

# Unanticipated effects of bad modeling assumptions

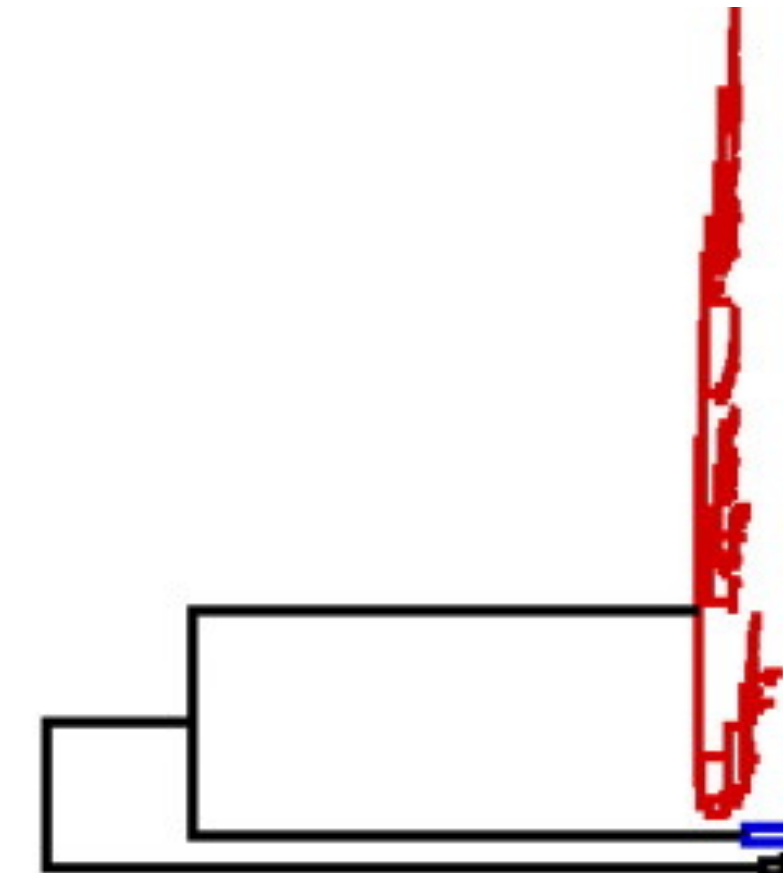
- Models that fail to account for significant shifts in selective pressures through lineages also significantly underestimate branch lengths
- An instructive example is long-range molecular dating of pathogens, where recent isolates (e.g., 30-50 years of sampling) are used to extrapolate the date when a particular pathogen had emerged
- This creates the situation when terminal branches in the tree have relatively high dN/dS (within-host level evolution), which deep interior branches have very low dN/dS (long term conservation)

- Using models that do not vary selection pressure across lineages yields a patently false “*too young*” estimate for the origin of **measles** (about 600 years ago)
- This estimate is refuted by clear historical records which suggest

that measles is at least 1,500-5,000 years old

- *This includes a treatise by a Persian physician Rhazes about **differential diagnosis of measles and smallpox** published circa 600 AD.*
- Same patterns found for corona-viruses, ebola, avian influenza and herpesvirus

A GTR +  $\Gamma_4$



B Lineage+Dual (two rate)

