

relationships between mice and retroviruses. In particular, we establish the presence of a more ancient ERV component in the murine genome, comprised of isolated, highly degraded insertions. These sequences evidence a transition in murine evolutionary history, beginning about one million years ago, wherein the ancient ERV families that have counterparts in humans and other large mammals were overtaken by a wave of newly acquired and/or transpositionally active ERVs.

A15 Rapid radiation of *treponema pallidum pertenu* in wild non-human primates

A. D  x,^{1,*} V.J. Schuenemann,² J.F. Gogarten,^{1,3,4} H.M. De Nys,^{1,5} K. Nieselt,⁶ M.A. Mayhew,⁷ F.H. Leendertz,¹ S. Calvignac-Spencer,¹ J. Krause,^{2,8,9}

¹Project Group Epidemiology of Highly Pathogenic Microorganisms, Robert Koch Institute, Berlin, Germany, ²Institute for Archaeological Sciences, Archaeo- and Palaeogenetics, University of T  bingen, T  bingen, Germany, ³Department of Biology, McGill University, Montreal, Canada, ⁴Max-Planck-Institute for Evolutionary Anthropology, Leipzig, Germany, ⁵UMI 233, Institut de Recherche pour le D  veloppement (IRD) and University of Montpellier, Montpellier, France, ⁶Center for Bioinformatics, University of T  bingen, T  bingen, Germany, ⁷Department of Science, Natural Resources and Outdoor Studies, University of Cumbria, Carlisle, UK, ⁸Max-Planck-Institute for the Science of Human History, Jena, Germany and ⁹Senckenberg Centre for Human Evolution and Palaeoenvironment, University of T  bingen, T  bingen, Germany

Bacteria of the species *Treponema pallidum* are causative agents of venereal syphilis (*Treponema pallidum pallidum*), Bejel (*T. p. endemicum*), and yaws (*T. p. pertenu*) in humans. We documented *Treponema pallidum* infections associated with disease in wild sooty mangabeys (*Cercocebus atys*) in Tai National Park, C  te d'Ivoire, and green monkeys (*Chlorocebus sabaues*) from Bijilo Forest Park, Gambia and Niokolo-Koba National Park, Senegal. To examine the evolutionary relatedness of these treponemes to those responsible for diseases in humans and for previously documented infections in baboons (*Papio papio*), we conducted a hybridization capture experiment to enrich *Treponema pallidum* DNA from samples collected from symptomatic individuals. This approach allowed us to sequence the full genomes of *Treponema pallidum* strains infecting sooty mangabeys ($n=2$) and green monkeys ($n=4$). Phylogenomic analyses revealed that all *Treponema pallidum* strains infecting non-human primates are most closely related to the sub-species *T. p. pertenu*. Strains infecting humans and non-human primates do not appear to be reciprocally monophyletic. The star-like phylogenetic branching pattern of the *T. p. pertenu* clade, with short basal branches receiving low statistical support, suggests a rapid initial radiation across humans and non-human primates. These results greatly broaden the known host range of *T. p. pertenu* and suggest the existence of a vast zoonotic reservoir that could possibly contribute to the failure of global eradication efforts.

A16 A distributed pan-viral typing framework

Michael Vilsker and Koen Deforche
Emweb, Herent, Belgium

With the increasing amount of DNA sequence information obtained from new sequencing methods, opening up the possibility for a complete viral screen of a host, there is an increasing need for the rapid and accurate identification of the virus types as well as their epidemiological background. While in the past, typing tools have been developed and made available (such as the Rega HIV-1 subtyping tool) and hosted at multiple sites, these tools require maintenance to track the ongoing evolution

of the virus. The setup and the maintenance of typing tools, which are often deployed at multiple sites, has been a challenge. Within the EU-funded VIROGENESIS project (Horizon 2020), the Rega typing tool framework is being redesigned to separate clearly the framework from the specifics for an individual tool. This will (1) enable an expert to independently and easily setup, create and maintain a typing tool for a new pathogen; (2) establish an online repository of typing tools/versions to which participants can push updates and from which up-to-date versions can be fetched to a distributed network of servers hosting the typing tools; and (3) create a pan-viral typing tool to identify the correct pathogen and which will allow further analysis of the sequences using the specialized typing tool for that pathogen. The transformation of the framework is expected to be completed by September 2016, at the same time co-evolving existing typing tools already available (including HIV, HCV, HTLV, Enterovirus, Norovirus), and new typing tools (including Chikungunya, Coronaviruses, Dengue virus, Zika virus) that are being designed by partners within the VIROGENESIS project. Within the VIROGENESIS project, the existing typing tools based upon the Rega typing tool framework will evolve into a distributed pan-viral typing tool.

A17 Molecular characteristics of hepatitis B virus (HBV) isolated from chronic hepatitis B patients in South Vietnam

Nguyen Hoa Trang,¹ Bui Thi Ton That,² Tran Thi Thanh Thanh,¹ Le Ngoc Chau,¹ Tran Tan Thanh,¹ Nghiem My Ngoc,² Nguyen Manh Hung,² Nguyen Van Vinh Chau,² Motiur Rahman,¹

¹Oxford University Clinical Research Unit, Wellcome Trust Major Overseas Programme, Ho Chi Minh City, Viet Nam and ²The Hospital for Tropical Diseases, Ho Chi Minh City, Viet Nam

Chronic infection with hepatitis B virus (HBV) is a major health problem due to its worldwide distribution and its long-term effects. Vietnam is country with a high HBV burden and the prevalence of chronic HBV infection in general population is 8.8–12.3%. In East Asia, the most common HBV genotypes are B and C. Despite high prevalence of HBV, data on HBV (genotype and subgenotype, virulence markers, drug resistance mutations and prevalence of recombinant strains) is limited in Vietnam. There are only few reports on HBV genotypes in Vietnam, mostly based on pre-S/S gene sequences. We have analyzed whole genome sequence of 98 HBV isolates from chronic HBV patients attending at Hospital for Tropical Disease, Ho Chi Minh City, Vietnam, who were under treatment for 1–6 years. HBV genome was amplified in 4 overlapping fragments (777–1,136 bp) and the amplicons were subjected to deep sequencing by using Illumina MiSeq system. Sequence assembly, genome analysis and phylogenetic analysis were performed within Geneious package. A sequence was assigned to a certain genotype and subgenotype if it was contained within a well-supported phylogenetic cluster (bootstrap value >75%) and the intra-genotypic nucleotide divergence was <7.5 and >4.5%, respectively. Mutations in Basal core promoter (BCP), pre-CORE, and CORE gene regions were determined by comparing with reference sequences. Finally, screening of minor (sub-consensus) variants was performed using the SNP detection tool available in Geneious. 1% frequency and 500-fold coverage were chosen as cut-off values. Among the isolates, 71.43% were genotype B, 27.55% were genotype C and one isolate was a recombinant (between B and C). Among genotype B isolates, 65 were subgenotype B4 (92.86%) and 5 were B2 (7.14%). 92.6% of subgenotype C belong to C1, 3.7% is subgenotype C2 and the remaining 3.7% to C3. Mutations G1752A, T1753C, G1757A, A1762G/T, G1764A and C1766G on BCP and CORE were found in 76 of 98

(77.55%) strains; and in 96.3% of subgenotype C (26/27) isolates. Stop codon mutation on Pre-Core, G1896A, was found in 23 of 98 patients (23.47%); among them, 20 of 23 (86.96%) are genotype B. Mutations related to antiviral drug resistance with frequency more than 1% in the reverse transcriptase (RT) domain of Polymerase gene were identified. 83.67% sequences have more than one mutation. 50% sequences have mutation at rtI169V/M, and all belong to sub-genotype B. Mutations at rtM204I/V and rtM250I were present in 23.47% and 11.22% of isolates (23 and 11 of 98, respectively), and was higher in sub-genotype C (P -values = 0.06 and <0.01, respectively). Our study provides high precision whole genome sequences of the Vietnamese HBV population and shows potential information related to treatment prediction for HBV infection as well as viral evolution. A further study with larger sample size and intensive analysis combined with clinical information promises more useful data for treatment and invention.

A18 Random amplification with next-generation sequencing to cover HIV and HCV full-length genomes

L. Cuypers,^{1,*} N. Conceição-Neto,¹ T. Dierickx,¹ Y. Schrooten,^{1,2} B. Vrancken,¹ E. Van Wijngaerden,^{3,4} F. Nevens,^{5,6} A.-M. Vandamme,^{1,7} J. Matthijnssens,¹ K. Van Laethem,^{1,2}

¹Department of Microbiology and Immunology, Rega Institute for Medical Research, Clinical and Epidemiological Virology, KU Leuven – University of Leuven, Leuven, Belgium, ²AIDS Reference Laboratory, University Hospitals Leuven, Leuven, Belgium, ³Department of Microbiology and Immunology, Clinical and Inflammatory Disorders, KU Leuven – University of Leuven, Leuven, Belgium, ⁴Department of General Internal Medicine, University Hospitals Leuven, Leuven, Belgium, ⁵Department of Clinical and Experimental Medicine, KU Leuven – University of Leuven, Leuven, Belgium Hepatology, ⁶Department of Gastroenterology and Hepatology, University Hospitals Leuven, Leuven, Belgium and ⁷Center for Global Health and Tropical Medicine, Microbiology Unit, Institute for Hygiene and Tropical Medicine, University Nova of Lisboa, Rua da Junqueira 100 Lisbon 1349-008, Portugal

Since both HIV and HCV are known to be highly variable and infections are treated with combination therapies that target different viral proteins, random amplification followed by next-generation sequencing could be an attractive alternative to Sanger sequencing for obtaining complete viral genomes. Moreover, one experimental approach could enable the simultaneous mapping of virus and human origin sequence reads, which is of particular interest for the detection of single nucleotide polymorphisms (SNPs) associated with disease progression and treatment response. A plasma sample obtained from a patient infected with HIV-1 recombinant subtype B and F1, and diagnosed with AIDS; and another sample from an HCV infected patient who progressed to fibrosis stage F1 were included in this pilot study. They were both characterized by a high viral load, i.e. >10M copies/ml for HIV and >3.5M IU/ml for HCV. Samples were directly extracted or pre-treated with homogenisation, centrifugation and filtration. Subsequently, purified nucleic acids were subjected to targeted (HIV-1 PR-RT and HCV NS3 protease) amplification and Sanger sequencing, or to random amplification followed by Illumina sequencing (~2M paired end reads attributed to each sample). Quality of the Illumina reads was checked using FastQC, followed by a *de novo* assembly of sample-specific contigs using the VICUNA and V-FAT software packages. In total, 0.3% ($n=5557$) and 0.02% ($n=421$) of the reads mapped to the HCV genome, respectively with or without virus enrichment steps. This sufficed for the *de novo* generated

contigs to cover 98.3 and 51.3% of the HCV genome with and without pre-treatment steps, respectively. Lower numbers of reads could be mapped for HIV-1 (821 versus 196 without virus enrichment), resulting in only 45.6 and 21.7% coverage of the full-genome. For both samples, the number of reads (<2M) was too low to map SNPs of interest within the human genome. Subject to adaptations in the protocol, such as improved enrichment strategies and deeper sequencing, this approach could be suitable for the surveillance of local HIV-1 and HCV epidemics. More extensive studies are required to investigate its usefulness in clinical research.

A19 Genetic markers for protease inhibitor drug resistance in regions outside of the protease gene

J. Giandhari,^{1,*} M. Gordon,² C. van Vuuren,³ P. Moodley,⁴ T. de Oliveira,¹

¹Africa Centre for Health and Population Studies, College of Health Sciences, University of KwaZulu-Natal, ²HIV Pathogenesis Programme, University of Kwa-Zulu Natal, Durban, South Africa, ³Military Hospital Division of Infectious Diseases, University of the Free State, South Africa and ⁴Department of Virology, National Health Laboratory Service, University of KwaZulu-Natal, Durban, South Africa

There are currently 6.8 million individuals infected with HIV-1 in South Africa. Nearly half of these individuals are on antiretroviral therapy. Protease inhibitors (PIs) form part of the South African national treatment guidelines as a first-line regimen for paediatric patients and a second-line regimen for adults. However, successful treatment is often hindered by the development of drug resistance. Resistance to PIs is characterised by a stepwise accumulation of mutations in the protease gene. However, studies have shown a low frequency of mutations in patients failing PIs. Regions outside of the protease, such as gag and more recently env, have been associated with PI drug resistance in the absence of protease mutations. We aim to identify genetic markers in the gag and envelope genes which are associated with PI treatment failure. Stored plasma samples from HIV-1-infected patients failing PI-based therapy ($n=500$) will be collected from collaborators in South Africa and whole genome sequences from PI-naïve patients will be downloaded from the Los Alamos Sequence Database. Failure will be defined as having two consecutive viral loads >1,000 cpm after being on a PI-based regimen for ≥6 months. The whole HIV-1 genome will be amplified from the PI failures and sequenced using Illumina Miseq. Sequences will be aligned to a reference and analysed for single nucleotide polymorphisms (SNPs) in all HIV genes using Geneious v 8.1.8 and GATK software application. Genome wide association analysis using PLINK will be performed to identify SNPs in PI treated patients that are associated with treatment failure. Episodic directional selection model such as MEDS and IFEL will be used to identify mutations that occur at specific amino acid positions and confer resistance to PIs. The previously mentioned methods allow the identification of drug resistance mutations without the need of baseline samples. Preliminary data from GWAS analysis on 26 patients failing PI-based therapy has identified codons in envelope and gag which are significantly associated with PI failure. Further examination of these sites in the viral minority population has shown that they are present at >2% of the viral population. By increasing our sample size we will obtain a more comprehensive and robust analysis of the role of regions outside the protease gene in PI resistance. Analysis of the control samples is currently being performed.