cd hit cluster 435/SAUSA300_1991/NE873/hypothetical protein 2 independent mutations 1139fs I45fs V392fs