cd hit cluster 60/hypothetical protein 2 independent mutations I12fs N778I