Omar Ahmed Mahmoud

Third year [Medical informatics Department]

Decoding Covid-19 with the SARS-CoV-2 Genome Analysis and

Visualization: Development and Usability study

Method

The SARS-CoV-2 Spike Protein Determines Cell and Tissue Tropism

The S ORF of human CoVs encodes the critical S protein that covers the surface of the viral particle and facilitates entry into cells. The S protein sequence has high homology to that of RaTG13 (94.9%) and also to SARS-CoV (84.74%); therefore, the SARS-CoV-2 S is well-adapted to bind human ACE2. SARS-CoV-2 has high affinity for human ACE2 which is the entry receptor. Homology with the remaining human CoVs that do not bind ACE2 and use a variety of different entry receptors is not surprisingly lower. S is a large protein of over 1200 amino acids that can be broken down into two subunits S1 and S2 by host cell enzymatic cleavage. The S1 subunit is the most external region and determines interaction with ACE2 through its RBM. Alignment of the SARS-CoV-2, SARS-CoV and RaTG13 RBM demonstrate homologous regions that specify interaction with ACE2. To infect cells, SARS-like CoVs must first bind with ACE2 on the surface of cells followed conformational changes which precedes membrane fusion events and uptake of the virus into cells. This interaction between SARS-CoV-2 and ACE2 defines the cells and tissues that the virus can infect and subsequent pathology.

SARS-CoV-2 Spike ORF Contains a Unique Insertion Encoding a Furin Cleavage Site

After receptor interaction through the S glycoprotein, CoVs utilise diverse host cell proteases for cleavage activation of virus-host cell membrane fusion and subsequent genome delivery. Inspection of the SARS-CoV-2 genome reveals the presence of a furin cleavage site at the S1/S2 junction of the S protein that is not found in other beta-CoVs of

any species. This is achieved by a 12-base insertion (cctcggcgggca) encoding PRRA amino acids. Protease cleavage of the S protein is necessary for initiating host cell invasion after ACE2 attachment and is usually achieved by the S2' cleavage site that is catalyzed by the host cell serine protease TMPRSS. Dual-protease cleavage of SARS-CoV-2 spike would therefore provide a fusion and entry advantage over that of related human CoVs, potentially priming S for optimal conformation and entry receptor interaction that could enhance replication and transmission. This concept is not without precedent, as manipulation of a porcine CoV enabled researchers to alter the trypsin-dependent protease cleavage site to be activated instead by furin, which enhanced infectivity of target cells. Although this suggests furin cleavage can be synthetically engineered, a separate study of proteolytic cleavage in the MERS-CoV spike protein emphasises that MERS-CoV is capable of adapting to various conditions and cleavage can be activated by either trypsin or furin proteases. Furin cleavage motifs were identified at S1/S2 and S2' in MERS both by sequence alignment and a furin cleavage prediction algorithm. This adaptability contradicts theories of genetic manipulation of the spike protein cleavage site, as SARS-like CoVs may have already been readily capable of naturally switching to furin-activated cleavage.

Result

SARS-CoV-2, a recently emerged CoV responsible for the current global pandemic Covid-19, first appeared in late 2019 in Wuhan, China. The virus is thought to have arisen from SARS-like CoVs in bats due to high similarities in genome sequence which are also shared with the prior SARS-CoV. The SARS-CoV-2 genome retains many features of endemic human CoVs but the critical determinant, the S protein, is sufficiently adapted to bind the human entry receptor ACE2 much more readily than SARS-CoV which is the most closely related human CoV. There is evidence that the viral genome is undergoing subtle evolution through mutation to enhance transmission and there is evidence for limited

attenuation that might weaken the virus. The scientific and medical community has mobilised in an unprecedented fashion to understand the virus at molecular and epidemiological levels, determine its pathological consequences, understand protective immunity and develop specific antivirals, vaccines and other treatments. However, as of writing, the pandemic is yet to be fully under control although some territories have had success. Currently, the only effective measures to restrict viral transmission are limiting social interactions, mass diagnostic testing and contact tracing applications. Further understanding of the genetics behind the virus' emergence and mechanisms of replication will be critical to generate specific therapeutics and protective vaccines to halt the continued spread. This knowledge will also assist and limit the future potential for new CoVs to emerge..