Replication files:

- R file (packages version: Sept 2017): runs sequence analysis
    - seqsubm: compute the optimal matching distances using substitution costs based on transition rates observed in the data and indel cost equal to 1 (see Appendix to see how optimal matching works)
    - seqdist: store results in distance matrix
    - hclust: hierarchical clustering ward method
    - as.clustrange: compute several clustering quality measures for a range of numbers of groups
    - summary: retrieve the best solutions from the function above according to each quality measure
    - seqdplot: plot the state distribution at each time point within each cluster
    - dissmfacw: runs a multi-factor analysis of variance from a dissimilarity matrix
- R log files: show results of sequence analysis
- Stata do file: runs multinomial analysis with clusters from sequence analysis
- Stata log file: shows results from multinomial analysis
- Stata datasets: datasets created with R
- Excel dataset: original dataset