Complete Genome Sequence of a Reference Stock of Simian Immunodeficiency Virus RNA (SIVmac251/32H/L28) Determined by Deep Sequencing

Adrian Jenkins, Claire Ham, Neil Almond, Neil Berry
Division of Virology, National Institute for Biological Standards and Control, South Mimms, United Kingdom

A reference preparation for simian immunodeficiency virus (SIV) RNA nucleic acid assays was characterized by complete genome deep sequencing. The entire coding sequence and flanking long terminal repeats, including minority species, were determined. This information will inform SIV research investigations and aid evaluation and development of amplification assays for SIV RNA quantification.

Model systems to evaluate putative vaccine strategies and pathogenesis studies of HIV/AIDS employ strains of simian immunodeficiency virus (SIV) in experimental design. In cynomolgus macaques (Macaca fascicularis), a high-titer bulk SIV positive plasma stock recovered at the peak of virus replication in vivo was developed into an SIV RNA reference panel (C1–C5), based on the SIVmac251/32H/L28 isolate (1, 2). These represent commutable materials for determination of SIV viral load (2). The genetic composition of the highest viral concentration member of this panel (C1) was determined by deep sequencing analysis. The genetic composition of this SIV RNA reference material has not been described; here, we report the previously undescribed complete genome sequence using deep sequencing approaches, which are analogous to those previously reported for HIV-1 and HIV-2 (3, 4).

The C1 preparation represents a 1/100 dilution of undiluted plasma, from which viral RNA was extracted using QiaAmp viral RNA minikits (Qiagen). Reverse transcription and amplification reactions were performed with the SuperScript III One-Step RT-PCR system with Platinum Taq HiFi (Invitrogen) using primer pairs designed to generate five overlapping amplicons. Primers were designated as follows: 1F: CAGTGGCTCTGGGAGAGCGTGCGTGG; 1R: CAAGCTTAGTACTACTGCTGCTCTCCTCAGAGAGAG; 2F: TAGCTGTCTTTATCCAGGAAG; 2R: TCTAATTAACCTATGGC; 3F: AAAATTGAAGCAGTGGCCATTAT; 3R: TACTTATGAGCTCTCGGGAACCT; 4R: GCGAGAAAACCCAAGAACCCTAG; 4F: GGCATAGCCTCAGAGATGTTTGG; 5F: CAGTCGCTCTGCGGAGAGGC.

The complete genome sequence of the C1 SIVmac251/32H/L28 reference used in this study is 10,276 bp long with a G+C content of 44.0%, representing the complete coding sequence with nine open reading frames (vif, vpr, vpx, tat, rev, env, pol, nef). This is the first report of the complete genome sequence of the SIVmm32H reference (D01065), and a consensus sequence was generated. All reads were mapped against the consensus to establish read depth and minority species.

The sequence of the SIV RNA reference material described here is 10,276 bp long with a G+C content of 44.0%, representing the complete coding sequence with nine open reading frames (gag, pol, vif, vpr, tat, rev, vpx, env, and nef), as well as the U5 region of the 5’-long terminal repeat (5’-LTR) and U3 and R regions of the 3’-LTR of the SIV genome. BLAST analysis (5) of the consensus RNA template-derived sequence further confirmed the highest similarity with the SIVmm32H proviral clone sequence (4) (total score 18,384, 98% identity, and 100% coverage). The mean read depth was 74,051.7-fold (± 87,072.4 standard deviation) with a minimum of 16,285-fold. There are 39 positions with a minority nucleotide differing from the reference sequence with a frequency >3% and a base call accuracy of 99.9% (>Phred 30). In addition, there is one insertion event and one deletion event.

This is the first report of the complete genome sequence of the SIV RNA reference panel (CFAR catalogue no. 100092) derived from a viral RNA template, which will augment the development and evaluation of SIV RNA amplification assays applied in SIV vaccine and pathogenesis studies.

Nucleotide sequence accession number. The complete genome sequence of the C1 SIVmac251/32H/L28 reference used in the reference panel reported here has been deposited in GenBank under the accession number KU892415.

FUNDING INFORMATION
This work, including the efforts of Neil Almond and Neil Berry, was funded by Medical Research Council (MRC) (G060007). This work was funded by U.K. Medical Research Council grant G060007 to N. Berry, N. Almond, D. Ferguson, and R. Stebbings. The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.
REFERENCES

1. Rud EW, Cranage M, Yon J, Quirk J, Ogilvie L, Cook N, Webster S, Dennis M, Clarke BE. 1994. Molecular and biological characterization of simian immunodeficiency virus macaque strain 32H proviral clones containing nef size variants. J Gen Virol 75:529–543. http://dx.doi.org/10.1099/0022-1317-75-3-529.

2. Ham C, Srinivasan P, Thorstensson R, Verschoor E, Fagrouche Z, Sernicola L, Ramos A, Titti F, Almond N, Berry N. 2010. International multicenter study to assess a panel of reference materials for quantification of simian immunodeficiency virus RNA in plasma. J Clin Microbiol 48:2582–2585. http://dx.doi.org/10.1128/JCM.00082-10.

3. Gall A, Morris C, Kellam P, Berry N. 2014. Complete genome sequence of the WHO International Standard for HIV-1 RNA determined by deep sequencing. Genome Announc 2(1):e01254-13. http://dx.doi.org/10.1128/genomeA.01254-13.

4. Ham C, Morris C, Berry N. 2016. Complete genome sequence of the WHO International Standard for HIV-2 RNA determined by deep sequencing. Genome Announc 4(1):e01626-15. http://dx.doi.org/10.1128/genomeA.01626-15.

5. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol 215:403–410. http://dx.doi.org/10.1016/S0022-2836(05)80360-2.