Video Article

Detection of Low Copy Number Integrated Viral DNA Formed by *In Vitro* Hepatitis B Infection

Thomas Tu[1,2], Stephan Urban[1,3]

[1]Department of Infectious Diseases, Molecular Virology, Heidelberg University Hospital

[2]Gastroenterology and Liver Research Laboratory, Ingham Institute

[3]German Center for Infection Research (DZIF)

Correspondence to: Thomas Tu at Thomas.Tu@med.uni-heidelberg.de

Abstract

Hepatitis B Virus (HBV) is a common blood-borne pathogen causing liver cancer and liver cirrhosis resulting from chronic infection. The virus generally replicates through an episomal DNA intermediate; however, integration of HBV DNA fragments into the host genome has been observed, even though this form is not necessary for viral replication. The exact purpose, timing, and mechanism(s) by which HBV DNA integration occurs is not yet clear, but recent data show that it occurs very early after infection. Here, the *in vitro* generation and detection of HBV DNA integrations are described in detail. Our protocol specifically amplifies single copies of virus-cell DNA integrations and allows both absolute quantification and single-base pair resolution of the junction sequence. This technique has been applied to various HBV-susceptible cell types (including primary human hepatocytes), to various HBV mutants, and in conjunction with various drug exposures. We foresee this technique becoming a key assay to determine the underlying molecular mechanisms of this clinically relevant phenomenon.

Video Link

The video component of this article can be found at https://www.jove.com/video/58202/

Introduction

HBV is a double-stranded DNA virus that can cause life-long chronic infection, leading to liver cirrhosis and hepatocellular carcinoma (HCC)[1,2,3]. While there are multiple molecular mechanisms driving HBV persistence[4] (*e.g.,* high stability of the epigenetic viral transcriptional template, evasion of immune surveillance, and low turnover of hepatocytes in the liver) and its associated risk of HCC initiation[5,6] (*e.g.,* chronic inflammation and activation of cellular stress pathways), HBV DNA integration into the host cell genome (a reported mechanism for both of these phenomena) has been poorly studied. A major reason is the lack of suitable *in vitro* infection systems for HBV that allow reliable detection of integration events. Here, we describe a recently-developed protocol for both *in vitro* generation and detection of HBV DNA integrations that can be used to elucidate the underlying molecular mechanisms and consequences thereof.

HBV replication and the formation of HBV DNA integration has been previously reviewed in detail[7]. Briefly, HBV enters hepatocytes using the Sodium Taurocholate Co-transporting Peptide (NTCP) as the main cellular receptor for infection[8,9]. The HBV nucleocapsids containing the HBV-relaxed circular DNA (rcDNA) genome enters the nucleus, where the rcDNA is converted into episomal covalently closed circular DNA (cccDNA). The nuclear cccDNA acts as the transcriptional template for viral mRNAs and pre-genomic RNA (pgRNA)[10]. HBV polymerase and pgRNA are then packaged into newly-formed nucleocapsids (composed of HBV core protein dimers). HBV pgRNA is reverse-transcribed within the nucleocapsid, resulting in either an rcDNA genome or a double-stranded linear DNA (dslDNA) genome[11,12]. These mature nucleocapsids containing HBV DNA genomes are finally enveloped and exported as virions.

Infection of hepatocytes by enveloped particles containing dslDNA molecules can result in viral integration into the host cell genome[13], leading to replication-incompetent forms of HBV DNA[7,14,15]. HBV DNA integration occurs at the site of chromosomal double-stranded DNA breaks[15]. Accumulating evidence suggests that each integration event occurs in an essentially random position within the host cell genome[16,17]. In addition, HBV DNA integration occurs somewhat rarely, at a rate of 1 per 10^4 cells[13,18,19,20]. Important questions regarding HBV DNA integration remain unanswered, particularly regarding the exact molecular pathways involved, the dependence on viral and host factors, the viral antigens expressed from integrated forms, and their possible contribution to viral persistence[7]. We have set up an *in vitro* model to shed light on some of these questions.

The rarity of HBV DNA integration events (both with respect to integration rate per cell and to the copy number of each unique integration) in *in vitro* HBV infection models make them challenging to detect. Cell mitosis is limited in our *in vitro* system, as dividing cells do not support efficient infection. Thus, unlike in patient liver tissues where significant clonal expansion of hepatocytes occurs[18,19,20], very few (1 to 2) copies of each integration are present in a given pool of cells infected *in vitro*. We have also found that integration mainly occurs during the initial infection of

hepatocytes (and not continuously in chronically-infected hepatocytes)[13]. Accordingly, HBV DNA integration cannot be increased by simply culturing cells for a longer period.

In general, methods previously used to detect integrated HBV DNA, including Southern blot hybridization[21,22,23,24,25], Alu-PCR[26,27], cassette-mediated ligation PCR[28], and whole genome sequencing[29,30,31,32,33], do not have the sensitivity to detect single-copies of integrations. We and other researchers have used inverse nested PCR (invPCR) to detect hepadnaviral DNA integration in duck, woodchuck, and human infected livers[13,14,18,19,34,35,36]. Others have introduced procedural modifications to the invPCR technique that may alter the generation of false-positive signals and restrict the ability for quantification[37]. If the assay is carried out as described in this protocol, invPCR represents a specific and sensitive assay that identifies and quantifies (in absolute copy numbers) multiple HBV DNA integrations. The HBV-cell DNA junction is sequenced at a single-base pair resolution, allowing bioinformatical studies of viral and host DNA sequences at the sites of integration[13].

We have previously described[38] results from invPCR on DNA of HBV-infected tissue extracted from a large range of sources, including snap-frozen liver tissues, paraffin-embedded liver sections, and extremely small tissue samples isolated by laser-microdissection[19]. This protocol outlines an updated version of an invPCR assay using DNA extracted from cell culture-derived tissues after *in vitro* infection, in which low copy numbers of integrations (1–2 copies per integration) are generated. The HBV DNA integrations formed *in vitro* resemble those found in patient tissues with respect to their distribution over the cellular genome and junction within the viral sequence that is integrated[13,16], suggesting a comparable pathway to within an infected liver.

Protocol

1. Cell Infection and DNA Extraction

1. Maintain cultured Huh7-NTCP cells[8,9] in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% v/v fetal bovine serum, 1x Penicillin/Streptomycin, and 2 mM L-glutamine.
   NOTE: This has previously been reported in detail[40].
2. Seed 2×10^5 cells/mL Huh7-NTCP cells into a 12-well plate with 1 mL of DMEM.
3. If testing cell treatments (*e.g.,* HBV inhibitors), apply them to the culture supernatant 4 h after seeding (1 day prior to HBV infection). For a negative control, apply 200 nM Myrcludex B (a potent HBV entry inhibitor[41]).
4. Use heparin-column purified[42] supernatant from HepAD38[43] as an inoculum to infect cells at 1,000 VGE/cell in 500 µL of culture media (DMEM supplemented with 10% v/v fetal bovine serum, 1x Penicillin/Streptomycin, 2 mM L-glutamine, and 2.5% v/v DMSO), containing 4% w/v polyethylene glycol 8000 [dissolved in 1x phosphate-buffered saline (PBS)].
5. Culture the cells in a 37 °C incubator (set at 5% CO_2 and 90% humidity).
6. Wash the cells twice with 1 mL of sterile 1x PBS at 16–24 h post-infection.
7. Replace the culture media every 2 days following HBV infection until harvest.
8. At day 3 post-infection, treat the cells with 5 µM Tenofovir disoproxil and 10 µM Lamivudine to limit production of HBV replicative intermediates that are amplifiable by invPCR.
9. At day 5 post-infection, trypsinize the cells with 200 µL of Trypsin-EDTA and resuspend them in 2 mL of DMEM (supplemented with 10% v/v fetal bovine serum, 1x Penicillin/Streptomycin, and 2 mM L-glutamine), also containing 5 µM Tenofovir disoproxil and 10 µM Lamivudine.
10. Transfer the cell suspensions to a 6-well plate to induce one round of mitosis (which has been reported to induce loss of HBV cccDNA[44] and other HBV DNA intermediates that are amplifiable by invPCR).
11. At day 7 post-infection, trypsinize the expanded cells in 400 µL of Trypsin-EDTA and resuspend the mixture in 1 mL of DMEM. Place the suspension in a 1.5 mL tube, pellet the cells by centrifugation at 500 x g for 5 min, and remove the supernatant by aspiration.
12. (Optional) Store the cell pellets at -20 °C until they are ready for DNA extraction.
13. Extract the DNA from the cell pellet using a DNA extraction kit, as per the manufacturer's instructions. Estimate the final DNA concentration using optical densitometry.
   NOTE: The DNA yield in a 100 µL elution volume is generally ~250-400 ng/µL.

2. Inversion of DNA

1. Aliquot ~1.5–2.5 µg of the total DNA extract from step 1 into a 200 µL PCR tube. Add the appropriate amount of restriction enzyme master-mix to result in a 40 µL reaction volume containing 1x digest reaction buffer (*e.g.*, CutSmart buffer) and 10 U *Nco*I-HF.
2. Mix the reactions thoroughly and spin down in a small tube centrifuge. Incubate the restriction enzyme reaction in a PCR machine at 37 °C for 1 h for optimal digestion efficiency.
3. Inactivate the restriction enzyme by incubating at 80 °C for 20 min.
4. Transfer the entire restriction enzyme reaction to a 1.5 mL tube. Add 400 µL of 1x T4 DNA ligase buffer and 500 U of T4 DNA ligase and mix it thoroughly. The large reaction volume encourages intra-molecular (as opposed to inter-molecular) ligation of the digested DNA fragments.
5. Incubate the ligation reaction at room temperature for 2 h to ensure complete ligation.
6. Inactivate the T4 DNA ligase at 70 °C for 20 min.
7. Add 10 µL of 10% w/v sodium dodecyl sulfate to ensure complete inactivation of the T4 ligase.
8. Mix the tubes by pulse vortexing and briefly spin down the reaction mix. Add NaCl to a final concentration of 100 mM and dextran (35–45 kDa) to a final concentration of 90 µg/mL. Mix the tubes by pulse vortexing and briefly spin down the reaction mix.
9. Add 900 µL of 100% ethanol and mix by inversion. Precipitate the DNA at -20 °C overnight.
10. Pellet the precipitated DNA by centrifugation at 14,000 x g for 15 min. Remove the supernatant by aspiration with a P200 pipette. Wash the pellet with 500 µL of 70% v/v ethanol and centrifugation at 14,000 x g for 15 min.
11. Remove the ethanol by aspiration with a P200 pipette. Air-dry the DNA pellet at room temperature for 20 min.
12. Dissolve the pellet in 20 µL of H_2O. Add 20 µL of a restriction enzyme master-mix to result in a 40 µL reaction volume containing 1x digest reaction buffer, 5 U of *Bsi*HKAI, and 5 U of *Sph*I-HF. Incubate the restriction enzyme reaction in a heat-block at 37 °C (the optimal temperature for *Sph*I-HF) for 1 h.

13. Briefly spin down the reaction mix. Incubate the restriction enzyme reaction in a heat-block at 65 °C (the optimal temperature for *Bsi*HKAI) for 1 h.
14. Briefly spin down the reaction mix.
15. Store the inverted DNA at -20 °C until the next step.

3. Nested PCR

1. Remove potential amplicons from a reusable silicon mat seal using a DNA degradation solution, rinse the mat thoroughly with DNA-free water, and air-dry it at room temperature while preparing the PCR mixture.
2. Prepare 1 mL of a 1x PCR mix containing the outer forward (5'-TTC GCT TCA CCT CTG CAC G-3') and reverse (5'-AAA GGA CGT CCC GCG CAG-3') primers at a concentration of 0.5 µM.
3. Add 170 µL of the 1x PCR mix to wells A1 and E1 in a 96-well PCR plate.
4. Add 120 µL of the 1x PCR mix to wells B1 to H1.
5. Add 10 µL of the inverted DNA from step 2 to wells A1 and E1 (2 different inverted samples can be analyzed on the same PCR plate). Mix the reaction in each well by gently pipetting each about 10 times using a P1000 set to 100 µL.
6. Serially dilute samples from wells A1 to D1 at a ratio of 1:3 by transferring 60 µL at each step. Mix the wells at each step by gently pipetting them about 10 times using a P1000 set to 100 µL. Avoid forming bubbles. Repeat step 3.6 for well E1, diluting down to well H1.
7. Aliquot 10 µL of the reaction mixture using a multi-channel pipette from wells A1-H1 into wells A2-H2, A3-H3, and so on until reaching wells A12-H12 of the 96-well plate. Cover the PCR plate with the dry silicon mat from step 3.1, pressing firmly.
8. Place the plate in a PCR machine and run the following program: 10 min at 95 °C; 35 cycles of 15 s at 95 °C, 15 s at 54 °C and 3 min at 72 °C; 7 min at 72 °C; then, hold the plate at room temperature.
9. Heat the pins of a 96-pin replicator to red-hot with a Bunsen burner and then cool for at least 5 min at room temperature.
10. Fill the wells of a second PCR plate with 10 µL of 1x PCR mix (containing a gel load-ready buffer) and the inner forward (5'-CGC ATG GAG ACC ACC GTG A-3') and reverse (5'-CAC AGC CTA GCA GCC ATG G-3') at 0.5 µM.
11. Carefully remove the silicon mat from the PCR plate of the 96-well plate from the first-round PCR, and avoid cross-contamination between wells.
12. Use the cooled replicator to transfer the PCR products of the 96-well plate from first-round PCR to the newly-aliquoted 96-well plate. Carry out the nested PCR using the same conditions as in step 3.8, except for changing the initial denaturation step at 95 °C from 10 min to 2 min.

4. PCR Product Isolation and Gel Extraction

1. Analyze the PCR products by gel electrophoresis using a 96-well with 1.3% w/v agarose gel. For a 100 mL agarose gel, run at 200 V for 10–15 min.
2. Excise the DNA bands from the agarose gels using disposable drinking straws. For each PCR product, place the straw and agarose gel plug into a 1.5 mL tube and trim the straw to size with scissors.
   NOTE: It is sufficient to isolate bands only from those dilutions in which single PCR products can be resolved.
3. (Optional) Store the tubes at -20°C for later extraction.
4. Eject the agarose plugs into each tube by squeezing on the end of the straw fragment. Add 300 µL of gel extraction buffer and 5 µL of gel extraction glass bead slurry to each tube.
5. Extract the PCR products as per the manufacturer's instructions for the gel extraction kit (except using half-volumes for washing steps) and elute the DNA from the beads with 30 µL of water.
6. Submit the purified DNA for Sanger sequencing (as per the supplier's instructions) with the Forward primer used in the second-round nested PCR.

5. Sequence Analysis

1. Confirm virus-cell DNA junctions by performing nucleotide BLAST analysis (using default settings aligning to the entire nucleotide collection).
   NOTE: Representative results are detailed in **Figure 3** below.
   1. If only partial alignment of the sequence is observed, trim the 5' HBV DNA sequence before re-running a BLAST analysis.

2. Apply stringent selection criteria to determine if a given sequence represents a true integration junction as follows:
   1. Include only sequences containing >20 bp of a cellular sequence for confident mapping to the cellular genome.
   2. Disregard any sequences that contain restriction sites of any enzymes used within 10 bp of the supposed virus-cell integration junction, as they likely represent *in vitro* ligation events rather than true integration junctions.
   3. Disregard any sequences with obvious dissimilar peak fluorescence levels on either side of the supposed integration junction in sequencing chromatographs, as these are likely artifacts generated by cross-contamination during the sequencing reaction or capillary chromatography.
   4. Define unique integration events as those with the exact HBV and cellular sequences at the virus-cell junction.
      NOTE: Repetition of unique integration events can result from clonal expansion of cells containing those integrations or cross-contamination during the PCR (which can be tested using negative-control reactions, further detailed in the representative results below). Repeated integrations into specific cellular sequences are expected to display different HBV terminal sites for each integration event, as non-homologous end-joining pathways are used[15].

3. Calculate the integration frequency by multiplying the dilution factor of inverted DNA templates by the number of virus-cell junctions detected at that dilution, followed by normalization to the amount of total DNA input into the inversion reaction. Generally, the integration frequency is on the order of 1:10^4 cells.

Representative Results

A schematic representation of the invPCR technique is shown in **Figure 1**. An example agarose gel electrophoresis of a successful invPCR is shown in **Figure 2** before and after the PCR product isolation. **Figure 3** shows the BLAST analysis output from step 5.1 in the cases of (i) amplification of virus-cell DNA junction and (ii) amplification of an HBV DNA replicative intermediate.

Expected results from positive controls: Huh7-NTCP cells infected with heparin-purified HBV stocks as described above (**Figure 1**) can serve as a positive control. In this instance, single PCR products should be achieved by the second or third 1:3 dilution of each sample (corresponding to ~10^4 cell equivalents per dilution, **Figure 2**). At these dilutions, approximately 50% of products will represent true virus-cell DNA junctions, while the other half represents amplification of HBV DNA intermediates (**Figure 3**).

In previous studies[13], we have found few instances of repeated virus-cell junctions, suggesting no cellular genomic sites of preferential HBV DNA integration. We also find that **>**80% of virus-cell junctions occur between nucleotides 1732 and 1832 (according to the nucleotide numbering of HBV genotype D, GenBank Accession #U95551.1), where the latter represents the right-hand terminus of the HBV dslDNA form. Sequences of true virus-cell junctions found through our method in *in vitro* experiments are publicly available on GenBank (Accession numbers MH057851 to MH058006).

Expected results from negative controls: Either uninfected Huh7-NTCP cells or inoculated Huh7-NTCP cells pre-treated with Myrcludex B can act as negative controls. InvPCR analysis of DNA extracted from these cells should produce no PCR products. Running these negative-control samples on the same PCR plate as positive samples allows for testing of cross-contamination during the nested-PCR process. The most stringent test for cross-contamination is the running of a negative-control sample in wells A-D and the positive sample in wells E-H on a 96-well plate. In this case, the most concentrated dilution of the positive sample is run in a row directly adjacent to the least concentrated dilution of the negative sample. If amplification products are observed in negative-control samples, institute the measures suggested in the discussion to minimize cross-contamination events.
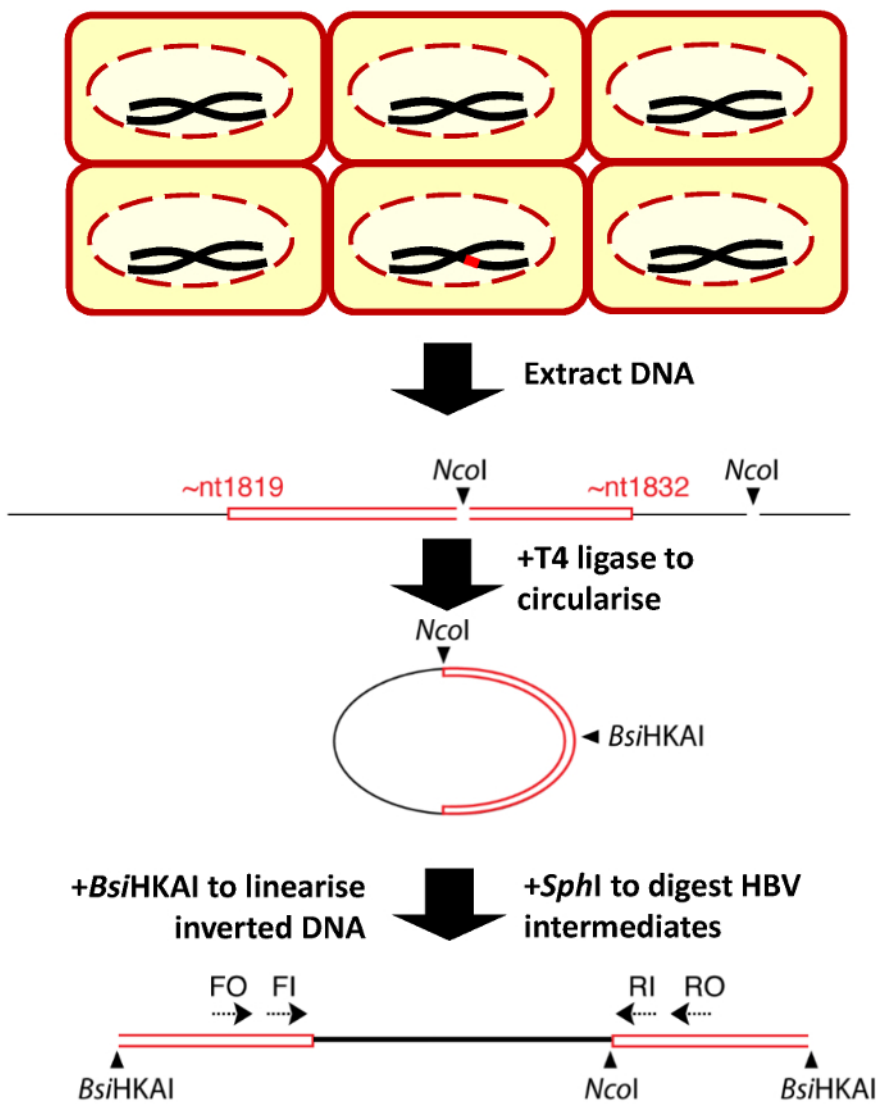
Figure 1: Schematic figure of the invPCR process to detect HBV DNA integrations. After HBV infection, only a minority of Huh7-NTCP cells contain HBV dsIDNA (red) integrated into the host cell genome (red), depicted in the top section. Total DNA is then extracted from the cells (step 1), inverted (step 2), and amplified by nested PCR (step 3). Shown are the relative positions of the restriction enzyme sites (*Nco*I and *Bsi*HKAI) in the excised virus-cell DNA junction. The figure is adapted from a previous publication[13]. Please click here to view a larger version of this figure.
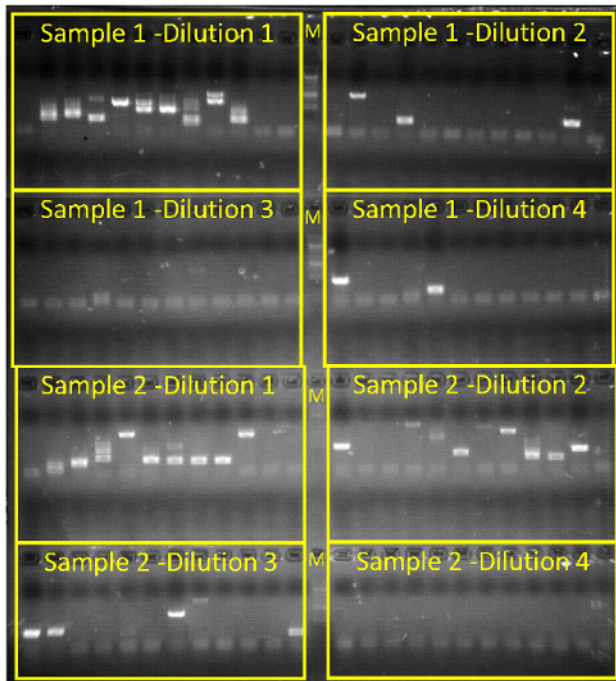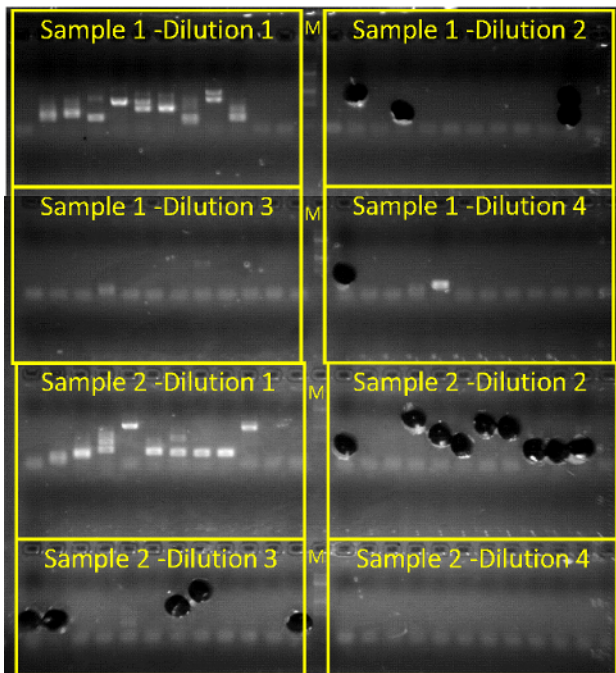
Figure 2: Agarose gel electrophoresis and subsequent gel extraction of a successful invPCR. "M" represents DNA marker ladder. Each row of 12 represents technical replicates at a single dilution on the PCR plate. Two inverted DNA samples can be run per PCR plate/agarose gel. The PCR products for each dilution should titrate out in a ~1:3 ratio. Extraction of the products from the two least concentrated dilutions where single bands can be easily isolated is generally sufficient, as HBV DNA integration rate is determined by end-point titration. Please click here to view a larger version of this figure.
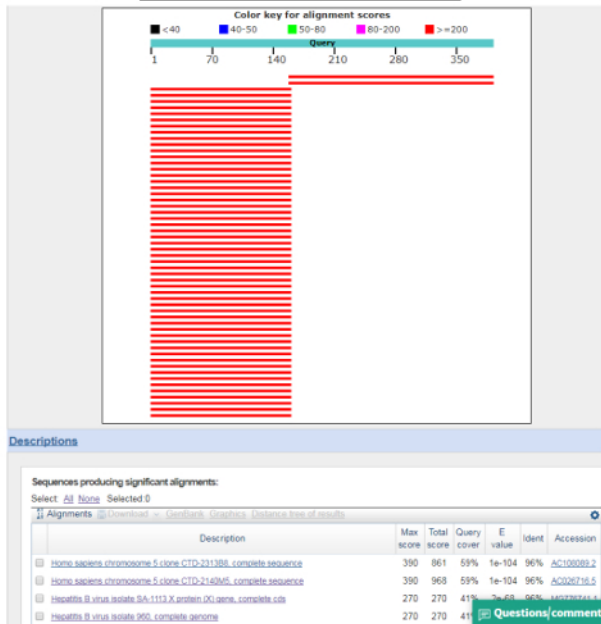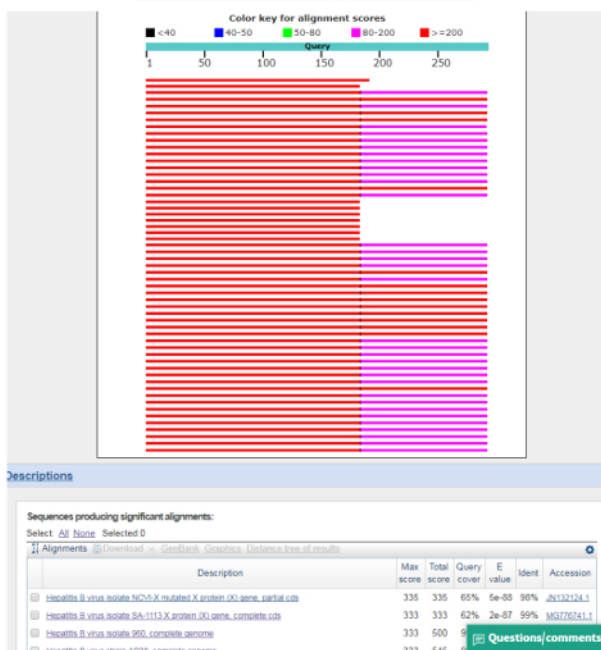
Figure 3: BLAST analysis of invPCR product sequences. As long tracts of sequence are generated from Sanger sequencing, the source of DNA sequences is generally unambiguous. Strong alignment to HBV and cellular genomes is observed in different areas of the FASTA sequence file generated from Sanger sequencing in the top panel, suggesting a true virus-cell DNA junction. In the bottom panel, the sequence aligns only to fragments of the HBV genome, suggesting a product generated by the amplification of a rearranged HBV DNA genome. Please click here to view a larger version of this figure.

November 2018 | 141 | e58202 | Page 7 of 10

Discussion

Before performing the protocol, it is important to note that this invPCR assay is a highly sensitive technique, capable of amplifying single copies of DNA template. Therefore, limiting contamination from PCR products is of utmost importance. General strategies to limit PCR product contamination include the following. (i) Establish physically separate areas for different steps of the method. Each area should have separate lab coats, and gloves should be changed when moving between these areas. We have listed these areas below in order from most to least likely to be contaminated: PCR product extraction and sequencing reaction set-up area (post-PCR); PCR template addition and flamed-pin transfer area (we have used PCR hoods with a decontaminating UV lamp with good results); DNA extraction and inversion area (pre-PCR); and a "DNA template-free" area used solely for preparing stock and PCR solutions. (ii) Be aware of air flow in the lab as a potential driver of cross-contamination. In particular, cross-contamination of PCR reactions during the flamed pin transfer step is most likely to occur and will lead to inaccurate quantification of integration frequency. PCR hoods can be used to limit these cross-currents. Negative control wells (*e.g.,* Myrcludex B-treated samples or no template controls) in the PCR can be also used to test for cross-contamination. (iii) Limit potential PCR contaminants on pipettes and work surfaces by regularly wiping them down with a DNA degradation solution.

The protocol specified here is arranged to detect integration of a known infectious HBV clone generated from the HepAD38 cell line[42]. If the inoculum used is of a different HBV DNA sequence (*e.g.,* from patient serum), then the HBV genome should be first sequenced to confirm compatibility of PCR primers and inversion design. In previously published studies[19], we have used 3 sets of primers that bind conserved HBV DNA sequences flanking the expected site of HBV DNA integration junctions. Other primer sequences and protocols have been described to determine the genomic sequence of HBV[44,45,46,47,48] and may work successfully.

Furthermore, different HBV-susceptible cells (*e.g.,* primary human hepatocytes, differentiated HepaRG, HepG2-NTCP cells) can be used; however, we have found that Huh7-NTCP cells provide the greatest signal-to-noise ratio, when considering the number of virus-cell DNA junctions that are amplified compared to the virus intermediate DNA that is amplified[13]. In particular, terminally differentiated cells such as primary human hepatocytes or differentiated HepaRG do not efficiently undergo mitosis, resulting in a high level of amplifiable HBV DNA replicative intermediates remaining within the cells. We have found that ~90% of amplified sequences represent HBV DNA rearrangements (and not integration events) in terminally-differentiated cells, compared to 70–50% in hepatoma cell lines[13]. These products are generally HBV DNA genomes containing large deletions in the surface and core open reading frames, or they can represent HBV quasi-species with an additional *Nco*I site prior to the *Sph*I site. The amplification of these HBV species (including a schematic diagram) have been described in detail previously[13].

There are several drawbacks to this method. Due to the laborious and multi-day nature of this protocol, invPCR is not suited to high-throughput analyses of a large number of samples. Furthermore, as it relies on limiting dilution titration, our method is not highly precise in quantification; although, it should easily measure log-level changes in integration frequency.

Moreover, our inversion protocol is only suited to detect integrations occurring between the DR2 and DR1 region of the HBV genome, as the majority of HBV integrations occur within this region[49]. NGS analysis of HBV patient tissues has shown that a large minority (up to ~50%) may also occur outside of this region[49]. New invPCR designs are theoretically able to detect these other integration sites, though they have not (to our knowledge) been carried out yet. Relatedly, due to the necessary restriction enzyme sites required downstream from the virus-cell junction sequence for the inversion reaction, invPCR does not detect all integrations occurring within the DR2 and DR1 regions of the HBV genome (*i.e.,* we have estimated that ~10% of all integrations are detectable using *in silico* simulations[13]). However, when applied to a focused batch of samples with small numbers of different treatments, invPCR is one of the only practical methods for detecting integrated HBV DNA at a single base-pair resolution.

We therefore envision applications of this method serving a key role in finding viral (via mutations of the HBV inoculum), cellular (through knockout or overexpression of specific cellular genes, or through application of various drugs), and environmental (*e.g.,* exposure to oxidative stress) factors that induce HBV DNA integration. With this method and newly developed HBV infection systems, we enable unprecedented control of these factors so that they can be well-isolated and better characterized. We also expect that this will lead to a key understanding of the cellular consequences of HBV DNA integration, including to what extent viral antigens (*e.g.,* HBx or HBsAg[50]) are expressed from the integrated form, what controls this expression, and whether or not HBV integration significantly changes the cellular phenotype towards a more pro-oncogenic state. The results of these future studies will have a profound impact on the therapeutic strategies used to treat chronic hepatitis B and on the basic understanding of the virus itself.

Disclosures

Acknowledgements

References

1. Lin, X., *et al.* Chronic hepatitis B virus infection in the Asia-Pacific region and Africa: review of disease progression. *Journal of Gastroenterology and Hepatology.* **20** (6), 833-843 (2005).
2. Iloeje, U. H., *et al.* Predicting cirrhosis risk based on the level of circulating hepatitis B viral load. *Gastroenterology.* **130** (3), 678-686 (2006).
3. Huang, Y. T., *et al.* Lifetime risk and sex difference of hepatocellular carcinoma among patients with chronic hepatitis B and C. *Journal of Clinical Oncology.* **29** (27), 3643-3650 (2011).
4. Nassal, M. HBV cccDNA: viral persistence reservoir and key obstacle for a cure of chronic hepatitis B. *Gut.* **64** (12), 1972-1984 (2015).
5. Tu, T., Budzinska, M. A., Shackel, N. A., Jilbert, A. R. Conceptual models for the initiation of hepatitis B virus-associated hepatocellular carcinoma. *Liver International.* **35** (7), 1786-1800 (2015).
6. Tu, T., Buhler, S., Bartenschlager, R. Chronic viral hepatitis and its association with liver cancer. *Biological Chemistry.* **398** (8), 817-837 (2017).
7. Tu, T., Budzinska, M. A., Shackel, N. A., Urban, S. HBV DNA Integration: Molecular Mechanisms and Clinical Implications. *Viruses.* **9** (4) (2017).
8. Yan, H., *et al.* Sodium taurocholate cotransporting polypeptide is a functional receptor for human hepatitis B and D virus. *eLife.* **1**, e00049 (2012).
9. Ni, Y., *et al.* Hepatitis B and D viruses exploit sodium taurocholate co-transporting polypeptide for species-specific entry into hepatocytes. *Gastroenterology.* **146** (4), 1070-1083 (2014).
10. Tuttleman, J. S., Pourcel, C., Summers, J. Formation of the pool of covalently closed circular viral DNA in hepadnavirus-infected cells. *Cell.* **47** (3), 451-460 (1986).
11. Staprans, S., Loeb, D. D., Ganem, D. Mutations affecting hepadnavirus plus-strand DNA synthesis dissociate primer cleavage from translocation and reveal the origin of linear viral DNA. *Journal of Virology.* **65** (3), 1255-1262 (1991).
12. Wang, G. H., Seeger, C. The reverse transcriptase of hepatitis B virus acts as a protein primer for viral DNA synthesis. *Cell.* **71** (4), 663-670 (1992).
13. Tu, T., Budzinska, M. A., Vondran, F. W. R., Shackel, N. A., Urban, S. Hepatitis B virus DNA integration occurs early in the viral life cycle in an in vitro infection model via NTCP-dependent uptake of enveloped virus particles. *Journal of Virology.* (2018).
14. Yang, W., Summers, J. Integration of hepadnavirus DNA in infected liver: evidence for a linear precursor. *Journal of Virology.* **73** (12), 9710-9717 (1999).
15. Bill, C. A., Summers, J. Genomic DNA double-strand breaks are targets for hepadnaviral DNA integration. *Proceedings of the National Academy of Sciences of the United States of America.* **101** (30), 11135-11140 (2004).
16. Budzinska, M., Shackel, N. A., Urban, S., Tu, T. Sequence Analysis of Integrated Hepatitis B Virus DNA during HBeAg-Seroconversion. *Emerging Microbes and Infections.* (2018).
17. Budzinska, M., Shackel, N.A., Urban, S., Tu, T. Cellular genomic sites of HBV DNA integration. *Genes.* (2018).
18. Mason, W. S., Liu, C., Aldrich, C. E., Litwin, S., Yeh, M. M. Clonal expansion of normal-appearing human hepatocytes during chronic hepatitis B virus infection. *Journal of Virology.* **84** (16), 8308-8315 (2010).
19. Tu, T., *et al.* Clonal expansion of hepatocytes with a selective advantage occurs during all stages of chronic hepatitis B virus infection. *Journal of Viral Hepatitis.* **22** (9), 737-753 (2015).
20. Mason, W. S., *et al.* HBV DNA Integration and Clonal Hepatocyte Expansion in Chronic Hepatitis B Patients Considered Immune Tolerant. *Gastroenterology.* **151** (5), 986-998 e984 (2016).
21. Shafritz, D. A., Shouval, D., Sherman, H. I., Hadziyannis, S. J., Kew, M. C. Integration of hepatitis B virus DNA into the genome of liver cells in chronic liver disease and hepatocellular carcinoma. Studies in percutaneous liver biopsies and post-mortem tissue specimens. *New England Journal of Medicine.* **305** (18), 1067-1073 (1981).
22. Hino, O., *et al.* Detection of hepatitis B virus DNA in hepatocellular carcinomas in Japan. *Hepatology.* **4** (1), 90-95 (1984).
23. Fowler, M. J., *et al.* Integration of HBV-DNA may not be a prerequisite for the maintenance of the state of malignant transformation. An analysis of 110 liver biopsies. *Journal of Hepatology.* **2** (2), 218-229 (1986).
24. Chen, J. Y., Harrison, T. J., Lee, C. S., Chen, D. S., Zuckerman, A. J. Detection of hepatitis B virus DNA in hepatocellular carcinoma: analysis by hybridization with subgenomic DNA fragments. *Hepatology.* **8** (3), 518-523 (1988).
25. Esumi, M., Tanaka, Y., Tozuka, S., Shikata, T. Clonal state of human hepatocellular carcinoma and non-tumorous hepatocytes. *Cancer Chemotherapy and Pharmacology.* **23 Suppl**, S1-3 (1989).
26. Murakami, Y., *et al.* Large scaled analysis of hepatitis B virus (HBV) DNA integration in HBV related hepatocellular carcinomas. *Gut.* **54** (8), 1162-1168 (2005).
27. Kimbi, G. C., Kramvis, A., Kew, M. C. Integration of hepatitis B virus DNA into chromosomal DNA during acute hepatitis B. *World Journal of Gastroenterology.* **11** (41), 6416-6421 (2005).
28. Tamori, A., *et al.* Alteration of gene expression in human hepatocellular carcinoma with integrated hepatitis B virus DNA. *Clinical Cancer Research.* **11** (16), 5821-5826 (2005).
29. Sung, W. K., *et al.* Genome-wide survey of recurrent HBV integration in hepatocellular carcinoma. *Nature Genetics.* **44** (7), 765-769 (2012).
30. Fujimoto, A., *et al.* Whole-genome sequencing of liver cancers identifies etiological influences on mutation patterns and recurrent mutations in chromatin regulators. *Nature Genetics.* **44** (7), 760-764 (2012).
31. Jiang, S., *et al.* Re-evaluation of the Carcinogenic Significance of Hepatitis B Virus Integration in Hepatocarcinogenesis. *Proceedings of the National Academy of Sciences of the United States of America.* **7** (9), e40363 (2012).
32. Jiang, Z., *et al.* The effects of hepatitis B virus integration into the genomes of hepatocellular carcinoma patients. *Genome Research.* **22** (4), 593-601 (2012).
33. Kan, Z., *et al.* Whole-genome sequencing identifies recurrent mutations in hepatocellular carcinoma. *Genome Research.* **23** (9), 1422-1433 (2013).
34. Summers, J., *et al.* Hepatocyte turnover during resolution of a transient hepadnaviral infection. *Proceedings of the National Academy of Sciences of the United States of America.* **100** (20), 11652-11659 (2003).

35. Mason, W. S., Jilbert, A. R., Summers, J. Clonal expansion of hepatocytes during chronic woodchuck hepatitis virus infection. *Proceedings of the National Academy of Sciences of the United States of America.* **102** (4), 1139-1144 (2005).

36. Mason, W. S., *et al.* Detection of clonally expanded hepatocytes in chimpanzees with chronic hepatitis B virus infection. *Journal of Virology.* **83** (17), 8396-8408 (2009).

37. Chauhan, R., Churchill, N. D., Mulrooney-Cousins, P. M., Michalak, T. I. Initial sites of hepadnavirus integration into host genome in human hepatocytes and in the woodchuck model of hepatitis B-associated hepatocellular carcinoma. *Oncogenesis.* **6** (4), e317 (2017).

38. Tu, T., Jilbert, A. R. Detection of Hepatocyte Clones Containing Integrated Hepatitis B Virus DNA Using Inverse Nested PCR. *Methods in Molecular Biology.* **1540**, 97-118 (2017).

39. Ni, Y., Urban, S. Hepatitis B Virus Infection of HepaRG Cells, HepaRG-hNTCP Cells, and Primary Human Hepatocytes. *Methods in Molecular Biology.* **1540**, 15-25 (2017).

40. Schulze, A., Schieck, A., Ni, Y., Mier, W., Urban, S. Fine mapping of pre-S sequence requirements for hepatitis B virus large envelope protein-mediated receptor interaction. *Journal of Virology.* **84** (4), 1989-2000 (2010).

41. Lempp, F. A., *et al.* Evidence that hepatitis B virus replication in mouse cells is limited by the lack of a host cell dependency factor. *Journal of Hepatology.* **64** (3), 556-564 (2016).

42. Ladner, S. K., *et al.* Inducible expression of human hepatitis B virus (HBV) in stably transfected hepatoblastoma cells: a novel system for screening potential inhibitors of HBV replication. *Antimicrobial Agents and Chemotherapy.* **41** (8), 1715-1720 (1997).

43. Allweiss, L., *et al.* Proliferation of primary human hepatocytes and prevention of hepatitis B virus reinfection efficiently deplete nuclear cccDNA in vivo. *Gut.* (2017).

44. Yang, Z. T., *et al.* Characterization of Full-Length Genomes of Hepatitis B Virus Quasispecies in Sera of Patients at Different Phases of Infection. *Journal of Clinical Microbiology.* **53** (7), 2203-2214 (2015).

45. Chook, J. B., *et al.* Universal Primers for Detection and Sequencing of Hepatitis B Virus Genomes across Genotypes A to G. *Journal of Clinical Microbiology.* **53** (6), 1831-1835 (2015).

46. Li, F., *et al.* Whole genome characterization of hepatitis B virus quasispecies with massively parallel pyrosequencing. *Clinical Microbiology and Infection.* **21** (3), 280-287 (2015).

47. Zhou, T. C., *et al.* Evolution of full-length genomes of HBV quasispecies in sera of patients with a coexistence of HBsAg and anti-HBs antibodies. *Scientific Reports.* **7** (1), 661 (2017).

48. Long, Q. X., Hu, J. L., Huang, A. L. Deep Sequencing of the Hepatitis B Virus Genome: Analysis of Multiple Samples by Implementation of the Illumina Platform. *Methods in Molecular Biology.* **1540**, 211-218 (2017).

49. Li, X., *et al.* The function of targeted host genes determines the oncogenicity of HBV integration in hepatocellular carcinoma. *Journal of Hepatology.* **60** (5), 975-984 (2014).

50. Wooddell, C. I., *et al.* RNAi-based treatment of chronically infected patients and chimpanzees reveals that integrated hepatitis B virus DNA is a source of HBsAg. *Science Translational Medicine.* **9** (409), (2017).