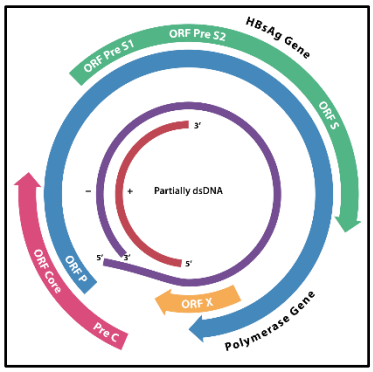
Vishvak Subramanyam
C&SB 195
8/25/21

HBV Genome Browser

Introduction

Hepatitis B is a vaccine-preventable liver infection caused by the Hepatitis B Virus (HBV). The vast majority of those infected encounter the infection during adulthood and develop self-limited, acute infection. However, 90% of those who acquire an infection during infancy, and 5% of those infected during adulthood develop chronic infections.[i] Chronic Hepatitis B (CHB) is a major global health burden highly endemic to Africa, East-Asia and parts of Central Europe, affecting 257 million people. [ii] This public health catastrophe contributes to 880,000 deaths annually due to complications resulting from persistent liver infection such as severe liver cirrhosis and hepatocellular carcinoma (HCC).

Although a preventative vaccine exists, it is not effective against those already infected. The current treatment paradigm for those with CHB involves immunomodulatory agents or antivirals.[iii] Although effective in inhibiting viral replication and reducing the likelihood of liver disease, these treatments rarely achieve complete viral clearance. There is a need for the development of curative therapies for CHB; however, features of the HBV genome and replication cycle make complete viral elimination difficult. The transcriptional template of the HBV virus, termed covalently closed circular DNA (cccDNA), is highly stable and the cause of persistent infection as it can integrate into the host genome, encoding the proteins necessary for HBV replication, secretion and reinfection. Even patients that clear an acute infection have detectable levels of cccDNA in their hepatocytes.[iv] In place of a "sterilizing cure", the concept of a "functional cure" has taken hold - defined as the loss of detectable HBV surface antigen (HBsAg).[v]

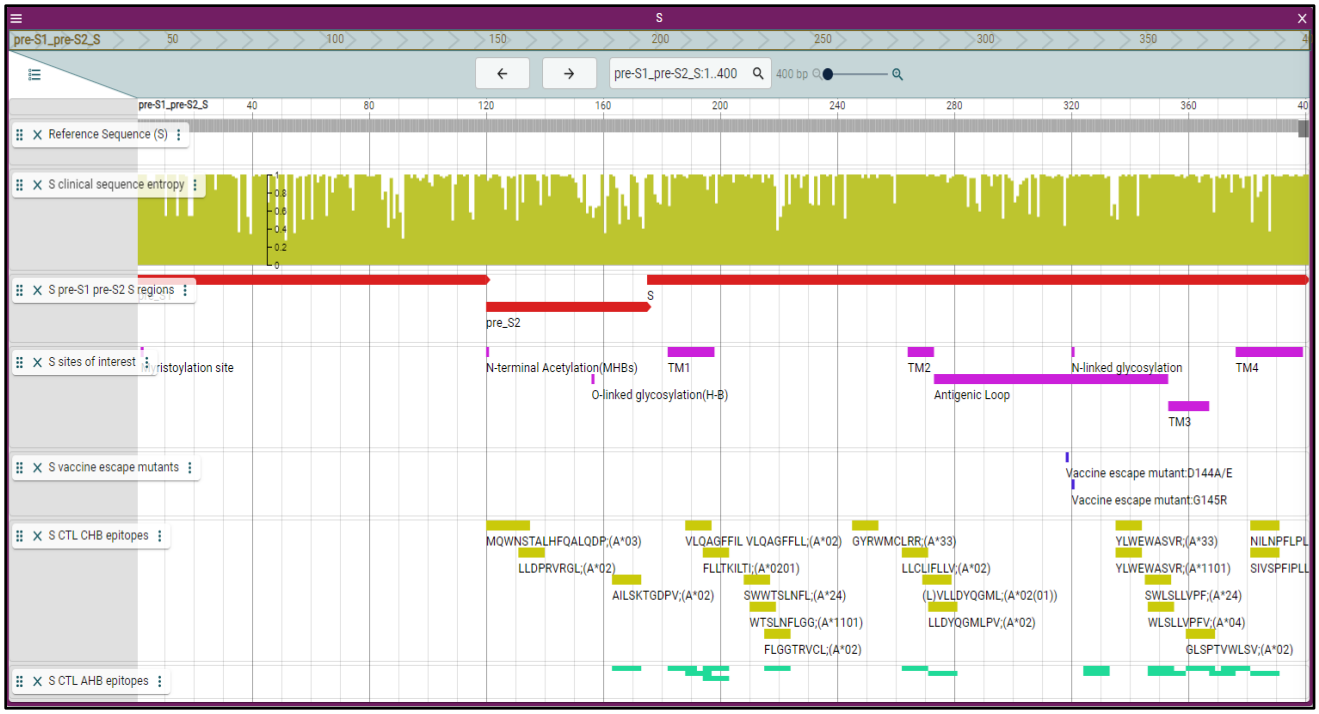


CD8+ T cells are widely regarded as the ultimate effectors of viral clearance in acute HBV infections.[v] However, in CHB infection, CD8+ T-cells are characterized by inhibited proliferation, differentiation and cytotoxic activity. Additionally, HBV is marked by many mechanisms of escape in chronic patients. These mechanisms include selective mutations of antibody and HLA-binding residues, masking of epitopes through N-linked glycosylation and downregulation of class I MHCs on hepatocytes.[vi]

Thus, HBV epitopes are of great interest for new therapeutics and are a core component of the HBV cure strategy at Gilead Sciences. The identification of highly conserved epitopes that induce a strong cytotoxic response, is crucial for the development of therapeutics. These may include epitope vaccines (DNA vaccines) to stimulate T-cell and B-cell responses, antibody treatments and, the genetic engineering of CD8+ to preferentially target conserved epitopes. This summer as a clinical virology intern supporting the HBV cure program at Gilead, I configured an HBV Genome Browser to visualize the epitopes, sites of interest and the conservation across the 4 HBV proteins. This browser can be used to identify regions of therapeutic interest through conservation tracks and biologically contextualize clinical sequences by identifying the binding sites/epitopes that are disrupted as a result of mutations.

HBV Genome Browser

Please see a temporary version of the browser available at: http://hbvgenomebrowser.org.s3-website-us-west-1.amazonaws.com/



The HBV Genome is 3.2kb long and contains 4 open reading frames that code for the 4 proteins of HBV: S (surface), Core (capsid protein), Pol (polymerase) and X (HBxAg). For each of these proteins, several informational tracks were configured:

Epitope tracks:
Based on extensive literature searches, HBV specific cytotoxic T-cell, helper T-cell and antibody epitopes were retrieved and categorized based on the disease stage that they were observed in (acute (AHB) vs chronic (CHB)). These were then configured into tracks for each protein of the genome browser (e.g.: S CTL CHB Epitopes).

Entropy tracks:
Shannon entropy was calculated for each residue within all publicly available GenBank clinical sequences (~7,000). Separate tracks were configured for genotypes A-D of HBV and a consensus track including all the genotypes were also configured for each protein. This track helps identify candidate epitopes as it shows HBV sequence conservation at each residue.

Sites of interest:
These tracks included receptor binding sites, glycosylation sites, vaccine escape mutants, etc. This track can be used to contextualize the mutations within clinical sequences and identify the viral/host mechanisms that are disrupted.

Sequence Analysis:

To provide an example of how this tool can be used by the clinical virology team, I analyzed clinical sequences from patients who were developing HCC (from an internal study). Deletions were observed in the pre-S2 region of the S protein in 30% of patients. When overlaying these sequences with the genome browser, we see that epitope "LLDPRVRGL" associated with the HLA-A*02 is disrupted. This indicates a possible mechanism of HBV to evade TCR binding and inhibit viral clearance in HCC patients. This epitope can then be looked up in the consolidated epitope database that I configured (see methods) to see the study that determined this epitope and the other HLA alleles that may be disrupted.

Conclusion

This HBV Genome browser is a tool to visualize the HBV Genome, its proteins, and associated epitopes. It can be used to analyze the mutations within clinical sequences and identify possible host mechanisms that may be disrupted as a result of indels within epitopes and binding regions. Additionally, it can be used to identify candidate epitopes that are highly conserved for cure development. This tool contextualizes HBV deep sequencing data by highlighting regions of therapeutic interest.

Methods:

Genome Browser

The genome browser was built using JBrowse2 which is a dynamic platform for genome visualization and analysis based on Node.js and JavaScript.[vii] It is an embeddable web-application that can be configured with many different types of tracks and hosted on any web hosting service.

HBV Epitopes

HBV epitopes were pulled from existing epitope databases such as Hepitopes[viii] (which contains HLA-A*02 HBV epitopes) and from the Immune Epitope Database (IEDB)[ix]. Additionally, antibody and T-helper epitopes were pulled from other published literature. These epitopes were consolidated into one database with 78 CHB/80 AHB T-cell epitopes and 65 antibody epitopes across all 4 proteins. These epitopes were then converted into general feature format (GFF), sorted using GenomeTools[x], indexed using Tabix[xi] and then loaded into the browser.

Entropy

Publicly available Genbank HBV sequences were collected for genotypes A-D. Entropy for the 4 protein sequences was calculated and converted into bigwig format to be used by JBrowse2. For each of the proteins, 5 entropy tracks are available: 4 for the four HBV genotypes and 1 consensus track that showed the conservation across all genotypes.

Web Hosting:

The genome browser was hosted using AWS s3 services.

---

[i] Yuen MF, Chen DS, Dusheiko GM, Janssen HLA, Lau DTY, Locarnini SA, Peters MG, Lai CL. Hepatitis B virus infection. Nat Rev Dis Primers. 2018 Jun 7;4:18035. doi: 10.1038/nrdp.2018.35. PMID: 29877316.

[ii] World Health Organization. Global Hepatitis Report, 2019 https://www.who.int/hepatitis/publications/global-hepatitis-report2017/en/ (WHO, 2017).

[iii] Salpini, Romina et al. "A snapshot of virological presentation and outcome of immunosuppression-driven HBV reactivation from real clinical practice: Evidence of a relevant risk of death and evolution from silent to chronic infection." *Journal of viral hepatitis* vol. 26,7 (2019): 846-855. doi:10.1111/jvh.13101

[iv] Iannacone, Matteo, and Luca G. Guidotti. "Immunobiology and Pathogenesis of Hepatitis B Virus Infection." *Nature News*, Nature Publishing Group, 17 May 2021, www.nature.com/articles/s41577-021-00549-4.

[v] Revill, Peter A., et al. "The Evolution and Clinical Impact of Hepatitis B Virus Genome Diversity." *Nature Reviews Gastroenterology & Hepatology*, vol. 17, no. 10, 2020, pp. 618–634., doi:10.1038/s41575-020-0296-6.

[vi] Lumley, Sheila F., et al. "Hepatitis B Virus Adaptation to the CD8 T Cell Response: Consequences for Host and Pathogen." *Frontiers*, Frontiers, 1 Jan. 1AD, www.frontiersin.org/articles/10.3389/fimmu.2018.01561/full.

[vii] Buels, Robert, et al. "JBrowse: a Dynamic Web Platform for Genome Visualization and Analysis." Genome Biology, vol. 17, no. 1, 2016, doi:10.1186/s13059-016-0924-1.

[viii] Lumley, Sheila, et al. "Hepitopes: A Live Interactive Database of HLA Class I Epitopes in Hepatitis B Virus." *Wellcome Open Research*, U.S. National Library of Medicine, www.ncbi.nlm.nih.gov/pmc/articles/PMC5142601/.

[ix] Vita R, Mahajan S, Overton JA, Dhanda SK, Martini S, Cantrell JR, Wheeler DK, Sette A, Peters B. The Immune Epitope Database (IEDB): 2018 update. Nucleic Acids Res. 2018 Oct 24. doi: 10.1093/nar/gky1006. [Epub ahead of print] PubMed PMID: 30357391.

[x] G. Gremme, S. Steinbiss and S. Kurtz.
GenomeTools: a comprehensive software library for efficient processing of structured genome annotations. IEEE/ACM Transactions on Computational Biology and Bioinformatics 2013, 10(3):645–656

[xi] Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, Li H, Twelve years of SAMtools and BCFtools**,** *GigaScience* (2021) 10(2) giab008 [33590861]