Unbiased optical mapping of telomere-integrated endogenous human herpesvirus 6

Darren J. Wight[a,1], Giulia Aimola[a], Amr Aswad[a], Chi-Yu Jill Lai[b], Christian Bahamon[b], Karl Hong[b], Joshua A. Hill[c,d], and Benedikt B. Kaufer[a,1]

[a]Institut für Virologie, Freie Universität Berlin, 14163 Berlin, Germany; [b]Bionano Genomics, San Diego, CA 92121; [c]Department of Medicine, University of Washington, Seattle, WA 98195-6420; and [d]Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Research Center, Seattle, WA 98109-1024

Edited by Thomas Shenk, Princeton University, Princeton, NJ, and approved October 27, 2020 (received for review June 10, 2020)

Next-generation sequencing technologies allowed sequencing of thousands of genomes. However, there are genomic regions that remain difficult to characterize, including telomeres, centromeres, and other low-complexity regions, as well as transposable elements and endogenous viruses. Human herpesvirus 6A and 6B (HHV-6A and HHV-6B) are closely related viruses that infect most humans and can integrate their genomes into the telomeres of infected cells. Integration also occurs in germ cells, meaning that the virus can be inherited and result in individuals harboring the virus in every cell of their body. The integrated virus can reactivate and cause disease in humans. While it is well established that the virus resides in the telomere region, the integration locus is poorly defined due to the low sequence complexity (TTAGGG)n of telomeres that cannot be easily resolved through sequencing. We therefore employed genome imaging of the integrated HHV-6A and HHV-6B genomes using whole-genome optical site mapping technology. Using this technology, we identified which chromosome arm harbors the virus genome and obtained a high-resolution map of the integration loci of multiple patients. Surprisingly, this revealed long telomere sequences at the virus–subtelomere junction that were previously missed using PCR-based approaches. Contrary to what was previously thought, our technique revealed that the telomere lengths of chromosomes harboring the integrated virus genome were comparable to the other chromosomes. Taken together, our data shed light on the genetic structure of the HHV-6A and HHV-6B integration locus, demonstrating the utility of optical mapping for the analysis of genomic regions that are difficult to sequence.

human herpesvirus 6 | iciHHV-6 | telomere integration | structural genomic mapping | virus integration

Human herpesvirus 6A and 6B (HHV-6A and HHV-6B) are closely related virus species that infect humans (1). HHV-6B infects almost all humans within the first years of life and causes the febrile illness *exanthema subitum* (2). Infection with HHV-6A is thought to occur later in life, but the epidemiology of the virus is poorly characterized. Like all herpesviruses, HHV-6A and HHV-6B establish latency upon primary infection, allowing the virus to persist in the host for life. In contrast to other human herpesviruses, HHV-6A and HHV-6B integrate their genomes into the telomere region of host chromosomes of latently infected cells (3–6). The viruses can also heritably integrate into germ cells, resulting in individuals and their offspring that harbor the virus in every nucleated cell (3, 7–11). This phenomenon is referred to as inherited chromosomally integrated HHV-6 (iciHHV-6). Importantly, HHV-6 reactivation from the integrated state is associated with various diseases, including graft-versus-host disease, encephalitis, and heart disease (12–18). More recently, a study indicated that telomeres carrying integrated HHV-6 are shorter and more unstable (19), which, in turn, could influence aging and/or diseases of individuals with iciHHV-6.

It remains unknown whether certain integration events or iciHHV-6 genomes at specific loci are responsible for, or contribute directly to, disease in humans. Although we can identify individuals that harbor iciHHV-6A and iciHHV-6B through qPCR (20), a major challenge in studying iciHHV-6 is that we cannot readily examine the integration site with high resolution. Integration loci of HHV-6A and HHV-6B were first detected by fluorescent in situ hybridization (FISH) (3), and the junction has since been sequenced using a PCR-based approach for three different integrations. Sanger sequencing of these PCR fragments indicated that the right direct repeat (DR-R) is fused to the subtelomeres with only a very short stretch of telomere sequences between the virus and host chromosome (4, 5, 19, 21).

However, FISH is a challenging and laborious technique, offering very low spatial resolution that does not allow us to directly observe the junction. Although PCR amplification and Sanger sequencing does offer sequence information, this approach requires previous knowledge of the chromosomal location of the virus genome (19, 22), is prone to amplification errors, and has not succeeded in determining structures of junctions. Next-generation sequencing (NGS) using various platforms has not succeeded in sequencing the junction, as either the coverage or the read length was not sufficient to resolve the complex and repetitive nature of the telomere region harboring the virus genome.

We therefore set out to develop an approach to study the virus–host junction because understanding its composition is key to understanding the effect of iciHHV-6 on human biology. Using a whole-genome optical mapping technology (23, 24), we were able to generate unbiased, high-resolution maps of HHV-6A

Significance

Low-complexity and repetitive elements are difficult to study using existing sequencing technologies. Determination of the boundaries and structures of the termini of inherited chromosomally integrated HHV-6 (iciHHV-6) genomes is particularly challenging, as it integrates into highly repetitive human telomeres. We therefore developed a genome imaging approach that revealed the chromosomal location of the virus, its orientation, the presence of long internal telomeres at the host–virus junction, and the lengths of the distal telomeres capping the integrated virus genome. This genome imaging approach has wide applications for mapping transposable elements or large viruses that integrate into low-complexity regions of host genomes.

Author contributions: D.J.W. and B.B.K. designed research; D.J.W., G.A., C.-Y.J.L., C.B., and K.H. performed research; A.A. and J.A.H. contributed new reagents/analytic tools; D.J.W., C.-Y.J.L., C.B., and K.H. analyzed data; and D.J.W., A.A., and B.B.K. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. Email: d.j.wight@gmail.com or b.kaufer@fu-berlin.de.

First published November 23, 2020.

and HHV-6B integration sites, which allowed the identification of the chromosomes harboring the integrated HHV-6A and HHV-6B genome in four different iciHHV-6 patient-derived cell lines. We also know, from our recent work on the evolutionary history of the virus, that most known integrations are derived from very few ancestral genome invasion events (25). This means that, for every sample we determine the chromosomal location for, this can be extrapolated to many orthologs without the expense and effort of further experimental work. Our results conclusively determine the orientation of the virus genome, as well as the length of the telomeric virus–chromosome junction and the distal telomeres on the end of the virus genome. Furthermore, the ability to determine the chromosomal location of the virus will be crucial to decipher the role of the integrated virus genome in human diseases.

Results

Genome Imaging of iciHHV-6. To map the integrated HHV-6A and HHV-6B genome and overcome the obstacle of repetitive sequences at the integration site, we used a genome imaging approach. High molecular weight DNA was isolated from four iciHHV-6 patient cell lines and labeled at direct labeling enzyme 1 (DLE-1) sites via covalent bonding of fluorophores at specific interspersed sequences throughout the genome (Fig. 1*A*). Labeled DNA was counterstained with a general DNA fluorescent marker, loaded into a flow cell, and electrophoresed through nanochannels. Fluorescence from labeled individual DNA molecules was scanned, resulting in long double-stranded DNA (dsDNA) maps that were aligned to the GRCh38 human genome. Based on the predicted DLE-1 fluorophore acceptor sites in the iciHHV-6 genomes (Fig. 1*B*), this process allowed the identification of HHV-6 integration sites at chromosomes 18q, 19q, Xp, and 18p (Fig. 2). The integration in chromosome Xp (Fig. 2*D*) represents a report of HHV-6 integration into a sex chromosome.

To validate our genome imaging approach, we performed FISH to confirm the location of each integration. We used probes against HHV-6 as well as chromosome-specific probes to confirm the integration loci for ici6A-1 and ici6B-1, which were in 18q and 19q, respectively (Fig. 3*A*). The size and structure of the virus-containing chromosomes of ici6B-2 and ici6B-3 were also consistent with Xp and 18p, respectively (Fig. 3*B*). The optical maps also revealed that all virus genomes had integrated with the DR-R facing toward the centromere.

HHV-6 Integration and Telomere Length. Next, we investigated the length of the telomeric repeats at the terminal ends of the HHV-6 genome and the virus–subtelomere junction. To characterize the subtelomere junction for these samples, we combined the results of optical mapping with existing NGS sequence data of the integrated virus (which cannot bridge the junction) and human genome. Using the median lengths of the individual optical map molecules, we calculated the size of the virus–host junctions (Fig. 4). Our data indicate that there are long telomere stretches of 5.0 kb to 13.3 kb between the subtelomeres and the virus genome (Fig. 4). For one integration site, the length of the internal telomere was shorter than the respective chromosome pair end, suggesting that HHV-6 integrated within the telomere repeats (Fig. 5). In the case of the other integration sites, length of these internal telomeres corresponds to physiologically normal human telomeres, suggesting that HHV-6 integration occurred at or near the end of the telomere. Consistently, the internal telomeres were as long as or even longer than the telomere of the sister chromosome lacking the virus genome (Fig. 5).

In addition, we assessed the length of the telomeres at the distal end of the iciHHV-6 genome in our patient cells by measuring the length of all of the individual molecules that extend out from the left DR (DR-L) and taking the median length (Fig. 6). In general, it is challenging to measure telomere length, due to the repetitive nature of telomeres. Quantification of the telomeres at the distal end of the virus genome for all iciHHV-6 individuals revealed lengths ranging from 8.9 kb to 14.7 kb (Fig. 6), consistent with normal human telomeres (26, 27). Taken together, this approach to virus integration mapping was able to quantify distal telomere lengths in iciHHV-6 individuals and showed they have sizes comparable with regular human telomeres.

Discussion

The HHV-6A and HHV-6B integration mechanism and any physiological/pathological consequences for iciHHV-6 individuals remains poorly understood (28–30). The ability to accurately determine the virus's location in the genome is a crucial first step toward addressing these open questions. We therefore applied an unbiased whole-genome optical mapping strategy that

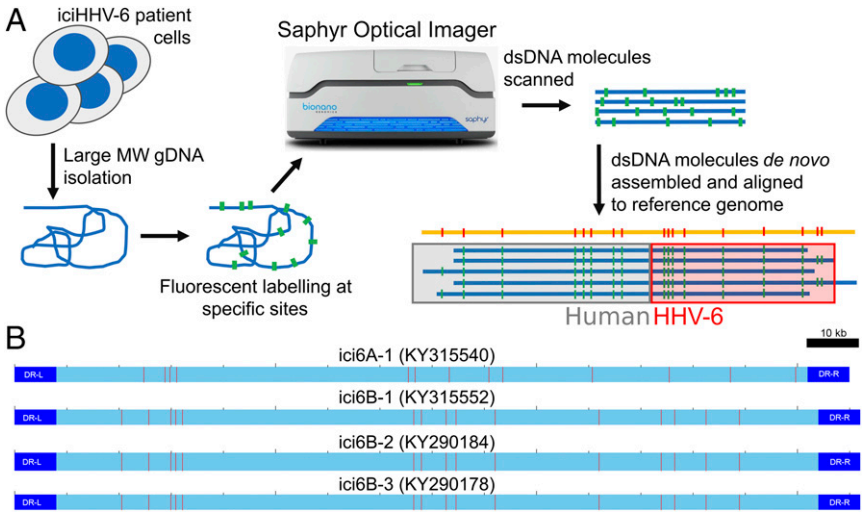

Fig. 1. Mapping iciHHV-6 integrations using genome imaging technology. (*A*) Simplified overview of the workflow to create whole-genome–length optical maps from patient cells with iciHHV-6. After alignment of the optical maps to the reference human genome, location, orientation, and information about the virus–host junction should be revealed. (*B*) DLE-I fluorophore acceptor sites in the four iciHHV-6 studied. Direct repeats (DR) are annotated as the first and final 8 kb of HHV-6 sequence for orientation purposes. gDNA: genomic DNA.

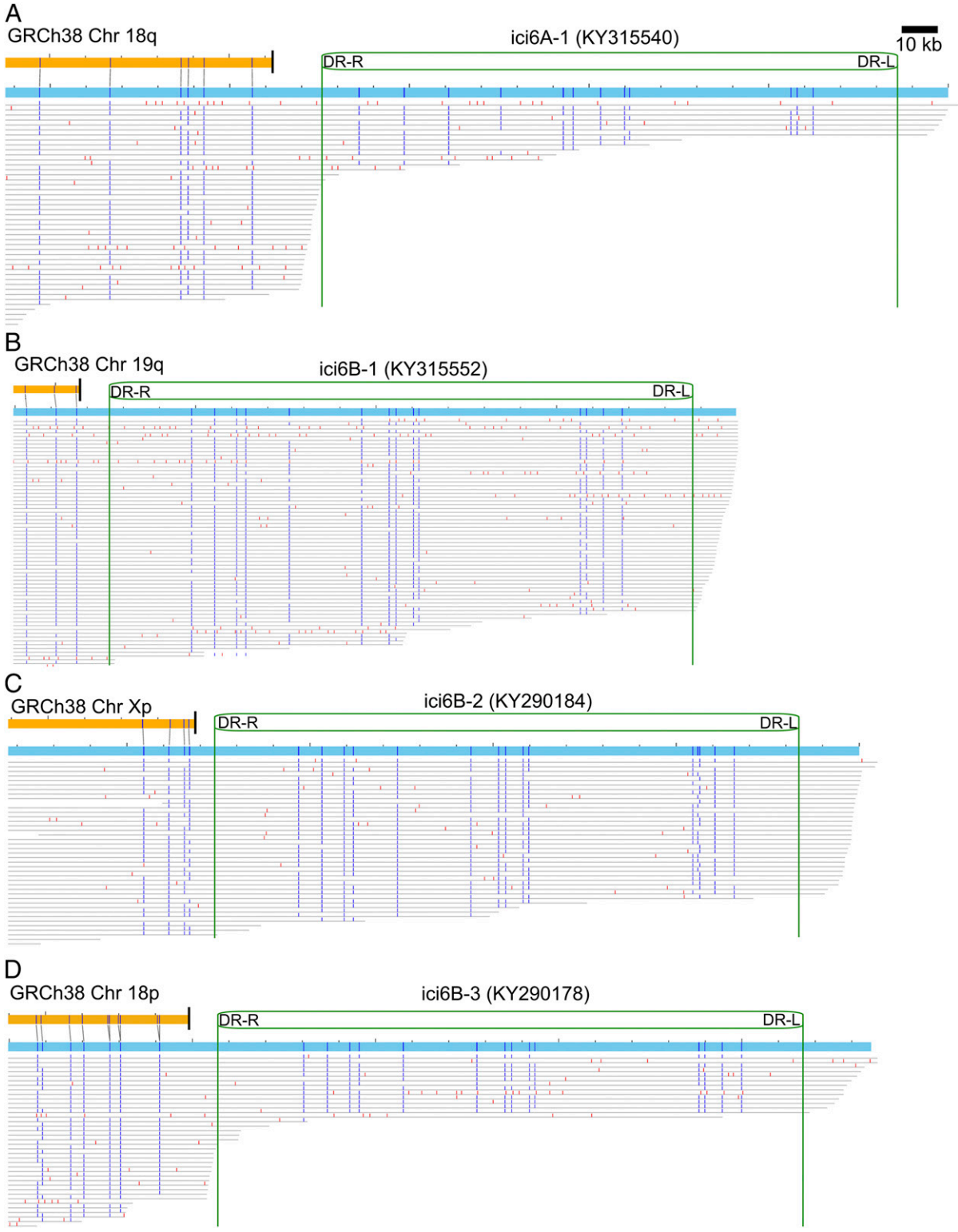

Fig. 2. Genome imaging reveals structure of telomere-integrated iciHHV-6. Optical consensus maps (light blue ruler) were constructed from the scanned dsDNA molecules and were aligned to human GRCh38 reference genome (orange ruler) with the terminal Ns removed. Shown are the integration sites on an assembled optical genome map for (*A*) ici6A-1 at 18q, (*B*) ici6B-1 at 19q, (*C*) ici6B-2 at Xp, and (*D*) ici6B-3 at 18p. Below the consensus maps are the individual dsDNA molecules that were used to build the map (gray lines). Blue or red marks indicate DLE-I sites (marked with fluorophores in the samples), with the red marks not matching to the reference genome. Green lines indicate the positions of the iciHHV-6 genomes, and the starts of the telomere repeats in the GRCh38 annotation are marked with a black line on the orange map. Virus genome size was inferred from the NGS data (45) and is annotated with green boxes and lines.

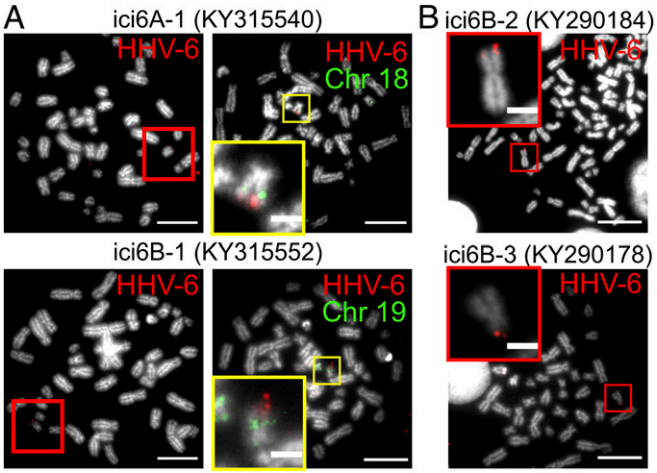

Fig. 3. Mapping iciHHV-6 integration using FISH. B-LCLs from patients carrying iciHHV-6A and iciHHV-6B were fixed and stained with FISH probes against the HHV-6 genome (red) with or without an additional probe against a specific human chromosome (green). (*A*) One iciHHV-6A and one iciHHV-6B integration that were mapped using dual FISH staining. (*B*) Due to lack of specific human chromosome probes, these iciHHV-6B samples could not be mapped to specific chromosomes. Red boxes highlight chromosomes with integrated HHV-6, and yellow marks out dual-FISH-stained chromosomes. (Scale bars: main images, 10 μm; zoomed images, 2 μm.)

produces "fingerprint"-like maps of the human chromosomes to identify and visualize the entire length of the telomere-integrated HHV-6, including the virus–subtelomere junction. This allowed us to identify the chromosomal arm in which the virus has integrated for each sample, and visually corroborate our results with FISH. Our data conclusively reveal that, in all samples studied, the virus is oriented with the DR-R situated at the virus–chromosome junction, and is consistent with previous data using a PCR-based approach (4, 5, 19, 21) and the current models of HHV-6 integration (1, 31). Contrary to previous findings, our data conclusively show that the virus–subtelomere junctions are extremely long, ranging from 5.0 kb to 13.3 kb (Fig. 4).

Many recent advances have been made to sequencing technology, especially in the long-read sequencing domain (third-generation sequencing). Notable technologies in this field are Oxford Nanopore Technology (ONT) and single-molecule real-time sequencing, which both use single-stranded DNA as a template and detect the specific bases. While these approaches mark a giant leap forward for genomics, challenges remain regarding the reproducibility of read lengths, high error rates, and low coverage (difficulty in obtaining overlapping sequencing reads). This was exemplified by a recent study that used ONT to sequence the iciHHV-6–host junction (32). Although successful, it is important to note that only two reads were obtained across a relatively small junction of 1.3 kb. Our results here have shown that, even in our small dataset, the junction region is far longer and that we obtained a much higher coverage compared to the ONT approach. Therefore, while sequencing approaches are a necessary element to the study of HHV-6 integration, optical mapping offers an additional source of reliable data.

We recently analyzed the evolutionary history of over 250 HHV-6 and iciHHV-6 sequences (25). Our phylogenetic reconstruction revealed that nearly 80% of the known iciHHV-6 sequences cluster in a handful of clades that represent single ancestral integration events, including KY315540, KY290178, and KY315552. Using optical mapping technology has thus enabled us to determine the chromosomal location of many more iciHHV-6 sequences in a more reliable and less laborious way than with FISH. Moreover, the technique opens the opportunity to study the differences between sites of the same ancestral integration, which may reveal changes to the junction that occurred over time, or as a result of variable environmental/genetic factors.

This study also demonstrates an extremely useful tool for examining similarly difficult-to-sequence regions of the genome. This approach has clear applications for telomere biology and endogenous viruses, transposable elements, and other low-complexity regions. These genomic elements/features are not only difficult to sequence from a biochemical standpoint but are notorious for being extremely challenging at the genomic assembly stage. We have shown that, although optical mapping technologies do not reveal the target sequence, we can combine these results with sequence data to provide a clearer picture than either approach can offer alone. For instance, the endogenous retrovirus (ERV) group HERV-K (HML-2) is a human ERV that is insertionally polymorphic; that is, not all people have the same insertions in the same loci (33). It has been shown that certain HML-2 insertions can influence nearby genes, such as the RASGRF2 involved in dopaminergic activity. People carrying an HML-2 insertion within an intron of RASGRF2 are significantly more likely to be intravenous drug users (34). This highlights the importance of being able to accurately map the location of neglected genomic elements. Since, like with HHV-6A and HHV-6B, we know the sequence of the HML-2 insertion, we could easily identify their optical mapping signature from these data.

The length of the virus–host junction that we show here provides important insights into the integration mechanism, which, thus far, remains poorly understood. The fact that internal telomeric repeats are much longer than the telomeric repeats present in the virus genome could indicate that the host telomere is the main contributor for the sequences in the junction. This suggests that the virus genome integration occurs either at the telomere ends or at least quite distant from the subtelomere (Fig. 4). Another interpretation is that the length of the junction is a side effect of the mechanism itself, whereby some telomere lengthening process is involved at the point of genome invasion.

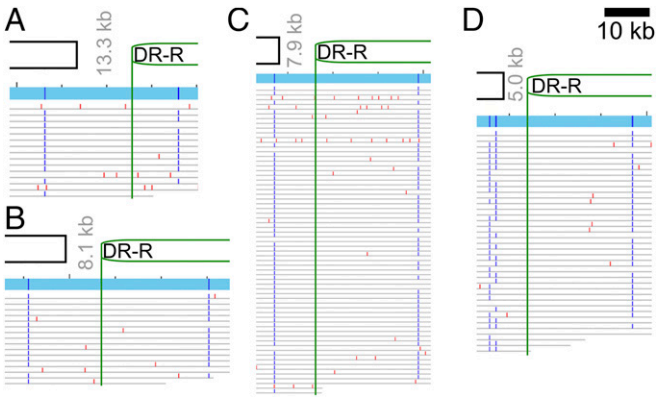

Fig. 4. Host–virus internal telomere from iciHHV-6 individuals. Detailed views of the host–virus junction with the molecules spanning these regions are shown for (*A*) ici6A-1 (KY315540) at 18q, (*B*) ici6B-3 (KY290178) at 18p, (*C*) ici6B-1 (KY315552) at 19q, and (*D*) ici6B-2 (KY290184) at Xp. In gray text are the median sizes from the host telomere repeats up to the virus genome. Below the annotations are the individual dsDNA molecules that were used to build the maps. Blue or red marks indicate DLE-I sites (marked with fluorophores in the samples), with the red marks not matching to the reference genome. Black boxes mark out the end of the subtelomere region of the host and the beginning of the telomeres in the GRCh38 human genome annotation. Virus genome positions are annotated with green boxes and lines.

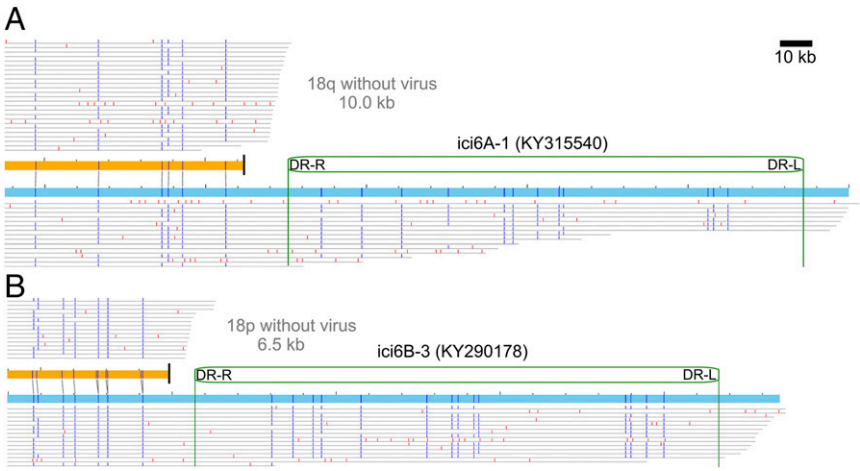

Fig. 5. HHV-6–host junction and native telomere length. As iciHHV-6 is a heterozygous event, the optical maps also contain the reads obtained for the terminal portion of the chromosome pair without HHV-6 integration. The telomere length of the chromosome without HHV-6 and the host–virus junction are shown for (*A*) ici6A-1 at 18q and (*B*) ici6B-3 at 18p. Median size of the telomere ends for the chromosomes without a virus integration are shown in gray, above the reference genome. Color scheme is the same as in Fig. 2.

It could also be that the virus preferentially targets certain chromosomal ends based on the length of their telomere.

The data also indicate that the telomeric repeat at the distal end of the virus must then act as the chromosome's telomere, and has undergone extension to a length corresponding to "normal" human telomeres. Human telomeres range in size from 8 kb to 15 kb (26, 27) and are maintained in germline cells by the telomerase enzyme complex (35–39). Consistent with this, we show that the distal telomeres of iciHHV-6 individuals are around 10 kb. This suggests that the imperfect telomere repeats on DR-L likely serve as a template for the formation of a neo-telomere at the distal end. The telomerase complex could mediate this process, as it has recently been shown to contribute to the integration process (40), but more work is required to fully understand this pathway. Important to note is that the cells used in this study were immortalized with Epstein–Barr virus (EBV), which can induce telomerase activity (41). This suggests that the length of the terminal telomeres measured in our experiments rather reflects their length in telomerase positive cells in the human body. Somatic cells in the human body do not exhibit telomerase activity and could have overall shorter telomeres.

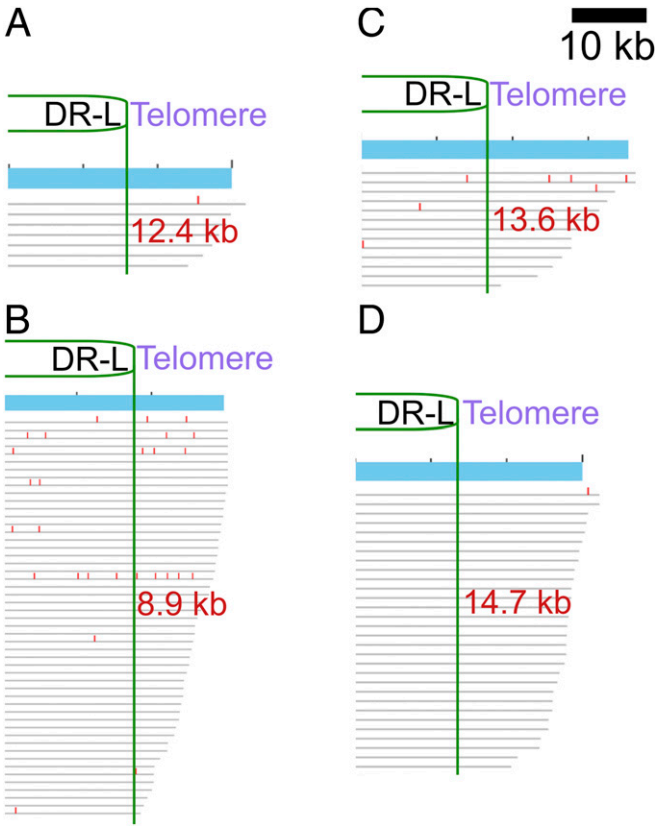

Fig. 6. Distal telomeres of iciHHV-6 chromosomes. Molecules making up the optical maps that cover the terminal parts of iciHHV-6 genomes are shown for (*A*) ici6A-1 (KY315540), (*B*) ici6B-1 (KY315552), (*C*) ici6B-3 (KY290178), and (*D*) ici6B-2 (KY290184). Median telomere length is shown for all external HHV-6 telomeres in red text.

An essential function of the telomere is to disguise the chromosome end from the DNA damage response, which is accomplished by the shelterin complex and self-hybridization of the single-stranded telomere end into the upstream telomere repeats (42, 43). It is therefore interesting to speculate that perhaps these internal telomere length differences between samples may mirror these self-hybridization points of the telomere overhang at the time HHV-6 integrated into the genome. A deeper look at a larger number of iciHHV-6 junctions in combination with analysis during the integration process is needed to confirm whether this is the case.

Materials and Methods

Patient Cells and Culture. Samples used for this study were derived from healthy stem cell donors (n = 3) and a patient with acute lymphoblastic leukemia (KY290178) who were identified and described previously (16). The donor with the X chromosome integration was a female. Peripheral blood mononuclear cells (PBMCs) from iciHHV-6 individuals were immortalized with EBV to obtain lymphoblastoid cell lines (B-LCLs). The B-LCLs were grown to a concentration of ~5 million cells per mL at 37 °C in Roswell Park Memorial Institute medium (RPMI) 1640 with L-glutamine culture medium (Gibco) supplemented with 15% fetal bovine serum (Gibco/Invitrogen Corp) and 1% 100 mM sodium pyruvate (Gibco). Cell lines were frozen in liquid nitrogen until further use. Use of these samples was approved by the Fred Hutchinson Cancer Research Center Institutional Review Board. The samples were deidentified prior to use in this study.

FISH. FISH staining of the HHV-6 genome and/or cellular chromosomes was performed using two specific set of probes: HHV-6 probes were labeled using Biotin-High prime (Sigma Aldrich), and detection of the probe signal

Table 1. Specifications and coverage of the data collected from the Bionano Saphyr system

Sample	Data collected (Gbp; molecules > 150 kbp)	Median molecule size (kbp)	Median genome map size (Mbp)*	Assembly size (Gbp)†
ici6A-1 (KY315540)	453	350.3	67.1	5.89
ici6B-1 (KY315552)	438	344.9	69.6	5.92
ici6B-2 (KY290184)	702	348.7	70.4	5.85
ici6B-3 (KY290178)	432	299.7	59.6	5.88

*Median genome map size: median size of the de novo assembled genome maps created from the individual molecules.
†Assembly size: total size of the diploid de novo assembly created from the generated genome maps.

was achieved using Cy3-Streptavidin (1:200; Roche); chromosome-specific probes were labeled using Digoxigenin-High prime (DIG-High prime) (1:1,000; Sigma Aldrich), and probe signal was detected with anti-DIG fluorescein isothiocyanate Fab fragments (GE healthcare). The FISH protocol was performed as described previously (30, 44).

DNA Preparation. High molecular weight dsDNA was isolated from viable human B-LCLs as outlined below. Cryopreserved cells were thawed in a 37 °C water bath, and a volume containing 1.5 million cells was centrifuged at 2,000 × *g* for 2 min at 4 °C. Each cell pellet was washed in 500 μL of cold Bionano DNA Stabilizing Buffer, containing 2% (vol/vol) DNA Stabilizer (Bionano Genomics) in Cell Buffer (Bionano Genomics). The washed cells were centrifuged a second time, as described above, and DNA extraction was performed using an SP Blood & Cell DNA Isolation Kit (Bionano Genomics), following the Prep SP Frozen Cell Pellet DNA Isolation Protocol. In short, the washed cell pellets were resuspended in DNA Stabilizing Buffer (Bionano Genomics) and lysed in the presence of detergents, proteinase K, and RNase A. High molecular weight dsDNA was bound to a silica disk, washed, and eluted.

Optical Molecule Generation. For each cell line, 750 ng of high molecular weight dsDNA was fluorescently labeled using a Bionano Prep DNA Labeling Kit-DLS according to the Prep Direct Label and Stain (DLS) Protocol. The enzyme DLE-1 (Bionano Genomics) was used to fluorescently tag the motif CTTAAG, and the dsDNA backbone was counterstained with DNA Stain (Bionano Genomics). Targeting at least 480 Gbp of data per sample, high-throughput imaging was performed using Bionano Genomics second-generation Saphyr Systems and chips with Instrument Control Software version 4.7.18339.1. Metadata regarding the optical molecules and alignments generated from these experiments can be found in Table 1.

Alignment of Molecules and Generation of Genome Maps. The HHV-6 genomes were in silico digested with motif sequence CTTAAG of DLE-1 enzyme. To detect the integration site of HHV-6, the labeled HHV-6 genomes were concatenated with human hg38 genome and used as a reference genome for subsequence analysis. The genome maps of each cell line were then generated via haplotype-aware de novo assembly of single-molecule reads with the following software suite: Bionano Access 1.4.3, Tools 1.4.3, Solve 3.4.1, and RefAligner 10026.10020rel. The HHV-6 integration sites were identified by regions in the cell line genome maps that align to both human hg38 genome and the viral genome maps.

Genetic Map Analysis. Genetic maps and aligned molecules were exported from Bionano Access software (v1.5) as png image files. HHV-6 genomes were annotated on the maps based on the distance to the last DLE-1 mark and the known length in kilobases. Pixel length measurements were performed in Fiji (ImageJ) and converted into kilobases using the measured scale. Median lengths of all molecules measured are displayed in the manuscript to reduce the influence of outlier measurements.

Data Availability. All study data are included in the article.

ACKNOWLEDGMENTS. We are grateful to Ann Reum for her technical assistance and to the Research Cell Bank at Fred Hutchinson Cancer Research Center for providing the samples. This study was supported by the Dharam Ablashi Pilot Grant and European Research Council Starting Grant Stg 677673 awarded to D.J.W. and B.B.K., respectively.

1. B. B. Kaufer, L. Flamand, Chromosomally integrated HHV-6: Impact on virus, cell and organismal biology. *Curr. Opin. Virol.* **9**, 111–118 (2014).
2. K. Yamanishi *et al*., Identification of human herpesvirus-6 as a causal agent for exanthem subitum. *Lancet* **1**, 1065–1067 (1988).
3. E. P. Nacheva *et al*., Human herpesvirus 6 integrates within telomeric regions as evidenced by five different chromosomal sites. *J. Med. Virol.* **80**, 1952–1958 (2008).
4. J. H. Arbuckle *et al*., The latent human herpesvirus-6A genome specifically integrates in telomeres of human chromosomes in vivo and in vitro. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 5563–5568 (2010).
5. J. H. Arbuckle *et al*., Mapping the telomere integrated genome of human herpesvirus 6A and 6B. *Virology* **442**, 3–11 (2013).
6. N. Osterrieder, N. Wallaschek, B. B. Kaufer, Herpesvirus genome integration into telomeric repeats of host cell chromosomes. *Annu. Rev. Virol.* **1**, 215–235 (2014).
7. K. Tanaka-Taya *et al*., Human herpesvirus 6 (HHV-6) is transmitted from parent to child in an integrated form and characterization of cases with chromosomally integrated HHV-6 DNA. *J. Med. Virol.* **73**, 465–473 (2004).
8. M. Daibata, T. Taguchi, Y. Nemoto, H. Taguchi, I. Miyoshi, Inheritance of chromosomally integrated human herpesvirus 6 DNA. *Blood* **94**, 1545–1549 (1999).
9. P. Hubacek *et al*., Prevalence of HHV-6 integrated chromosomally among children treated for acute lymphoblastic or myeloid leukemia in the Czech Republic. *J. Med. Virol.* **81**, 258–263 (2009).
10. L. Potenza *et al*., Prevalence of human herpesvirus-6 chromosomal integration (CIHHV-6) in Italian solid organ and allogeneic stem cell transplant patients. *Am. J. Transplant.* **9**, 1690–1697 (2009).
11. K. N. Ward, H. N. Leong, A. D. Thiruchelvam, C. E. Atkinson, D. A. Clark, Human herpesvirus 6 DNA levels in cerebrospinal fluid due to primary infection differ from those due to chromosomal viral integration and have implications for diagnosis of encephalitis. *J. Clin. Microbiol.* **45**, 1298–1304 (2007).
12. H. Miura *et al*., Late-phase human herpesvirus 6B reactivation in hematopoietic stem cell transplant recipients. *Transpl. Infect. Dis.* **20**, e12916 (2018).
13. L. R. Wang, L. J. Dong, M. J. Zhang, D. P. Lu, The impact of human herpesvirus 6B reactivation on early complications following allogeneic hematopoietic stem cell transplantation. *Biol. Blood Marrow Transplant.* **12**, 1031–1037 (2006).
14. M. Montejo *et al*., Encephalitis caused by human herpesvirus-6 in a liver transplant recipient. *Eur. Neurol.* **48**, 234–235 (2002).
15. J. A. Hill *et al*., Human herpesvirus 6B reactivation and delirium are frequent and associated events after cord blood transplantation. *Bone Marrow Transplant.* **50**, 1348–1351 (2015).
16. J. A. Hill *et al*., Outcomes of hematopoietic cell transplantation using donors or recipients with inherited chromosomally integrated HHV-6. *Blood* **130**, 1062–1069 (2017).
17. A. Gravel, C. B. Hall, L. Flamand, Sequence analysis of transplacentally acquired human herpesvirus 6 DNA is consistent with transmission of a chromosomally integrated reactivated virus. *J. Infect. Dis.* **207**, 1585–1589 (2013).
18. U. Kühl *et al*., Chromosomally integrated human herpesvirus 6 in heart failure: Prevalence and treatment. *Eur. J. Heart Fail.* **17**, 9–19 (2015).
19. Y. Huang *et al*., Human telomeres that carry an integrated copy of human herpesvirus 6 are often short and unstable, facilitating release of the viral genome from the chromosome. *Nucleic Acids Res.* **42**, 315–327 (2014).
20. P. E. Pellett *et al*., Chromosomally integrated human herpesvirus 6: Questions and answers. *Rev. Med. Virol.* **22**, 144–155 (2012).
21. J. H. Arbuckle, P. G. Medveczky, The molecular biology of human herpesvirus-6 latency and telomere integration. *Microbes Infect.* **13**, 731–741 (2011).
22. J. Tweedy *et al*., 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**, 19 (2016).
23. Y. Mostovoy *et al*., A hybrid approach for de novo human genome sequence assembly and phasing. *Nat. Methods* **13**, 587–590 (2016).
24. W. B. Jiao *et al*., Improving and correcting the contiguity of long-read genome assemblies of three plant species using optical mapping and chromosome conformation capture data. *Genome Res.* **27**, 778–786 (2017).
25. A. Aswad *et al*., Evolutionary history of endogenous human herpesvirus 6 reflects human migration out of Africa. *Mol. Biol. Evol.*, 10.1093/molbev/msaa190 (2020).
26. R. C. Allshire, M. Dempster, N. D. Hastie, Human telomeres contain at least three types of G-rich repeat distributed non-randomly. *Nucleic Acids Res.* **17**, 4611–4627 (1989).
27. R. M. Cawthon, Telomere measurement by quantitative PCR. *Nucleic Acids Res.* **30**, e47 (2002).
28. N. Wallaschek *et al*., The telomeric repeats of human herpesvirus 6A (HHV-6A) are required for efficient virus integration. *PLoS Pathog.* **12**, e1005666 (2016).

29. N. Wallaschek, A. Gravel, L. Flamand, B. B. Kaufer, The putative U94 integrase is dispensable for human herpesvirus 6 (HHV-6) chromosomal integration. *J. Gen. Virol.* **97**, 1899–1903 (2016).
30. D. J. Wight *et al*., Viral proteins U41 and U70 of human herpesvirus 6A are dispensable for telomere integration. *Viruses* **10**, 656 (2018).
31. G. Aimola, G. Beythien, A. Aswad, B. B. Kaufer, Current understanding of human herpesvirus 6 (HHV-6) chromosomal integration. *Antiviral Res.* **176**, 104720 (2020).
32. X. Liu *et al*., Endogenization and excision of human herpesvirus 6 in human genomes. *PLoS Genet.* **16**, e1008915 (2020).
33. J. H. Wildschutte *et al*., Discovery of unfixed endogenous retrovirus insertions in diverse human populations. *Proc. Natl. Acad. Sci. U. S. A.* **113**, E2326–E2334 (2016).
34. T. Karamitros *et al*., Human endogenous retrovirus-K HML-2 integration within *RASGRF2* is associated with intravenous drug abuse and modulates transcription in a cell-line model. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 10434–10439 (2018).
35. E. H. Blackburn *et al*., Recognition and elongation of telomeres by telomerase. *Genome* **31**, 553–560 (1989).
36. G. B. Morin, The human telomere terminal transferase enzyme is a ribonucleoprotein that synthesizes TTAGGG repeats. *Cell* **59**, 521–529 (1989).
37. S. B. Cohen *et al*., Protein composition of catalytically active human telomerase from immortal cells. *Science* **315**, 1850–1853 (2007).
38. K. M. Eisenhauer, R. M. Gerstein, C. P. Chiu, M. Conti, A. J. Hsueh, Telomerase activity in female and male rat germ cells undergoing meiosis and in early embryos. *Biol. Reprod.* **56**, 1120–1125 (1997).
39. W. E. Wright, M. A. Piatyszek, W. E. Rainey, W. Byrd, J. W. Shay, Telomerase activity in human germline and embryonic tissues and cells. *Dev. Genet.* **18**, 173–179 (1996).
40. S. Gilbert-Girard *et al*., Stabilization of telomere G-quadruplexes interferes with human herpesvirus 6A chromosomal integration. *J. Virol.* **91**, e00402-17 (2017).
41. L. Ding *et al*., Latent membrane protein 1 encoded by Epstein–Barr virus induces telomerase activity via p16INK4A/Rb/E2F1 and JNK signaling pathways. *J. Med. Virol.* **79**, 1153–1163 (2007).
42. R. M. Stansel, T. de Lange, J. D. Griffith, T-loop assembly in vitro involves binding of TRF2 near the 3′ telomeric overhang. *EMBO J.* **20**, 5532–5540 (2001).
43. J. D. Griffith *et al*., Mammalian telomeres end in a large duplex loop. *Cell* **97**, 503–514 (1999).
44. B. B. Kaufer, Detection of integrated herpesvirus genomes by fluorescence in situ hybridization (FISH). *Methods Mol. Biol.* **1064**, 141–152 (2013).
45. A. L. Greninger *et al*., Comparative genomic, transcriptomic, and proteomic reannotation of human herpesvirus 6. *BMC Genomics* **19**, 204 (2018).