Confounding factors:
Many cases in which the different clusters are not formed by diff PDBs, but rather by diff CHAINS. Ex:
clst 1: 1BD3_A 1BD3_C 1BD4_A 1BD4_C
clst 2: 1BD3_B 1BD3_D 1BD4_B 1BD4_D

Cases in which multiple chains from one PDB are the elements that form the entire sequence group

More common: cases in which different clusters do not really differ biologically (manual annotation).

Overlap with ASD:
num_eligible_predicted_allo_pdbs: 976
num_matches_to_ASD: 131
total num entries in PDB: 82114 (as of 2012-06-15)
total num ASD PDBs: 9611
Fract of asd in pdb: 11.7%
num_matches_to_ASD: 131
num_EXPECTED_matches_to_ASD: 114 = (9611/82114) * 976