22nd International BioInformatics Workshop on Virus Evolution and Molecular Epidemiology

A1 Pairwise diversity and tMRCA as potential markers of HIV infection recency

Sikhulile Moyo,[1,2] Eduan Wilkinson,[3] Alain Vandormael,[3] Rui Wang,[4] Jia Weng,[4] Kenanao P. Kotokwe,[2] Simani Gaseitsiwe,[2,4] Rosemary Musonda,[2,4] Joseph Makhema,[2,4] Max Essex,[2,4] Susan Engelbrecht,[1,5] Tulio de Oliveira,[3,6,7] and Vladimir Novitsky[2,4]

[1]Division of Medical Virology, Stellenbosch University, Tygerberg, South Africa, [2]Botswana-Harvard AIDS Institute Partnership, Gaborone, Botswana, [3]Wellcome Trust Africa Centre for Health and Population Studies, Doris Duke Medical Research Centre, Nelson R Mandela School of Medicine, University of KwaZulu-Natal, Durban, South Africa, [4]Harvard T.H. Chan School of Public Health, Boston, MA, USA, [5]National Health Laboratory Services (NHLS), Tygerberg Coastal, South Africa, [6]Research Department of Infection, University College London, London, UK and [7]College of Health Sciences, University of KwaZulu-Natal, Durban, South Africa

Intra-host HIV-1 diversity increases linearly over time. We assessed the extent to which mean pairwise distances and the time to the most recent common ancestor (tMRCA) inferred from intra-host HIV-1 C env sequences were associated with the estimated time of HIV infection. Data from a primary HIV-1 C infection study in Botswana were used for this analysis ($N = 42$). A total of 2,540 HIV-1 C env gp120 V1C5 viral sequences were generated by single genome amplification and sequencing, with an average of sixty-one viral sequences per participant and eleven sequences per time point per participant. Raw pairwise distances were calculated for each time point and participant using the ape package in R software. The tMRCA was estimated using phylogenetic inference implemented in BEAST v1.8.2. Pairwise distances and tMRCA were significantly associated with the estimated time since HIV infection (both $P < 0.001$). Taking into account multiplicity of HIV infection strengthened these associations. HIV-1 C env-based pairwise distances and tMRCA can be used as potential markers for HIV recency. However, the tMRCA estimates demonstrated no advantage over the pairwise distances estimates.

A2 Optimization of the results generated by large-scale sequencing for the study of drug resistance in HIV infection: A systematic review

J. A. Fernández-Caballero,[1] N. Chueca,[1] E. Poveda,[2] and F. García[1]

[1]Department of Clinical Microbiology, Instituto de Investigación Biosanitaria (IBIS)-Complejo Hospitalario Universitario de Granada, Granada, Spain and [2]Division of Clinical Virology, Instituto de Investigación Biomédica de A Coruña (INIBIC)-Complexo Hospitalario Universitario de A Coruña (CHUAC), Sergas, Universidade da Coruña (UDC), Spain

Next-generation sequencing (NGS) approaches are now used in many clinical diagnostic laboratories for the routine diagnosis of resistance to antiretrovirals approved for the treatment of HIV infection. As some of NGS platforms may be a source of sequencing error, it is necessary to improve currently available protocols and implement bioinformatics tools that may help to correctly identify the presence of resistance mutations with clinical impact. In this study, we reviewed all studies dealing with software or methods aiming to decrease these errors, published during the period 2006–2016. We considered, as bioinformatic strategies, software aiming to delete or detect sequencing errors, and as protocol improvements, those changes in PCR temperature profiles and/or reagent concentration aiming to minimize sequencing errors. We used a combination of non-MeSH and MeSH terms related to error correction and NGS sequence filtering. All abstracts of papers available through January 2006 and June 2016 were reviewed. Our search identified 611 studies, we finally selected seven papers that met all the eligibility criteria, three of which dealt with protocol modifications and four with bioinformatics aiming to eliminate errors. Some studies are mainly focused on improving protocols for decreasing the magnitude of errors during the polymerase change reaction. Other studies propose specific bioinformatics tools, able to reach both a 93–98 per cent reduction of indels (insertions/deletions) and a sensitivity and specificity close to 100 per cent in single nucleotide polymorphism variant calling. Moreover, error rates decreased from a median value (95% CI) of 0.2 per cent (0.008–0.4) before processing to 0.06 per cent (0.05–0.08) after using a bioinformatic tool. All the software did not incur in a high loss in the number of reads. New protocols and bioinformatics tools that improve the accuracy of NGS results must be considered for correct analysis of HIV resistance mutations. We recommend using bioinformatic software to filter short and low-quality sequences, and using high fidelity polymerases.

A3 Brazilian network for HIV drug resistance surveillance: An investigation of pre-treatment drug resistance transmission chains

N. M. C. Véras,[1] A. F. Pires,[1,2] A. Tanuri,[3] and A. S. Benzaken[1]

[1]Ministério da Saúde, Secretaria de Vigilância em Saúde, Departamento de Vigilância, Prevenção e Controle das Infecções Sexuais Transmissíveis, HIV/Aids e Hepatites Virais, Setor Administrativo Federal Sul (SAF Sul), Trecho 2, Bloco F, Ed. Premium, Torre I, Auditório, Brasília-DF, CEP 70.070-600, Brazil, [2]Programa de Pós-Graduação em Saúde Coletiva, Faculdade de Medicina, Faculdade de Ciências de Saúde, Universidade de Brasília (UnB), Brasília-DF, Brazil and [3]Laboratório de Virologia Molecular, Departamento de Genética, Instituto de Biologia,