nonsynonymous substitutions within HIV-1 infected individuals. Among the 25,251 polymorphic codon sites analysed, FUBAR revealed that 189-fold more were detectably evolving under persistent negative selection than were evolving under persistent positive selection. Three specific codon sites within the genes celA2b, katG, and cyp138 were identified by MEDS as displaying significant evidence of evolving under directional selection influenced by HIV-1 co-infection. All three genes encode proteins that may indirectly interact with human proteins that, in turn, interact functionally with HIV proteins. Unexpectedly, epitope encoding regions were enriched for sites displaying weak evidence of directional selection influenced by HIV-1. Although the low degree of genetic diversity observed in our *M. tuberculosis* dataset means that these results should be interpreted carefully, the effects of HIV-1 on epitope evolution in *M. tuberculosis* may have implications for the design of *M. tuberculosis* vaccines that are intended for use in populations with high HIV-1 infection rates.

A20 Deep sequencing reveals viral evolution in GAG within protective HLA Alleles B*57: 02, B*58: 01, and B*7 supertype individuals acutely infected with HIV-1 subtype C in Durban, South Africa

K. Gounder,[1,2] V. Naidoo,[1,2] N. Padayachi,[1,2] Q. Mthethwa,[1,2] D. Dilernia,[3] E. Hunter,[3] B. Walker,[1,4] and T. Ndung'u[1,2]

[1]HIV Pathogenesis Programme, Doris Duke Medical Research Institute, University of KwaZulu-Natal, Durban 4013, South Africa, [2]KwaZulu-Natal Research Institute for Tuberculosis and HIV (K-RITH), [3]Emory Vaccine Center at Yerkes National Primate Research Center, Emory University, Atlanta, GA, USA and [4]Ragon Institute of Massachusetts General Hospital, MIT and Harvard, Boston, Cambridge, MA, USA

Transmission of cytotoxic T cell escape variants and the timing and frequency of CTL-mediated viral escape following acute HIV-1 infection can profoundly influence disease course, but comprehensive analysis of CTL epitopes restricted by protective HLA class I alleles is lacking. We evaluated the transmission of CTL immune escape variants and immune selection over one-year following acute HIV-1 infection for epitopes restricted by the B*7 supertype and protective HLA-B*57 and HLA-B*58: 01 alleles. HIV-1 uninfected women were screened twice weekly for HIV-1 RNA by finger prick blood draw. Six females were identified possessing the HLA-B*7 supertype [HLA-B*07 ($n = 2$), HLA-B*39: 01 ($n = 2$), HLA-B*42: 01 ($n = 1$), and HLA-B*42: 02 ($n = 1$)] while six women possessed the protective HLA-B*57: 02 ($n = 1$) and HLA-B*58: 01 ($n = 5$) class I alleles. Plasma samples were available at baseline (Fiebig I–III) and at two to six subsequent time points thereafter over one year of infection. Deep sequencing of near full-length HIV-1 genomes was performed using the Illumina MiSeq platform. Amplicons were molecularly barcoded, pooled, and sequenced resulting in >250-fold coverage. A subset of limiting-dilution full-length HIV-1 genome amplicons were also sequenced by PacBio at ~1,000-fold coverage. Sequence analysis was performed using Geneious v8.1.8 (Biomatters Ltd). In transmitted/founder viruses, the four known Gag CTL epitopes presented by HLA-B*57/B*58: 01 were wildtype in two of the six participants, with the remaining four participants showing evidence of CTL-mediated pre-adaptation in at least one epitope. By one-year post-infection, de novo CTL-mediated selection was observed in all six subjects, but never in all four epitopes. About five of six participants experienced escape within the immunodominant TW10 epitope. Of the five known Gag epitopes presented by the HLA-B*7 supertype, all six participants showed evidence of a CTL variant in at least three epitopes within the transmitted virus sequence. The TL9 epitope remained wild type in five of six participants with no evidence of CTL escape up to one-year post-infection. Transmitted escape variants remained virtually unchanged throughout the follow up period except in one participant who showed evidence of slow reversion in the HLA-B*57/58 ISW9 epitope. Deep sequencing reveals extensive transmission of pre-adapted CTL variants and slow immune selection within immunodominant epitopes restricted by protective HLA class I alleles.

A21 Evolvability of HIV-1 is influenced by codon pair usage

M. Nevot, M. Parera, G. Martrus and M. A. Martínez

Irsicaixa, Hospital Germans Trias i Pujol, Badalona, Spain

HIV-1 populations, like other RNA viruses, are established as a closely related mutant spectrum or cloud, which have an impact on virus evolvability, fitness, and virulence. The influence of codon choice in population diversity and evolvability of HIV-1 remains poorly explored. Here, we compared the development of HIV-1 resistance to protease inhibitors (PIs) of wild-type (WT) virus and a synthetic virus (MAX) carrying a codon-pair re-engineered protease sequence with thirty-eight (13 per cent) synonymous mutations. WT and MAX viruses replicated indistinguishably in MT-4 cells or PBMCs. To explore the evolvability of the codon pair re-coded protease, WT and MAX viruses were subjected to serial passages with the selective pressure of PIs [atazanavir (ATV) and darunavir (DRV)]. After the same number of successive passages in MT-4 cells in the presence of PIs, WT and MAX viruses developed phenotypic resistance to PIs (IC50 14.63 ± 5.39 nM and 21.26 ± 8.67 nM, for ATV; and IC50 5.69 ± 1.01 μM and 9.35 ± 1.89 for DRV, respectively). Sequence clonal analysis showed the presence, in both viruses, of previously described resistance mutations to ATV and DRV. However, a different resistance variant repertoire appeared in the MAX virus protease when compared to WT. The G16E substitution was only observed in the WT protease while the L10F, L33F, K45I, G48L, and L89I substitutions were only detected in the re-coded MAX protease population. The influence of the G48L mutation, which is extremely rare in vivo, on viral fitness was assessed by introducing this variant in the WT background. In the absence of drug, no differences on viral fitness were observed between the WT and MAX viruses carrying the G48L mutation. In order to detect minority variants in the quasispecies that could explain the emerge of G48L mutation, deep sequencing analysis using the Illumina MiSeq benchtop deep sequencer was performed. The implication of this mutation in the re-coded context is being explored considering that a particular sequence space can delineate the evolution of their mutant spectrum.

A22 Spatio-temporal history of the HIV-1 circulating recombinant form 35_AD in Afghanistan and Iran

S. Eybpoosh,[1,2] E. Mostafavi,[2] K. Azadmanesh,[3] and A.-A. Haghdoost[2,4]

[1]Research Centre for Emerging and Reemerging Infectious Diseases, Pasteur Institute of Iran, Tehran, IR Iran, [2]HIV/STI Surveillance Research Center, and WHO Collaborating Center for HIV Surveillance, Institute for Futures Studies in Health, Kerman University of Medical Sciences, Kerman, Iran, [3]Department of Virology, Pasteur Institute of Iran, Tehran, IR, Iran and [4]HIV/STI Surveillance Research Center, and WHO Collaborating Center for HIV Surveillance, Institute for Futures Studies in Health, Kerman University of Medical Sciences, Kerman, Iran

Circulating recombinant form 35_AD (CRF35_AD) has an important role in the epidemiological profile of Afghanistan and Iran. Despite the presence of this clade in Afghanistan and Iran for

over a decade, our understanding of its origin and dissemination patterns is limited. In this study, we employed a Bayesian phylogeographic approach to reconstruct the spatio-temporal dispersion pattern of this clade in Afghanistan and Iran for the first time. We performed a secondary data analysis on eligible HIV-1 CRF35_AD (gag and pol) sequences available in the Los Alamos HIV database (432 sequences available from Iran, 16 sequences available from Afghanistan, and a single CRF35_AD-like pol sequence available from USA). Sequences were excluded prior to analysis if they showed evidence of incorrect subtype assignment, frameshift, or drug resistance mutations, and/or stop codon positions. Subtype assignment was confirmed by maximum likelihood phylogenetic analysis. In order to reconstruct the spatio-temporal history of CRF35_AD, we used discrete Bayesian phylogeographic model in BEAST v1.8.1. Between-country viral dispersion rates were tested with Bayesian Stochastic Search Variable Selection method as implemented in SPREAD v1.0.7, and were considered as significant when Bayes factor values were >3. We checked the robustness of the key parameter estimates through a sensitivity analysis, using different priors and data subsets. The findings suggested that CRF35_AD sequences were genetically similar to parental sequences from Kenya and Uganda, and to a set of subtype A1 sequences available from Afghan refugees living in Pakistan. Our results also showed that across all phylogenies, Afghan and Iranian CRF35_AD sequences formed a monophyletic cluster (posterior clade credibility> 0.9). The divergence date of this cluster was estimated to be between 1990 and 1992. Within this cluster, a bidirectional dispersal of the virus was observed across Afghanistan and Iran. We could not clearly identify if Afghanistan or Iran first established or received this epidemic, as the root location of this cluster could not be robustly estimated. Three CRF35_AD sequences from Afghan refugees living in Pakistan nested among Afghan and Iranian CRF35_AD branches. However, the CRF35_AD-like sequence available from USA diverged independently from Kenyan subtype A1 sequences, suggesting that it may not be a true CRF35_AD lineage. The CRF35_AD viruses from Afghanistan, Iran, and Afghan refugees living in Pakistan seem to constitute a single epidemic, with multiple genetic exchanges among these populations. The date of onset for this epidemic (1990–1992) coincides with the rise of heroin production in Afghanistan (1970s). This highlights the potential role of drug trafficking in epidemic ignition in this region. Mass migration of Afghan refugees and illegal workers to Iran may be other possible contributors to among-country virus transmission.

A23 Large phylogenetic clusters highlight the HIV-1 epidemic in Canadian at risk populations

Rupert Capina,[1] Francois Cholette,[1] Eric Enns,[1] Christina Daniuk,[1] Tracy Taylor,[1] James Brooks,[1] P. Richard Harrigan,[2] and Paul Sandstrom[1]

[1]National Microbiology Laboratory at JC Wilt Infectious Diseases Research Centre, Winnipeg, MB, Canada and [2]BC Centre for Excellence in HIV/AIDS, Vancouver, BC, Canada

Since 1998, Canadian provinces and the Public Health Agency of Canada collaborated to monitor the country's HIV epidemic to form the Canadian HIV Strain and Drug Resistance Surveillance Program (SDR). The program traditionally reported incidence rates and prevalence of subtypes and drug resistance among newly diagnosed and not yet treated individuals. However, modern methods in phylogenetics have not yet been implemented in SDR. Here, we attempt to further characterize the HIV-1epidemic in Canadian populations experiencing high diagnosis rates by using phylogenetic clustering to deduce transmission dynamics. HIV-1 pol sequencing and genotyping was conducted on specimens submitted to the SDR program from treatment-naive individuals newly diagnosed with HIV from 2004 to 2013. REGAv3 and COMET were used for subtyping and Sierra, the Stanford Algorithm web service, was used for drug resistance genotyping. Phylogenetic clusters were inferred using patristic distances generated from bootstrap resampled trees (FastTree2). Statistical analyses were done using R. After quality filtering, 1,009 specimens were successfully sequenced and genotyped. Of those, 907 (89.9 per cent) formed a cluster of two or more. Overall, fifty-six clusters were inferred: three clusters ($n > 100$), six clusters ($n = 29–74$), seven clusters ($n = 8–15$), thirteen clusters ($n = 3–7$), and twenty-seven clusters ($n = 2$). We investigated predictors of clustering and found that people who self-reported inject drugs significantly clustered ($P = 0.006$), independent of other epidemiological variables such as age, sex, geographical regions, and risk behaviors. Phylogenetic clustering is a valuable tool to enhance HIV surveillance and monitoring efforts. Timely identification and investigation of clusters can inform focused prevention interventions. Effective use of HIV drug resistance genotype data for public health action will require revising information flows of the current provincial surveillance system building upon recommended clinical laboratory testing practices.

A24 Role of phylogenetic analysis in epidemiological case definitions during an outbreak of HIV-1 in people who inject drugs in Ireland

L. Dunford,[1] A. Waters,[1] M. Neary,[1] J. Dean,[1] C. Giese,[2] D. Igoe,[2] C. Hurley,[3] K. O'Donnell,[2] M. Fitzgerald,[3] and C. De Gascun,[1,4] the HIV Epidemiology Subgroup

[1]National Virus Reference Laboratory, University College Dublin, Dublin, Ireland, [2]HSE Health Protection Surveillance Centre, Dublin, Ireland, [3]Department of Public Health, Dublin, Ireland and [4]School of Medicine, University College Dublin, Dublin, Ireland

In 2015, an upsurge in acute HIV-1 subtype B infections was observed in Ireland among people who inject drugs, the majority of whom was homeless. The epidemiological investigation identified a significant association with injection of a new cathinone derivative, colloquially known as 'snow blow'. Concatenated HIV-1 polymerase and protease partial gene sequences (881 nucleotides; $n = 48$) were aligned with all subtype B patient sequences analysed at the NVRL from 2000 to 2015 ($n = 918$) and appropriate reference sequences using Bioedit v7.05. A neighbour-joining guide tree was constructed under a Kimura-2-Parameter model. Directed by this initial analysis, a maximum likelihood tree was constructed with a smaller number of more related Irish reference sequences ($n = 274$) under a HKY model of evolution and a gamma distribution. Statistical support was provided by bootstrapping with 1,000 replicates. Trees were constructed using PAUP* software version 4.0 beta10 and Mega version 7. The sequences from cases under investigation clustered within larger transmission networks of Irish people who inject drugs. More refined phylogenetic analyses confirmed that 79 per cent of the cases fell into four distinct clusters; cluster 1 ($n = 16$), cluster 2 ($n = 16$), cluster 3 ($n = 3$), and cluster 4 ($n = 3$), with high bootstrap support for each cluster (>75 per cent). There were ten outliers which branched outside the four clusters. The phylogenetic analysis largely supported the epidemiological investigation and the majority of epidemiologically linked cases were found to be contained within the same genetic clusters. In addition, this analysis identified two further possible cases and also eliminated three more 'cases'