HCV reference genome using BWA. Primary maximum likelihood phylogenetic analyses were performed using ClustalW and IQTREE to infer the phylogenetic relationships of the sequenced samples in the context of complete genome sequences of the same genotype. NGS sample preparation method of HCV from blood plasma was established. Complete genome sequences of HCV could be obtained with variable coverage depending on the viral load of plasma samples. No significant reduction of host DNA proportion in DNase treated samples in comparison to the controls was observed. The new sequences clustered within the Los Alamos National Laboratory database-deposited HCV subtype 4d samples. The method can be used to obtain full-length sequences of HCV from nucleic acid samples not previously planned for NGS. No improvement was observed when DNase pre-treatment of nucleic acids extracted from blood plasma was performed.

A21 Retrospectively describing hepatitis C virus transmission dynamics and tracking HCV transmission networks in real-time for strategic elimination interventions

J. Koopsen,[1] C. Russell,[2] M. van der Valk,[3] and J. Schinkel[1]

[1]Lab of Clinical Virology, Department of Medical Microbiology, Academic Medical Center, Amsterdam, The Netherlands, [2]Lab of Applied Evolutionary Biology, Department of Medical Microbiology, Academic Medical Center, Amsterdam, The Netherlands and [3]Division of Infectious Diseases, Department of Internal Medicine, Academic Medical Center, Amsterdam, The Netherlands

Despite impressive uptake of direct acting antivirals for hepatitis C virus (HCV) in the Netherlands among HIV/HCV co-infected men who have sex with men (MSM), HCV transmission continues, especially among patients previously successfully treated for HCV. The incidence of reinfection occurs at the extremely high rate of 15 per 100 person-years. Clearly, more sophisticated methods are necessary to identify the sources and timing of new HCV infections among MSM. The aim of this research is to phylogenetically characterize HCV transmission dynamics within MSM-specific networks in order to provide a solid base for targeted interventions to monitor, control, and eventually stop the ongoing transmission of HCV among HIV-infected MSM and to prevent further spread of HCV to the community at large. The methodology that will be used is two-fold. Firstly, it concerns setting up a real-time monitoring system to track the HCV epidemic using phylogenetic tools and open-source software from http://nextstrain.org. Secondly, several phylogenetic methods will be used to retrospectively identify transmission clusters in Amsterdam and define epidemiological characteristics, including the directionality of transmission and the size and introduction dates of the clusters. This means that cluster cut-off points will have to be calculated. This research will result in a web-based molecular surveillance tool to monitor the persistence of endemic clades, emergence of new clades, and transmission clusters in 'real time', which, combined with clinical and epidemiological data, will be used for targeted interventions. The surveillance tool will be based on the open-source software from nextstrain.org. Secondly, by retrospectively describing the HCV transmission clusters in terms of introduction dates and subsequent dynamics, we may be able to better predict the future dynamics of the different clusters. High-resolution viral sequencing will allow us to identify the source and timing of (new) HCV infections and follow the trajectory of these MSM-specific lineages through the MSM population. Real-time insight in transmission networks using a web-based molecular surveillance tool will identify key targets for rapid interventions, awareness campaigns, and testing strategies. This can be used to prevent further spread to HIV-negative MSM and to control and eventually eliminate HCV from the MSM population.

A22 Phylogenetic clustering of hepatitis C virus infection among people who inject drugs in Baltimore

Oluwaseun Falade-Nwulia,[1] Jada Hackman,[2] Shruti Mehta,[1] Mark Sulkowski,[1] Carl Latkin,[1] David Thomas,[1] Zach Downing,[1] Rachel Latanich,[1] Gregory Kirk,[1] Oliver Laeyendecker,[2] and Stuart Ray[1]

[1]Johns Hopkins University, Baltimore, MD, USA and [2]National Institute of Allergy and Infectious Diseases, Baltimore, MD, USA

The availability of effective, oral direct acting antivirals for hepatitis C virus (HCV) treatment has fueled optimism for HCV elimination through treatment as prevention (TasP) among people who inject drugs (PWID). Identifying characteristics of individuals in transmission networks would provide critical information for the development and implementation of effective, targeted HCV TasP strategies. The AIDS linked to the IntraVenous Experience (ALIVE) cohort has followed PWID in Baltimore since 1988. Sequencing of the HCV core/E1 region (342 nucleotides) was performed on HCV viremic samples from the most recent study visit attended by ALIVE participants between August, 2005 and December, 2016. Outgroup sequences were retrieved from GenBank through a BLAST search for HCV sequences similar to study sequences to support identification of 'local clusters' and were aligned to study sequences using Clustal O. Phylogenetic trees were inferred for each of HCV subtype 1a and 1b separately through maximum likelihood analysis implemented in the MEGA X software using the Tamura-Nei model with gamma distribution and invariant sites. Nucleotide substitution model selection was based on the corrected Akaike information criterion scores of various models in MEGA. Robustness of the resulting tree was assessed by bootstrapping with 1,000 replicates. Clusters were identified using ClusterPicker software (70% bootstrap threshold and 0.05 maximum genetic distance threshold). Sensitivity analyses were performed by varying the genetic distance threshold between 0.025 and 0.05 to determine the effect on identification of factors associated with clustering. HCV infection clustering was defined as >2 participants with HCV genome sequences satisfying 70 per cent bootstrap and 0.05 genetic threshold distance requirement for sequence similarity. Logistic regression was used to assess sociodemographic factors associated with being in an HCV cluster. Among 512 HCV genotype 1 viremic PWID, HCV subtype prevalence was 83 per cent genotype 1a and 17 per cent genotype 1b. The median age of participants was 54 years, 68 per cent male, 87 per cent Black, and 38 per cent HIV infected. Overall, 9 per cent ($n=44$) were grouped into 21 clusters, consisting of 20 pairs and 1 triad. Of the 425 genotype 1a and 87 genotype 1b samples evaluated, 8 per cent ($n=33$) and 13 per cent ($n=11$) respectively, were in clusters. In unadjusted analyses, membership in a cluster, was associated with younger age (odds ratio (OR) 1.5 [95% confidence interval (CI) 1.1–2.1] per 10 year age decrease); female sex (OR 2.8 [95% CI 1.5–5.3]), HIV infection (OR 4.9 [95% CI 2.5–9.9]), and living in East Baltimore (versus outside East Baltimore, OR 2.0 [95% CI 1.0–3.9]). In adjusted analyses, female sex (OR 2.0 [95% CI 1.0–3.9] and HIV infection (OR 5.4 [95% CI 2.6–11.1] remained independently associated with being in an HCV infection cluster. HIV-infected PWID and their networks should be prioritized for HCV treatment and prevention interventions given an increased likelihood of transmission in these groups.

A23 Population level diversification of hepatitis C viral strains over time among people who inject drugs in Baltimore, MD

J. Hackman,[1] O. Falade-Nwulia,[2] S. Mehta,[2] Z. Downing,[2] G. Kirk,[2] S. Ray,[2] D. Thomas,[2] and O. Laeyendecker[1]

[1]National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD, USA and [2]Division of Infectious Diseases, Johns Hopkins University School of Medicine, Baltimore, MD, USA

Hepatitis C virus (HCV) infection occurs in 30–90 per cent of people who inject drugs (PWID). Although cure rates can exceed 95 per cent, treatment access is limited and approximately 400,000 people die each year due to complications of chronic infection. A temporal analysis of cluster networks among PWID can be used to inform strategies to interdict transmission. In Baltimore, PWID have been recruited for The AIDS Linked to the IntraVenous Experience (ALIVE) cohort. A demographic questionnaire was administered and recorded for baseline and recent participants. Viral RNA underwent PCR with primers targeting the core and envelope-1 protein (CE1) and sequenced via Sanger sequencing. Sequences with >400 bp reads and Q-scores >370 were used for downstream analysis resulting in 322 ALIVE baseline participants (1988–9) and 548 recently diagnosed subjects enrolled approximately two decades later (2005–16). Cluster networks were rendered with a threshold of 4 per cent in MicrobeTRACE, and statistical analyses were performed in R Studio. Of the 1988–9 subjects, the majority (259/317, 81.7%) were a part of cluster. There were nine clusters and fifty-eight singletons, with two large clusters containing most sequences of genotype 1a (73.5%). Two decades later, a minority of recently diagnosed individuals (235/512, 44.1%) were part of a cluster. There were seventeen clusters with 286 singletons with two large clusters containing 1a genotype

individuals (21.5%). Additional clustering was done by parsing the two datasets by subtype 1a ($n = 714$) and 1b ($n = 151$). The genotype 1a network demonstrates a majority, 65.8 per cent, of participants in clusters. Moreover, two large clusters can be observed with baseline participants towards the center and recent participants on the outskirts indicative of high linkage at baseline. The genotype 1b network produced a single large cluster but subclusters were observed. The sequences between the two time points co-mingled but subclusters were also observed. Interestingly, the two large clusters from 1988 to 1989 were still evident in the 2005–16 viral sequences. We observed greater cluster diversity in more recently diagnosed individuals, indicative of a less connected network of individuals sharing transmission risk, though major viral strains did persist over time in this cohort.

A24 Phylogeographic analysis of hepatitis A virus in Russia

A. Koreshova,[1] G. Bazykin,[1,2] and A. Neverov[3]

[1]Center for Data-Intensive Biomedicine and Biotechnology, Skolkovo Institute of Science and Technology, Moscow, 143026, Russia, [2]Sector for Molecular Evolution, Kharkevich Institute of Information Transmission Problems of the Russian Academy of Sciences, Moscow, 127051, Russia and [3]Central Research Institute for Epidemiology, Moscow, 111123, Russia

Hepatitis A virus (HAV) is a positive-stranded RNA virus, a member of Picornaviridae, and a representative of genus Hepatovirus. It is unique among picornaviruses with regards to its hepatotropism, structure, and life cycle. HAV is spread via the fecal-oral route as a non-enveloped particle, while, in the blood the virus circulates in an envelope formed from the host cell membrane. HAV causes acute hepatitis in adults and is usually asymptomatic in children <6 years of age. The clinical features include fever, malaise, anorexia, nausea, abdominal discomfort, dark urine, and jaundice, all of which usually last >2 months. There is no evidence of chronic liver disease or persistent infection following acute infection. Due to its mode of transmission, HAV prevails in areas with low hygiene standards but does not give rise to epidemics because most people are infected at an early age and derive a life-long immunity. Thus, HAV infection has more impact on countries with higher socio-economical level where it is mostly registered as an outbreak in adults, which is the case in Russia. One feature distinguishing HAV from other picornaviruses is its remarkably slow mutation rate. HAV genotyping is typically carried out using highly variable regions VP1/2A and 2C/3A. Recently, it was shown that resolution provided by short fragments is not enough for reliable results. Unfortunately, previous research in HAV phylogeography was carried out only on these short sequences and did not include Russia or CIS territories. HAV comprises six genotypes, of which I and III are most frequent in humans and are both divided into A and B subgenotypes. Preliminary phylodynamic analysis of 80 highly variable region sequences (carried out by A. Neverov) has shown a particular pattern of geographical distribution of HAV genotypes in Russia. There are only two subgenotypes widely spread: IA predominates in the European part of Russia, and IIIA is found mainly in the Asian part. However, the history of HAV spread in Russia remains unclear. We hypothesized that IIIA subgenotype originated from India, while IA subgenotype came later from Europe and is still expanding. The Central Research Institute of Epidemiology kindly provided us with the unique collection of HAV isolates obtained from more than 30 subjects of the Russian Federation, as well as a number of isolates obtained from CIS countries. Samples (>500 isolates) were collected from 1999 to 2015 and characterized by one or both of the two most variable fragments of HAV genome (VP1/2A and 2C regions). The dataset includes 145 unique sequences of 2C/3A region, length ~650 bp, and 243 sequences of VP1/2A region, length ~400 bp. For each sample, date and location of collection are indicated. Whole-genome sequences of HAV from GenBank database were also used. They were aligned with MUSCLE, and the target 2C/3A and VP1/2A fragments were extracted. Partial HAV sequences from GenBank were not added to the analysis due to too little overlap with our sequences. Initial phylogeographic analysis was carried out in BEAST. Results were checked with the Tracer program, and the Spread3 package was used to visualize the results of the phylogeographic analysis in continuous space [16]. The BEAST output supports the hypothesis that IIIA subgenotype originated from India, whereas the situation with the IA subgenotype remains unclear. The reason for this might be either poor sampling of the Mediterranean area and Middle East in our analysis or low precision provided with variable fragments. The next step is to obtain full-genome sequences of approximately 100 of our samples to increase resolution and make use of hundreds of partial sequences of HAV genomes available in GenBank.

A25 Impact of polymorphism in the hepatitis B surface gene on human leukocyte antigen (HLA) class II

W.-T. Choga,[1] M. Anderson,[1,2] B.-B. Phinius,[1] T. Mbangiwa,[1,3] T.-G. Bell,[4] K.-K. Seatla,[1,3] R.-M. Musonda,[1,5] S. Moyo,[1,5] J.-T. Blackard,[6] and S. Gaseitsiwe[1,5]

[1]Botswana Harvard AIDS Institute Partnership, Gaborone, Botswana, [2]Department of Biological Sciences, University of Botswana, Gaborone, Botswana, [3]Department of Medical Laboratory Sciences, Faculty of Health Sciences, University of Botswana, Gaborone, Botswana, [4]HDVRU, Department of Internal Medicine, School of Medicine, University of Witwatersrand, Johannesburg, South Africa, [5]Department of Immunology and Infectious Diseases, Harvard T.H. Chan School of Public Health, Boston, MA, USA and [6]University of Cincinnati College of Medicine, Cincinnati, OH, USA

There is still no cure for chronic hepatitis B virus infection (CHBV), a major cause of liver cancers and related malignancies. Elucidating the role of CD4+ T-helper cells in activating immunological responses that clear antigenic peptides during primary HBV infection holds a potential strategy for developing potent vaccines. Since the strength of CD4+ T cell responses is dictated by binding of viral epitopes to class-II human leukocyte antigens (HLAs), we hypothesize that the quality of immunological responses in CHBV patients is influenced by host genetics and HBV genotypes. Here, ninety-two non-recombinant complete HBV surface-gene proteins (PreS1/S) from Botswana were sequenced (genotype A 44(47.8%); D 48(52.2%)) and 15-mer binding epitopes restricted to nine HLA-class II molecules (DRB5/1) were mapped *in silico*. The HLAs used have high population coverage in Botswana. The total predicted epitopes per HLA were 94-(genotype A) and 105-(genotype D) for PreS1, 42 (A and D) for PreS2, and 105 (A and D) for S. Epitope densities (binding peptides to total epitopes) were 3 per cent and 6 per cent (PreS1A&D), 4 per cent and 2 per cent (PreS2A&D), and 23 per cent and 22 per cent (S1A&D). SA&D proteins had most polytopes: CPGYRWMCLRRFII66-81, PGYRWMCLRRFIIF67-82, GYRWMCLRRFIIFL68-83, and YRWMCLRRFIIFLF69-84 binding to 5 (55.6%) HLAs (DRB1*0101/0701/1101/1501 and DRB5*0101) used. HLA-DRB*0101 bound the most epitopes, and the least were bound by HLA-DRB*0302/0701/0401 for both genotypes. PreS1D polytope: PAFRANTANPDWDFN32-46 binds to DRB1*0101/0401/1302 and PreS2 polytopes: TAFHQALQDPRVRG6-19 and AFHQALQDPRVRGL7-20 bind to DRB1*010/1501 alleles. Non-synonymous mutations impair peptide-HLA binding when assessed as combinations of > 2. The least active HLAs may be associated with CHBV and vice-versa for HBV clearance, thus the algorithm may be used to predict HBV prognosis for different haplotypes. The results favor the use of epitopes from S protein as broad genotype vaccine. This study highlights the need to explore further the mechanisms of PreS1 and its effect on the immune system.

A26 Molecular epidemiology of hepatitis E virus in Ireland 2016

Charlene Bennett, Linda Dunford, Suzie Coughlan, Cillian De Gascun, and Joanne O'Gorman

National Virus Reference Laboratory, University College Dublin, Belfield, Dublin 4, Ireland

Foodborne viruses such as hepatitis E virus (HEV) pose an increasing risk to public health and to confidence in Irish food. Hepatitis E has been acknowledged as a significant pathogen of likely zoonotic transmission, with pork products and shellfish being implicated as potential sources. The European Food Safety Authority has recommended that systematic strain typing of viruses in humans, animals, and food commodities is needed to improve understanding of etiological agents and foodborne transmission pathways, in particular for HEV. The dominant autochthonous genotype of HEV in Europe is genotype 3, thought to be associated with consumption of contaminated food, specifically pork products. However, little is known about the epidemiology of HEV in Ireland. In 2016, HEV became a notifiable disease in Ireland. Following this, as part of the Department of Agriculture, Food and the Marine-funded FoVIRA study, the molecular epidemiology of HEV in Irish clinical samples has been characterized for the first time. HEV RNA-positive clinical