Estimation of hepatitis B virus cccDNA persistence in chronic infection using within-host evolutionary rates

Katrina A. Lythgoe[1,2], Sheila F. Lumley[3,4], Jane A. McKeating[5+] and Philippa C. Matthews[3,4,6+]

[1] Big Data Institute, University of Oxford, Old Road Campus, Oxford, OX3 7LF.
[2] Department of Zoology, University of Oxford, Medawar Building, South Parks Road, Oxford OX1 3SY.
[3] Nuffield Department of Medicine, University of Oxford, Medawar Building, South Parks Road, Oxford OX1 3SY.
[4] Department of Infectious Diseases and Microbiology, Oxford University Hospitals NHS Foundation Trust, John Radcliffe Hospital, Headley Way, Oxford OX3 9DU.
[5] Nuffield Department of Medicine Research Building, University of Oxford, Oxford, OX3 7LF.
[6] NIHR Biomedical Research Centre, John Radcliffe Hospital, Headley Way, Oxford OX3 9DU.

[+] Denotes equal contribution

Corresponding author: Katrina Lythgoe, Email: katrina.lythgoe@bdi.ox.ac.uk

Running header: cccDNA persistence during HBV infection

Key words: hepatitis B virus, cccDNA, persistence, evolution, dynamics, reservoir, modelling

Author contributions:
KAL: Conceptualisation, Formal analysis, Funding acquisition, Methodology, Writing – original draft, Writing – Review & Editing
SFL: Formal analysis, Investigation, Writing – original draft
JAM: Conceptualisation, Writing – Review & Editing
PCM: Conceptualisation, Writing – Review & Editing

Abstract

Hepatitis B virus (HBV) infection is a major global health problem with over 240 million infected individuals at risk of developing progressive liver disease and hepatocellular carcinoma. HBV is an enveloped DNA virus that establishes its genome as an episomal, covalently closed circular DNA (cccDNA) in the nucleus of infected hepatocytes. Currently available standard-of-care treatments for chronic hepatitis B (CHB) include nucleos(t)ide analogues (NA) that suppress HBV replication but do not target the cccDNA and hence rarely cure infection. There is considerable interest in determining the lifespan of cccDNA molecules to design and evaluate new curative treatments. We took a novel approach to this problem by developing a new mathematical framework to model changes in evolutionary rates during infection which, combined with previously determined within-host evolutionary rates of HBV, we used to determine the lifespan of cccDNA. We estimate that during HBe-antigen positive ($HBeAg^{POS}$) infection the cccDNA lifespan is 61 (36-236) days, whereas during the $HBeAg^{NEG}$ phase of infection it is only 26 (16-81) days. We found that cccDNA replicative capacity declined by an order of magnitude between $HBeAg^{POS}$ and $HBeAg^{NEG}$ phases of infection. Our estimated lifespan of cccDNA is too short to explain the long durations of chronic infection observed in patients on NA treatment, suggesting that either a sub-population of long-lived hepatocytes harbouring cccDNA molecules persists during therapy, or that NA therapy does not suppress all viral replication. These results provide a greater understanding of the biology of the cccDNA reservoir and can aid the development of new curative therapeutic strategies for treating CHB.

INTRODUCTION

Hepatitis B Virus (HBV) is a global health problem, with more than 240 million chronically infected individuals at risk of developing liver fibrosis, cirrhosis, and hepatocellular carcinoma [1]. HBV is the prototypic member of the *hepadnaviruses*, a family of small, enveloped hepatotropic viruses with a partially double-stranded relaxed circular DNA (rcDNA) genome that replicates via reverse transcription of pregenomic RNA (pgRNA). Following the infection of hepatocytes, the rcDNA genome is imported to the nucleus and converted to covalently closed circular DNA (cccDNA), that provides the transcriptional template for pregenomic and subgenomic RNAs. Newly synthesized pgRNA is assembled into nucleocapsids that undergo reverse transcription to generate rcDNA, which is subsequently enveloped and released as infectious virions. Alternatively, capsids can be redirected to the nucleus to replenish and maintain the episomal pool of cccDNA and this intracellular amplification pathway, together with the long half-life of cccDNA, contributes to viral persistence [2,3].

Chronic hepatitis B (CHB) is usually treated with nucleos(t)ide analogues (NAs) that inhibit the reverse transcription of pgRNA to rcDNA. However, these therapies do not target the cccDNA reservoir [4] and viremia rebounds once treatment is stopped, even when peripheral levels of viral DNA have remained undetectable for months or years [5]. Effective anti-viral drugs have revolutionized treatments for hepatitis C virus and there is a growing impetus to identify curative therapies for HBV [6]. Despite its central role in the HBV life cycle our understanding of the viral and host factors that regulate cccDNA abundance and half-life is limited. cccDNA half-life is the time for the number of copies in the liver to reduce by half and will depend on a number of factors, including cccDNA 'lifespan' (the time an individual cccDNA molecule persists) and genesis of new cccDNA via extra-cellular virus infection or intra-cellular amplification [7,8] (Fig.1a). A recent *in vitro* study reported a 40 day half-life of HBV cccDNA in infected HepG2 cells [2] that provides an estimated lifespan of 58 days. However, our current knowledge of cccDNA lifespan in the human liver is unknown.

Double-stranded DNA viruses typically have low mutation rates, but since rcDNA is generated via error prone reverse transcription in the *hepadnaviridae*, they have higher mutation rates than other DNA viruses (Fig 1A) [9]. The estimated mutation rate for avian hepadnavirus is $2x10^{-5}$ substitutions per site per infected cell (s/s/c) [10], and an upper limit of $8.7x10^{-5}$ s/s/c has been estimated for HBV [11], similar to estimates for RNA retroviruses [12]. The evolutionary rate measures how quickly mutations become fixed in a population over a period of time [13]. Strikingly, the evolutionary rate of HBV is much lower than for RNA viruses with similar mutation rates [14–16]. Different mechanisms could explain this observation, including the biological constraint of multiple overlapping reading frames in the HBV genome [14], limited viral population size in the liver, or long cccDNA lifespan [14,15]. Both evolutionary constraint and population size should only influence the rate of evolution of variants that experience selection [17]. For neutral or near neutral mutations, the long cccDNA lifespan provides the simplest explanation for the low evolutionary rate of HBV [15].

We propose the within-host evolutionary rate of HBV can be used to estimate cccDNA lifespan. We developed a novel mathematical model to determine the relationship between HBV evolutionary rate and the lifespan of cccDNA, and combined with published mutation and evolutionary rates [10,14,16], we inferred the lifespan of cccDNA during different phases of CHB. To the best of our knowledge, these are the first estimates of cccDNA lifespan in treatment naïve subjects and provide important insights into the HBV reservoir that will be valuable for the design and evaluation of future treatment interventions.

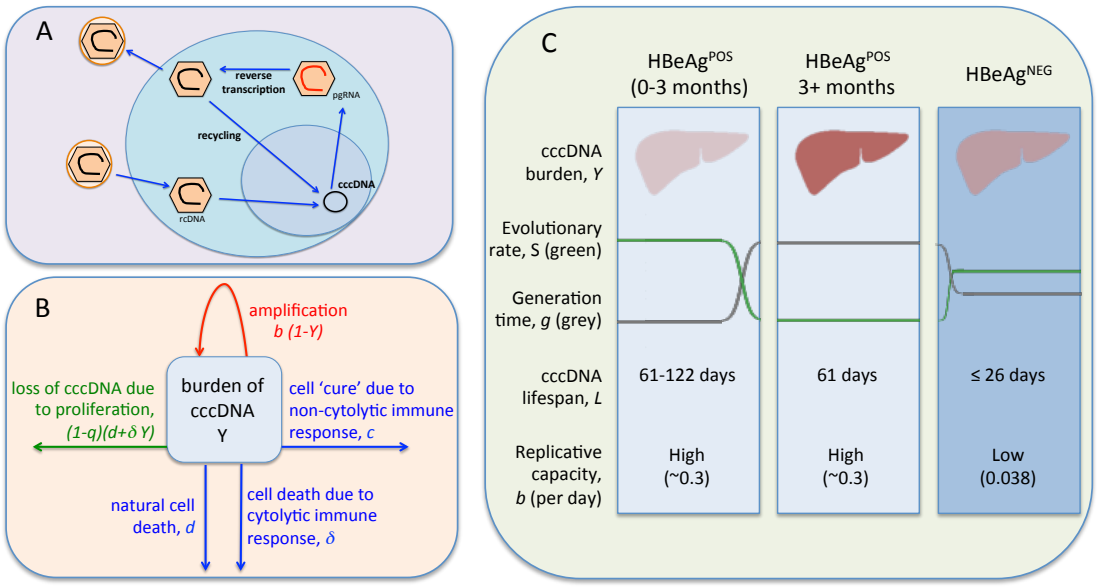

Fig 1: **Summary of the HBV life cycle, mathematical model, and estimated cccDNA lifespan.** **A:** Simplified HBV replication cycle. A virus particle containing relaxed circular DNA (rcDNA) enters a hepatocyte (blue circle) and is uncoated. The rcDNA is transported to the nucleus (purple circle) and repaired to generate cccDNA. This cccDNA is the transcriptional template for all viral RNAs, including pre-genomic (pgRNA), which is transported to the cytoplasm, encapsidated, and converted into rcDNA by error-prone reverse transcription. The encapsidated rcDNA can be transported back into the nucleus to form more cccDNA (intra-cellular amplification), or enveloped and released as virions that can infect hepatocytes (extra-cellular amplification). **B**: Structure of the mathematical model. This is a single compartment model representing the burden of cccDNA in the liver, Y, over the course of infection. The cccDNA burden can increase due to amplification (intra- and extra-cellular), where b is a measure of the within-host replicative capacity of cccDNA. cccDNA can be cleared from the liver due to natural cell death, at rate d, cytolytic immune responses at rate δ, and non-cytolytic immune responses at rate c. Proliferation can also result in loss of cccDNA at rate $(1-q)(d+\delta Y)$, where q is probability that an individual cccDNA survives mitosis. **C**: Representation of the model dynamics and key results, where the numbers give the most likely values inferred by fitting the mutation and evolutionary rates to the model. The darker the colours on the figure the higher the cccDNA burden (reds) and the stronger the immune response (blues).

RESULTS

We developed a mathematical model describing the number of cccDNA molecules in the liver that is independent of infected cell frequency, and accounts for intra- and extra-cellular cccDNA amplification and loss of cccDNA during hepatocyte mitosis (Fig 1B and Methods, Eqs 1 and 2). Using this model we derived expressions for the viral generation time, defined as the typical time for one cccDNA molecule to generate another cccDNA molecule at time t since infection, $g(t)$ (Eq 5), and the neutral rate of evolution at time $S(t)$ (Eq 6). At equilibrium, we show that the lifespan of cccDNA, L_E, is equal to the virus generation time, g_E, which is given by the neutral mutation rate divided by the neutral rate of evolution, μ/S_E (Eq 8). The notation used throughout is given in Table 1.

Table 1: Notation used to describe the model

$N(t)$, N_E	Number of cccDNA molecules at time t since infection and at equilibrium
K	Maximum possible number of cccDNA molecules
$Y(t)$, Y_E	cccDNA burden in the liver at time t since infection and at equilibrium, expressed as a proportion of the maximum possible number of cccDNA, $Y(t)=N(t)/K$, $Y_E= N_E/K$
$\rho(t)$	Proliferation rate of cccDNA at time t since infection
$g(t)$, g_E, $g_{Y<<1}$	Generation time of cccDNA at time t since infection, at equilibrium, and when few cells are infected
$S(t)$, S_E	Neutral rate of evolution at time t since infection and at equilibrium
L_E , $L_{Y<<1}$	Life span of cccDNA molecules at equilibrium and when few cells are infected
β	Per capita replicative capacity, defined as the per capita growth rate of cccDNA when few cells are infected and in the absence of infected cell death or loss of cccDNA due to non-cytolytic immune responses.
R_0	The basic reproductive rate of cccDNA (the number of cccDNA molecules a single cccDNA produces in its lifetime in an otherwise uninfected population of hepatocytes)
$b= \beta K$	Replicative capacity of cccDNA (a rescaled measure of the per capita replicative capacity)
d	Natural death rate of hepatocytes
δ	Additional death rate of infected hepatocytes due to cytolytic immune responses
c	Loss rate of cccDNA due to non-cytolytic immune responses
q	Probability that a cccDNA molecule survives mitosis
μ	Mutation rate of cccDNA (substitutions per site per reproduction)

Lifespan of cccDNA

The lifespan of cccDNA molecules most likely changes over the course of HBV infection, and will be influenced by host and viral factors [18] , including the rate of hepatocyte proliferation [19,20]. Early in infection cccDNA is transcriptionally active and translation of pre-core/pgRNA results in detectable levels of hepatitis B e antigen (HBeAg) in the periphery that associates with high HBV DNA levels (viral load - VL) [21]. In later stages of infection after seroconversion and genesis of anti-HBe antibodies

there is a loss of HBeAg and more efficient immune targeting of infected cells [22], leading to a reduction in VL and a shortening of cccDNA lifespan. This HBeAg^{NEG} phase of infection is often associated with the emergence of precore mutations that limit HBeAg expression [23].The higher hepatocyte death rates during HBeAg^{NEG} CHB infection will induce hepatocyte proliferation [21]. Although the extent to which cccDNA is lost during hepatocyte mitosis is uncertain [8], unless all cccDNA episomes survive mitosis, the increased proliferation rate of infected cells will shorten the average lifespan of cccDNA [20,24,25].

From the published estimates for the mutation [10] and evolutionary rates of HBV [14,16], we inferred the probability distributions for cccDNA lifespan during HBeAg^{POS} and HBeAg^{NEG} phases of CHB using Eq 8. The predicted lifespan of cccDNA during HBeAg^{POS} infection, and when VLs are high and stable, is 61 days (5th-95th percentiles 36-236 days; Fig 2, orange line). In contrast, during HBeAg^{NEG} infection the lifespan of cccDNA is estimated at only 26 days (16-81 days; Fig 2, blue line).

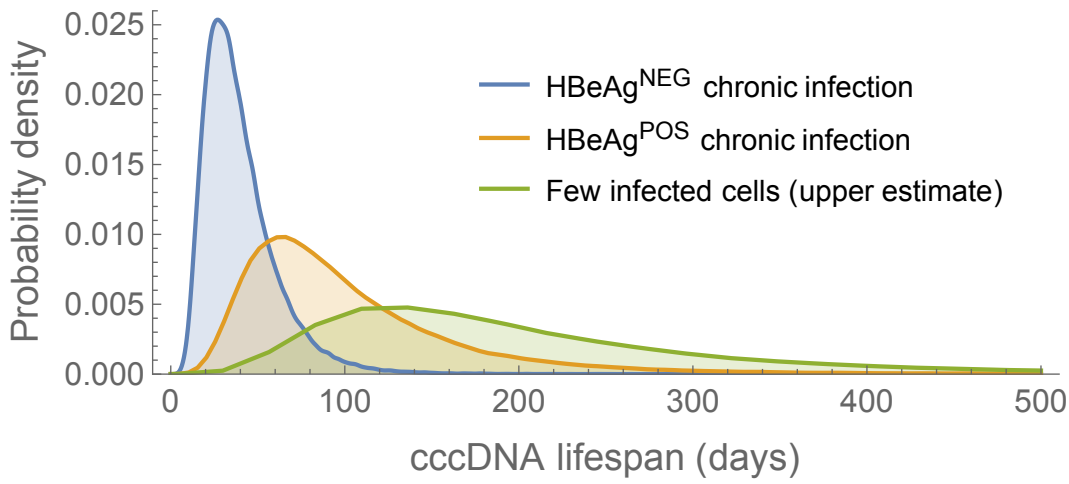

Fig 2. Probability distributions for cccDNA lifespan during different stages of HBV infection. The distributions for cccDNA in stable HBeAg^{POS} and HBeAg^{NEG} chronic infection are based on the neutral mutation rate and rate of neutral evolution (orange and blue lines, respectively). If the cccDNA burden during HBeAg^{NEG} infection is not stable, but gradually falling (i.e. the basic reproduction number, R_0, is less than one) the lifespan will be slightly less than inferred here. The upper estimate reflects the maximum likely cccDNA lifespan when few cells are infected, based on the neutral rate of evolution during HBeAg-postive infection and assuming no cccDNA survives mitosis (q=0; green line).

The shorter lifespan of cccDNA during HBeAg^{NEG} compared to HBeAg^{POS} infection can be explained by higher rates of cccDNA clearance (Eq 9). This may reflect changes in the immune environment due to HBe-antigen seroconversion that is associated with increased cytolytic and non-cytolytic immune responses (δ and c respectively). Mutational changes in the virus that limit HBeAg expression may also affect HBV replication and stability of cccDNA [23]. Increased host immune responses during HBeAg^{NEG} infection could push the basic reproduction number, R_0, of cccDNA below one (Eq 3) due to the higher clearance rates of cccDNA molecules, and also due to

reduced replicative capacity, *b*, of cccDNA. If R_0<1, the number of cccDNA will not reach a stable level but will continually decline. In this non-equilibrium situation the lifespan of cccDNA may be less than our inferred 26 days since the viral generation time will be greater than the lifespan of cccDNA (Fig 3 and methods).

Our model suggests cccDNA lifespan can be up to two times longer when few cells are infected compared to when most cells are infected cells (see methods; compare Eqs 9 and 10). When few cells are infected there is less cell death due to cytolytic immune responses, a lower rate of hepatocyte proliferation to maintain the number of hepatocytes, and consequently reduced loss of cccDNA via mitosis of infected cells. This is of more than theoretical interest, because when estimating how long it will take to deplete the cccDNA reservoir on treatment, it is the lifespan of cccDNA when relatively few cells are infected that is important since treatment is known to reduce the cccDNA load. The maximum expected cccDNA lifespan, corresponding to $HBeAg^{POS}$ infection, few infected cells, and no cccDNA surviving mitosis, is 123 days (71-472 days; Fig 2, green line). Reports for duck hepatitis B virus (DHBV) show a high proportion of cccDNA survives mitosis [25]. In contrast, for HBV recent experimental [8,24] and modelling [20] results suggest that relatively few cccDNA molecules survive mitosis, making this longer lifespan a reasonable expectation.

Dynamics of the mathematical model

To demonstrate the behaviour of our model we present examples of the dynamics when no cccDNA survives mitosis (q=0, Fig 3; see S1 Fig for model dynamics when *q*=1). We used parameters that are compatible with our estimated cccDNA generation times (61 days during $HBeAg^{POS}$ infection and 26 days during $HBeAg^{NEG}$ infection). Since hepatocytes are long-lived we defined the natural death rate as *d*=0.002 per day throughout and, for simplicity, we set *c*=0 under the assumption that cytolytic responses have greater antiviral activity than non-cytolytic responses. We assume a neutral mutation rate $\mu=2\times10^{-5}$ s/s/c [10]. The model dynamics when *q*=1 are similar to the case where *q=0*, apart from the lifespan of cccDNA in the early stages of infection is predicted to be higher if q=1 (see below). A graphical representation of the results is given in Fig 1C, and a summary of the parameters in Table 2.

<u>$HBeAg^{POS}$ infection</u>

The replicative capacity of cccDNA, *b,* was chosen to be 0.3/day so that the peak number of cccDNA molecules in the liver is reached at approximately 3 months since infection, in line with reported observations [26]. The death rate of infected cells due to cytolytic immune responses, δ, was determined assuming a cccDNA generation time at equilibrium of 61 days, and solving Eq 9 for δ (giving δ=0.006 per day if *q*=0; the associated R_0 is 30).

Under these assumptions, during the first few months of infection the cccDNA burden (number of cccDNA divided by the maximum number of cccDNA) increases rapidly, leading to a short viral generation time predicted by the model of 3.3 days (Eq 11, Fig 3, S1 Fig). A recent study estimated an eclipse period of approximately 3

Table 2: Parameters used to describe the model dynamics[1,2]

	Model 1 No *cccDNA survives mitosis*		Model 2 *All cccDNA survives mitosis*	
	HBeAg^{POS}	HBeAg^{NEG}	HBeAg^{POS}	HBeAg^{NEG}
q	0	0	1	1
g	g_E=61	g_E=26 [3]$g_{Y<<1}$=26 [3]$g_{Y<<1}$=26	g_E=61	g_E=26 [3]$g_{Y<<1}$=26 [3]$g_{Y<<1}$=26
R_0	30.0	13.6 1.0 0.7	37.5	7.9 1.0 0.7
b	0.3	0.3 0.038 0.038	0.3	0.3 0.038 0.038
d	0.002	0.002	0.002	0.002
δ	0.006	0.018 0.034 0.050	0.014	0.036 0.036 0.052
c	0	0	0	0
μ	$2x10^{-5}$	$2x10^{-5}$	$2x10^{-5}$	$2x10^{-5}$

[1]All rates are given per day, and generation times are listed as days
[2]Where three values are given these refer to the alternative parameters used for the different trajectories presented in Fig 3 and S1 Fig (top parameter, black line; middle parameter, blue line; bottom parameter, orange line).
[3]Since R0<=1, a non-zero equilibrium is not reached.

days for a newly infected cell to produce viral particles [2], so our estimated viral generation time seems reasonable. This short generation time of cccDNA during early infection contrasts with the long cccDNA lifespan (dots in Fig 3, S1 Fig; note the longer lifespan predicted if *q=0* compared to *q*=1). The neutral rate of evolution is also predicted to be high during this early stage of infection due to the short generation time.

As infection progresses, the viral generation time increases due to fewer susceptible target cells (Eq 5), in line with results in epidemiology [27], and this in turn reduces the evolutionary rate (Eq 6). This dependency of evolutionary rate on epidemiological dynamics has been noted in a previous simulation study on within-host viral infection [28], but is generally an underappreciated factor influencing evolutionary rates. At equilibrium, the estimated viral generation time and cccDNA lifespan are the same, and it is this equivalency that enables us to determine these parameters from the neutral rate of evolution, independent of the parameters of the model (see Methods). Due to the long lifespan of infected hepatocytes, a high cccDNA burden is reached in the model. This is in line with observations that most hepatocytes are infected at peak infection [29].

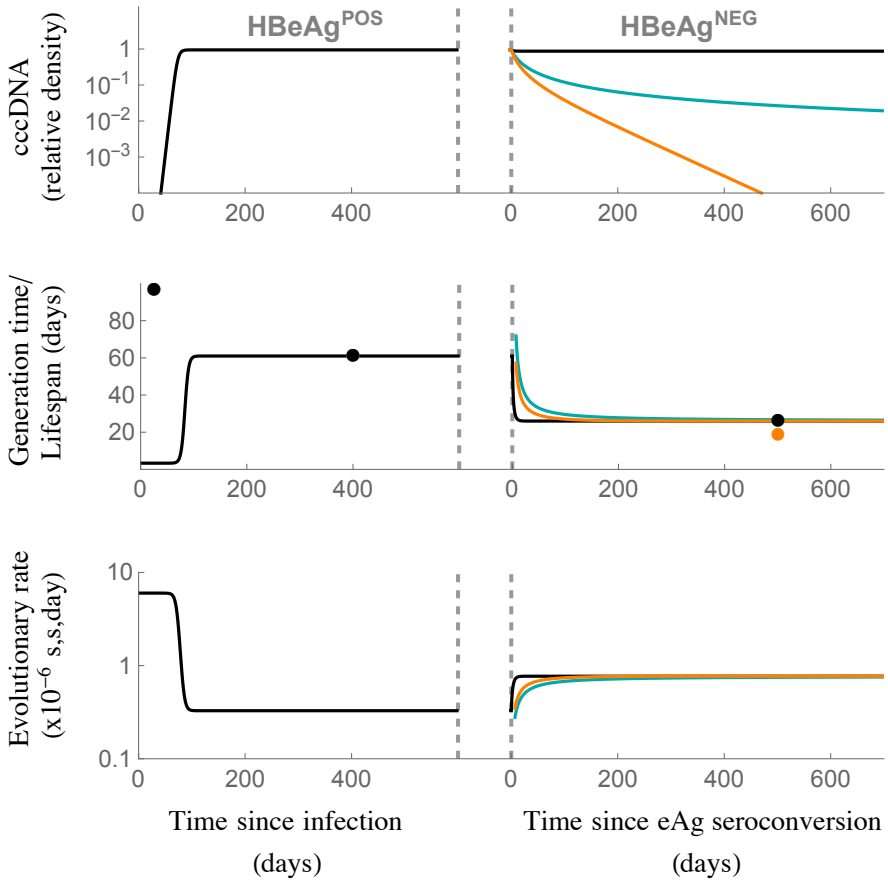

Fig 3. Population and evolutionary dynamics of cccDNA for the within-host model. We assumed no cccDNA survives mitosis (q=0; see S1 Fig for the case where q=1). The model was parameterised assuming the generation time, g, during chronic HBeAg^{POS} infection is 61 days, and during chronic HBeAg^{NEG} infection is 26 days, in line with our predictions for cccDNA generation time *in vivo*. The top panel shows cccDNA burden, where 1 represents the maximum. The middle panel shows the viral generation time (lines) and cccDNA lifespan during key stages of infection (dots, derived from Eqs 9 and 10). The bottom row shows the evolutionary rates. Black line: replicative capacity during HBeAg^{NEG} infection remains the same as during HBeAg-postive infection (b_{eAg+}= b_{eAg-}=0.3 per day). Blue line: replicative capacity falls to b_{eAg-}=0.038 per day during HBeAg^{NEG} infection, and R_0=1. Orange line: replicative capacity falls to b_{eAg-}=0.038 per day and R_0=0.7. See Table 1 for all other parameters.

HBeAg^{NEG} infection

We assumed the transition from HBeAg^{POS} to HBeAg^{NEG} occurs after an arbitrary amount of time after HBeAg^{POS} equilibrium is reached and associates with a reduced cccDNA generation time from 61 to 26 days. If this reduced generation time is not accompanied by a decrease in replicative capacity, only a modest fall in the cccDNA burden is predicted (Fig 3, blue line; δ =0.018 and b=0.3). However, this is inconsistent with *in vivo* infections, where the number of cccDNA molecules and extracellular HBV DNA levels (VL) typically decline by several orders of magnitude after transition to HBeAg^{NEG} infection [30,31]. If the cccDNA burden is low (Y<<1), then replicative capacity, b, is estimated by the reciprocal of the generation time ($b=1/g$; Eq 5). For a generation time of 26 days, this gives b=0.038 per day, leading to an

estimated 10-fold reduction in the ability of cccDNA to reproduce during HBeAg^{NEG} compared to HBeAg^{POS} infection. In Fig 3, the orange line shows the model dynamics given this decline in b, and when R_0=1 during HBeAg^{NEG} infection (i.e. $\delta = 0.034$ per day and b=0.038 per day). In this case, the cccDNA burden falls at a relatively modest rate. Perhaps more likely is that $R_0 < 1$ and the number of cccDNA molecules continues to decline. The green line shows the dynamics if R_0=0.7 (δ =0.050). However, even with this modest increase in δ, the number of cccDNA is predicted to fall rapidly.

The difficulty in explaining low but steady VL using standard within-host virus models, and the sensitivity of VL to model parameters when R_0 is close to one, have been acknowledged previously, particularly in relation to HIV-1 infections [32–34]. Possible explanations for the low numbers of cccDNA during HBeAg^{NEG} infection and low rates of spontaneous cure include the existence of a small number of hepatocytes that are susceptible to infection, resulting in low numbers of cccDNA molecules even if R_0 is high [32], or the existence of a metapopulation-type partitioned structure in the liver, which enables the cccDNA to persist when R_0 is low [34].

Estimated time to eradicate cccDNA on treatment

When few cells are infected, the inferred cccDNA lifespan is 123 days during HBeAg^{POS} infection if q=0. Even with this longer estimate for cccDNA lifespan, if there are 10^{12} cccDNA molecules at the start of treatment (see methods), we would expect the reservoir to be depleted after less than ten years of treatment (Eq 13, Fig 4A). Moreover, if treatment is initiated during HBeAg^{NEG} CHB the time to eradicate cccDNA is predicted to be even faster (only 1.5 years) with a lifespan of 26 days, and a lower number of cccDNA molecules ($2x10^9$) in the liver at the start of treatment. However, these predictions are in stark contrast to what is observed in the clinic, where a high proportion of individuals remain infected after many years of continuous treatment [35] and there is no appreciable difference in treatment mediated cure in HBeAg^{NEG} or HBeAg^{POS} patients [36,37]. The discrepancy may arise due to our estimated cccDNA lifespan being too short. An estimated lifespan of 236 days during HBeAg-postive CHB still lies within our 95% confidence interval, and would give a time to eradication, and hence sterilizing cure between 18 and 36 years (Eq 13). However, this does not explain the long time to eradicate cccDNA during HBeAg^{NEG} infection. Alternative explanations include ongoing (albeit reduced) cccDNA amplification during NA treatment (b>0) [37,38], or the presence of a long-lived subset of infected hepatocytes [24,39].

To evaluate these two scenarios, we modelled cccDNA dynamics in CHB patients on treatment assuming different levels of viral replication (Fig 4B) or a subset of long-lived cells (Fig 4C, S2 Fig). The dynamics of cccDNA are sensitive to the amount of replication, making it unlikely that ongoing amplification alone explains the failure of treatments to eliminate cccDNA. Apart from a narrow range of replicative capacities, either a high and steady cccDNA burden, or relatively rapid cccDNA elimination, is predicted on treatment. The existence of a long-lived population of infected hepatocytes is more robust to differences in model parameters, with a gradual increase in the time to eradicate cccDNA as the death rate of long-lived cells is

increased, making a long-lived population a more parsimonious explanation for the slow decline in the HBV reservoir. However, since the decay dynamics of the reservoir on treatment can be complex, and differ between individuals [38], a combination of factors most likely explains the clinical observations.

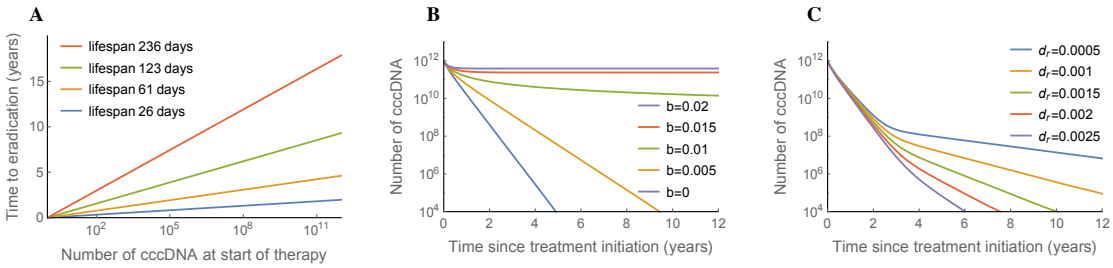

Fig 4. Effect of NA treatment predicted by the model. A: Predicted time for cccDNA to be eradicated in the absence of any cccDNA reproduction (b=0). B: cccDNA dynamics whilst on treatment, assuming some residual reproduction. For all runs d=0.002 per day, δ=0.006 per day, c=0, q=0. C: cccDNA dynamics on treatment, assuming no residual reproduction (b=0) but 0.1% cccDNA is long-lived, for different death rates of long-lived cells, d_r per day. For normal cells d=0.002, δ=0.006,c=0, q=0, and for long-lived cells δr=0, c_r=0, q=0. The maximum number of cccDNA was assumed to be 10^{12}, and all model runs were started at equilibrium in the absence of treatment (b=0.3).

DISCUSSION

We provide a new model to estimate the HBV cccDNA lifespan based on reported mutation and within-host evolutionary rates [10,14,16]. The lifespan of cccDNA is an important component of the half-life of the cccDNA reservoir, which describes how the population of cccDNA molecules in an individual declines over time. We predict an average cccDNA lifespan of 61 days during HBeAg^{POS} CHB compared to only 26 days in the HBeAg^{NEG} phase of infection. Although estimates for the mutation and evolutionary rates for HBV are associated with high levels of uncertainty, our predicted lifespan is in agreement with *in vitro* studies showing a 40 day half-life of HBV cccDNA [2] and an estimated half-life of 33-57 days in woodchucks and ducks *in vivo* [40,41]. As far as we are aware, this is the first time cccDNA lifespan has been estimated during untreated infection. The lower lifespan during HBeAg^{NEG} infection is consistent with a study in which VL data during therapy was fitted to a mathematical model, concluding that the turnover of infected cells is higher if therapy is initiated during HBeAg^{NEG} infection [22], although our predictions for cccDNA persistence are longer [22] .

The shorter cccDNA lifespan during HBeAg^{NEG} CHB may reflect host immune responses, with our model suggesting a doubling of the clearance rate compared to HBeAg^{POS} infection. However, this increased clearance rate is predicted to have a modest effect on the total number of cccDNA molecules. As well as inferring the lifespan of cccDNA, we inferred cccDNA replicative capacity (a combined measure of intra and extra-cellular amplification). Our results predict an approximate ten-fold reduction in replicative capacity between HBeAg^{POS} and HBeAg^{NEG} phases of infection. This can explain the lower cccDNA levels reported in HBeAg^{NEG} CHB [30,31], and is consistent with observations that the replicative capacity of cccDNA in the HBeAg^{NEG}

phase of infection is reduced compared to HBeAg^{POS} infection [30]. This may reflect immune control at the level of the viral epigenome, but without cell death [42].

Our estimates for cccDNA lifespan have implications for curative treatment strategies. If NA therapy inhibits all cccDNA amplification, we would predict HBV to be cured after 1 to 10 years of continuous treatment. However, this is not observed in the clinic, with only 1% of individuals clearing HBsAg each year [35]. Possible explanations for this discrepancy are that NAs do not inhibit all intra- and extra-cellular amplification [37,38], or the existence of long-lived infected cells [24,39]. Our model is consistent with the presence of long-lived infected cells providing the most parsimonious explanation for sustained infection on treatment. There is growing evidence that there is negligible intra-cellular cccDNA amplification in human HBV infection [8], and since NA treatment will inhibit the genesis of viral particles this will prevent extra-cellular amplification. Furthermore, the dynamics of cccDNA clearance is sensitive to the assumed amplification rates, and therefore if amplification alone explains the dynamics we would expect to see a proportion of individuals clearing infection within 1-2 years of starting treatment. The presence of long-lived HBV infected cells has parallels with the HIV reservoir, where long-lived latent-infected CD4^{+} T cells prevent cure [43]. Distinguishing between residual amplification and long-lived infected cells will help define the expected impact of treatment strategies that prevent cccDNA replication, compared to those directly targetting cccDNA. As HBV evolution will only occur if there is cccDNA amplification, it may be possible to distinguish between these two mechanisms by measuring the rate of cccDNA evolution whilst on treatment.

Our estimates of cccDNA persistence and amplification provide insights into mechanisms underlying CHB and will inform our understanding of how spontaneous or therapeutic clearance may be achieved. Given different infection profiles among individuals, and limited datasets available for our model, the confidence intervals of our estimations are wide. Our analysis exemplifies the power of modelling as a tool to inform therapeutic interventions and highlights the need for genomic studies to determine HBV evolutionary rates in CHB.

METHODS

To derive estimates of HBV cccDNA lifespan using the neutral mutation rate and the rate of evolution we developed a deterministic mathematical model describing the dynamics of cccDNA during the course of treatment naïve CHB. We used this model to derive expressions for viral generation time and neutral rate of evolution, both of which are predicted to change during the course of infection. Finally, we derived expressions for the lifespan of cccDNA during (i) stable CHB and (ii) when the proportion of infected cells is low, as would be expected in early stages of infection or in the first few months of NA treatment.

A within-host model of HBV dynamics

HBV cccDNA can replicate via intra-cellular and extra-cellular routes (Fig 1A), with a reported copy number between 1-50 molecules within a single hepatocyte nucleus [2,24,44–47] (the higher estimates tend to be for DHBV and lower estimates for human

HBV [8]). Since cccDNA can be lost during mitosis, we modelled the number of cccDNA copies in the liver, rather than the number of infected cells. To do this, we implicitly assume that viral production is proportional to the number of cccDNA molecules. This is a reasonable assumption since VL has been reported to associate with increasing cccDNA copy numbers [30,48].

We describe the number of copies of cccDNA in the liver at time *t* since infection, *N(t)* as:

$$\frac{dN(t)}{dt} = \beta N(t)[K - N(t)] - (d + \delta + c)N(t) - \rho(t)(1 - q)N(t) \quad \text{(Eq1a)}$$

where the first term describes the increase in cccDNA due to intra- and extra-cellular amplification. We assume that the rate of increase is density dependent, with a maximum per capita growth rate β per day and a maximum possible number of cccDNA, *K*. We assume *K* is constant since proliferation ensures the number of hepatocytes in the liver remains stable during infection [21], and since the maximum number of copies of cccDNA that can persist within each hepatocyte is virally controlled [48–50].

The second term describes the rate at which cccDNA is lost due to the natural death of hepatocytes and the host immune response, under the assumption that cccDNA is randomly distributed among infected hepatocytes. We assume that hepatocytes, and therefore cccDNA, have a natural death rate *d* per day. Infected hepatocytes (and hence cccDNA) have an additional death rate δ per day due to cytolytic immune responses, and cccDNA is lost at rate *c* per day due to non-cytolytic immune responses.

The final term describes the loss of cccDNA due to cell proliferation. Uninfected and infected hepatocytes are assumed to proliferate at the rate $\rho(t)$ per day, and hence cccDNA will be exposed to proliferation at rate, $\rho(t)$, with a probability *q* that a cccDNA molecule will survive mitosis. Since the maximum possible number of cccDNA, *K,* is constant, proliferation and cell death are balanced, hence:

$$\rho(t)K = dK + \delta N(t) \quad \text{(Eq1b)}$$

A complete expression for the dynamics of *N(t)* can be found by solving Eq 1b for $\rho(t)$ and substituting into Eq 1a. To simplify further, we consider the cccDNA burden in the liver, $Y(t) = N(t)/K$, rather than the total number of cccDNA molecules, giving us:

$$\frac{dY(t)}{dt} = b\,[1 - Y(t)]Y(t) - (d + \delta + c)Y(t) - (1 - q)[d + \delta Y(t)]Y(t) \quad \text{(Eq2)}$$

where $b = \beta K$ is a rescaled measure of cccDNA replicative capacity. From this equation we can calculate the basic reproductive rate of cccDNA, R_0, which is defined as the number of new cccDNA molecules a single cccDNA molecule will produce in a susceptible population of hepatocytes:

$$R_0 = \frac{b}{d(2-q)+(\delta+c)} \quad \text{(Eq3)}$$

If R_0<1, then the infection cannot be sustained in the long term. At equilibrium, the cccDNA burden is given by:

$$Y_E = Max\left[0, \frac{b-d(2-q)-(\delta+c)}{b+\delta(1-q)}\right] \quad \text{(Eq4)}$$

which is equivalent to the cccDNA burden during stable chronic infection. Our model considers the number of cccDNA molecules independent of their distribution within cells. This is similar to the "single copy" modelling assumption used in [25], in which only one cccDNA molecule can persist in a cell, and which was shown to produce almost identical dynamics to one in which multiple copies of cccDNA are explicitly modelled within infected cells [25].

An expression for the neutral rate of HBV evolution

In a large well-mixed viral population, and in the absence of selection, the rate of evolution at time *t* is given by $S(t) = \mu/g\,(t)$, where μ is the (neutral) mutation rate, measured per site per viral generation, and $g(t)$ is the generation time [51]. For our within-host model of HBV infection, *g* is equivalent to the typical amount of time it takes for one cccDNA molecule to replicate another molecule. This is similar to the meaning of generation time in demography and epidemiology [27,52,53], and which from Eq 2 is given by:

$$g(t) = 1/[b(1 - Y(t))] \quad \text{(Eq5)}$$

At time *t* since initial infection, the neutral substitution rate is therefore given by:

$$S(t) = \frac{\mu}{g(t)} = \mu b[1 - Y(t)] \quad \text{(Eq6)}$$

Since intra- and extra-cellular amplification involve an error-prone reverse transcription step, we have assumed they have similar mutation rates. Substituting Y_E into Eq 6, we can find an expression for the neutral rate of evolution rate at equilibrium:

$$S_E = \frac{\mu b[c+(2-q)(d+\delta)]}{b+\delta(1-q)} \text{ for } Y_E > 0 \quad \text{(Eq7)}$$

Lifespan of cccDNA during steady state infection

In our model, at equilibrium the generation time of HBV will be equal to the typical cccDNA lifespan, L_E. At equilibrium the number of cccDNA molecules remains constant, and therefore the rate at which cccDNA is produced is equal to the rate at which cccDNA is lost due to infected cell death, non-cytolytic clearance of cccDNA, and proliferation of infected cells. Since the reciprocal of the production rate is equal to the generation time, and the reciprocal of the rate cccDNA is lost is the typical lifespan of cccDNA, at equilibrium, viral generation time and cccDNA lifespan are

identical ($g_E = L_E$). This relationship holds because of our assumption of constant death rate and hence exponentially distributed lifetimes of cccDNA [27]; see [27,52] for how this changes for different distributions.

Using the equivalence of g_E and L_E, the lifespan of cccDNA at equilibrium can be determined from the mutation and neutral evolution rates by rearranging the first part of Eq 6:

$$L_E = \frac{\mu}{S_E} \qquad \text{(Eq8)}$$

Substituting the expression for S_E from Eq 7 into Eq 8, we can write an expression for the lifespan of cccDNA at equilibrium based on the model parameters:

$$L_E = \frac{b+\delta(1-q)}{b[c+(d+\delta)(2-q)]} \quad \text{for } Y_E > 0 \qquad \text{(Eq9)}$$

The lifespan of cccDNA when few cells are infected

If infection increases the death rate of hepatocytes, then the level of proliferation (to replace eliminated cells) will be larger the more cells are infected. Consequently, the lifespan of cccDNA when few cells are infected (e.g. during early phases of infection or during spontaneous clearance of infection, or after prolonged successful suppressive treatment) may differ from the lifespan during HBeAg^{POS} or HBeAg^{NEG} steady state infection. By setting *Y<<1* in equation 2, we can derive an expression for cccDNA lifespan when the copy number or burden is low:

$$L_{Y\ll 1} \approx \frac{1}{c+d(2-q)+\delta} \qquad \text{(Eq10)}$$

Comparing the expressions for L_E and $L_{Y\ll 1}$, we can see that if all cccDNA survives mitosis (*q=1*) or infection has a minimal effect on the death rate of infected cells (δ=0), then cccDNA lifespan remains unchanged during infection (as long as *d* and *c* don't change). However, if these conditions are not met, then the lifespan of cccDNA when few cells are infected, $L_{Y\ll 1}$, can be up to double the lifespan during chronic stable infection, L_E, for identical model parameters (e.g. when *q=c=d=0*, and *b>>* δ).

As we noted above, the cccDNA lifespan is only equivalent to the generation time at equilibrium. Using equation 5, when few cells are infected, the generation time is given by:

$$g_{Y\ll 1} \approx \frac{1}{b} \qquad \text{(Eq11)}$$

This has also been observed in the epidemiological literature [27]. Combining equations 3, 10 and 11 we see that:

$$R_0 = \frac{L_{Y\ll 1}}{g_{Y\ll 1}} \qquad \text{(Eq12)}$$

If $R_0 > 1$ and few cells are infected (i.e. the number of cccDNA is increasing) the life expectancy of cccDNA will be greater than the viral generation time, whereas if $R_0 < 1$ the life expectancy will be less than the viral generation time. This might be the case if, for example, increased immune responses associated with HBeAg^{NEG} infection push R_0 below one.

Estimating the generation time and lifespan of cccDNA from within-host evolutionary rates.

During stable chronic infection, the lifespan of cccDNA, L_E, equals the viral generation time, g_E, with $g_E = \mu/S_E$ (Eq 6). Although the mutation rate of HBV has not been determined, for avian hepadnavirus it has been estimated at $2x10^{-5}$ s/s/c (in the range $0.8x10^{-5}$ to $4.5x10^{-5}$; [10]). Since we are interested in the neutral rate of evolution, we assume that a third of all mutations in non-overlapping reading frames are synonymous, and that synonymous mutations are neutral or nearly neutral [54], giving a neutral mutation rate of around $0.67x10^{-5}$ s/s/c ($0.3x10^{-5}$ to $1.5x10^{-5}$) in non-overlapping reading frames. To incorporate the uncertainty associated with this estimate, we assumed the probability of the true mutation rate is log-normally distributed with mean $10^{-5.2}$ and standard deviation $10^{0.2}$.

Using longitudinal HBV sequence data, rates of evolution for non-overlapping regions of the genome were generated using a relaxed clock method, inferring $16.1x10^{-8}$ ($8.1\ x10^{-8}$, $25.5\ x10^{-8}$) substitutions per site per day (s/s/day) for HBeAg^{POS} and $38.9\ x10^{-8}$ ($27.2\ x10^{-8}$, $51.5\ x10^{-8}$) for HBeAg^{NEG} chronic infection (the numbers in brackets give the 5% and 95% highest posterior density (HPD) intervals; see Table 5 in [14]). In a separate study, using data from [55], the synonymous rate of evolution in non-overlapping genomic regions was estimated as half of the overall rate of evolution [16]. Assuming synonymous mutations are neutral, and that the ratio of synonymous to nonsynonymous evolutionary rates is constant during infection, we therefore take the neutral within-host rates of evolution during the HBeAg^{POS} and HBeAg^{NEG} phases of infection to be half the rates of evolution reported in [14] for non-overlapping reading frames. This gives a neutral rate of evolution of $8.0x10^{-8}$ ($4.0x10^{-8}$, $12.7x10^{-8}$) s/s/day during the HBeAg^{POS} phase, and $19.5\ x10^{-8}$ ($13.6\ x10^{-8}$, $25.8\ x10^{-8}$) s/n/day during the HBeAg^{NEG} phase. We assumed the probability distributions of these rates are normally distributed, with the standard deviation calculated using the difference between the estimated rate and the 5% HPD.

We randomly sampled from each of the probability distribution functions (PDFs) for the mutation rate and substitution rates, and used these values to calculate the generation time of cccDNA during HBeAg^{POS} and HBeAg^{NEG} CHB. This was repeated 100,000 times, from which the probability distributions for cccDNA generation time during HBeAg^{POS} and HBeAg^{NEG} chronic infection were estimated using the built in SmoothKernalDistribution function in *Mathematica* [56]. Assuming the number of cccDNA rapidly reaches equilibrium during HBeAg^{POS} and HBeAg^{NEG} infection, the virus generation will provide an approximation of the cccDNA lifespan during stable chronic infection (Fig 3).

Time to cccDNA eradication on treatment

Apart from when treatment is first initiated, the number of infected cells on treatment will be relatively low. Assuming eradication in our model is achieved when fewer than one cccDNA molecule persists, and there is no cccDNA replication whilst on treatment, the time to eradication can be approximated by:

$$t_{erad} \approx L_{Y \ll 1}\, Ln[N_{init}] \qquad \text{(Eq13)}$$

where N_{init} is the number of cccDNA when therapy is initiated and *Ln* is the natural logarithm. To determine reasonable values for N_{init}, we multiplied the number of hepatocytes in a human liver by the number of cccDNA per hepatocyte during untreated infection. There are about 1.4×10^8 hepatocytes per gram of human liver [57], and an adult human liver is around 1.5kg, giving approximately 2×10^{11} hepatocytes in total. In a recent study, an average of 6.3 copies of cccDNA per hepatocyte were found during chronic $HBeAg^{POS}$ infection, and 0.01 per hepatocyte during $HBeAg^{NEG}$ infection [30], which gives a total of approximately 1×10^{12} copies of cccDNA during $HBeAg^{POS}$ infection and 2×10^9 copies of cccDNA during $HBeAg^{NEG}$ infection.

Model assuming a subset of long-lived hepatocytes

If a proportion, α, of hepatocytes are long-lived, the dynamics of cccDNA in 'normal' infected cells, *Y[t],* and in long-lived infected cells *Z[t]*, are given by:

$$\frac{dY(t)}{dt} = b\left[1 - \frac{Y(t)}{1-\alpha}\right][Y(t) + Z(t)] - (1-q)[(d+\delta)Y(t) + (d_r + \delta_r)Z(t)]Y(t) - (d + \delta + c)Y(t)$$

$$\frac{dZ(t)}{dt} = b\left[1 - \frac{Z(t)}{\alpha}\right][Y(t) + Z(t)] - (1-q)[(d+\delta)Y(t) + (d_r + \delta_r)Z(t)]Z(t) - (d_r + \delta_r + c_r)Z(t)$$

where d_r, δ_r and c_r represent the natural death rates, cytoloytic death rate and clearance rates of the very long-lived cccDNA.

ACKNOWLEDGEMENTS

Thank you to Lorenzo Pellis, Fabian Zoulim and Helen Fryer for their helpful comments on earlier versions of this work. KAL is supported by The Wellcome Trust and the Royal Society grant number, 107652/Z/15/Z. JAM is supported by EU 2020 Research and Innovation Programme Consortia HEP-CAR under grant agreement No.667273, MRC MR/R022011/1 and The Wellcome Trust IA 200838/Z/16/Z. PM is supported by The Wellcome Trust, grant number 110110/Z/15/Z.

REFERENCES

1. World Health Organisation. Global hepatitis report, 2017. *WHO* (2017).
2. Ko, C. *et al.* Hepatitis B virus genome recycling and de novo secondary infection events maintain stable cccDNA levels. 1231–1241 (2018). doi:10.1016/j.jhep.2018.08.012
3. Urban, S., Schulze, A., Dandri, M. & Petersen, J. The replication cycle of hepatitis B virus. *J. Hepatol.* **52**, 282–284 (2010).

4. Lucifora, J. & Protzer, U. Review Attacking hepatitis B virus cccDNA – The holy grail to hepatitis B cure. *J. Hepatol.* **64**, S41–S48 (2016).
5. Nassal, M. HBV cccDNA: viral persistence reservoir and key obstacle for a cure of chronic hepatitis B. *Gut* **64**, 1972–1984 (2015).
6. Revill, P. A. *et al.* A global scientific strategy to cure hepatitis B. *Lancet Gastroenterol Hepatol* **4**, 545–548 (2019).
7. Boyd, A. *et al.* Decay of ccc-DNA marks persistence of intrahepatic viral DNA synthesis under tenofovir in HIV-HBV co-infected patients. *J. Hepatol.* **65**, 683–691 (2016).
8. Tu, T. & Urban, S. Virus entry and its inhibition to prevent and treat hepatitis B and hepatitis D virus infections. *Curr. Opin. Virol.* **30**, 68–79 (2018).
9. Mcnaughton, A. L. *et al.* Insights From Deep Sequencing of the HBV Genome-Unique, Tiny, and Misunderstood. *Gastroenterology* **156**, 384–399 (2019).
10. Pult, I., Abbott, N., Zhang, Y. Y. & Summers, J. Frequency of spontaneous mutations in an avian hepadnavirus infection. *J. Virol.* **75**, 9623–32 (2001).
11. Pereira-Gómez, M., Bou, J. V., Andreu, I. & Sanjuán, R. Lamivudine/Adefovir treatment increases the rate of spontaneous mutation of hepatitis B Virus in patients. *PLoS One* **11**, 1–12 (2016).
12. Sanjuán, R., Nebot, M. R., Chirico, N., Mansky, L. M. & Belshaw, R. Viral mutation rates. *J. Virol.* **84**, 9733–48 (2010).
13. Pybus, O. G., Belshaw, R. & Sanjua, R. Viral mutation and substitution : units and levels. 430–435 (2011). doi:10.1016/j.coviro.2011.08.004
14. Harrison, A. *et al.* Genomic analysis of hepatitis B virus reveals antigen state and genotype as sources of evolutionary rate variation. *Viruses* **3**, 83–101 (2011).
15. Lythgoe, K. A., Gardner, A., Pybus, O. G. & Grove, J. Short-Sighted Virus Evolution and a Germline Hypothesis for Chronic Viral Infections. **xx**, 1–13 (2017).
16. Vrancken, B., Suchard, M. A. & Lemey, P. Accurate quantification of within- and between-host HBV evolutionary rates requires explicit transmission chain modelling. *Virus Evol.* **3**, 1–9 (2017).
17. Lanfear, R., Kokko, H. & Eyre-walker, A. Population size and the rate of evolution. *Trends Ecol. Evol.* **29**, 33–41 (2014).
18. Locarnini, S. & Zoulim, F. Molecular genetics of HBV infection. *Antivir. Ther.* **15**, 3–14 (2010).
19. Murray, J. M. & Goyal, A. In silico single cell dynamics of hepatitis B virus infection and clearance. *J. Theor. Biol.* **366**, 91–102 (2015).
20. Goyal, A. The Role of Infected Cell Proliferation in the Clearance of Acute HBV Infection in Humans. 23–38 (2017). doi:10.3390/v9110350
21. Mason, W. S. *et al.* HBV DNA Integration and Clonal Hepatocyte Expansion in Chronic Hepatitis B Patients Considered Immune Tolerant. *Gastroenterology* **151**, 986-998.e4 (2016).
22. Ribeiro, R. M. *et al.* Hepatitis B Virus Kinetics under Antiviral Therapy Sheds Light on Differences in Hepatitis B e Antigen Positive and Negative Infections. *J. Infect. Dis.* **202**, 1309–1318 (2010).
23. Mason, W. S., Litwin, S. & Jilbert, A. R. Immune selection during chronic hepadnavirus infection. *Hepatol. Int.* **2**, 3–16 (2008).

24. Allweiss, L. *et al.* Proliferation of primary human hepatocytes and prevention of hepatitis B virus reinfection ef fi ciently deplete nuclear cccDNA in vivo. 542–552 (2018). doi:10.1136/gutjnl-2016-312162
25. Reaiche-Miller, G. Y. *et al.* Duck hepatitis B virus covalently closed circular DNA appears to survive hepatocyte mitosis in the growing liver. *Virology* (2013). doi:10.1016/j.virol.2013.08.014
26. Whalley, B. S. A. *et al.* Kinetics of Acute Hepatitis B Virus Infection in Humans. **193**, (2001).
27. Scalia, G., Svensson, Å., Asikainen, T. & Giesecke, J. Some model based considerations on observing generation times for communicable diseases. *Math. Biosci.* **223**, 24–31 (2010).
28. Scholle, S. O., Ypma, R. J. F., Lloyd, A. L. & Koelle, K. Viral substitution rate variation can arise from the interplay between within-host and epidemiological dynamics. *Am. Nat.* **182**, 494–513 (2013).
29. Kajino, K. *et al.* Woodchuck hepatitis virus infections: very rapid recovery after a prolonged viremia and infection of virtually every hepatocyte. *J. Virol.* **68**, 5792–803 (1994).
30. Testoni, B. *et al.* Serum hepatitis B core-related antigen (HBcrAg) correlates with covalently closed circular DNA transcriptional activity in chronic hepatitis B patients. *J. Hepatol.* **70**, 615–625 (2019).
31. Downs, L. O. *et al.* Bimodal distribution and set point HBV DNA viral loads in chronic infection : retrospective analysis of cohorts from the UK and South Africa. *medRxiv* (2020). doi:http://dx.doi.org/10.1101/2020.01.08.20016972
32. Bonhoeffer, S., Funk, G. A., Gunthard, H. F., Fischer, M. & Muller, V. Glancing behind virus load variation in HIV-1 infection. *Trends Microbiol.* **11**, 499–504 (2003).
33. Müller, V., Marée, A. F. M. & De Boer, R. J. Small variations in multiple parameters account for wide variations in HIV-1 set-points: a novel modelling approach. *Proc. R. Soc. B* **268**, 235–42 (2001).
34. Lythgoe, K. A., Pellis, L. & Fraser, C. Large Variations in HIV-1 Viral Load Explained by Shifting-Mosaic Metapopulation Dynamics. 1–25 (2016). doi:10.1371/journal.pbio.1002567
35. Downs, L. O. *et al.* Electronic health informatics data to describe clearance dynamics of hepatitis B surface antigen (HBsAg) and e antigen (HBeAg) in chronic hepatitis B virus infection. *MBio* **10**, 1–15 (2019).
36. Wong, D. K. H. *et al.* Reduction of hepatitis B surface antigen and covalently closed circular DNA by nucleos(t)ide analogues of different potency. *Clin. Gastroenterol. Hepatol.* **11**, 1004-1010.e1 (2013).
37. Boyd, A. *et al.* Decay of ccc-DNA marks persistence of intrahepatic viral DNA synthesis under tenofovir in HIV-HBV co-infected patients. *J. Hepatol.* **65**, 683–691 (2016).
38. Dahari, H., Shudo, E., Ribeiro, R. M. & Perelson, A. S. Modeling complex decay profiles of hepatitis B virus during antiviral therapy. *Hepatology* **49**, 32–38 (2009).
39. De Crignis, E. *et al.* Human liver organoids; a patient-derived primary model for HBV infection and related hepatocellular carcinoma. *bioRxiv* 568147 (2019). doi:10.1101/568147

40. Addison, W. R. *et al.* Half-life of the duck hepatitis B virus covalently closed circular DNA pool in vivo following inhibition of viral replication. *J. Virol.* **76**, 6356–6363 (2002).
41. Zhu, Y. *et al.* Kinetics of Hepadnavirus Loss from the Liver during Inhibition of Viral DNA Synthesis. *J. Virol.* **75**, 311–322 (2001).
42. Volz, T. *et al.* Impaired Intrahepatic Hepatitis B Virus Productivity Contributes to Low Viremia in Most HBeAg-Negative Patients. 843–852 (2007). doi:10.1053/j.gastro.2007.06.057
43. Palmer, S. *et al.* Low-level viremia persists for at least 7 years in patients on suppressive antiretroviral therapy. **105**, 3879–3884 (2008).
44. Kajino, K. *et al.* Woodchuck hepatitis virus infections: very rapid recovery after a prolonged viremia and infection of virtually every hepatocyte. *J. Virol.* **68**, 5792–803 (1994).
45. Zhang, Y. *et al.* Single-cell analysis of covalently closed circular DNA copy numbers in a hepadnavirus-infected liver. **100**, (2003).
46. Li, M., Sohn, J. A. & Seeger, C. Distribution of Hepatitis B Virus Nuclear DNA. *J. Virol.* **92**, 1–13 (2017).
47. Jilbert, A. R. *et al.* Rapid resolution of duck hepatitis B virus infections occurs after massive hepatocellular involvement. *J Virol* **66**, 1377–1388 (1992).
48. Lenhoff, R. J. & Summers, J. Construction of avian hepadnavirus variants with enhanced replication and cytopathicity in primary hepatocytes. *J. Virol.* **68**, 5706–5713 (1994).
49. Summers, J., Smith, P. M. & Horwich, A. L. Hepadnavirus Envelope Proteins Regulate Covalently Closed Circular DNA Amplification. **64**, 2819–2824 (1990).
50. Summers, J., Smith, P. M., Huang, M. J. & Yu, M. S. Morphogenetic and regulatory effects of mutations in the envelope proteins of an avian hepadnavirus. *J. Virol.* **65**, 1310–7 (1991).
51. Kimura, M. *The Neutral Theory of Molecular Evolution*. (Cambridge University Press, 1985).
52. Svensson, Å. A note on generation times in epidemic models ˚. *Mathematic* **208**, 300–311 (2007).
53. Nishiura, H. Time variations in the generation time of an infectious disease: Implications for sampling to appropriately quantify transmission potential. *Math. Biosci. Eng.* **7**, 851–869 (2010).
54. Cuevas, J. M., Domingo-Calap, P. & Sanjuán, R. The Fitness Effects of Synonymous Mutations in DNA and RNA Viruses. *Mol. Biol. Evol.* **29**, 17–20 (2011).
55. Lin, Y.-Y. *et al.* New Insights into the Evolutionary Rate of Hepatitis B Virus at Different Biological Scales. *J. Virol.* **89**, 3512–3522 (2015).
56. *Mathematica Version 10.3*. (Wolfram Research, Inc, 2015).
57. Sohlenius-Sternbeck, A. K. Determination of the hepatocellularity number for human, dog, rabbit, rat and mouse livers from protein concentration measurements. *Toxicol. Vitr.* **20**, 1582–1586 (2006).