



Figure S2. Clustering of nearest neighbour “species” by Unifrac distances. Sample label code: 1<sup>st</sup> letter represents sample designation (M = non-STAT; S = STAT); 2<sup>nd</sup> and 3<sup>rd</sup> letters represent specimen type (No = nasal swab, NS = nasopharyngeal swab, NW = nasopharyngeal washing, TA = tracheal aspirate); 4<sup>th</sup> number represents age group of patient (1 = 0.01-0.50 years, 2 = 0.51-3.0 years, 3 = 3.01 - 9.0 years, 4 = 9.01-20.0 years, 5 = 20.01-50.0 years, 6 = 50.01-65.0 years, 7 = >=65.01 years); Flu## represents the sample ID.