]. In HHV1, repeats of non-zip code sequences ("a" sequences) are found at the ends of both L and S segments [26]; inverted repeats not involving "a" sequences are found to have the ability to substitute for "a" [27]. The results with HHV1 raise the possibility that zip code inverted repeats may be involved in segment inversion in the viruses where they occur.

In contrast to the inverted repeat symmetry observed with zip codes in the S segment of some viruses, it was rare to find similar inverted zip code repeats in the L segment despite the fact that the L segment is also found to invert (see Figures 1 and 2). Inverted repeat L segment zip codes were observed only in MDV1.

Zip codes near herpesvirus DNA ends

The presence of zip codes near the ends of herpesvirus genomes is

of interest as it suggests zip codes may be involved in two functions, uptake of virus DNA into the host cell nucleus and circularization of the DNA once it has entered. In all herpesviruses it is considered that the virus DNA enters the host cell nucleus by the pathway established for HHV1 [32-34]. Uptake begins when a DNA-containing virus capsid binds to the cytoplasmic face of a nuclear pore. DNA is then threaded out of the capsid, through the nuclear pore complex and into the nucleoplasm. The affinity of an end zip code for the inner aspect of a nuclear pore suggests itself as a way the zip code could prevent DNA from exiting the nucleus after it has entered. As each zip code has an affinity for only a subset of nuclear pores [16], it is expected that the involvement of a zip code in DNA uptake would be most pronounced in those nuclear pores recognized by the end zip code.

Circularization of herpesvirus DNA is found to occur promptly after it enters the host cell nucleus [35,36]. By binding DNA to the inner aspect of a nuclear pore, a zip code would be well suited to facilitate circularization by bringing the two genome ends into close physical proximity for end ligation. Because of the zip code specificity for particular nuclear pores [16], the role of zip codes would be enhanced if the same zip code were near both genome ends.

Examples of this are HHV2 (MRS'), MDV1 (MRS'), EHV4 (Put3B) and MeHV1 (Put3B) among the alphaherpesviruses (see Figure 1), and MHV2 (MRS') and HHV6B (MRS') among the betaherpesviruses (see Figure 2).

Acknowledgment

I thank Jennifer Thompson for help in the early stages of this investigation.

Supplementary Data

Supplementary Table 1 Information contained in Supplementary Table 1 is available as a searchable MySQL database. Send request to author.

References

- Davison AJ (2010) Herpesvirus systematics. *Vet Microbiol* 143: 52-69.
- Pellet PE, Roizman, B (2013) Herpesviridae, *Fields Virology* (6th edn) Lippincott Williams & Wilkins, Philadelphia.
- McGeoch DJ, Cook S, Dolan A, Jamieson FE, Telford EA (1995) Molecular phylogeny and evolutionary timescale for the family of mammalian herpesviruses. *J Mol Biol* 247: 443-458.
- McGeoch DJ, Cook S (1994) Molecular phylogeny of the alphaherpesvirinae subfamily and a proposed evolutionary timescale. *J Mol Biol* 238: 9-22.
- Grünewald K, Desai P, Winkler DC, Heymann JB, Belnap DM, et al. (2003) Three-dimensional structure of herpes simplex virus from cryo-electron tomography. *Science* 302: 1396-1398.
- Miller G, Lipman L (1973) Comparison of the yield of infectious virus from clones of human and simian lymphoblastoid lines transformed by Epstein-Barr virus. *J Exp Med* 138: -1412.
- Booy FP, Newcomb WW, Trus BL, Brown JC, Baker TS, et al. (1991) Liquid-crystalline, phage-like packing of encapsidated DNA in herpes simplex virus. *Cell* 64: 1007-1015.
- Hayward GS, Jacob RJ, Wadsworth SC, Roizman B (1975) Anatomy of herpes simplex virus DNA: evidence for four populations of molecules that differ in the relative orientations of their long and short components. *Proc Natl Acad Sci U S A* 72: 4243-4247.
- Delius H, Clements JB (1976) A partial denaturation map of herpes simplex virus type 1 DNA: evidence for inversions of the unique DNA regions. *J Gen Virol* 33: 125-133.
- Roizman B, Knipe DM, Whitley RJ (2013) Herpes simplex viruses. *Fields Virology*, (6edn) Lippincott Williams & Wilkins, Philadelphia.
- Sinclair J (2008) Human cytomegalovirus: Latency and reactivation in the myeloid lineage. *J Clin Virol* 41: 180-185.
- Thorley-Lawson DA, Hawkins JB, Tracy SI, Shapiro M (2013) The pathogenesis of Epstein-Barr virus persistent infection. *Curr Opin Virol* 3: 227-232.
- Ahmed S, Brickner DG, Light WH, Cajigas I, McDonough M, et al. (2010) DNA zip codes control an ancient mechanism for gene targeting to the nuclear periphery. *Nat Cell Biol* 12: 111-118.
- Sood V, Brickner JH2 (2014) Nuclear pore interactions with the genome. *Curr Opin Genet Dev* 25: 43-49.
- Light WH, Brickner DG, Brand VR, Brickner JH (2010) Interaction of a DNA zip code with the nuclear pore complex promotes H2A.Z incorporation and INO1 transcriptional memory. *Mol Cell* 40: 112-125.
- Brickner DG, Ahmed S, Meldi L, Thompson A, et al. (2010) Transcription factor binding to a DNA zip code controls interchromosomal clustering at the nuclear periphery. *Dev Cell* 22: 1234-1246.
- Kalverda B, Pickersgill H, Shloma VV, Fornerod M (2010) Nucleoporins directly stimulate expression of developmental and cell-cycle genes inside the nucleoplasm. *Cell* 140: 360-371.
- Capelson M, Liang Y, Schulte R, Mair W, Wagner U, et al. (2010) Chromatin-bound nuclear pore components regulate gene expression in higher eukaryotes. *Cell* 140: 372-383.
- Guan Q, Haroon S, Bravo DG, Will JL, Gasch AP (2012) Cellular memory of acquired stress resistance in *Saccharomyces cerevisiae*. *Genetics* 192: 495-505.
- Badis G, Chan ET, van, L, Pena-Castillo BH, Tillo D, et al. (2008) A library of yeast transcription factor motifs reveals a widespread function for Rsc3 in targeting nucleosome exclusion at promoters. *Mol Cell* 32: 878-887.
- Dolan A, Jamieson FE, Cunningham C, Barnett BC, McGeoch DJ (1998) The genome sequence of herpes simplex virus type 2. *J Virol* 72: 2010-2021.
- McGeoch DJ, Dolan A, Donald S, Rixon FJ (1985) Sequence determination and genetic content of the short unique region in the genome of herpes simplex virus type 1. *J Mol Biol* 181: 1-13.
- McGeoch DJ, Dalrymple MA, Davison AJ, Dolan A, Frame MC, et al. (1988) The complete DNA sequence of the long unique region in the genome of herpes simplex virus type 1. *J Gen Virol* 69: 1531-1574.
- McVoy MA, Ramnarain D (2000) Machinery to support genome segment inversion exists in a herpesvirus which does not naturally contain invertible elements. *J Virol* 74: 4887.
- Ben-Porat T (1981) Replication of herpesvirus DNA. *Curr Top Microbiol Immunol* 91: 81-107.
- Mocarski ES, Post LE, Roizman B (1980) Molecular engineering of the herpes simplex virus genome: insertion of a second L-S junction into the genome causes additional genome inversions. *Cell* 22: 243-255.
- Weber PC, Challberg MD, Nelson NJ, Levine M, Glorioso JC (1988) Inversion events in the HSV-1 genome are directly mediated by the viral DNA replication machinery and lack sequence specificity. *Cell* 54: 369-381.
- Silva L, Oh HS, Chang L, Yan Z, Triezenberg SJ, et al. (2012) Roles of the nuclear lamina in stable nuclear association and assembly of a herpesviral transactivator complex on viral immediate-early genes. *M Bio* 3.
- Silva L, Cliffe A, Chang L, Knipe DM (2008) Role for A-type lamins in herpesviral DNA targeting and heterochromatin modulation. *PLoS Pathog* 4: e1000071.
- Kosz-Vnenchak M, Coen DM, Knipe DM (1990) Restricted expression of herpes simplex virus lytic genes during establishment of latent infection by thymidine kinase-negative mutant viruses. *J Virol* 64: 5396-5402.
- Kramer MF, Chen SH, Knipe DM, Coen DM (1998) Accumulation of viral transcripts and DNA during establishment of latency by herpes simplex virus. *J Virol* 72: 1177-1185.
- Batterson W, Furlong D, Roizman B (1983) Molecular genetics of herpes simplex virus. VIII. further characterization of a temperature-sensitive mutant defective in release of viral DNA and in other stages of the viral reproductive cycle. *J Virol* 45: 397-407.
- Copeland AM, Newcomb WW, Brown JC (2009) Herpes simplex virus replication: roles of viral proteins and nucleoporins in capsid-nucleus attachment. *J Virol* 83: 1660-1668.

34. Pasdeloup D, Blondel D, Isidro AL, Rixon FJ (2009) Herpesvirus capsid association with the nuclear pore complex and viral DNA release involve the nucleoporin CAN/Nup214 and the capsid protein pUL25. J Virol 83: 6610-6623.
35. Garber DA, Beverley SM, Coen DM (1993) Demonstration of circularization of herpes simplex virus DNA following infection using pulsed field gel electrophoresis. Virology 197: 459-462.
36. McVoy MA, Nixon DE, Adler SP (1997) Circularization and cleavage of guinea pig cytomegalovirus genomes. J Virol 71: 4209-4217.

Author Affiliations

[Top](#)

Department of Microbiology, Immunology and Cancer Biology, Immunology and Cancer Biology, University of Virginia School of Medicine, USA

Submit your next manuscript and get advantages of SciTechnol submissions

- ❖ 50 Journals
- ❖ 21 Day rapid review process
- ❖ 1000 Editorial team
- ❖ 2 Million readers
- ❖ More than 5000
- ❖ Publication immediately after acceptance
- ❖ Quality and quick editorial, review processing

Submit your next manuscript at • www.scitechnol.com/submission