

Coronavirus Diversity, Phylogeny and Interspecies Jumping

Patrick CY Woo

The SARS epidemic has boosted interest in research on coronavirus biodiversity and genomics. Before 2003, there were less than 10 coronaviruses with complete genomes available. After the SARS epidemic, up to August 2015, there was an addition of more than 40 coronaviruses with complete genomes sequenced. These include two human coronaviruses (human coronavirus NL63 and human coronavirus HKU1) and at least 30 other mammalian and avian coronaviruses. Two novel lineages in *Betacoronavirus* and a novel genus (*Deltacoronavirus*) have been discovered. The diversity of coronaviruses is a result of the infidelity of RNA dependent RNA polymerase, high frequency of homologous RNA recombination and the large genomes of coronaviruses. Among all hosts, the diversity of coronaviruses is most evidenced in bats and birds, which may be a result of their species diversity, ability to fly, environmental pressures and habits of roosting and flocking. In general, bat coronaviruses are the gene pools of *Alphacoronavirus* and *Betacoronavirus* lineages B, C and D whereas bird coronaviruses are the gene pools of *Gammacoronavirus* and *Deltacoronavirus*. However, recent evidence also suggested that rodent coronaviruses could be the gene pools of *Betacoronavirus* lineage A. With the increasing number of coronaviruses, more and more closely related coronaviruses from distantly related animals have been observed, which were results of recent interspecies jumping and may be the cause of disastrous outbreaks of zoonotic diseases.