cd hit cluster 56/SAUSA300_1750/NE703/hypothetical protein 2 independent mutations 1798L Y601N M599fs