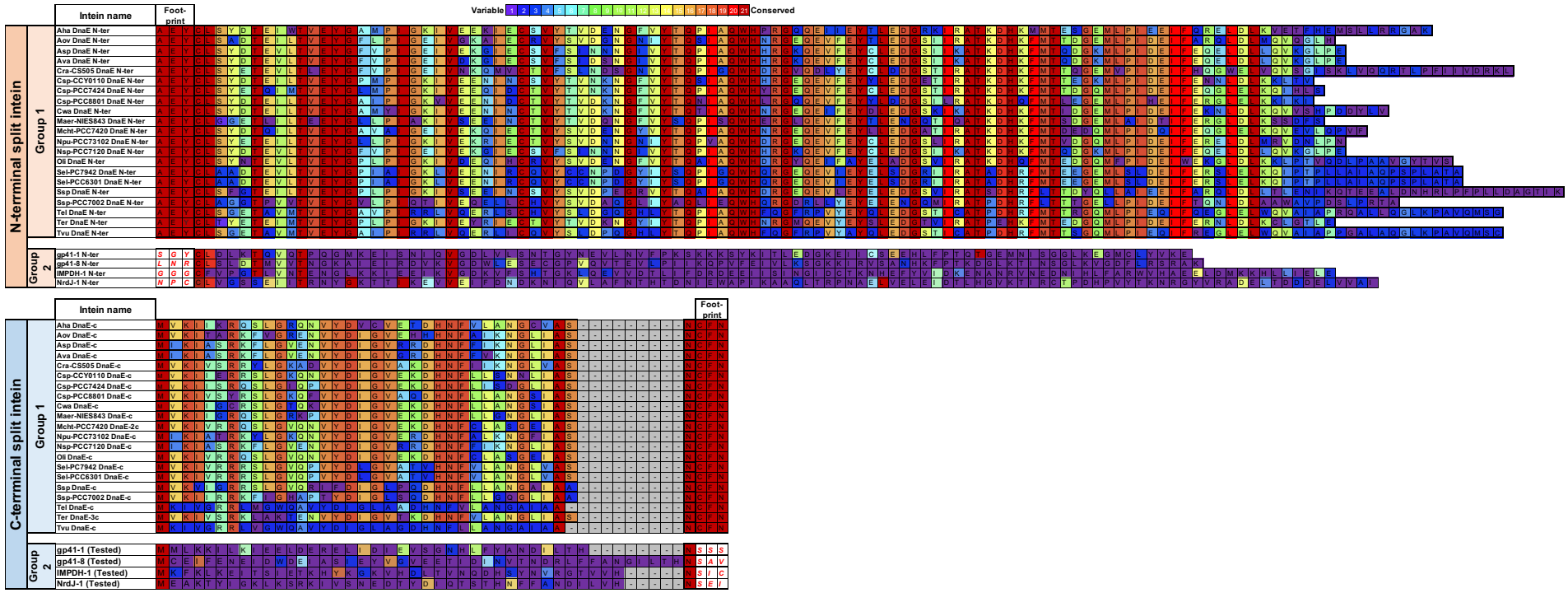

Extended data Table 1: Protein sequence of the 23 split inteins with amino acid conservation.
The intein library was divided into two subgroups. Group 1 contains 19 split intein pairs that share high sequence homology including the 6-residue footprint (i.e AEY and CFN at respectively the N- and the C-terminal end). Group 2 was composed of four candidates that are unique with 6-residue foot-prints for each pair (i.e. SGY/SSS for gp41.1, LNR/SAV for gp41.8, GGG/SIC for IMPDH and NPC/SEI for Nrdj1).