Copy Number Heterogeneity, Large Origin Tandem Repeats, and Interspecies Recombination in Human Herpesvirus 6A (HHV-6A) and HHV-6B Reference Strains

Alexander L. Greninger,a,b Pavitra Roychoudhury,a,b Negar Maksous,a Derek Hanson,c Jill Chase,a Gerhard Krueger,f Hong Xie,a Meei-Li Huang,a Lindsay Saunders,a Dharam Ablashi,e David M. Koelle,a,b,c,d Linda Cook,a Keith R. Jeromea,b

aDepartment of Laboratory Medicine, University of Washington, Seattle, Washington, USA
bFred Hutchinson Cancer Research Institute, Seattle, Washington, USA
cDepartment of Medicine, University of Washington, Seattle, Washington, USA
dDepartment of Global Health, University of Washington, Seattle, Washington, USA
eHHV-6 Foundation, Santa Barbara, California, USA
fDepartment of Pathology and Laboratory Medicine, University of Houston, Houston, Texas, USA

ABSTRACT Quantitative PCR is a diagnostic pillar for clinical virology testing, and reference materials are necessary for accurate, comparable quantitation between clinical laboratories. Accurate quantitation of human herpesvirus 6A/B (HHV-6A/B) is important for detection of viral reactivation and inherited chromosomally integrated HHV-6A/B in immunocompromised patients. Reference materials in clinical virology commonly consist of laboratory-adapted viral strains that may be affected by the culture process. We performed next-generation sequencing to make relative copy number measurements at single nucleotide resolution of eight candidate HHV-6A and seven HHV-6B reference strains and DNA materials from the HHV-6 Foundation and Advanced Biotechnologies Inc. Eleven of 17 (65%) HHV-6A/B candidate reference materials showed multiple copies of the origin of replication upstream of the U41 gene by next-generation sequencing. These large tandem repeats arose independently in culture-adapted HHV-6A and HHV-6B strains, measuring 1,254 bp and 983 bp, respectively. The average copy number measured was between 5 and 10 times the number of copies of the rest of the genome. We also report the first interspecies recombinant HHV-6A/B strain with a HHV-6A backbone and a 5.5-kb region from HHV-6B, from U41 to U43, that covered the origin tandem repeat. Specific HHV-6A reference strains demonstrated duplication of regions at U1/U2, U87, and U89, as well as deletion in the U12-to-U24 region and the U94/U95 genes. HHV-6A/B strains derived from cord blood mononuclear cells from different laboratories on different continents with fewer passages revealed no copy number differences throughout the viral genome. These data indicate that large origin tandem duplications are an adaptation of both HHV-6A and HHV-6B in culture and show interspecies recombination is possible within the Betaherpesvirinae.

IMPORTANCE Anything in science that needs to be quantitated requires a standard unit of measurement. This includes viruses, for which quantitation increasingly determines definitions of pathology and guidelines for treatment. However, the act of making standard or reference material in virology can alter its very accuracy through genomic duplications, insertions, and rearrangements. We used deep sequencing to examine candidate reference strains for HHV-6, a ubiquitous human virus that can reactivate in the immunocompromised population and is integrated into the human genome in every cell of the body for 1% of people worldwide. We found large tandem repeats in the origin of replication for both HHV-6A and HHV-6B that are selected for in culture. We also found the first interspecies recombinant between
HHV-6A and HHV-6B, a phenomenon that is well known in alphaherpesviruses but to date has not been seen in betaherpesviruses. These data critically inform HHV-6A/B biology and the standard selection process.

**KEYWORDS** human herpesvirus 6, HHV-6, HHV-6A, HHV-6B, copy number, secondary standards, Z29, GS, origin of replication, origin, tandem repeat, nanopore, nanopore sequencing, direct repeat, quantitative PCR, international standard, reference materials

Human herpesvirus 6B (HHV-6B) is a ubiquitous human virus with human exposure levels of >90% by the age of 2 years as measured by serological assays performed worldwide (1, 2). Both HHV-6A and HHV-6B establish chronic infections in the majority of infected individuals, leading to asymptomatic persistent viral shedding (3, 4). Exanthema subitum is the most common HHV-6B-related infection seen after a primary exposure in 6-month- to 3-year-old children (5). Less frequently the virus can result in seizures, gastrointestinal and respiratory symptoms, thrombocytopenia, hepatitis, colitis, and central nervous system (CNS) infections (6). Additionally, both HHV-6A and HHV-6B have been shown to integrate into host chromosomes in the telomere regions and to be passed from parents to their children as inherited chromosomally integrated HHV-6A/B (iciHHV-6A/B) (7–9). The potential to detect HHV-6A/B in genomic DNA from these patients, as well as possible reactivation from the integrated HHV-6A/B, makes the diagnosis of HHV-6A/B infection from serum or plasma viral load testing challenging (10, 11).

Several prototypic strains of HHV-6A/B have been identified and utilized, including GS, U1102, SIE, LHV, Z29, and HST (12–16). Until very recently, there were fewer than 200 HHV-6A/B sequences in GenBank, including only 3 complete genomes. HHV-6A and -6B genomes have approximately 90% nucleotide identity to each other and about 50% similarity to the most closely related human betaherpesvirus, HHV-7. The genome is approximately 160 to 170 kb and contains many of the gene and regulatory elements present in the genomes of other betaherpesviruses. Despite the close sequence identity between HHV-6A and -6B, no interspecies recombinants have been isolated in nature to date. In comparison, interspecies herpes simplex virus 1 (HSV-1) and HSV-2 recombination is relatively common, even though they share significantly less sequence homology (17, 18). Recently, two large-scale efforts have sequenced more than 150 nearly full-length HHV-6B genomes from four continents to add to existing HHV-6A/B genomes (19, 20).

Accurate and sensitive real-time PCR assays that detect and quantify HHV-6A/B are critical to diagnosis and monitoring of the variety of manifestations of HHV-6A/B-associated disease. Recently, several quantitative cutoffs have been proposed that are associated with end-organ disease or iciHHV-6A/B status (21–23). A review of the PCR methods used in 46 recently published papers (2014 to 2017) revealed the use of 17 different primer sets at multiple locations throughout the genome (locations U6, U12, U13, U22, U27, U31, U32, U38, U41, U57, U65, U66, U67, U69, U90, U95, and U100), with only the U31, U65, and U66 primers used more than twice. Not surprisingly, cross-laboratory proficiency testing has shown differences in quantitation as high as 4 log units (24, 25). Previous studies have identified the critical role that standardized materials play in the ability to establish clinical viral load cutoffs and assay sensitivity and to compare results between laboratories (24, 26–30). Finally, effective primer designs have been significantly limited by the lack of available DNA sequences.

Previous work has reported linear amplification of the lytic origin of replication of HHV-6B strain Z29 to be associated with increased viral passage and increased replication in vitro (31, 32). In an effort to determine whether the available cultured “reference” strains have undergone significant changes during culture similar to that seen with the recently produced WHO BK and JC strains, we obtained 15 strains from the HHV-6 Foundation repository and used shotgun sequencing to obtain full-length genomes and estimates of copy numbers (33, 34). We found 9 of the 15 strains had
high-copy-number tandem-repeat amplifications in the origin of replication, including
the above-mentioned HHV-6B strain Z29, albeit with different breakpoints than previ-
ously described, as well as the first described origin amplification in an HHV-6A strain.
We also describe the first HHV-6A/HHV-6B interspecies recombinant. Other HHV-6A/B
reference materials had multiple loci with copy number variations of up to 20 times.
These duplications, deletions, and rearrangements may impact the utility of the strains
for the production of standard materials for PCR testing. Changes in the genomes of the
strains in culture may have an impact on the results of current and future studies
utilizing these materials.

RESULTS

Large tandem repeats covering the origin of replication in both HHV-6A and
HHV-6B strains. In order to obtain single nucleotide resolution and copy number
measurement for HHV-6A/B type strain reference materials, we sequenced an HHV-6A
GS strain obtained from the HHV-6 Foundation and a HHV-6B Z29 strain obtained from
the NIH AIDS repository. Libraries of the Z29 and GS strains were each prepared twice
and sequenced to an average depth of 76× and 453×, respectively. The HHV-6B Z29
strain contained a homogeneous 983-bp-long tandem repeat (Fig. 1A). Copy number
estimates based on relative coverage at the edges of the repeats across multiple library
preparations indicated an average of 11 to 13 copies of the repeat were present.
Mapping of the edges of the Z29 origin tandem repeat gave different repeat break-
points than previously described for what was possibly a different passage lineage (31).
All Z29 strains sequenced in this study had an additional 123 nucleotides at the 5’
end of the repeat and an extra 4 nucleotides at the 3’ end of the repeat compared to
the previously described repeat to make the 983-bp tandem repeat (31).

HHV-6A strain GS contained a heterogeneous tandem repeat that covered 1,254 bp
of the HHV-6A reference genome (Fig. 1B). Copy number estimates based on relative
coverage at the edges of the repeats indicated an average of 10 to 12 copies of the
repeat were present. The most common tandem repeat present included deletions of
193 bp and 2 bp, along with an insertion of 189 bp, based on the HHV-6A reference
genome, giving a mode length of 1,254 bp.

To demonstrate that the copy number heterogeneity we found in the type strains
of HHV-6A/B are also present in commercially available quantitative clinical reference
materials, we also performed shotgun DNA sequencing on an HHV-6A GS strain and an
HHV-6B Z29 strain from Advanced Biotechnology Inc. (ABI). These strains had tandem
repeats that were similar in size at the origin of replication present in the GS strain
obtained from the HHV-6 Foundation and Z29 strains obtained from the NIH AIDS
repository (Fig. 1C). However, the Z29 origin tandem repeat in the commercial refer-
ence material was present at approximately twice the copy number observed in Z29
from the NIH AIDS repository. Intriguingly, the HHV-6A GS strain quantitative secondary
standard material also contained a 4-fold increase in coverage, covering the U90 and
U91 genes and the N-terminal two-thirds of the U86 gene. The U91 end of the repeat
contained a complex rearrangement with the U100-direct repeat (DR) intergenic region
266 nucleotides 5’ of the beginning of the annotated DR repeat region. The HHV-6B
Z29 secondary standard contained three large rearrangements associated with the
presence of junctional reads—one between U12 and U20 constituting ~20 to 30% of
the genomes present, one between U73 and R3 repeat region constituting ~50% of the
genomes present, and one between U86 and U95 representing ~25% of the genomes
present. Thus, the copy number of HHV-6B between U86 and R3 was 4-fold lower and
between U73 to U86 and R3 to U95 representing 2-fold lower than the rest of the genome.
Quantitative PCR (qPCR) analysis of origin and U32 loci confirmed the deep-sequencing
data, demonstrating a range between 1 and 10× higher copy numbers of the origin
than the U32 housekeeping gene (Fig. 2). Of note, an early-passage GS strain had equal
copy numbers of the origin and U32 genes, while the later-passage GS-HSB2 strain had
a 5.8-fold higher copy number at the origin, consistent with what had been described
previously in Z29 (31). Both HHV-6A GS and HHV-6B Z29 strains that did not have large
FIG 1 Representative coverage maps of HHV-6B Z29 and HHV-6A GS reference strains. Shotgun DNA-sequencing reads from cultured virus were mapped to the NCBI HHV-6B and HHV-6A reference genomes, NC_000898 and NC_001664, respectively. The green stacked lines indicate the gene models for the respective viral species. (A) HHV-6B strain Z29 yielded a homogeneous 983-bp tandem repeat that was present at approximately 12.5 times higher coverage than the rest of the genome. Sequences at the 5’ and (Continued on next page)
deletions or rearrangements outside the origin repeat had direct-repeat copy numbers approximately two-thirds that of the average coverage of the unique region when aligned with a reference genome with two direct repeats (Table 1). No differences in the lengths of direct-repeat regions were apparent in our strains, as was found in the high-passage-number HHV-6A strain U1102 (35).

Previous analysis of 125 HHV-6B genomes obtained from clinical specimens revealed no tandem repeats across the origin of replication (20). To confirm the origin tandem repeat present in HHV-6B strain Z29, we performed PCR amplification and fragment analysis by gel electrophoresis across the sequence tandem repeat with two separate primer sets. As a control, we performed PCR across the origin of replication in an HHV-6B PCR-positive patient specimen. The patient specimen demonstrated the presence of a single copy of the locus, while the Z29 strain contained an amplification ladder, consistent with a population of virus with different numbers of multiple tandem repeats present at the locus (Fig. 3A). We also performed amplification-free long-read nanopore sequencing on DNA extracted from HHV-6B strain Z29 in culture. Across 6,369 nanopore reads with an average read length of 3,235 nucleotides, we recovered two reads that contained more than one copy of the origin tandem repeat (Fig. 3B). One read contained three tandem repeats of the origin repeat, while the other contained two repeats. Neither read spanned the entire tandem repeat, consistent with the longer complex repeats previously observed by Southern blotting of multiple Z29 strains in culture (31).

Interspecies recombination between HHV-6A and HHV-6B in strain DA. Whole-genome sequencing of the DA strain revealed a hybrid genome indicative of interspecies recombination between HHV-6A and HHV-6B strains. The DA strain genome overall showed closer sequence identity to HHV-6A than to HHV-6B strains but included an
Analysis of the DA strain sequence yielded consensus recombination breakpoints at bp 67314 and bp 72667 of the HHV-6A U1102 reference genome (NC_001664) due to the presence of the oriLyt repeat. Only a 61-bp fragment with an oriLyt repeat had 6 variant sites compared to HHV-6B. The DA strain origin-binding protein U73 gene matched with perfect identity that of the HHV-6A GS strain by BLASTN analysis. Of note, the DA strain origin-binding protein U73 gene more closely matched HHV-6A U73 than HHV-6B U73 (99.5% versus 97.1% pairwise nucleotide identity). Analysis of HHV-6A/B strains sequenced in this study

| Species Strain Cell line | No. of copies/ml | No. of trimmed reads | No. of HHV-6 reads | Unique/DR coverage | Accession no. |
|--------------------------|------------------|----------------------|-------------------|--------------------|---------------|
| HHV-6A GS HSB2           | 3.30E9           | 2,091,792            | 47,508            | 54×10/36×          | MF994822      |
| HHV-6A GS-early CBMC     | 3.30E4           | 308,104              | 66,878            | 60×10/37×          | KC465951      |
| DA HSB2                  | 2.60E9           | 3,787,217            | 30,508            | 31×19/             | MF994820      |
| CO-1 HSB2                | 3.60E9           | 1,913,207            | 164,835           | 165×10/7×          | MF994815      |
| CO-2 HSB2                | 2.60E9           | 1,278,582            | 118,403           | 135×8/4×           | MF994816      |
| CO-3 HSB2                | 5.60E8           | 896,950              | 54,464            | 65×3/9×            | MF994817      |
| CO-4 HSB2                | 5.70E8           | 6,692,042            | 42,596            | 35×3/4×            | MF994818      |
| CO-7 HSB2                | 2.20E8           | 10,887,157           | 54,046            | 43×4/2×            | MF994819      |
| SIE PHA-stimulated CBMC  | 1.20E10          | 1,015,200            | 49,307            | 63×36×             | MF994828      |
| ABI-HHV6A (GS) Unknown   | 6.41E5           | 1,644,798            | 225,145           | 165×66/            | MF994813      |
| HHV-6B Z29 SupT1         | 2.00E6           | 2,436,488            | 62,770            | 90×52/             | MF994829      |
| HST MT4                  | 1.10E7           | 10,666,502           | 28,251            | 33×19/             | MF994824      |
| HST PHA-stimulated CBMC  | 1.70E9           | 1,155,336            | 29,365            | 39×25/             | MF994823      |
| KYO PHA-stimulated CBMC  | 7.00E8           | 949,314              | 28,626            | 38×25/             | MF994825      |
| ENO PHA-stimulated CBMC  | 3.30E9           | 1,116,480            | 39,004            | 51×32/             | MF994821      |
| MAR PHA-stimulated CBMC  | 2.30E8           | 801,650              | 26,594            | 34×21/             | MF994826      |
| NAK PHA-stimulated CBMC  | 3.60E6           | 5,617,132            | 42,955            | 44×33/             | MF994827      |
| ABI-HHV6B (Z29) Unknown  | 1.50E5           | 1,200,052            | 112,632           | 104×10/            | MF994814      |

*Sequenced previously (21); reads received from L. Flamand.
had equivalent mixed variant allele frequencies at 25 loci based on read mapping to the unique region of the HHV-6A reference genome (NC_001664) (Table 2). No other variants were isolated between the HHV-6A CO4 and CO7 strains, suggesting the strains are identical.

Reduced-passage HHV-6A/B strains do not contain large tandem duplications or deletions. To find the best standard for HHV-6A/B clinical testing and to better understand the source of the origin tandem repeat, we sequenced four HHV-6B strains and one HHV-6A strain that were isolated from phytohemagglutinin (PHA)-stimulated cord blood mononuclear cells (CBMC). Four HHV-6B strains (HST, ENO, KYO, and NAK) were isolated from Japanese exanthem subitum patients, one HHV-6B strain (MAR) was isolated from an asymptomatic French child, and one HHV-6A strain (SIE) was isolated from an Ivory Coast patient with adult T-cell leukemia (37–39). Interestingly, all six strains contained minimal levels of copy number heterogeneity, with an average coefficient of variation of coverage in the unique region of 21.0% compared with 119% averaged over the HHV-6A GS and HHV-6B Z29 type strains (Fig. 7). The lack of origin amplification in these CBMC-passaged strains was also confirmed by qPCR (Fig. 2). Without the tandem repeat in the origin, the only source of copy number differences in the cord blood mononuclear cell-passaged strains was direct-repeat coverage at approximately two-thirds of that in the unique region when aligned with a reference genome with two direct repeats, consistent with active replication and sequencing of mostly single DRs containing HHV-6A/B genomes (Fig. 7 and Table 1) (40, 41). Of note,
FIG 4  HHV-6A strain DA shows genomic evidence of interspecies recombination between HHV-6A and HHV-6B strains. (A) Tandem repeat of the origin for strain DA showing a Z29-like length of 983 bp. Gene annotations are depicted in green below the genome. (B) RDP4 scan recombination analysis demonstrating two recombination breakpoints at nucleotides 67,194 and 72,681 of the HHV-6A DA genome. (C) Mapping of a selection of reads to the HHV-6B Z29 reference genome at the U41 locus, with reads depicted in gray and disagreements with the HHV-6B Z29 reference genome highlighted in black. (D) Nucleotide sequence of the 61-bp region of the tandem repeat that most closely matches the HHV-6A reference genome (NC_001664), with base agreements depicted in gray and disagreements with the HHV-6A and HHV-6B reference genomes highlighted in black. (E) Phylogenetic tree analysis of the recombination region supports the HHV-6B-like nature of the U41 locus.
the decreased coverage in the direct-repeat region relative to the unique region was present in all the strains sequenced in this study.

**DISCUSSION**

We showed the presence of copy number heterogeneity at multiple genomic loci across HHV-6A and HHV-6B culture isolates that are used as reference material for clinical assay development and normalization, as well as basic-science virology work. The most prominent copy number difference was a large tandem repeat in the origin.
Recent analysis of 125 HHV-6B genomes and 10 HHV-6A sequences spanning the HHV-6B origin region revealed no large tandem repeats in the origin of replication among HHV-6B strains. The presence of large copy number differences in HHV-6A and HHV-6B strains was associated with passing in immortalized cell lines, as strains passed in primary cell lines and cord blood mononuclear cells did not carry genomic duplications and deletions and early-passage virus contained fewer tandem repeats than late-passage virus.

Recent analysis of 125 HHV-6B genomes and 10 HHV-6A sequences spanning the U41 origin region revealed no large tandem repeats in the origin of replication among clinical isolates (20). Previous work had shown that a large heterogeneous tandem repeat present at the oriLyt in HHV-6B Z29 strains associated with higher passage numbers conferred a growth advantage in vitro and that tandem HHV-6B origin elements were more efficiently replicated in vitro (31, 32). Our data show that this is a

| Nucleotide no. | Virus strain | Sequencing depth | Variant frequency (%) | Polymorphism type | Gene | Change | Amino acid change | CDS codon no. | Protein effect |
|---------------|--------------|-----------------|----------------------|------------------|------|--------|------------------|--------------|---------------|
| 15942         | HHV-6 CO4    | 22              | 31.80                | SNP (transition) | C→T | None   |                  |              |               |
| 15942         | HHV-6 CO7    | 40              | 45.00                | SNP (transition) | C→T | None   |                  |              |               |
| 16978         | HHV-6 CO4    | 35              | 40.00                | SNP (transversion) | U7  | G→T   | 38 None          |              |               |
| 16978         | HHV-6 CO7    | 51              | 43.10                | SNP (transversion) | U7  | G→T   | 38 None          |              |               |
| 18877         | HHV-6 CO4    | 33              | 48.50                | SNP (transversion) | U10 | G→A   | 362 Substitution |              |               |
| 18877         | HHV-6 CO7    | 32              | 59.40                | SNP (transversion) | U10 | C→A   | 425 Substitution |              |               |
| 19258         | HHV-6 CO4    | 30              | 50.00                | SNP (transition) | U11 | T→C   | 774 Substitution |              |               |
| 19258         | HHV-6 CO7    | 26              | 50.00                | SNP (transition) | U11 | T→C   | 774 Substitution |              |               |
| 29716         | HHV-6 CO4    | 19              | 36.80                | SNP (transversion) | U19 | T→A   | 368 Substitution |              |               |
| 29716         | HHV-6 CO7    | 15              | 53.30                | SNP (transversion) | U19 | T→A   | 368 Substitution |              |               |
| 36965         | HHV-6 CO7    | 47              | 34.00                | Insertion (tandem repeat) | U26 | (A)8→(A)9 | 282 Frame shift |              |               |
| 36965         | HHV-6 CO4    | 27              | 37.00                | Insertion (tandem repeat) | U26 | (A)8→(A)9 | 282 Frame shift |              |               |
| 40589         | HHV-6 CO4    | 42              | 31.00                | SNP (transition) | U28 | C→T   | 282 None         |              |               |
| 40589         | HHV-6 CO7    | 42              | 50.00                | SNP (transition) | U28 | C→T   | 282 None         |              |               |
| 58704         | HHV-6 CO4    | 21              | 33.30                | SNP (transition) | U38 | G→A   | 295 None         |              |               |
| 69117         | HHV-6 CO4    | 33              | 45.50                | SNP (transition) | U42 | A→G   | 494 None         |              |               |
| 69117         | HHV-6 CO7    | 58              | 50.00                | SNP (transition) | U42 | A→G   | 494 None         |              |               |
| 69133         | HHV-6 CO7    | 60              | 48.30                | SNP (transition) | U42 | A→G   | 489 Substitution |              |               |
| 70212         | HHV-6 CO4    | 49              | 51.00                | SNP (transition) | U43 | C→T   | 465 Substitution |              |               |
| 70212         | HHV-6 CO7    | 39              | 56.40                | SNP (transition) | U43 | C→T   | 465 Substitution |              |               |
| 74768         | HHV-6 CO4    | 42              | 66.70                | SNP (transition) | U45 | T→C   | 151 Substitution |              |               |
| 74768         | HHV-6 CO7    | 32              | 68.80                | SNP (transition) | U45 | T→C   | 151 Substitution |              |               |
| 92258         | HHV-6 CO7    | 45              | 33.30                | SNP (transition) | U57 | C→T   | 552 Substitution |              |               |
| 92258         | HHV-6 CO4    | 30              | 36.70                | SNP (transition) | U57 | C→T   | 552 Substitution |              |               |
| 106203        | HHV-6 CO4    | 21              | 33.30                | SNP (transition) | U70 | G→A   | 214 None         |              |               |
| 106203        | HHV-6 CO7    | 45              | 48.90                | SNP (transition) | U70 | G→A   | 214 None         |              |               |
| 106566        | HHV-6 CO4    | 32              | 37.50                | SNP (transition) | U70 | T→A   | 335 Substitution |              |               |
| 106566        | HHV-6 CO7    | 39              | 59.00                | SNP (transition) | U70 | T→A   | 335 Substitution |              |               |
| 107612        | HHV-6 CO4    | 28              | 57.10                | SNP (transition) | U72 | G→A   | 234 Substitution |              |               |
| 107612        | HHV-6 CO7    | 40              | 60.00                | SNP (transition) | U72 | G→A   | 234 Substitution |              |               |
| 107916        | HHV-6 CO4    | 37              | 43.20                | SNP (transition) | U73 | G→A   | 464 None         |              |               |
| 107916        | HHV-6 CO7    | 40              | 55.00                | SNP (transition) | U73 | G→A   | 464 None         |              |               |
| 110489        | HHV-6 CO4    | 33              | 42.40                | SNP (transition) | U73 | A→G   | 722 Substitution |              |               |
| 110489        | HHV-6 CO7    | 46              | 43.50                | SNP (transition) | U73 | A→G   | 722 Substitution |              |               |
| 112286        | HHV-6 CO4    | 44              | 40.90                | SNP (transition) | U74 | T→C   | 551 Substitution |              |               |
| 112286        | HHV-6 CO7    | 42              | 61.90                | SNP (transition) | U74 | T→C   | 551 Substitution |              |               |
| 113466        | HHV-6 CO4    | 25              | 60.00                | SNP (transversion) | U76 | G→T   | 598 Substitution |              |               |
| 113466        | HHV-6 CO7    | 46              | 65.20                | SNP (transversion) | U76 | G→T   | 598 Substitution |              |               |
| 121571        | HHV-6 CO4    | 11              | 63.60                | Deletion          | U79 | -T    | 410 Frame Shift  |              |               |
| 121571        | HHV-6 CO7    | 16              | 68.80                | Deletion          | U79 | -T    | 410 Frame Shift  |              |               |
| 123584        | HHV-6 CO4    | 22              | 36.40                | SNP (transversion) | U83 | A→T   | 22 Substitution  |              |               |
| 123584        | HHV-6 CO7    | 42              | 42.90                | SNP (transversion) | U83 | A→T   | 22 Substitution  |              |               |
| 123591        | HHV-6 CO4    | 20              | 30.00                | SNP (transversion) | U83 | T→C   | S→P 22 Substitution |              |               |
| 123591        | HHV-6 CO7    | 44              | 45.50                | SNP (transversion) | U83 | T→C   | S→P 22 Substitution |              |               |

*Matched loci with variant allele frequencies between 30 and 75%, coverage of >10× at nucleotides 8,500 to 150,000 of NC_001664 for HHV-6A strains CO-4 and CO-7.*

SNP, single nucleotide polymorphism.
general feature of HHV-6A/B and that a heterogeneous larger tandem repeat is also present in multiple laboratory-adapted HHV-6A strains at the oriLyt. We showed that increased culture passage was associated with reduced copy numbers of two additional large genomic loci in HHV-6A CO strains. Unfortunately, increased culture passage numbers and passage in immortalized cell lines were correlated in our study, making it difficult to pinpoint the exact role of passage number versus cell line type. Based on the data from the early-passage strain of GS and previous work done in Z29, the data are most consistent with increased copy numbers being associated with increased passage numbers rather than passage in immortalized cell lines (31). We also hypothesize that an increased copy number of the origin confers a growth advantage for HHV-6A/B, as it arose independently in both HHV-6A and HHV-6B type strains and is associated and maintained with increasing passage numbers, although more experiments are required to formally characterize the mechanism of this phenomenon. With the discovery of the origin repeats in HHV-6A strain GS, transgenic experiments can now be performed in the HHV-6A bacterial artificial chromosome system, in addition to plasmid-based replicon experiments used to characterize origins (41).

Copy number variability at multiple genomic loci was also reflected in HHV-6A/B reference material that is used to normalize quantitative values for clinical assay development. It is intriguing to speculate whether HHV-6A/B strains missing tens of kilobases from the U12-to-U20/U24 or U73-to-R3 regions are replication competent. No doubt these viruses were grown at high multiplicities of infection, and it is highly likely virus with deletions could be complemented with intact virus. Nevertheless, the U12-to-U20 and U12-to-U24 deletion viruses (which also include 4 C-terminal residues of U25) in the HHV-6B Z29 ABI reference material and CO4-7 strains are notable, as none of these genes have been shown to be essential for HHV-6A/B replication (16, 42, 43). The U20 to U24 genes are Roseolovirus-specific immunomodulatory genes and have been shown to be dispensable for viral replication in vitro (42, 43). By analogy to the human cytomegalovirus (HCMV) genome, the HHV-6A U14 to U17 proteins are HCMV

![FIG 7](image-url) HHV-6A/B strains sequenced from PHA-activated cord blood mononuclear cells revealed no major copy number differences. Coverage maps from four HHV-6B strains from Japan (NAK, KYO, ENO, and HST), one HHV-6B strain from France (MAR), and one HHV-6A strain from Ivory Coast (SIE) are shown. The major copy number difference in these strains is the reduced coverage in the direct-repeat region, likely due to sequencing of HHV-6A/B strains with only one direct repeat. The direct-repeat coverage difference was present in most of the strains sequenced in this study (Table 1).
UL25-like proteins, which are not required for HCMV replication (44). The boundaries of the deletion may not be due entirely to chance, as U11 is an essential gene for HHV-6A replication (45). HHV-6A with U94 deleted is replication competent, albeit with a significant growth disadvantage when tested in vitro (46). Replication of the origin may help rescue such replication defects, although it is still unclear why such large deletions covering genes required for replication are recovered in high-passage-number material.

The U73-to-R3 deletion is more difficult to explain, as many of the genes that are encompassed by the deletion are required for replication in other herpesviruses. We note that repeated library preparations with different library preparation technologies (Nextera XT transposase and Kapa HyperPlus nuclease) gave the same rearrangement. As commercial material, its passage history is not as completely clear. Although not previously formally tested in HHV-6A/B, the helicase-primase complex (containing U74 and U77 in HHV-6A/B) is essential for HSV and HCMV replication, as is the HSV UL9 gene, encoding the homolog of the HHV-6A/B U73 origin-binding protein (44). The HSV UL2 homolog of the HHV-6A/B U75 uracil deglycosylase is not required for replication in vitro (47). HHV-6A/B U82 encodes glycoprotein L, which is required in HSV and HCMV replication (44). HHV-6A/B U90 encodes glycoprotein L, which is required in HSV and HCMV replication (44). The HCMV homologs (UL123 and UL124) of HHV-6A/B U90 and U91 are not required for viral replication (44).

Interestingly, no low-passage-number HHV-6A/B isolate grown in cord blood mononuclear cells showed tandem repeats, despite growing to high titer. These low-passage-number cord blood isolates may provide the best material for HHV-6A/B standard development. Instead, the isolates showed direct-repeat copy numbers that were approximately 60 to 70% that of the unique region. This copy number difference was present in the majority of strains sequenced in this study and suggests that the majority of HHV-6A/B strains sequenced contained only one direct-repeat region repeat, which has previously been seen in HHV-6A/B replication (35, 41, 48, 49).

We also showed the first genomic evidence of interspecies recombination between HHV-6A and HHV-6B strains along a >5.5-kb segment containing the oriLyt and the U42 gene, as well as portions of the U41 and U43 genes. Interspecies recombination is a relatively common feature of the alphaherpesviruses HSV-1 and HSV-2 but has not been described for any other human herpesviruses (17). The origin of the DA recombinant is not entirely clear. Our sequencing results for DA strain recombination are most consistent with a model in which cocultivation of an HHV-6A strain with a laboratory-adapted HHV-6B Z29 strain containing an oriLyt repeat resulted in recombination between the two strains. Strains GS and Z29 were both cultured in cord blood lymphocytes in the laboratory at the same time strain DA was being cultured. The DA strain oriLyt repeat is the exact same length as that of the Z29 strain. However, the recombinant strain also showed unique HHV-6A-like sequence in its HHV-6B oriLyt tandem repeat. Although this sequence fell outside the minimal origin of replication, it is possible that the oriLyt tandem-repeat sequence may have evolved to more effectively interact with HHV-6A replication proteins contained in the rest of the genome. No specific HHV-6B-like sequences were found in the DA strain U73 origin-binding protein to indicate reciprocal U73 evolution to match the HHV-6B-like origin sequences. Despite the high sequence similarity between HHV-6A and HHV-6B, recent analysis of 130 HHV-6B genomes from clinical and iciHHV-6B isolates revealed no evidence of interspecies recombination but widespread intraspecies recombination (20). We note that the ability of HSV-1 and HSV-2 to recombine was first noted in vitro, before the discovery of clinical strain recombinants, which are widespread (17, 50, 51). Whether interspecies recombination will ever be detected in clinical strains of HHV-6A/B remains to be seen.

Previous work from our group demonstrated the loss of almost one-third of the BK and JC polyomavirus genomes in up to 90% of viral species present in multiple viral stocks, including a WHO international standard, likely due to viral passage in simian virus 40 (SV40) T-antigen-immortalized cell lines (33, 34). While clinical PCR tests for HHV-6A/B are unlikely to target the oriLyt region, normalization of quantitative clinical HHV-6A/B testing to any of the loci found here at increased or decreased copy numbers...
could affect quantitation, not least because multiple viral populations were present in many of the reference materials tested. Based on the deletions detected in HHV-6A/B clinical reference materials, primers targeting genes between U25 and U40 and between U43 and U73 would be expected to show no copy number differences. We note that many of the published diagnostic primers target U32 and U67, which both fall within this range of core genes (52). Efforts are under way at the National Institute for Biological Standards and Controls in the United Kingdom to prepare WHO international standard material for both HHV-6A and 6B. We continue to recommend the use of next-generation sequencing to obtain genome-wide single nucleotide resolution copy number measurements in order to validate viral reference materials used in clinical virology and basic-science laboratories around the world.

MATERIALS AND METHODS

HHV-6A/B Reference Strain Copy Number Journal of Virology

TABLE 3 Primer pairs used in this study

| Primer    | Sequence                  |
|-----------|---------------------------|
| PCR of HHV-6B |                           |
| HHV6_68082F | GTTTCACCTAAGCAGGCAG       |
| HHV6_70190R | GAAACGTCAACTCTAAACATGAAAG |
| qPCR       |                           |
| HHV6_ CO4-142970F | GGCAACAAAATTCTCAATATTGGAT |
| HHV6_ CO4-143228R | ATTTGATGTGTACGTTTCCG |
| HHV6A-origin-F | CAGGCGGTGGTGGTTCATTC    |
| HHV6B-origin-F | CGGAGGCCTGGCTTACATCC     |
| HHV6-common-origin-R | GCCTCGACCTTTTTAAATCC  |
| HHV6-quartU32-F | CTTTCATCAACATACTCCTGTTT  |
| HHV6-quartU32-R | GCGTCGACTCCGTATTGATA    |
| HHV6-U32-probe | TGGTCAGTCTATGACGCAGCATTC |

Illumina sequencing library preparation. DNA was extracted from culture isolates using the Zymo viral DNA kit. DNA-sequencing libraries were prepared from 50 ng of genomic DNA using quarter volumes of the Kapa HyperPrep kit with 7 min of fragmentation time and 12 cycles of dual-indexed TruSeq adapter PCR (20). The libraries were sequenced on 2 × 300-bp, 1 × 190-bp, and/or 1 × 192-bp runs on an Illumina MiSeq. The sequences were quality and adapter trimmed using BBduk (http://jgi.doe.gov/data-and-tools/bbtools/) and de novo assembled using SPAdes (58), and contigs were aligned with reference HHV-6A (NC_001664) and HHV-6B (NC_000898) genomes and visualized using Geneious v9.1. Read mapping for copy number analysis was performed against the HHV-6A (NC_001664) and HHV-6B (NC_000898) reference genomes, using the Geneious read mapper with 10% allowed gaps per read, word length of 18, and 20% maximum mismatches per read and with structural variant, insertion, and gap finding allowed.

qPCR confirmation. Quantitative PCR to estimate copy numbers between the origin of replication and the HHV-6A/B U32 locus was performed in 20-μl reaction mixtures using the SsoAdvanced Universal SYBR green SuperMix (Bio-Rad). Three 10-fold dilutions of DNA template from HHV-6A/B strains were tested using quantU32 F-R and species-specific origin primers (Table 3) with cycling conditions of 95°C for 30 s and 40 cycles of 95°C for 5 s and 60°C for 30 s. For the U95 deletion in the CO strains, CO4-142970F and CO-143228R primers (Table 3) were used under the same cycling conditions with
HST-CBMC for origin). Quantities are listed in Table 1.

**PCR and electrophoretic analysis.** PCR across the origin tandem repeat was performed using 1 ng of template genomic DNA in 20-µl total volume reaction mixtures using 10 pmol of each primer and the NEB HiFi Q5 DNA polymerase according to the manufacturer’s instructions. PCR primer sequences are listed in Table 3. PCR mixtures were analyzed with the Genomic DNA ScreenTape assay on an Agilent 2400 TapeStation to visualize the electrophoretic mobility of the PCR products.

**Amplification-free nanopore sequencing.** Nanopore libraries were created using the SQK-RAD002 kit according to Oxford Nanopore Technologies. Amplification-free tagmented libraries were run according to Oxford Nanopore protocols v1.3.24 on a singular Mk1 (R9.4) FLO-MIN106 flow cell. Nanopore reads were mapped to the HHV-6B Z29 reference genome (NC_000898) using the LASTZ and Geneious Aligner to screen for origin-containing reads (S6, 57).

**Accession number(s).** The sequences from this study are available in NCBI GenBank (accession numbers MF994813 to MF994829) and associated with NCBI BioProject 338014.

**REFERENCES**

1. Salahuddin SZ, Ablashi DV, Markham PD, Josephs SF, Sturzenegger S, Kaplan M, Halligan G, Biberfeld P, Wong-Staal F, Kramerksy B. 1986. Isolation of a new virus, HBLV, in patients with lymphoproliferative disorders. Science 234:596–601. https://doi.org/10.1126/science.2876520.

2. Ablashi DV, Agut H, Alvarez-Lafuente R, Clark DA, Dhurwasth S, Diluka D, Flamand L, Frenkel N, Gallo R, Gompels UA, Hollberg P, Jacobsen S, Luppi M, Lusso P, Malnati M, Medveczky P, Mori Y, Pellett PE, Pritchett JC, Yamanishi K, Yoshikawa T. 2008. Classification of HHV-6A and HHV-6B as distinct viruses. Arch Virol 153:863–870. https://doi.org/10.1007/s00705-008-1092-5.

3. Fillet AM, Raphael M, Visse B, Audouin J, Poirel L, Agut H. 1996. Controlled study of human herpesvirus 6 detection in acquired immunodeficiency syndrome-associated non-Hodgkin’s lymphoma. French Study Group for HHV-Associated Tumors. J Med Virol 45:106–112.

4. Cone RW, Huang ML, Hackman RC, Corey L. 1996. Coinfection with human herpesvirus 6A variants A and B in lung tissue. J Clin Microbiol 34:877–881.

5. Zerr DM, Meier AS, Selke SS, Frenkel LM, Huang ML, Wald A, Roops MP, Nygu L, Bornemann R, Morrow RA, Corey L. 1995. A population-based study of primary human herpes virus 6 infection. N Engl J Med 332:768–776. https://doi.org/10.1056/NEJMoa9502207.

6. Agut H, Bonnafous P, Gautheret-Dejean A. 2015. Laboratory and clinical aspects of human herpesvirus 6 infections. Clin Microbiol Rev 28:313–335. https://doi.org/10.1128/CMR.00122-14.

7. Pellett PE, Ablashi DV, Ambros PF, Agut H, Caserta MT, Descanps V, Flamand L, Gautheret-Dejean A, Hall CB, Kambite RT, Kuehl U, Lassner D, Lautenschlager I, Loomis KS, Luppi M, Lusso P, Medveczky PM, Montoya JG, Mori Y, Ogata M, Pritchett JC, Rogez S, Seto E, Ward KN, Yoshikawa T, Razonable RR. 2012. Chromosomally integrated human herpesvirus 6: questions and answers. Rev Med Virol 22:144–155. https://doi.org/10.1002/rev.715.

8. Daibata M, Taguchi T, Nemoto Y, Taguchi H, Miyoshi I. 1999. Inheritance of chromosomally integrated human herpesvirus 6 DNA. Blood 94:1545–1549.

9. Krug LT, Pellett PE. 2014. Roseolovirus molecular biology: recent advances. Curr Opin Virol 9:170–177. https://doi.org/10.1016/j.coviro.2014.01.004.

10. Clark DA, Ward KN. 2008. Importance of chromosomally integrated HHV-6A and -6B in the diagnosis of active HHV-6 infection. Herpes 15:28–32.

11. Clark DA. 2016. Clinical and laboratory features of human herpesvirus 6 chromosomal integration. Clin Microbiol Infect 22:333–339. https://doi.org/10.1016/j.cmi.2015.12.022.

12. Tweedy J, Spyrou MA, Donaldson CD, Depledge D, Breuer J, Gompels UA. 2015. Complete genome sequence of the human herpesvirus 6A strain AJ from Africa resembles strain GS from North America. Genome Announc 3:e01498-14. https://doi.org/10.1128/genomeA.01498-14.

13. Tweedy J, Spyrou MA, Pearson M, Lassner D, Kuhl U, Gompels UA. 2016. Complete genome sequence of germline chromosomally integrated human herpesvirus 6A and analyses integration sites define a new human endogenous virus with potential to reactivate as an emerging infection. Viruses 8:E19. https://doi.org/10.3390/v8010019.

14. Dominguez G, Dambaugh TR, Stamay FR, Dhurwasth S, Inoue N, Pellett PE. 1999. Human herpesvirus 6B genome sequence: coding content and comparison with human herpesvirus 6A. J Virol 73:8040–8052.

15. Gompels UA, Nicholas J, Lawrence G, Jones M, Thomson BJ, Martin MED, Efstratiou S, Craxton M, Macaulay H. 1995. The DNA sequence of human herpesvirus-6: structure, coding content, and genome evolution. Virology 209:29–51. https://doi.org/10.1006/viro.1995.1228.

16. Isegawa Y, Mukai T, Nakano K, Kagawa M, Chen J, Mori Y, Sunagawa T, Kawanishi K, Sashihara J, Hata A, Zou P, Kosuge H, Yamanishi K. 1999. Comparison of the complete DNA sequences of human herpesvirus 6A variants A and B and a J Virol 73:8035–8043.

17. Koelle DM, Norberg P, Fitzgibbon MP, Russell RM, Greninger AL, Huang ML, Stensland L, Jing L, Magaret AS, Diem K, Selke S, Xie H, Celum C, Lingappa JR, Jerome KR, Wald A, Johnston C. 2017. Worldwide circulation of HSV-2 × HSV-1 recombinant strains. Sci Rep 7:40804. https://doi.org/10.1038/srep40804.

18. Burrel S, Boutolleau D, Ryu D, Agut H, Merkel K, Leendertz FH, Calvignac-Spencer S. 2017. Ancient recombination events between human herpes simplex viruses. Mol Biol Evol 34:1713–1721. https://doi.org/10.1093/molbev/msx113.

19. Zhang E, Bell AJ, Willkie GS, Suarez NM, Batini C, Veal CD, Arrondiz-Castillo I, Neumann R, Cotton VE, Huang Y, Porteous DJ, Jarrett RF, Davison AJ, Royle NJ. 2017. Inherited chromosomally integrated human herpesvirus 6 genomes are ancient, intact and potentially able to reactivate from telomeres. J Virol 91:e01137-17. https://doi.org/10.1128/JVI.01137-17.

20. Greninger AL, Knudsen GM, Roychoudhury P, Hanson DJ, Sandmaier BM, Hansen JA, Magaret AS, Hill JA, Jerome KR. 2018. Comparative genomic, transcriptomic, and proteomic reannotation of human herpesvirus 6. BMC Genomics 19:204.

21. Boutolleau D, Fernandez C, Andre E, Ibbert-Marcell B-M, Milpied N, Agut H, Gautheret-Dejean A. 2003. Human herpesvirus (HHV)-6 and HHV-7: two closely related viruses with different infection profiles in stem cell transplantation recipients. J Infect Dis 187:179–186. https://doi.org/10.1086/376677.

22. Hill JA, Magaret AS, Hall-Sedlak R, Mikhailova A, Huang ML, Sandmaier BM, Hansen JA, Jerome KR, Zerr DM, Boecht M. 2017. Outcomes of hematopoietic cell transplantation using donors or recipients with inherited chromosomally integrated HHV-6B. Blood 130:1062–1069. https://doi.org/10.1182/blood-2017-03-775759.

23. Ogata M, Kikuchi H, Satou T, Kawano R, Ikewaki J, Kohno K, Kashima K, Ohtsuka E, Kadota J-I. 2006. Human herpesvirus 6 DNA in plasma after hematopoietic cell transplantation using donors or recipients with inherited chromosomally integrated HHV-6. Blood 108:1062–1069. https://doi.org/10.1182/blood-2005-10-42531.

24. Hayden RT, Sun Y, Tang L, Procop GW, Hillyard DR, Pinsky BA, Young SA, Caliendo AM. 2017. Progress in quantitative viral load testing: variability

May 2018 Volume 92 Issue 10 e00135-18 jvi.asm.org
and impact of the WHO Quantitative International Standards. J Clin Microbiol 55:423–430. https://doi.org/10.1128/JCM.02044-16.

25. de Pagter PJ, Schuerman R, de Vos NM, Mackay W, van Loon AM. 2010. Multicenter external quality assessment of molecular methods for detection of human herpesvirus 6. J Clin Microbiol 48:2536–2540. https://doi.org/10.1128/JCM.01145-09.

26. Hayden RT, Shahbazian MD, Valsamakis A, Boonyaratakanokkit J, Cook L, Pang XL, Preiksaitis JK, Schönbrunner ER, Caliendo AM. 2013. Multicenter evaluation of a commercial cytomegalovirus quantitative standard: effects of commutability on interlaboratory concordance. J Clin Microbiol 51:3811–3817. https://doi.org/10.1128/JCM.02036-13.

27. Hayden RT, Yan X, Wick MT, Rodriguez AB, Xiong X, Ginocchio CC, Mitchell MJ, Caliendo AM, College of American Pathologists Microbiology Resource Committee. 2012. Factors contributing to variability of quantitative viral PCR results in proficiency testing samples: a multivariate analysis. J Clin Microbiol 50:3337–345. https://doi.org/10.1128/JCM.01495-15.

28. Hayden RT, Preiksaitis J, Tong Y, Pang X, Sun Y, Tang L, Cook L, Pounds S, Fryer J, Caliendo AM. 2015. Commutability of the first World Health Organization international standard for human cytomegalovirus. J Clin Microbiol 53:3325–3333. https://doi.org/10.1128/JCM.00906-15.

29. Holden MJ, Mađei RM, Minor P, Kalman LV. 2011. Molecular diagnostics: harmonization through reference materials, documentary standards and proficiency testing. Expert Rev Mol Diagn 11:741–755. https://doi.org/10.1586/erm.11.50.

30. Stemay FR, Dominguez G, Black JB, Dambaugh TR, Pellett PE. 1995. Intragenomic linear amplification of human herpesvirus 6B oriLyt sug- gests acquisition of oriLyt by transposition. J Virol 69:589–596.

31. Dewhurst S, Krenitsky DM, Dykes C. 1994. Human herpesvirus 6B origin: sequence diversity, requirement for two binding sites for origin-binding protein, and enhanced replication from origin multimers. J Virol 68:6799–6803.

32. Greninger AL, Bateman AC, Atienza EE, Wendent S, Makhsoos N, Jerome KR, Cook L. 2017. Copy number heterogeneity of JC virus standards. J Clin Microbiol 55:834–835. https://doi.org/10.1128/JCM.00237-16.

33. Bateman AC, Greninger AL, Atienza EE, Limaye AP, Jerome KR, Cook L. 2017. Quantification of BK virus standards by quantitative real-time PCR and droplet digital PCR is confounded by multiple virus populations in the WHO BKV international standard. Clin Chem 63:761–769. https://doi.org/10.1373/chinchem.2016.265512.

34. Borenstein R, Zeigerman H, Frenkel N. 2010. The DR1 and DR6 first exons of human herpesvirus-6A (GS strain) isolated from North America. Genome Announc 1:e00012-13. https://doi.org/10.1128/GA.00012-13.

35. Krueger GR, Sander C, Hoffmann A, Barth A, Koch B, Braun M. 1991. Isolation of human herpesvirus-6 (HHV-6) from patients with collagen vascular diseases. In Vivo 5:217–225.

36. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pyriev RR, et al. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19:455–477. https://doi.org/10.1089/cmb.2012.0021.