We discussed four specific features of 2019-nCoV:

1. Highly mutated receptor binding domain (RBD), including around key residues when comparing 2019-nCoV to a highly related bat SARS-like CoV (RaTG13; 96% identity).
    a. Nothing unusual. Comparisons were made between between SARS and SARS-like CoVs that were of similar divergence as nCoV and RaTG13. Similar levels of diversity were observed in the RBD, showing that this domain in general is highly variable, which is likely due to strong positive selection for receptor binding.
2. Reversion of gain-of-function site in the RBD to that seen in SARS.
    a. Nothing unusual. The reversion of F (observed in RaTG13) to Y (observed in SARS and nCoV) is only a single base-pair change (A > T transversion). Given that the RBD is variable in general, this is not unusual.
3. Gain of BamHI restriction site in the 3' end of the spike protein of nCoV. Sequence upstream of site is somewhat variable and sequence after is conserved.
    a. Probably not unusual. The gain of the BamHI site in nCoV is the result of a single synonymous transition (T > C) that happens frequently in RNA viruses. The 3' sequence following this site is conserved not only between nCoV and RaTG13, but also more broadly across similar viruses. The site could be used to insert different versions of the spike protein gene into nCoV, but no specific data suggest that it is utilized as such.
4. Gain of furin cleavage site and O-linked glycans.
    a. Unclear. This is the first time an optimal furin cleavage site has been observed in a betacoronavirus and it is additionally coupled to a gain of O-linked glycans. Several different scenarios could explain how this was gained:
        i. Natural selection, plausibly in a non-bat reservoir / intermediate host.
        ii. Repeated passage of virus in tissue culture.
        iii. Specific engineering of the site.

In summary, after considering all things above, the only thing that remains perplexing about 2019-nCoV is the fact that it has a furin site with O-linked glycans in the spike protein between S1 and S2. It is impossible to distinguish whether this was gained due to e.g., evolution or passage, and the data is consistent with either scenario. Specific engineering is also a possibility, but appears less likely as that would require significant amounts of molecular work utilizing an uncommon virus backbone.

Below we briefly outline different scenarios for how nCoV may have originated.

1. Bioweapon. Highly unlikely and there is no data supporting this hypothesis.
2. Specific engineering. Unlikely as this would require significant amounts of work utilizing uncommon and currently unknown backbones of SARS-like bat CoVs. For this type of work, there are preexisting backbones that could have been utilized, but they clearly were not.
3. Tissue culture passage. The data is consistent with this scenario, although no specific hypothesis exists for how the furin site was gained, but could be due to passage in tissue culture. The virus could have been released via accidental infection of researcher(s).
4. Spillover from animal host. The data is consistent with this scenario, although no specific hypothesis exists for how the furin site was gained in an animal host. However, even though the furin site has

not been observed in these viruses previously, virus evolution coupled with strong selective pressure (possibly in an intermediate host) would be capable of creating such a domain.