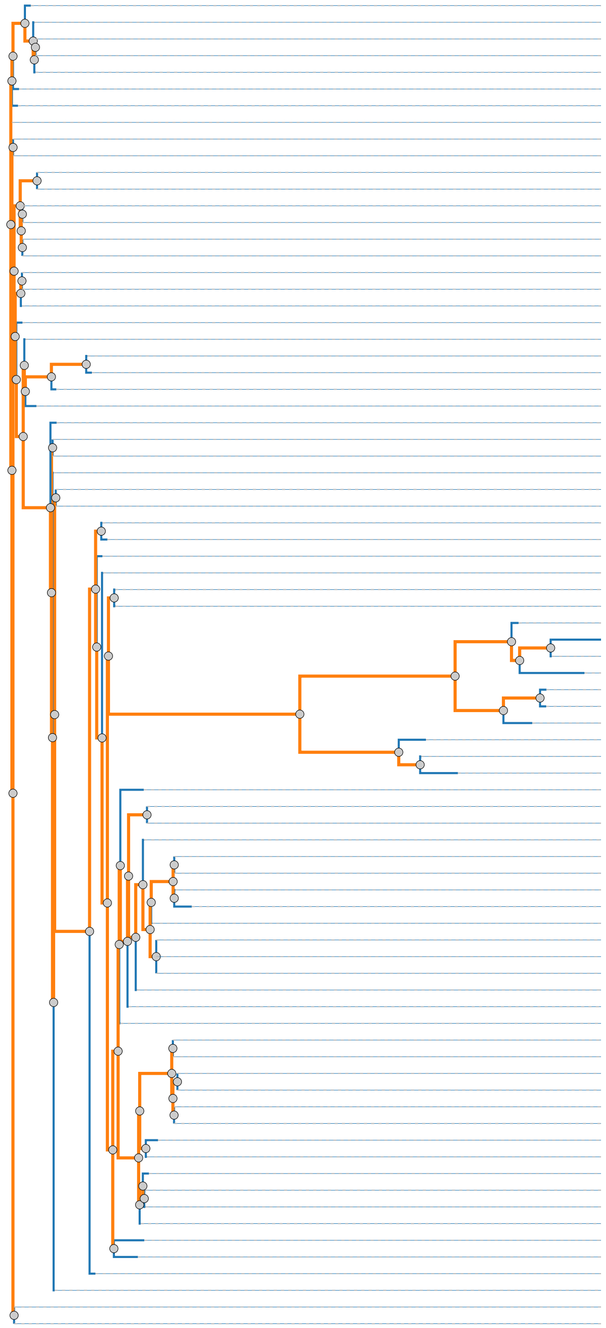
Interpreting dN/dS for intra-host and intra-species pathogen

- **dN/dS** can be estimated for all sorts of sequence data (e.g., it has been done for cancer SNP data)
- Traditional interpretation of dN/dS is based on the assumption that **substitution ~ fixation**
- Not the same for intra-species / intra-host pathogens
  - Much of variation is due to polymorphism, or even dead-end mutations
- This is because selection has not had a chance to “filter” mutations (except for patently deleterious ones)
- This often manifests as differences in selective “regimes” between tips and internal branches



- Partition a pathogen tree into **terminal** and **internal** branches
- **Terminal branches** potentially include “dead-end” lineages, i.e. those which are maladaptive
- **Internal branches** include at least one “*transmission*” (intra-species) or “*replication*” (intra-host) events: stronger action of selection
- Focusing on a subset of branches can allow one to interpret dN/dS more precisely