Ann Ist Super Sanità 2022 | Vol. 58, No. 1: 1-5
www.iss.it/annali

Supplementary Materials for

First detection of SARS-CoV-2 lineage A.27 in Sardinia, Italy

Alessandra Lo Presti, Ferdinando Coghe, Angela Di Martino, Sara Fais, Riccardo Cappai, Manuela Marra, Maria Carollo, Marco Crescenzi, Germano Orrù, Giovanni Rezza and Paola Stefanelli

Corresponding author:
Alessandra Lo Presti, Dipartimento Malattie Infettive, Istituto Superiore di Sanità, 00161 Rome, Italy. E-mail: alessandra.lopresti@iss.it.

Published on
Ann Ist Super Sanità 2022 Vol. 58, No. 1: 1-5
DOI: 10.4415/ANN_22_01_01

This PDF file includes:
Details on methodology

DETAILS ON METHODOLOGY

Whole-genome sequencing analysis

- The Ion AmpliSeq™ SARS-Cov-2 Research Panel consists of two primer pair pools that target 237 amplicons specific to the SARS-Cov-2 coronavirus (ThermoFisher Scientific). The panel, with the amplicon length range of 125-275 bp, provides 99% coverage of the SARS-Cov-2 genome and covers all potentials serotypes.
- The pooled library were clonally amplified on Ion Sphere™ Particles (ISPs) in an emulsion PCR using the Ion 520-530™ Kit-OT2 (Thermo Fisher scientific). Enrichment of positive Ion Spheres (ISPs) was performed using DynaBeadsMyOne streptavidin C1 beads (Invitrogen). Purified templated ISPs were then loaded onto the Ion semiconductor 530 chip and sequenced using the Ion S5 System. (Thermo Fisher scientific). Results were finally analyzed using The Ion Torrent Suite (Thermo Fisher scientific).
- The IRMA (Iterative Refinement Meta-Assembler) Assembly method is routinely used to process genome sequence data derived from the large volume of surveillance specimens (such as those characterized at CDC) (https://wonder.cdc.gov/amd/flu/irma/irma.html). IRMA was designed for the robust assembly, variant calling, and phasing of highly variable RNA viruses. Currently, IRMA is deployed with modules for influenza, ebolavirus and coronavirus (https://wonder.cdc.gov/amd/flu/irma/irma.html). The IRMA method works by the iterative optimization of read gathering and assembly. As with all reference-based assembly, reads are included in assembly when they match consensus template sets; however, IRMA provides for on-the-fly reference editing, correction, and optional elongation without the need for additional reference selection, and focuses on quality control, error correction, indel reporting, variant calling and variant phasing. This increases both read depth and breadth. Additional details are reported in Shepard S.S. et al., 2016 - Viral deep sequencing needs an adaptive approach: IRMA, the iterative refinement meta-assembler. BMC Genomics volume 17, Article number: 708 (2016).

Phylogenetic analysis

- Multiple sequence alignment is an important step in comparative and evolutionary analyses of sequences and in phylogenetic inference. MAFFT (Multiple Alignment using Fast Fourier Transform) is a high speed and high performance multiple sequence alignment program. MAFFT assume that the input sequences are all homologous, that is, descended from a common ancestor. The complete methodology is described in: Kazutaka Katoh and Daron M. Standley. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. Mol. Biol. Evol. 30(4):772-780 doi:10.1093/molbev/mst010.
- After obtaining the sequence alignment, it was manually edited with the Bioedit software (Hall, T.A. 1999. BioEdit: A User-Friendly Biological Sequence Alignment Editor and Analysis Program for Windows 95/98/NT. Nucleic Acids Symposium Series, 41, 95-98).
- The edited alignment has been analysed to choose the evolutionary model that best fitted the data (nucleotide substitution model selection) and to elaborate the phylogenetic tree through IQ-TREE.
- IQ-TREE is an efficient and versatile phylogenetic software for maximum likelihood analysis of large phylogenetic data, which can determine both the best-fit substitution model for the data and the tree reconstruction (Minh B.Q. and Schmidt H.A., 2020. IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. Mol. Biol. Evol., 37:1530-1534). Support for the tree topology and clades was estimated with the bootstrap test (1000 bootstrap replicates).