JCI The Journal of Clinical Investigation

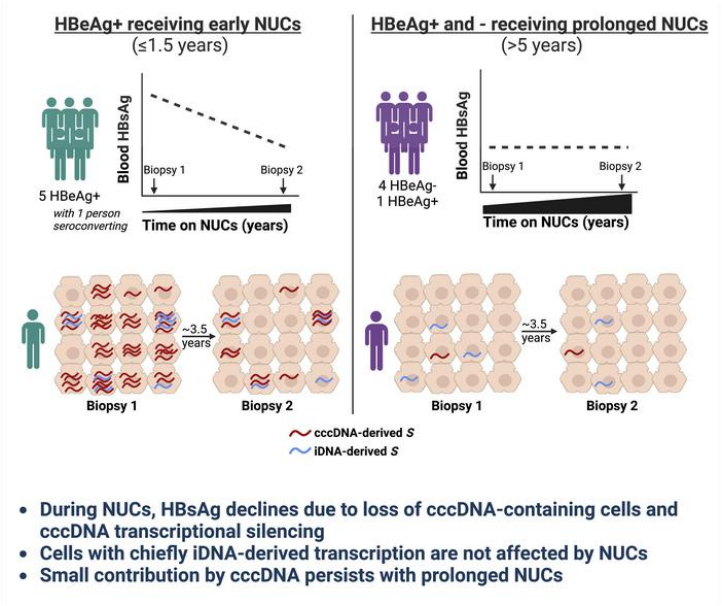
Transcription of hepatitis B surface antigen shifts from cccDNA to integrated HBV DNA during treatment

Maraake Taddese, … , Chloe L. Thio, Ashwin Balagopal

J Clin Invest. 2025. https://doi.org/10.1172/JCI184243.

Research In-Press Preview Infectious disease Virology

Graphical abstract



Transcription of hepatitis B surface antigen shifts from cccDNA to integrated HBV DNA during treatment

Maraake Taddese[1], Tanner Grudda[2], Giulia Belluccini[3], Mark Anderson[4], Gavin Cloherty[4], Hyon S. Hwang[1†], Monika Mani[1], Che-Min Lo[1], Naomi Esrig[1], Mark S. Sulkowski[1], Richard K. Sterling[5], Yang Zhang[6], Ruy M. Ribeiro[3], David L. Thomas[1], Chloe L. Thio[1, 2*‡], Ashwin Balagopal[1‡]

Affiliations:

[1] Department of Medicine, Johns Hopkins University School of Medicine; Baltimore, MD 21205, USA.

[2] Department of Molecular Microbiology and Immunology, Johns Hopkins Bloomberg School of Public Health; Baltimore, MD 21205, USA.

[3]Los Alamos National Laboratory; Los Alamos, New Mexico 87545, USA.

[4]Abbott Diagnostics; Abbott Park, Illinois 60064, USA.

[5]Divison of Gastroenterology, Hepatology, and Nutrition, Virginia Commonwealth University; Richmond, VA 23298, USA.

[6]Division of Gastrointestinal and Hepatic Pathology, Joint Pathology Center; Silver Spring, MD 20910, USA.

Footnotes:

†Current affiliation for HSH is University of Maryland Baltimore; Baltimore, Maryland, 21201, USA.

‡CLT and AB contributed equally to the work.

Corresponding Authors:

Chloe L Thio, cthio@jhmi.edu; 855 N Wolfe St, 5th Floor, Baltimore, MD 21205; ph:+1-410-614-6088

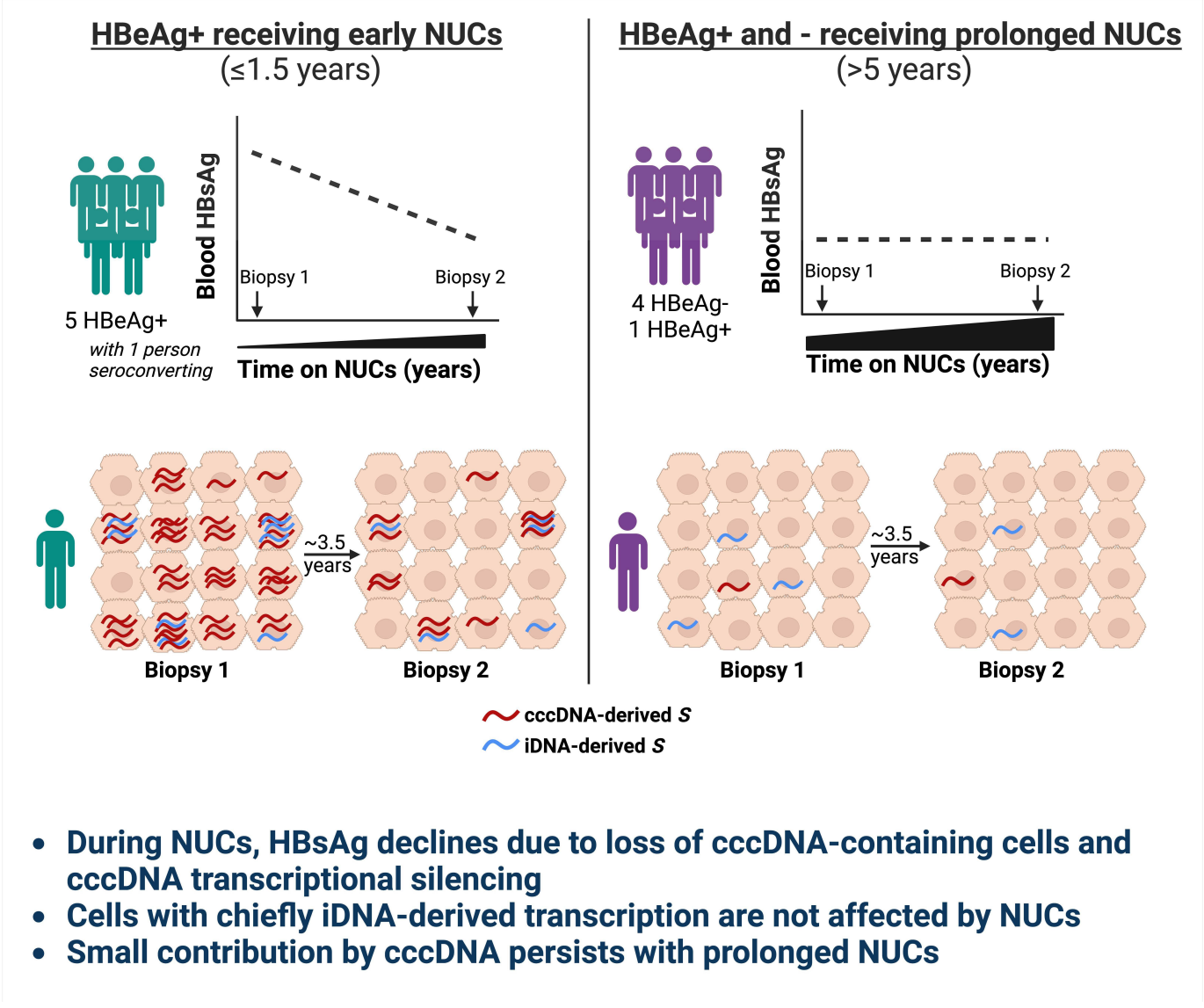
Ashwin Balagopal, abalago1@jhmi.edu; 855 N Wolfe St, 5th Floor, Baltimore, MD 21205; ph

Conflict of interest statement: Authors T.G., C.L.T., H.H., and A.B. share a patent that is used to discriminate the source of HBsAg as either cccDNA- or iDNA-derived. M.A., and G.C. are employees and shareholders of Abbott Laboratories. R.K.S. gets research support from Roche, Abbott, Abbvie, and Gilead.

Abstract

The cornerstone of functional cure for chronic hepatitis B (CHB) is hepatitis B surface antigen (HBsAg) loss from blood. HBsAg is encoded by covalently closed circular DNA (cccDNA) and HBV DNA integrated into the host genome (iDNA). Nucleos(t)ide analogues (NUCs), the mainstay of CHB treatment, rarely lead to HBsAg loss, which we hypothesized was due to continued iDNA transcription despite decreased cccDNA transcription. To test this, we applied a multiplex droplet digital PCR that identifies the dominant source of HBsAg mRNAs to 3436 single cells from paired liver biopsies from ten people with CHB and HIV receiving NUCs. With increased NUC duration, cells producing HBsAg mRNAs shifted from chiefly cccDNA to chiefly iDNA. This shift was due to both a reduction in the number of cccDNA-containing cells and diminished cccDNA-derived transcription per cell; furthermore, it correlated with reduced detection of proteins deriving from cccDNA but not iDNA. Despite this shift in the primary source of HBsAg, rare cells remained with detectable cccDNA-derived transcription, suggesting a source for maintaining the replication cycle. Functional cure must address both iDNA and residual cccDNA transcription. Further research is required to understand the significance of HBsAg when chiefly derived from iDNA.

Graphical Abstract



- **During NUCs, HBsAg declines due to loss of cccDNA-containing cells and cccDNA transcriptional silencing**
- **Cells with chiefly iDNA-derived transcription are not affected by NUCs**
- **Small contribution by cccDNA persists with prolonged NUCs**

INTRODUCTION

Chronic hepatitis B (CHB), which affects an estimated 300 million people worldwide, is a leading cause of hepatocellular carcinoma and end-stage liver disease. Although CHB can be treated with pegylated interferon α (PEG-IFN-α) and/or nucleos(t)ide analogues (NUCs), treatment typically only suppresses production of HBV DNA and rarely leads to loss of hepatitis B surface antigen (HBsAg) from blood, which is essential for meeting the current guidelines of a functional cure (1). Surface (*S*) mRNAs that are translated into HBsAg are transcribed from two sources—the covalently closed circular DNA (cccDNA) and HBV DNA that is integrated into the host genome (iDNA). It is important to understand the dominant source of transcription during treatment to uncover how HBsAg persists in most people despite suppression of HBV DNA production. Moreover, the dynamics of the HBsAg source during treatment are key for targeting novel therapies and understanding deficiencies of existing therapies. Such data are limited because existing clinical blood markers do not distinguish between these sources, so liver biopsies are required.

CHB proceeds through phases with a primary distinguishing feature being the presence of hepatitis B e antigen (HBeAg), which usually represents higher levels of HBV replication. *S* is transcribed primarily from cccDNA in HBeAg-positive CHB and from iDNA in HBeAg-negative CHB (2–4). However, in both HBeAg-positive and -negative CHB, we and others have reported that NUC treatment is associated with transcriptional silencing of cccDNA as well as a decrease in the overall number of infected hepatocytes (5–10). Although a decline in blood HBsAg would be expected with cccDNA transcriptional silencing and reductions in the number of infected cells, most people with CHB taking NUCs do not show substantial declines in HBsAg levels. Indeed, HBsAg loss during NUCs is rare (11, 12). These paradoxes strongly

indicate that *S* transcripts from iDNA along with persistent low level S transcription from cccDNA may maintain production of HBsAg during NUCs. This hypothesis is supported by our previous study of bulk liver from people with HIV and CHB. Individuals with chiefly cccDNA-derived transcription were more likely to have declines in blood HBsAg levels during NUCs compared to people with chiefly iDNA-derived transcription (2). To better understand the cellular and molecular dynamics underlying HBsAg production during NUCs, we employed single-cell laser capture microdissection (scLCM) and a droplet digital PCR (ddPCR) to paired biopsies from individuals with HIV and CHB who received varying durations of NUC treatment. We linked our findings to viral antigens in liver and serological testing of HBsAg in blood that are predicted to be differentially present during chiefly cccDNA- vs iDNA-derived transcription.

RESULTS

Study participants

All ten participants were males with a median age of 49 years at biopsy 1 (Table 1). Participants were stratified by the duration of NUC exposure at biopsy 1 into an *early group* and a *prolonged group*. In the five *early group* participants, the median NUC duration was 3 weeks (range 0-1.5 years) prior to biopsy 1 with a median (range) plasma HBV DNA level of 6.6 (4.2-8.6) $\log_{10}$ IU/mL that declined to 1.6 (UD-2.9) $\log_{10}$ IU/mL at biopsy 2, which was obtained a median of 3.7 years after biopsy 1. All five of the *early group* participants had a maximum decline in blood quantitative HBsAg (qHBsAg) of ≥ 0.5 $\log_{10}$ IU/ml during the study period relative to the level at biopsy 1 (median [range] decline 1.2 [0.5-2.0] $\log_{10}$ IU/mL) (Table 1, Supp Fig. 2). Four participants in this group were HBeAg-positive throughout the study while one participant seroconverted to HBeAg-negative shortly after biopsy 1. The five participants in the *prolonged group* had received NUCs for a median (range) of 7 (5-8) years prior to biopsy 1.

All except one participant had undetectable HBV DNA at both biopsies and had < 0.5 $\log_{10}$ IU/mL decline in qHBsAg during the study period relative to the level at biopsy 1 (Table 1, Supp Fig. 2, Supp Fig. 3). Four of those participants were HBeAg-negative and one was HBeAg-positive throughout.

cccDNA-derived transcription diminishes during NUCs

To understand changes in HBV transcription during NUCs in both groups, we first determined the proportion of cells with HBV transcripts in biopsy 1 and 2 from each participant by applying our scLCM/multiplex ddPCR (see Methods) approach to single-cell equivalents. Our multiplex ddPCR assay targets two regions of the HBV transcriptome: (i) the middle of the *S* gene (**mid-HBV**), which is present in both cccDNA- and iDNA-derived transcripts and captures transcripts that are translated into HBsAg, as well as longer transcripts (e.g., pregenomic RNA [pgRNA]) and (ii) the 3' terminus of cccDNA-derived transcripts (**3'-HBV**) that is absent in iDNA-derived transcripts because these have HBV-human hybrid chimeric junctions that lie in the domain between DR2 and the 3' canonical polyadenylation signal (PAS) (Fig. 1) (4). This strategy is supported by previously reported transcription maps and our own sequencing demonstrating that the vast majority of HBV-human hybrid junctions are indeed located upstream of our 3'-HBV assay, and therefore would not be detected by the assay (Supp Fig. 1 and Supp Methods).

After discarding cell fragments that did not meet our quality control, we examined a median (range) of 176 (85-271) cells in each biopsy, totaling 3,436 cells (see Methods). The proportion of cells with HBV transcripts (mid-HBV, 3'-HBV, or both) at biopsy 1 was higher in the *early* than *prolonged* participant group (median 98% vs 11%) (Fig. 2). In the *early* group, at biopsy 2, after a median 3.7 years on NUCs, these proportions declined in all participants with

one participant declining by 45%. In contrast, in the *prolonged* group, the proportions remained stable. Notably, the overall decrease in proportions of cells containing HBV transcripts in the *early* group by biopsy 2 made them appear to be more similar to the *prolonged* group at biopsy 1, although these proportions remained significantly greater in the *early* group at biopsy 2 compared to the *prolonged* group at biopsy 1 (median 7 years on NUCs) (median 58% vs 11%) (P=0.008). This result may reflect the longer duration on NUCs in the *prolonged* group biopsy 1 than the *early* group biopsy 2, but may also have a contribution from HBeAg status.

After characterizing changes in proportions of cells with HBV transcription during NUCs, we next focused on understanding whether there was a change in the *abundance* of viral transcripts (i.e., the mid-HBV and 3'-HBV transcripts) in actively-transcribing individual cells between biopsies, as we have previously shown for pgRNA (5–7). Thus, for each cell studied in each biopsy, we determined the quantity of the mid-HBV and 3'-HBV transcripts and calculated a median value per biopsy. In the *early* group participants, these values for mid-HBV quantities per cell decreased from a median (range) of 56 (4-232) copies/cell to 8 (0-92) copies/cell between biopsies 1 and 2 ($P<10^{-5}$) (Supp Fig. 4). Similarly, we also found that cccDNA-derived transcripts, as defined by the 3'-HBV amplicon, decreased from a median (range) of 112 (8-200) copies/cell at biopsy 1 to 0 (0-136) copies/cell at biopsy 2 ($P<10^{-5}$) (Supp Fig. 5). In contrast, in the *prolonged* group, the amount of HBV transcription per cell was low at both biopsies: the median was below detection for both the mid-HBV and 3'-HBV amplicons; therefore, at least 50% of cells did not contain these transcripts in any of the *prolonged* group participant biopsies. Because we tested hundreds of cells in each individual, we could still observe that the quantity per cell of the mid-HBV amplicon exhibited a small but significant decline (P=0.0001) in the *prolonged* group that was not seen with the 3'-HBV amplicon (P=0.2). Taken together, these

data suggest that changes in HBsAg during the first few years of treatment are likely due to both a decrease in the number of infected cells and a down-regulation of cccDNA-derived transcription in individual cells.

iDNA-derived transcription is maintained during NUCs

Since cccDNA transcription is downregulated and the number of infected hepatocytes decreases during NUCs, we tested the hypothesis that iDNA transcription continues contributing to the maintenance of circulating HBsAg during NUCs. To test this hypothesis, we focused on *cells actively transcribing S* in each biopsy and classified each cell as having chiefly cccDNA-derived, iDNA-derived, or mixed (both cccDNA and iDNA-derived) transcription, as determined by the iDNA transcriptional index (iDNA-TI) (see Methods) (2). The iDNA-TI is the ratio of the quantities of mid-HBV to 3'-HBV amplicons: a value ≤1 indicates chiefly cccDNA-derived transcription, >1 indicates mixed transcription, and only mid-HBV transcripts (no detectable 3'HBV transcripts) indicates chiefly iDNA-derived transcription.

Using this index, we determined the proportion of cells actively transcribing S with chiefly cccDNA-derived, iDNA-derived, or mixed transcription in each biopsy. Interestingly, in the *early* group, the proportion with chiefly iDNA-derived transcription enriched in every participant ranging from a 5% (HB11) to 85% (HB7) increase from biopsy 1 to biopsy 2 (Fig. 3, blue, $P<0.05$ for each participant, Supp Table). In contrast, cells with mixed transcription declined in 4/5 participants in the *early* group by 5% (HB7) to 44% (HB6) from biopsy 1 to biopsy 2 (Fig 3, purple, $P<0.05$ for all except HB7, Supp Table). Whereas the total number of cells actively transcribing S decreased during NUCs (Fig 3, smaller inner circle), the *proportion* of these cells with chiefly cccDNA-derived transcription was more variable across individuals: two showed a decrease (HB3 by 14% and HB7 by 80%) and the other three showed increases

(HB2 by 5%, HB6 by 27%, and HB11 by 33%). However, for two of the participants, HB2 and HB6, when studying the *total cell* population, the change in cccDNA-derived transcription between biopsies was nearly negligible because of the decrease in the number of infected cells. These data support that in the ~3-4 years between biopsies in the *early* group, cells with primarily iDNA-derived transcription were likely enriched from the diminished pool of hepatocytes with mixed transcription. In other words, early after NUCs, cells with both mixed transcription from iDNA and cccDNA undergo cccDNA transcriptional silencing during NUCs, yielding viral transcripts that are chiefly iDNA-derived.

This shift had already occurred in the *prolonged* group biopsies where fewer cells were transcriptionally active and per-cell transcription was dampened overall in comparison to *early* group participants. Further, at both biopsies in the *prolonged* group, very few cells appeared to have mixed transcription (median of 0% in both biopsies) supporting our findings from the *early* group that cells with mixed transcription decline with NUCs (Fig. 3). Notably, the proportion of cells with chiefly iDNA-derived transcription exceeded those with chiefly cccDNA-derived transcription in all *prolonged* group participants at both biopsies. The proportion of transcriptionally-active cells that were classified as chiefly iDNA-derived declined numerically in all participants ranging from 5% (HB12) to 35% (HB8) (but this was only significant in HB8, $P=0.002$, with $P>0.05$ for others, Supp Table) from biopsy 1 to biopsy 2, indicating that iDNA-derived transcription decays slowly if at all with long-term NUCs.

In the *prolonged* group participants, chiefly cccDNA-derived cells minimally enriched from biopsy 1 to biopsy 2 and was only significant in HB8 (Supp Table). It is important to note that both these changes were minimal in the context of the total cell population, and likely reflected relative changes between cell type rather than increases in total numbers of chiefly

cccDNA-derived cells. When summarizing the early and the prolonged NUC groups, we observed that as the total number of actively transcribing cells diminish, the chiefly iDNA-derived cell component takes an increasing fraction of actively transcribing cells (Fig 3). While chiefly iDNA-derived cells are expected in the prolonged NUC group since they are mostly HBeAg-negative, it was also true of the early NUC group at biopsy 2, who mostly remained HBeAg-positive at biopsy 2. Thus, cells with iDNA-derived transcription become dominant during prolonged NUCs, accounting for the majority of the continued HBsAg despite years of antiviral treatment. We also note that irrespective of HBeAg status and duration of NUCs, cccDNA-derived transcription persists at low levels in all people.

The transition from chiefly cccDNA-derived to chiefly iDNA-derived transcription is best exemplified by the *early* group participant (HB7) who was on NUCs the longest at biopsy 1 (1.5 years). The participant had a nearly complete transition from chiefly cccDNA-derived to chiefly iDNA-derived transcription between biopsies (7/98 [7.1%] to 80/87 [92%] cells with chiefly iDNA-derived transcription at biopsies 1 and 2, respectively, $P<0.0001$). By biopsy 2, HB7 resembled persons in the *prolonged* group both in terms of the source of viral transcription and in the duration of NUC therapy, although HB7 did not have a change in HBeAg status. Demonstrating this shift in HB7, the median (range) abundance of the mid-HBV amplicon was similar between biopsies (4 [0-457] and 8 [0-120] copies/cell at biopsies 1 and 2, respectively) while the median (range) abundance of 3' HBV declined from 8 (0-955) copies/cell to 0 (0-16) copies/cell, consistent with a reduction in cccDNA-derived transcription but maintenance of iDNA-derived transcription (Supp Fig. 4, 5).

Modeling decay of transcriptionally active cells

Focusing on cells actively transcribing S, we used our single cell data to develop mathematical models describing the decay of the fractions of cells transcribing S from chiefly cccDNA, mixed, or chiefly iDNA. We evaluated multiple models to describe empirically the dynamics of these different cell populations (see Methods), where the decay results from a balance of production of infected cells (new infections or cell division) and loss of infected cells (e.g., cell death). We found that the observed decays were best fit by independent and different rates in the *early* and *prolonged* groups. That is, the decay observed in the *prolonged* group was not consistent with long-term continuation of the decay seen in the *early* group. The most salient features were an initial slow decline (half-life of ~10 years) followed by an unexpected slow increase in the fraction of cells actively transcribing S from chiefly cccDNA at late times (doubling time of 7 years) (Fig. 4A). It is worth noting that the large variability in the data for the initial slow decline makes the estimated half-life less precise. We also observed a rapid decay in cells with mixed transcription in the *early* group, (half-life of 2.7 years) followed by a near loss in their detection in the *prolonged* group (Fig. 4B). This was accompanied by a fast increase in the *early* group in the fraction of cells actively transcribing S from chiefly iDNA, with a doubling time of ~1 year (Fig. 4C). These findings were substantiated by modeling the decay of the fraction of *total* cells exhibiting any 3'HBV amplicon or any mid-HBV amplicon. In the *early* group, the fraction of cells expressing the mid-HBV amplicon had a half-life of 5 years, which slowed down slightly to about 7.8 years in the *prolonged* group (Supp Fig. 6B). On the other hand, the half-life of cells with 3'-HBV amplicons mirrored that of the mixed population, 2.7 years. In contrast, there was essentially no decay of the fraction of cells with 3'-HBV amplicons in the *prolonged* group; indeed, we observe that these cells had a slow, but non-significant increase (doubling time of 9.2 years) (Supp Fig. 6A), which may be a consequence of

ongoing cccDNA transcription resulting in low level of ongoing infection (Fig 3). Taken together with the other decay rates, the observed increase in the fraction of cells with chiefly iDNA-derived transcription is likely due to a decrease in the population of cells with mixed transcription: as cells in the mixed population undergo cccDNA transcriptional silencing, these cells chiefly transcribe from iDNA.

Linking viral protein production in liver to the source of viral transcription

With an understanding of changes in HBV transcription in hepatocytes, we next sought to reconcile HBsAg protein production with the source of viral transcription, since historically immunohistochemistry of HBsAg correlates inconsistently with HBV replication (13). Transcriptional maps reported previously indicate that the full suite of viral proteins is not likely to be produced from iDNA (14, 15) even though rare integrations contain nearly the entire coding sequence of the HBV genome (Fig. 1), especially because the genomic organization of the circular HBV genome is disrupted in the linear iDNA. Specifically, the promoter and enhancer regions for the core gene are separated from the core gene by HBV-human chimeric junctions (Fig. 1 and Supp Fig. 1). Therefore, because the core protein is largely derived from cccDNA, we hypothesized that viral transcription and core protein production are correlated during cccDNA-derived but not during iDNA-derived transcription. To test this, we predicted that the proportion of cells with positive HBsAg staining by immunohistochemistry (IHC) would correlate with the proportion of cells with any cccDNA-derived transcription only in tissues that were positive for HBcAg by IHC. Of the 20 biopsies, 11 had detectable HBcAg, which included all of the *early* group and one *prolonged* group participant (Fig. 5A). As expected, we observed a correlation in HBcAg-positive biopsies between the proportion of transcriptionally-active cells with any evidence of cccDNA-derived transcription and the percentage of cells positive for

HBsAg staining (R^2=0.78, P=0.004). In contrast, among HBcAg-negative biopsies, there was no observable association (R^2=-0.084, P=0.83) (Fig. 5B).

Source of transcription affects changes in blood viral markers

We next investigated whether peripheral blood markers can provide insight into the intrahepatic source of transcription. In this and prior studies, we and others have shown that NUCs are associated with silencing of cccDNA transcription but not iDNA transcription (5–10) Therefore, we hypothesized that there would be a greater decline in blood qHBsAg between biopsies when most cells at biopsy 1 have cccDNA-derived rather than iDNA-derived transcription. To test this hypothesis, we determined the maximum qHBsAg decline between biopsy 1 and any point after biopsy 1. Among the *early* group, a median (range) of 96% (53%-100%) of cells had some cccDNA-derived transcription at biopsy 1: these participants had a median (range) of 1.2 $\log_{10}$ (0.5-2.0) IU/mL reduction in qHBsAg. In contrast, among *prolonged* group participants, the median (range) proportion of cells with cccDNA-derived transcription was 2.3% (2.2%-7.8%) at biopsy 1 and qHBsAg values did not change appreciably (median [range] 0.3 [0.06-0.35] $\log_{10}$ IU/mL decline) (Fig. 6). These results demonstrate that when most cells exhibit cccDNA-derived transcription, qHBsAg levels decline with NUC treatment.

We also separately explored whether HBsAg isoforms in blood were correlated with the source of transcription. Transcription maps from HBeAg-negative individuals, who transcribe chiefly from iDNA, suggest that the region encoding for (L)arge-HBsAg is integrated less frequently than those encoding (M)iddle- and (Sm)all-HBsAg (Fig. 1) (4), and its promoter may have greater dependence on EnhII than the promoters for the latter isoforms (16). Thus, we predicted that persons with chiefly cccDNA-derived transcription would be more likely to produce L-HBsAg than persons with chiefly iDNA-derived transcription. Supporting this

hypothesis, we found that the proportion of transcriptionally-active cells with any cccDNA-derived transcription, as determined by our iDNA-TI, correlated with L-HBsAg quantities in blood at time of biopsy (R^2=0.66, P=0.004) (Supp Fig. 7A, B).

Conversely, we tested whether the proportion of HBsAg that is Sm-HBsAg could be inferred from the proportion of cells with cccDNA- vs iDNA-derived transcription. Since our assay does not directly measure Sm-HBsAg and because L-HBsAg comprises a small portion of the HBsAg isoforms, we estimated the relative proportion of Sm-HBsAg by determining the ratio of M-HBsAg to (T)otal-HBsAg (which captures all isoforms) such that a higher M-HBsAg:T-HBsAg (M/T) ratio corresponds to lower proportions of Sm-HBsAg (17, 18). We found that the *early* group had a lower proportion of Sm-HBsAg compared to the *prolonged* group (P=0.03) (Supp Fig. 8A). Furthermore, liver biopsies with chiefly iDNA-derived transcription correlated inversely with the M/T ratio (R^2=-0.66, P=0.003), consistent with higher Sm-HBsAg proportions in blood at the time of biopsy in people with chiefly iDNA-derived transcription (Supp Fig. 8B). Taken together, these observations suggest that in persons with *prolonged* NUCs, more of the circulating HBsAg is made up of Sm-HBsAg, which may be driven by iDNA-derived transcription. Collectively, our results support that decreases in the number of cells with cccDNA-derived transcription during NUCs are associated with decreases in both qHBsAg and in L- compared to Sm-HBsAg.

DISCUSSION

Applying high-resolution molecular tools to paired liver biopsies from ten people co-infected with HIV and HBV, we demonstrate that with longer NUC duration *S* transcripts are maintained primarily by iDNA transcription albeit with a small but persistent contribution from cccDNA: this observation was observed irrespective of HBeAg status. Early NUC exposure in

303 people with chiefly cccDNA-derived transcription (who were HBeAg positive) was associated
304 with a decline in blood HBsAg levels. Further, whereas people with chiefly cccDNA-derived
305 transcription had detectable viral antigens in liver and blood, these were not detected in people
306 with chiefly iDNA-derived transcription. Although participants with *prolonged* NUC exposure
307 had fewer transcriptionally-active cells than those with *early* NUC exposure, a higher proportion
308 of these transcriptionally-active cells in the *prolonged* group had iDNA-derived rather than
309 cccDNA-derived transcription. Consistent with this, amongst transcriptionally-active cells, we
310 observed a multi-phase decay showing a slow decline in cells with predominantly cccDNA-
311 derived transcription and a relatively fast increase in cells with iDNA-derived *S*. Most
312 importantly, participants with mainly iDNA-derived transcription at biopsy 1 had minimal
313 changes in blood qHBsAg with continued NUC treatment, consistent with the lack of predicted
314 activity of NUCs on iDNA transcription.

315 NUCs prevent new infection events, resulting in a diminished number of infected cells,
316 but they are also associated with cccDNA transcriptional silencing (5–10). Although the precise
317 mechanism of NUC-associated cccDNA transcriptional silencing is yet to be determined, it may
318 be due tohost factors that target cccDNA, mutations induced in the viral genome, epigenetic
319 changes, or some combination of these. Thus, it has been hypothesized that blood HBsAg during
320 prolonged NUCs is maintained primarily by iDNA-derived HBsAg. In this study, we uncover the
321 single-cell composition of HBV-infected hepatocytes that produce and maintain HBsAg levels
322 during NUCs: we observe a global transition from primarily cccDNA-derived to primarily
323 iDNA-derived transcription with longer NUC duration. Whereas the *prolonged* group
324 participants were largely HBeAg-negative, this transition occurred even in the *early* group
325 participants who did not seroconvert their HBeAg status. It seems likely that this transition is due

to cccDNA-derived transcription dominating *S* production until the majority of the cccDNA is either silenced or those hepatocytes decay during NUCs, yielding detectable iDNA *S* transcription. This is also consistent with the near disappearance of cells with mixed transcription in *prolonged* treated individuals: we hypothesize that the cccDNA in cells with mixed transcription is progressively silenced, yielding a greater proportion of cells that have chiefly iDNA-derived transcription that enrich that population. It is also notable that in the *prolonged* group the proportions of hepatocytes with either iDNA-derived or cccDNA-derived transcription were relatively stable, suggesting that either these cells evade immune or senescent clearance due to inadequate antigen presentation or that low-level replication maintains these cells. It is also important to note that residual cccDNA transcription with prolonged NUC treatment may replenish both cccDNA and iDNA in hepatocytes. Thus, in prolonged NUC treatment, HBsAg levels are a result of the balance between the natural decay of hepatocytes and replenishment of cccDNA and iDNA.

Our modeling revealed multi-phase decays that are consistent with several subpopulations of infected cells decaying at different rates. We found evidence of two populations of cells with cccDNA transcription– the first, containing cells transcribing from both cccDNA and iDNA (mixed), decayed relatively quickly when NUCs were started (half-life=2.7 years) and largely disappeared with longer NUCs; the second population, consisting of cells transcribing chiefly from cccDNA, also initially decayed when NUCs were started but then were maintained with longer NUC duration and, surprisingly, showed slower decay than cells with chiefly iDNA-derived transcription. This longer-lived subpopulation of chiefly cccDNA-transcribing cells that persists after years of NUCs is consistent with a reservoir of infection that has persistent, low-level transcription, which is also supported by another study demonstrating

ongoing evolution of HBV in people suppressed with long-term NUCs (19). Another possibility for the apparent slower decay of chiefly cccDNA-transcribing cells compared to mixed cells arises if cccDNA transcriptional silencing is incomplete: in other words, when cccDNA is partially silenced in mixed cells, they result in cells with chiefly iDNA-derived transcription, whereas when cells with chiefly cccDNA-derived transcription are partially silenced, they still yield cells with chiefly cccDNA-derived transcription but with a lesser magnitude of viral transcription. A third population of cells that we found were those with chiefly iDNA-derived transcription that increased quickly as a proportion of transcriptionally-active cells after NUC initiation, and then showed slow decay with longer NUC duration. Taken together, these results are consistent with a loss in cells with chiefly cccDNA-derived transcription early during NUCs due to i) prevention of new infections and ii) cccDNA transcriptional silencing. This is followed by a natural slow decline of all infected cells due to cell death. Since cells with chiefly iDNA-derived transcription were largely found in the absence of viral antigens such as HBcAg, this slowly decaying population may represent infected cells that present a restricted suite of viral antigens, allowing them an added measure of evasion from immune surveillance (20). Although multi-phase decay supports the hypothesis of distinct subpopulations of infected hepatocytes decaying at different rates, it is possible that NUC-associated transcriptional silencing of cccDNA leads to changes (i.e., increase) in the half-life of cells that contain viral genomes. An alternate explanation for differential decay may be differences in hepatocyte renewal mechanisms. One study demonstrated that higher ploidy hepatocytes are less likely to divide into daughter cells, thus allowing multi-nucleated hepatocytes to persist longer (21). These infected cells may be a long-lived population thus serving as a quiescent reservoir for ongoing replication. Since repeated mitosis of cccDNA-containing hepatocytes has been suggested as a

potential strategy for cccDNA clearance, a slower rate of cell division in a subpopulation of cells containing cccDNA may explain its long-term persistence (20, 22).

There has been great interest in understanding how intrahepatic markers of transcription are reflected in peripheral blood biomarkers, particularly HBsAg, since functional cure requires its loss. Earlier studies found that blood qHBsAg correlated well with intrahepatic cccDNA levels prior to treatment initiation, but these parameters correlated poorly after extended NUCs (11, 13, 23, 24). Our findings help to explain the lack of correlation between measures of HBV replication that center on cccDNA quantities and blood qHBsAg levels in people taking prolonged NUCs, since cccDNA in this scenario is largely transcriptionally silent while HBsAg is mostly derived from iDNA. We extend these observations by demonstrating that in participants with a high proportion of cells with cccDNA-derived transcription, there was a decline in qHBsAg with NUCs, whereas that was not the case in participants with chiefly iDNA-derived transcription. This is consistent with other studies showing that the presence of integrated *S* at baseline in people who are NUC-naive was associated with poorer responses in qHBsAg decline (25, 26). It is worth noting that we and others have observed that transcriptionally active integrations decrease during NUC therapy; however, this pattern remains consistent with HBsAg trends (Supp Fig. 1) (27, 28). We also demonstrate that HBsAg isoforms may reflect intrahepatic HBV transcription, since the proportion of cells transcribing from cccDNA correlated with amounts of L-HBsAg in blood. Similarly, qHBsAg decline during NUCs, which is greater with cccDNA transcription, was also greater in people with more L-HBsAg. Conversely, though iDNA effectively transcribes PreS2/S, the proportion of Sm-HBsAg appeared higher in the *prolonged* group than in the *early* group. Pfefferkorn et al. found that in people who had HBsAg loss with NUCs, L- and M-HBsAg declined before total HBsAg, which

may reflect the silencing or elimination of hepatocytes with chiefly cccDNA-derived transcription (29). Thus, we present mechanistic and observational support that blood biomarkers reveal clues about intrahepatic transcription, although further work is needed to validate these as clinically useful biomarkers and to potentially explore others.

We encountered several challenges during this study. First, despite the intensive single-cell investigation, the core biopsy is a fraction of the whole liver, so there is the possibility of sampling error. It is encouraging that our findings correlate closely with blood markers and with liver tissue staining, supporting that the cells were representative of the liver. A second limitation is that these individuals were all from one geographic region, males, and with HIV receiving TDF. Single-cell studies will need to be expanded into other geographic regions, females, people on entecavir, and people without HIV to confirm our findings. A third limitation is that we focused on active transcription by studying cccDNA-derived and iDNA-derived transcripts rather than the total number of cccDNA molecules or integrations. While we have reported on the former earlier, the absence of direct measurements of iDNA was intentional, since our aim was to explain the persistence of HBsAg, which must involve transcriptionally active iDNA. It is possible that the dynamics of total integrations, many of which might be transcriptionally inactive, may be distinct compared to the dynamics of only the transcriptionally-active iDNA examined here (30). A fourth limitation is that heterogeneity exists regarding the timing of NUC initiation relative to biopsy 1 and NUC adherence between biopsies in the early group (Supp. Fig. 3). However, despite this, S transcription is dominated by cccDNA early during NUCs. A fifth limitation is that our modeling was based on five individuals in both the *early* and *prolonged* groups, which is partially mitigated by the large number of cells studied. Lastly, all of the *early* group participants were HBeAg-positive and only one *prolonged* group

participant was HBeAg-positive, so our conclusions about NUC duration enhancing iDNA-derived HBsAg cannot be fully distinguished from HBeAg-negative status. Overall, we observed that by biopsy 2, the *early* group proportions of transcriptionally active cells began to resemble biopsy 1 from the *prolonged* group. This was best exemplified by HB7 in the *early* group, who had been on NUC the longest and had the same transcriptional characteristics as the HBeAg-negative biopsies in that group by single-cell analysis, supporting our conjecture that increasing time on NUCs enhances iDNA-derived HBsAg irrespective of HBeAg status. Nevertheless, it will be important to verify our findings in more people with CHB at different stages of disease and different durations of treatment.

By focusing on single HBV-infected hepatocytes as the unit of functional cure, we offer a granular understanding of how the source of *S* transcription changes from cccDNA to iDNA during NUC treatment. We further show that, consistent with the shift toward iDNA-derived transcription, there are decreases in cccDNA-derived viral antigens such as intrahepatic HBcAg and blood L-HBsAg during prolonged NUC treatment, offering insights about these markers in people with chiefly cccDNA-derived but not iDNA-derived transcription. Importantly, during long-term NUC therapy, we still found occasional cells with low-levels of cccDNA-derived transcription that persisted. Modeling revealed that achieving a functional cure will require addressing the slowly decaying iDNA-derived *S* transcripts and permanently silencing cccDNA transcription. Moreover, clinical trials of existing and emerging agents should incorporate liver biopsies to understand how the complex viral reservoir of infected hepatocytes responds to treatment.

METHODS

Sex as a biological variable

441 Our study exclusively examined male individuals because samples from female
442 individuals were not available. We do not expect these findings to vary by sex.

443 *Study participants*

444 This study included previously collected paired core liver biopsies from ten males with HIV
445 and CHB (20 total biopsies) enrolled in the HIV-HBV Cohort Ancillary Study of the Hepatitis B
446 Research Network at Johns Hopkins University (31). The paired liver biopsies were obtained a
447 median of 3.7 years apart. At the time of biopsy 1, four individuals had HBeAg-negative CHB
448 and six individuals had HBeAg-positive CHB, including one individual who underwent HBeAg
449 seroconversion about 6 months after biopsy 1 (HB11). Participants were on varying durations of
450 nucleos(t)ide analogue (NUC) treatment (see Results; Table 1) as part of their antiretroviral
451 therapy (ART), which included tenofovir disoproxil fumarate (TDF), emtricitabine (FTC), or
452 entecavir (ETV). We classified people by their exposure to NUCs prior to biopsy 1: an *early*
453 *group* and a *prolonged group* (defined in Table 1). HBV DNA levels between biopsy 1 and 2
454 are shown in Supp Figure 3.

455 *Single-cell laser capture microdissection (scLCM) and DNA/RNA extraction*

456 At the time of biopsy, liver tissues were immediately placed into neutral optimal temperature
457 cutting media and stored in liquid nitrogen until use. Tissues were cryosectioned at 10 μm
458 thickness onto PEN membrane slides; single-hepatocyte equivalents were individually isolated in
459 a grid fashion, as previously described (5–7). Each hepatocyte was deposited into a
460 microcentrifuge tube with proprietary lysis buffer (ZR-Duet DNA/RNA MiniPrep kit, Zymo
461 Research). RNA and DNA were separately extracted, including a DNase I in-column digestion to
462 purify RNA, as previously described (5–7). Total complementary DNA (cDNA) was synthesized

463 from RNA using Oligo (dT) and random hexamer priming with the Superscript IV First-Strand
464 Synthesis System. An abundant host cytoplasmic RNA, 7SL, was measured in every capture
465 using real-time reverse-transcription quantitative PCR (RT-qPCR) to assess for cell
466 fragmentation. Captures within 1 standard deviation below a negative control cycle threshold
467 were excluded from the final analysis, as previously described (5–7)

468 After the quality control assessment to filter out cell fragments, a median of 172 single-cell
469 equivalents were analyzed per biopsy with 3,436 hepatocytes analyzed in 20 biopsies. For this
470 study, a cell was considered transcriptionally active if there were detectable ddPCR targets (Fig.
471 1).

472 *Multiplex ddPCR for cccDNA-derived and iDNA-derived S transcripts*

473 Our multiplex ddPCR assay targets two amplicons along the HBV transcriptome (Fig. 1A):
474 we exploit the abundance of viral-human chimeric junctions in mRNA that derives from iDNA
475 that frequently lack the canonical poly A signal (PAS). The **mid-HBV** amplicon (nt. 253-418)
476 captures transcripts derived from both cccDNA and iDNA (specifically the 3.5kb, 2.4kb, and
477 2.1kb transcripts) whereas the **3'-HBV** amplicon (nt. 1774-1881) is just upstream of the PAS
478 (around 40 nt) and largely captures cccDNA-derived transcripts rather than iDNA-derived
479 transcripts (Fig. 1), as previously described (2). cDNA along with primers/probes targeting the
480 mid-HBV and 3'-HBV ddPCRs were run using the following cycling parameters: 1 cycle of
481 94°C for 10 minutes, 40 cycles of 94°C for 30 seconds and 57°C for 1 minute, 1 cycle of 98°C
482 for 10 minutes, 1 cycle of 12°C for 10 minutes, and 1 continuous cycle at 4°C until reading, as
483 previously described (5–7). Plates were read on the QX200 Droplet Reader (Bio-Rad), which
484 provides results of the copies of each amplicon per microliter of reaction.

Analysis

We defined cells as transcriptionally active if they had positive droplets for either the mid-HBV or 3'-HBV assays. For each transcriptionally active cell, we calculated an iDNA transcriptional index (iDNA-TI) using the ratio of the number of copies of the mid-HBV to 3'-HBV amplicons, as previously described (2). Since the 3'-HBV amplicon only detects cccDNA-derived transcripts, we classified viral transcription in each cell as chiefly cccDNA-derived (iDNA-TI£1), mixed cccDNA- and iDNA (iDNA-TI>1), or chiefly iDNA-derived (only mid-HBV+).

Immunohistochemistry (IHC)

Glass slides were stained at The Johns Hopkins Pathology Center with immunohistochemical stains for hepatitis B core (HBcAg) and surface antigens, and these were processed and stained as previously described (6, 7).

Serological assessment

Frequent plasma sampling was performed in tandem with biopsies and tested for quantitative HBsAg (qHBsAg) and quantitative HBeAg (qHBeAg) using the Roche Diagnostics Elecsys platform according to the manufacturer's instructions. HBV viral load was determined using the Roche COBAS TaqMan assay with an LOD. Units for both were reported as $\log_{10}$ IU/mL.

Statistics and modeling

Treatment groups were compared to one another using the Wilcoxon's rank-sum test. To compare the levels of different transcripts per cell (mid-HBV or 3'HBV) across people and time

(biopsy 1 and 2), we used linear mixed effects models (with participant as the random factor) using R package lmerTest v.3.1.3 (32), because of the hierarchical repeated nature of the data (measurements for multiple cells within individuals). Comparing levels of transcripts *within* people between biopsies 1 and 2, we used a pairwise Wilcoxon's rank sum test in R ("pairwise.wilcox.test"). Spearman's correlation coefficient analyses were performed to assess relationships between variables. Fisher's uncondExact 2 × 2 function was used to analyze changes in proportions of cells within individuals. All statistical tests were done using R Statistical Software v. 4.2.2 (33). A P value < 0.05 was considered statistically significant.

We calculated decline rates of the fraction of cells infected with predominance of different transcripts using non-linear mixed-effects models with the software Monolix (Lixoft, Antony, France) (34). We tested three different models to describe the decay data across all individuals (early and prolonged treatment). In the first model, we assumed that the decay of the fraction of cells followed a single exponential decay across all individuals (including those in the early and prolonged groups), which implies a single population of infected cells decaying over the time span of the analyses. The second model assumed that the decay was biexponential across all, which can be interpreted as two populations of cells decaying at different rates from the start of treatment. The first population is dominant at early times and decays faster, whereas the second, slower, cell population dominates at later times. The third model assumed that the early and prolonged treatment groups had separate independent single exponential decays, which could indicate that there are two different (at least in terms of decay rates) populations of cells early and later in treatment. The models were compared with the corrected Bayesian Information Criteria (cBIC), as provided by Monolix, where lower cBIC corresponds to a better fit (35). Note that we calculated the fraction of infected cells (with mid-HBV or 3'-amplicons) over all cells.

529 Since we can assume that the total number of hepatocytes is stable within each individual (the
530 liver size is tightly controlled), the decay rate of the fraction of cells is equivalent to the decay
531 rate of total cells containing those mRNA species. On the other hand, proportions of cells in
532 different transcriptional classes (e.g. chiefly cccDNA) were calculated over infected cells, i.e.,
533 transcriptionally active cells. For completeness, we also tested a simpler linear decay model,
534 which did not accurately describe the data. The best model (lower cBIC) for almost all cases was
535 the one assuming two independent single exponential decays for the *early* and *prolonged* groups,
536 which is what we present in the Results.

537 *Study approval*

538 This study was reviewed and approved by the Office of Human Subjects Research
539 Institutional Review Board, IRB-3, Baltimore, Maryland, USA. Participants gave written
540 informed consent for use of their tissues for research purposes through the Hepatitis B Research
541 Network.

542 *Data availability*

543 Data can be accessed through Vivli (www.search.vivli.org).

544

Author Contributions:

MT, DLT, CLT, and AB conceptualized the study. MT, TG, GB, MA, HSH, MM, CL, YZ, RMR, CLT, and AB developed the methodology. MT, TG, GB, MA, MM, CL, NE, YZ, RMR, CLT, AB carried out the investigation. MT, GB, MA, RMR, CLT, AB conducted analyses. MT, GB, CLT, AB contributed to visualization and generated figures. GC, MSS, RKS, RMR, CLT, AB acquired funding. MT, TG, GB, MA, GC, HSH, MM, CL, NE, MSS, RKS, YZ, RMR, DLT, CLT, AB wrote the manuscript draft or reviewed and edited it.

Acknowledgements: We thank the study participants for their time and for the liver biopsies. We also thank the study coordinators for facilitating. We would like to thank Yasmeen S. Saad, Kristina Zambo, and Jessica Mimms for their contributions to sample processing and figure annotations. The graphical abstract was created with Biorender.com.

Funding: This work was supported by National Institute of Health**: R01 AI116868** (RMR and AB)**, R56 AI138810** (CLT and AB), **R21 AI157760** (CLT and AB), **K24 DA034621** (MSS), **R01 DK094818** (RKS), **T32 AI007417** (TG) and **F31 AI179435** (TG)

References

1. Ghany MG, et al. Guidance on treatment endpoints and study design for clinical trials aiming to achieve cure in chronic hepatitis B and D: Report from the 2022 AASLD-EASL HBV-HDV Treatment Endpoints Conference. *Hepatology*. 2023;78(5):1654–1673.

2. Grudda T, et al. Integrated hepatitis B virus DNA maintains surface antigen production during antiviral treatment. *Journal of Clinical Investigation*. 2022;132(18). https://doi.org/10.1172/JCI161818.

3. Meier M-A, et al. Ubiquitous expression of HBsAg from integrated HBV DNA in patients with low viral load. *J Hepatol*. 2021;75(4):840–847.

4. Wooddell CI, et al. RNAi-based treatment of chronically infected patients and chimpanzees reveals that integrated hepatitis B virus DNA is a source of HBsAg. *Sci Transl Med*. 2017;9(409). https://doi.org/10.1126/scitranslmed.aan0241.

5. Balagopal A, et al. Single Hepatocyte Hepatitis B Virus Transcriptional Landscape in HIV Coinfection. *J Infect Dis*. 2020;221(9):1462–1469.

6. Balagopal A, et al. Single hepatocytes show persistence and transcriptional inactivity of hepatitis B. *JCI Insight*. 2020;5(19). https://doi.org/10.1172/jci.insight.140584.

7. Thio CL, et al. Hepatitis B e Antigen-Negative Single Hepatocyte Analysis Shows Transcriptional Silencing and Slow Decay of Infected Cells With Treatment. *J Infect Dis*. 2023;228(9):1219–1226.

8. Lebossé F, et al. Quantification and epigenetic evaluation of the residual pool of hepatitis B covalently closed circular DNA in long-term nucleoside analogue-treated patients. *Sci Rep*. 2020;10(1):21097.

9. Kim SC, et al. Efficacy of antiviral therapy and host–virus interactions visualised using serial liver sampling with fine-needle aspirates. *JHEP Reports*. 2023;5(9):100817.

10. Lai C-L, et al. Reduction of covalently closed circular DNA with long-term nucleos(t)ide analogue treatment in chronic hepatitis B. *J Hepatol*. 2017;66(2):275–281.

11. Chevaliez S, et al. Long-term hepatitis B surface antigen (HBsAg) kinetics during nucleoside/nucleotide analogue therapy: Finite treatment duration unlikely. *J Hepatol*. 2013;58(4):676–683.

12. Lok ASF. Toward a Functional Cure for Hepatitis B. *Gut Liver*. [published online ahead of print: March 27, 2024]. https://doi.org/10.5009/gnl240023.

13. Thompson AJV, et al. Serum hepatitis B surface antigen and hepatitis B e antigen titers: Disease phase influences correlation with viral load and intrahepatic hepatitis B virus markers. *Hepatology*. 2010;51(6):1933–1944.

14. Tu T, et al. HBV DNA Integration: Molecular Mechanisms and Clinical Implications. *Viruses*. 2017;9(4):75.

15. van Buuren N, et al. Targeted long-read sequencing reveals clonally expanded HBV-associated chromosomal translocations in patients with chronic hepatitis B. *JHEP Reports*. 2022;4(4):100449.

16. Yuh CH, Ting LP. The genome of hepatitis B virus contains a second enhancer: cooperation of two elements within this enhancer is required for its function. *J Virol*. 1990;64(9):4281–4287.

17. Rodgers MA, et al. Characterization of HBV surface antigen isoforms in the natural history and treatment of HBV infection. *Hepatol Commun*. 2023;7(4). https://doi.org/10.1097/HC9.0000000000000027.

18. Bazinet M, et al. HBsAg isoform dynamics during NAP-based therapy of HBeAg-negative chronic HBV and HBV/HDV infection. *Hepatol Commun*. 2022;6(8):1870–1880.

19. Yu T, et al. Evidence of Residual Ongoing Viral Replication in Chronic Hepatitis B Patients Successfully Treated With Nucleos(t)ide Analogues. *J Infect Dis*. 2023;227(5):675–685.

20. Mason WS, et al. Hepatocyte turnover in transient and chronic hepadnavirus infections. *J Viral Hepat*. 2007;14(s1):22–28.

21. Heinke P, et al. Diploid hepatocytes drive physiological liver renewal in adult humans. *Cell Syst*. 2022;13(6):499-507.e12.

22. Tu T, et al. Mitosis of hepatitis B virus-infected cells in vitro results in uninfected daughter cells. *JHEP Reports*. 2022;4(9):100514.

23. Manesis EK, et al. Hepatitis B surface antigen: Relation to hepatitis B replication parameters in HBeAg-negative chronic hepatitis B. *J Hepatol*. 2011;55(1):61–68.

24. Lin LY, et al. Relationship between serum hepatitis B virus DNA and surface antigen with covalently closed circular DNA in HBeAg-negative patients. *J Med Virol*. 2010;82(9):1494–1500.

25. Hu B, et al. Integration of hepatitis B virus S gene impacts on hepatitis B surface antigen levels in patients with antiviral therapy. *J Gastroenterol Hepatol*. 2018;33(7):1389–1396.

26. Erken R, et al. Quantified integrated hepatitis B virus is related to viral activity in patients with chronic hepatitis B. *Hepatology*. 2022;76(1):196–206.

27. Hsu Y-C, et al. Inhibition of Viral Replication Reduces Transcriptionally Active Distinct Hepatitis B Virus Integrations With Implications on Host Gene Dysregulation. *Gastroenterology*. 2022;162(4):1160-1170.e1.

28. Yu X, et al. Spatial transcriptomics reveals a low extent of transcriptionally active hepatitis B virus integration in patients with HBsAg loss. *Gut*. 2024;73(5):797–809.

29. Pfefferkorn M, et al. Composition of HBsAg is predictive of HBsAg loss during treatment in patients with HBeAg-positive chronic hepatitis B. *J Hepatol*. 2021;74(2):283–292.

30. Chow N, et al. Effect of Antiviral Treatment on Hepatitis B Virus Integration and Hepatocyte Clonal Expansion. *Clinical Infectious Diseases*. 2023;76(3):e801–e809.

31. Lisker-Melman M, et al. HBV transcription and translation persist despite viral suppression in HBV-HIV co-infected patients on antiretroviral therapy. *Hepatology*. 2023;77(2):594–605.

32. Kuznetsova A, Brockhoff PB, Christensen RHB. **lmerTest** Package: Tests in Linear Mixed Effects Models. *J Stat Softw*. 2017;82(13). https://doi.org/10.18637/jss.v082.i13.

33. R Core Team. R: A Language and Environment for Statistical Computing [preprint]. 2021. https://www.R-project.org/. Accessed June 6, 2024.

34. Lavielle M. *Mixed Effects Models for the Population Approach*. Chapman and Hall/CRC; 2014.

35. Burnham KP, Anderson DR. *Model Selection and Multimodel Inference*. New York, New York: Springer; 2004.

36. Shieh F-S, et al. ChimericSeq: An open-source, user-friendly interface for analyzing NGS 633 data to identify and characterize viral-host chimeric sequences. PLoS One. 634 2017;12(8):e0182843.

652 **Figures and Tables**

653 **Table 1. Participant characteristics**

	Early					Prolonged				
Participant ID	**HB2**	**HB6**	**HB11[A]**	**HB3**	**HB7**	**HB12**	**HB9**	**HB4**	**HB8**	**HB10**
Age (years)	46	53	41	28	47	47	57	51	54	57
Time on NUCs	0	2 weeks	3 weeks	7 months	1.5 years	5 years	6 years	7 years	8 years	8 years
Interval between biopsies (years)	3.6	3.7	3.7	2.7	3.7	2.8	3.7	3.6	3.7	3.5
HBeAg status	Pos/Pos	Pos/Pos	Pos/Neg	Pos/Pos	Pos/Pos	Neg/Neg	Neg/Neg	Pos/Pos	Neg/Neg	Neg/Ne
CD4+ T cell count (cells/uL)	153	655	629	390	299	267	718	399	631	557
HIV-1 RNA (cp/mL)	358	UD	52027	54	UD	UD	UD	UD	UD	UD
HBV DNA (log_{10} IU/mL) @ biopsy 1	8.6	8.5	6.6	4.8	4.2	UD	UD	1.6	UD	UD
HBV DNA (log_{10} IU/mL) @ biopsy 2	UD	1.6	UD	2.9	2.9	UD	UD	UD	UD	UD
qHBsAg (log_{10} IU/mL) @ biopsy 1	5.3	5.8	4.4	5.0	2.8	4.3	2.6	3.3	3.2	3.1
qHBsAg (log_{10} IU/mL) @ biopsy 2	3.3	4.6	3.7	5.1	2.3	4.3	2.4	3.0	2.9	2.8

654 **Table 1.** Participant characteristics at biopsy 1, unless otherwise stated. Participants are
655 presented in order of treatment duration at biopsy 1. Abbreviations: NUCs, nucleos(t)ide
656 analogues; HBeAg, hepatitis B e antigen; HIV-1, human immunodeficiency virus 1; HBV,
657 hepatitis B virus; qHBsAg, quantitative hepatitis B surface antigen. Note, all participants were
658 adult Black men. [A]Seroconverted to HBeAg-negative CHB after biopsy 1.

659

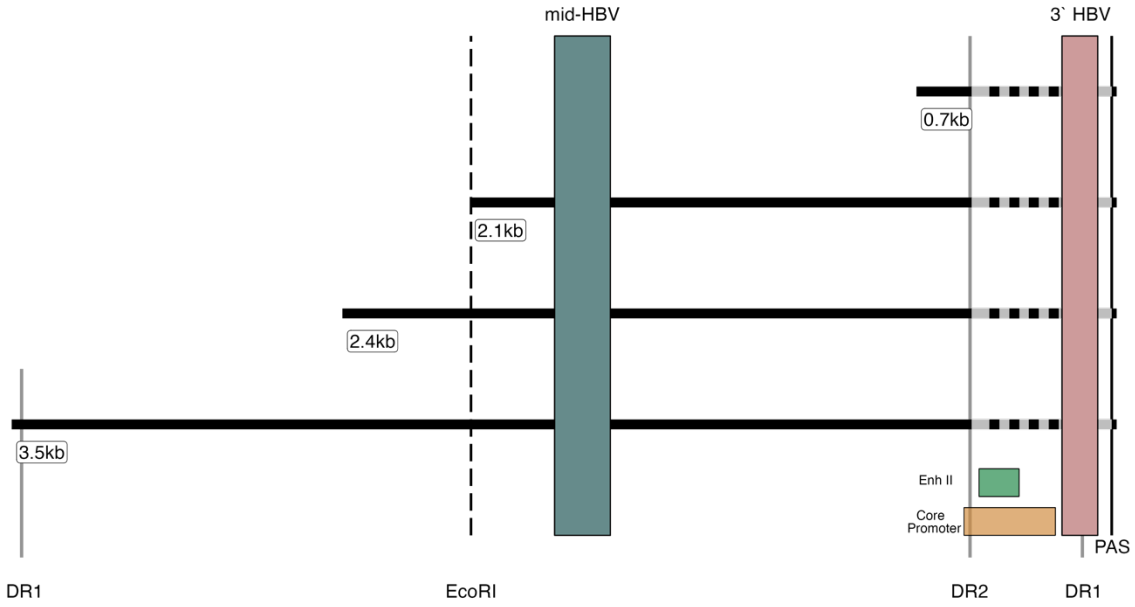
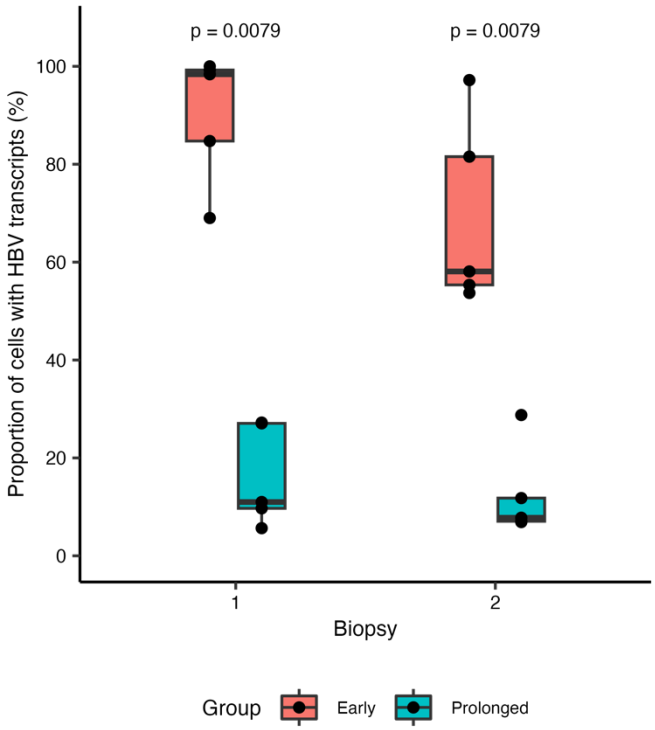


Figure 1. Transcriptional map with human-virus chimeric breakpoints from liver tissue. Horizontal lines depict the four canonical HBV mRNAs produced by each of the open reading frames. The variable chimeric virus-host regions are displayed as hashed lines at the 3' end of transcripts. Solid vertical lines show the positions of the DR2, DR1, and the canonical poly A signal (PAS), respectively. The colored boxes depict the two ddPCR target amplicons: mid-HBV and 3'HBV. The locations of the HBV Enhancer II (EnhII) and core promoter regions are displayed as green and orange boxes, respectively. Dotted line represents EcoRI cut site which is included for reference.

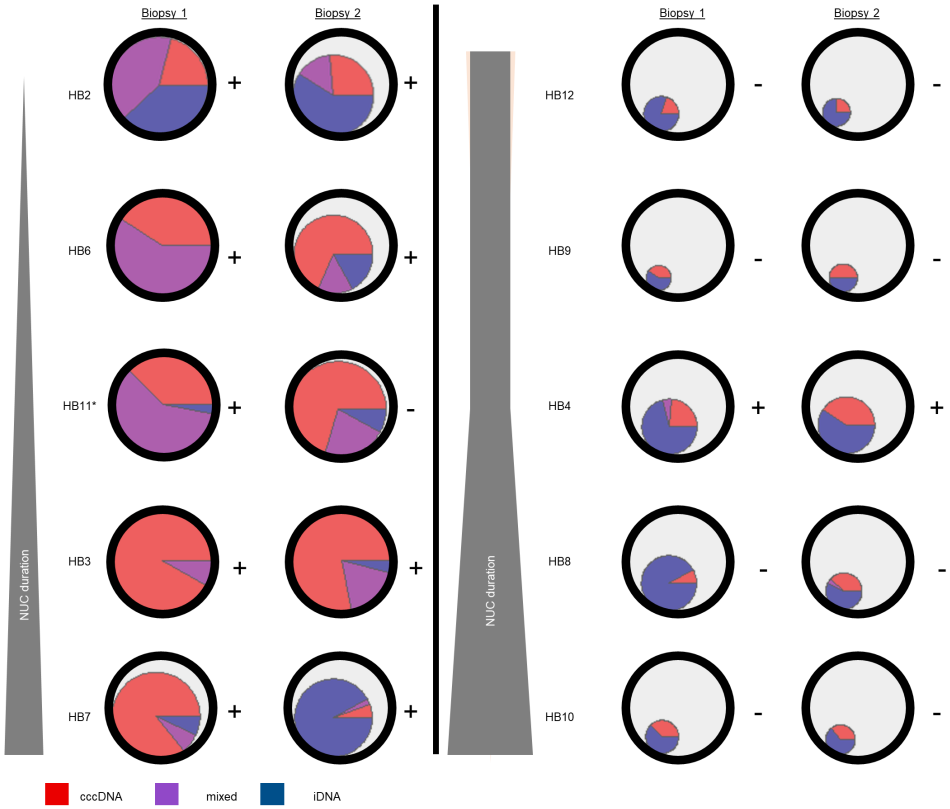


670

671 **Figure 2. Lower proportion of cells with HBV transcripts with prolonged NUC treatment.**
672 The proportion of all analyzed cells with detectable transcripts (positive for either mid-HBV, 3'-
673 HBV, or both) in each biopsy is shown where each dot represents a person at their respective
674 biopsies. The box spans the first and third quartiles with a horizontal line representing the
675 median. The tails correspond to the minimum and maximum of that respective group. Wilcoxon
676 rank-sum test used to compare between groups. Red, *early* group; blue, *prolonged* group.

677

678



679

680 **Figure 3. Changes in the cellular source of HBV transcription in paired biopsies during**
681 **early and prolonged NUC treatment.** Shown are the proportion of hepatocytes that are
682 transcriptionally active: the larger circles, outlined in black, are fixed in size. The inner inscribed
683 pie chart depicts the proportion of hepatocytes that were found to be transcriptionally active.
684 When the inner pie chart fills the entire larger circle it indicates that 100% of hepatocytes are
685 transcriptionally active, whereas smaller proportions are denoted by their respective areas. Each
686 pie chart is subdivided by color to indicate the proportion of transcriptionally active cells that
687 either have chiefly cccDNA-derived transcription (red), chiefly iDNA-derived transcription
688 (blue), or mixed transcription (purple). HBeAg status (positive [+] or negative [-]) at the time of
689 each biopsy is indicated to the right of each circle. Participants are ordered from top to bottom by
690 the duration of NUC therapy prior to biopsy 1, as indicated by the vertical gray wedges to the left

691 of all participants. P-values for each change are shown in the Supplementary Table. HB11*
692 indicates the individual who seroconverted HBeAg between biopsies.

693

694

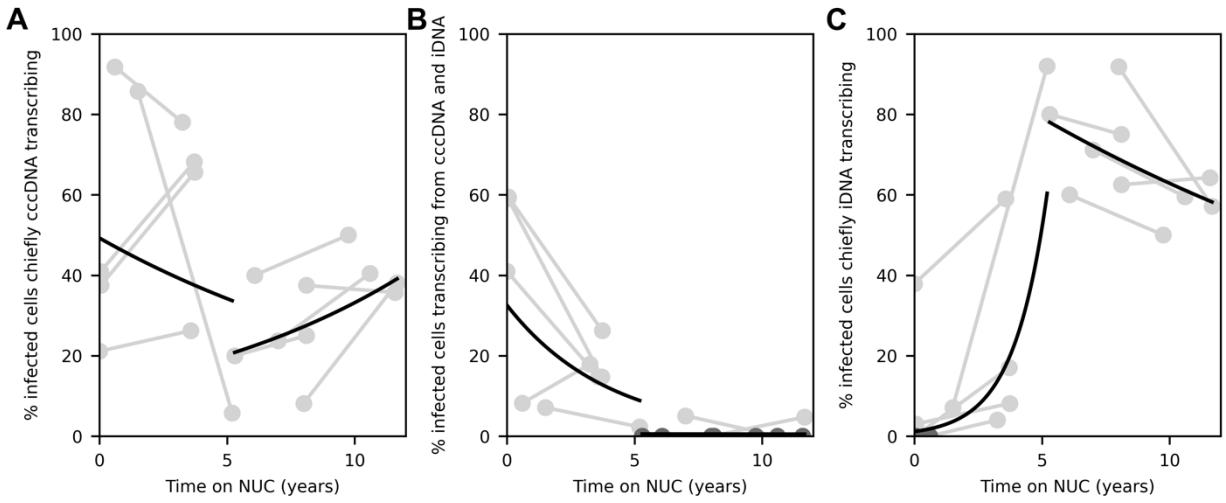


Figure 4. Among cells actively transcribing viral RNA there are different decay rates depending on the transcriptional source. Each grey line represents a participant with points corresponding to biopsies 1 and 2. The time on NUCs (years) is plotted against the proportion of cells actively transcribing S **(A**) chiefly from cccDNA (iDNA-TI≤1), **(B**) from both cccDNA and iDNA (iDNA-TI>1), and (**C**) chiefly from iDNA (only Mid-HBV positive). Mixed-effects modeling was used to generate rates across the group, represented by the black lines (see Methods).

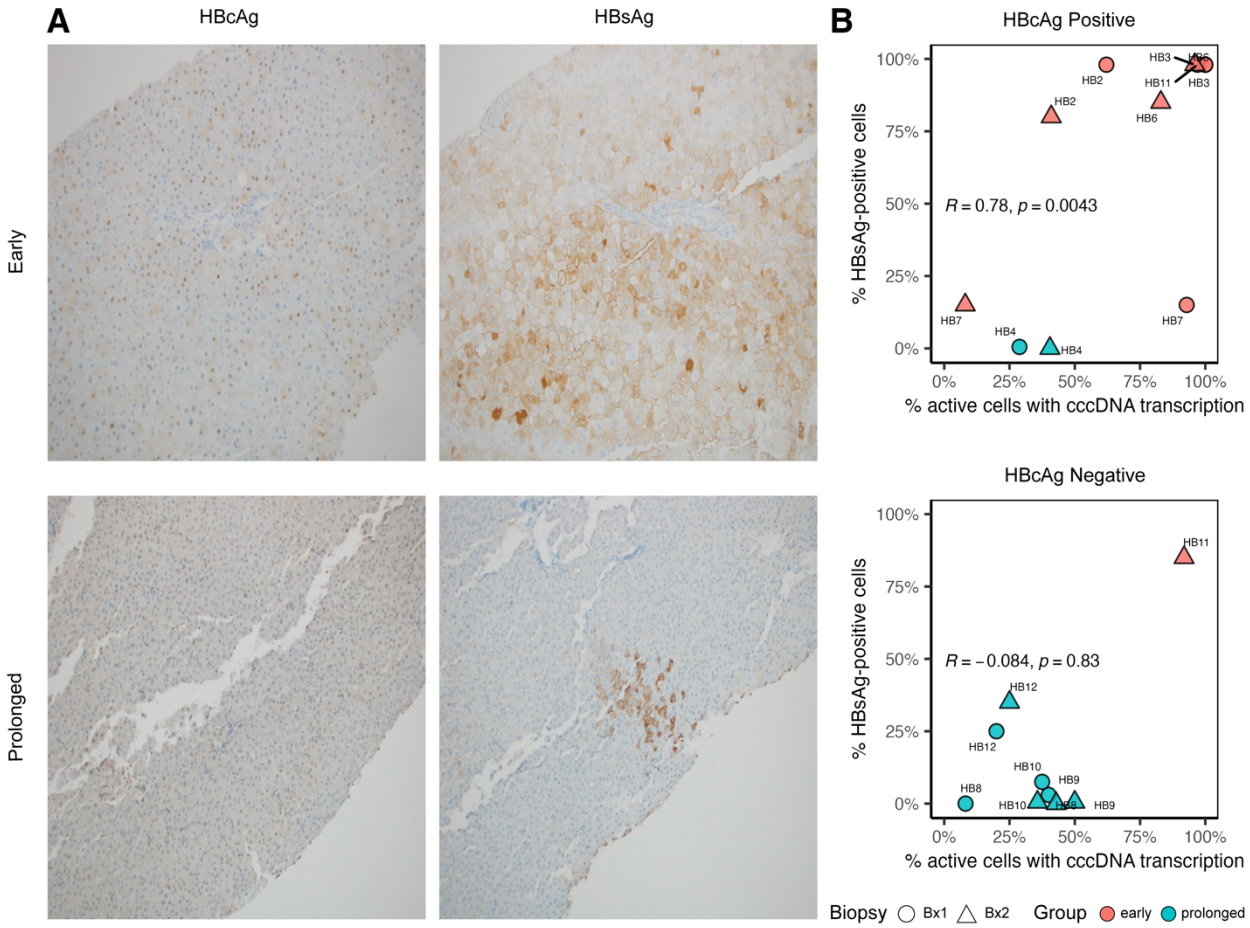


Figure 5. cccDNA-derived transcription drives intrahepatic viral antigen production early after NUCs but not after prolonged NUCs. Immunohistochemistry was used to stain for hepatitis B core antigen (HBcAg) and hepatitis B surface antigen (HBsAg) in each biopsy and the amount of staining was quantified by a pathologist who was blinded to participant identity. (**A**) Shown are representative HBcAg (**left**) and HBsAg (**right**) staining images at biopsy 1 for one participant in the early treatment group, HB6, (**top**) and one in the prolonged group, HB12, (**bottom**). (**B**) After stratifying by HBcAg-positive biopsies (**top**) and HBcAg-negative biopsies (**bottom**), we correlated the percentage of *transcriptionally-active* cells with cccDNA-derived transcripts (including chiefly cccDNA & mixed) with the percentage of cells positive for HBsAg staining. Red and blue dots are displayed on the plot to represent early and prolonged

716 individuals, respectively. Spearman's correlation coefficients and associated P values are shown.

717 Abbreviations: Bx, biopsy.

718

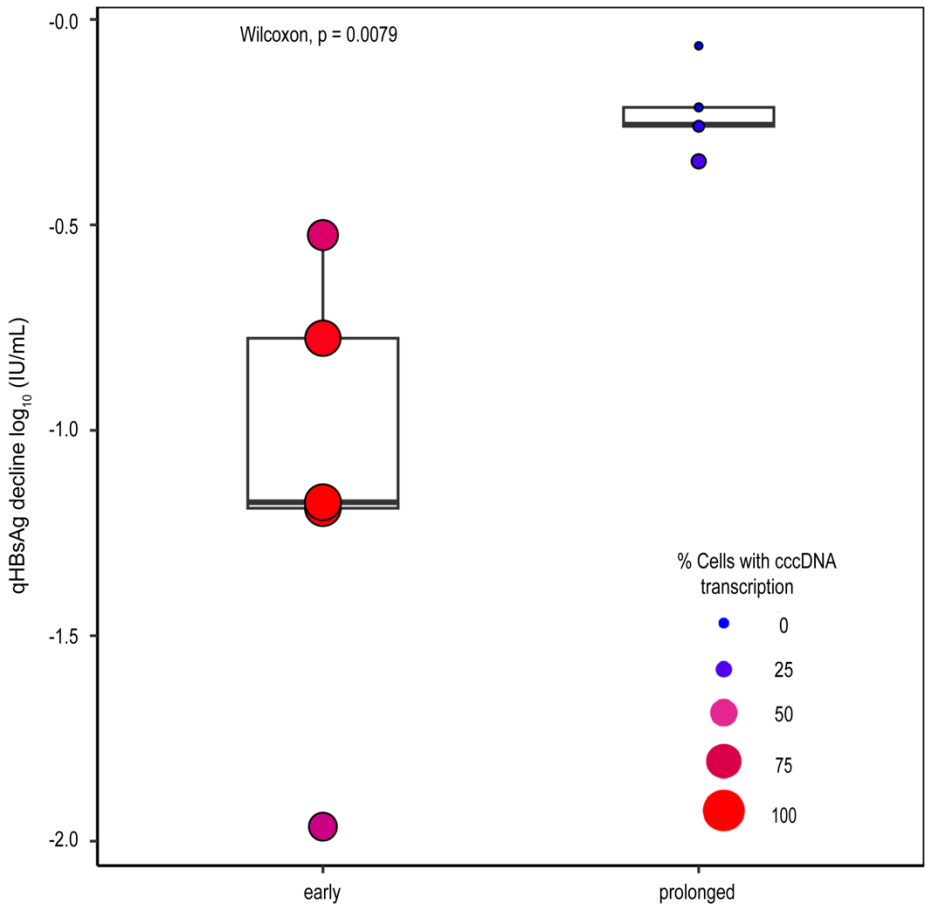


Figure 6. Participants with higher proportions of cells with cccDNA-derived transcription at biopsy 1 had larger declines in blood qHBsAg with NUC treatment. Y-axis represents the maximum decline of qHBsAg after NUCs relative to measures taken at biopsy 1 (Supp Fig. 2) while the size and color of the point indicate the proportion of all cells with any cccDNA transcription (including chiefly cccDNA & mixed) at biopsy 1. The box spans the first and third quartiles, and the horizontal line representing the median. The tails correspond to the minimum and maximum of that respective group. Wilcoxon rank-sum test was used to compare qHBsAg declines between *early* and *prolonged* groups.