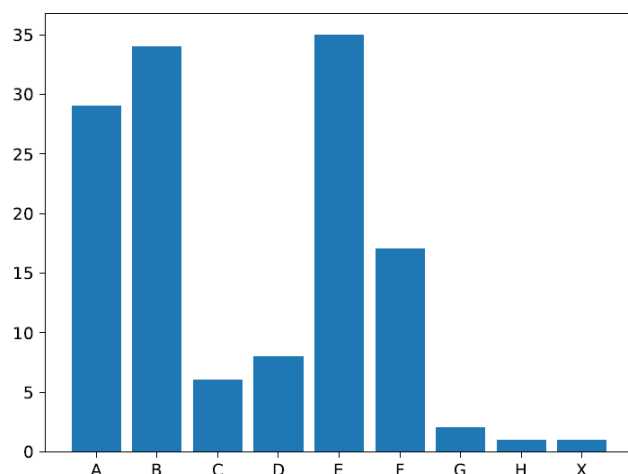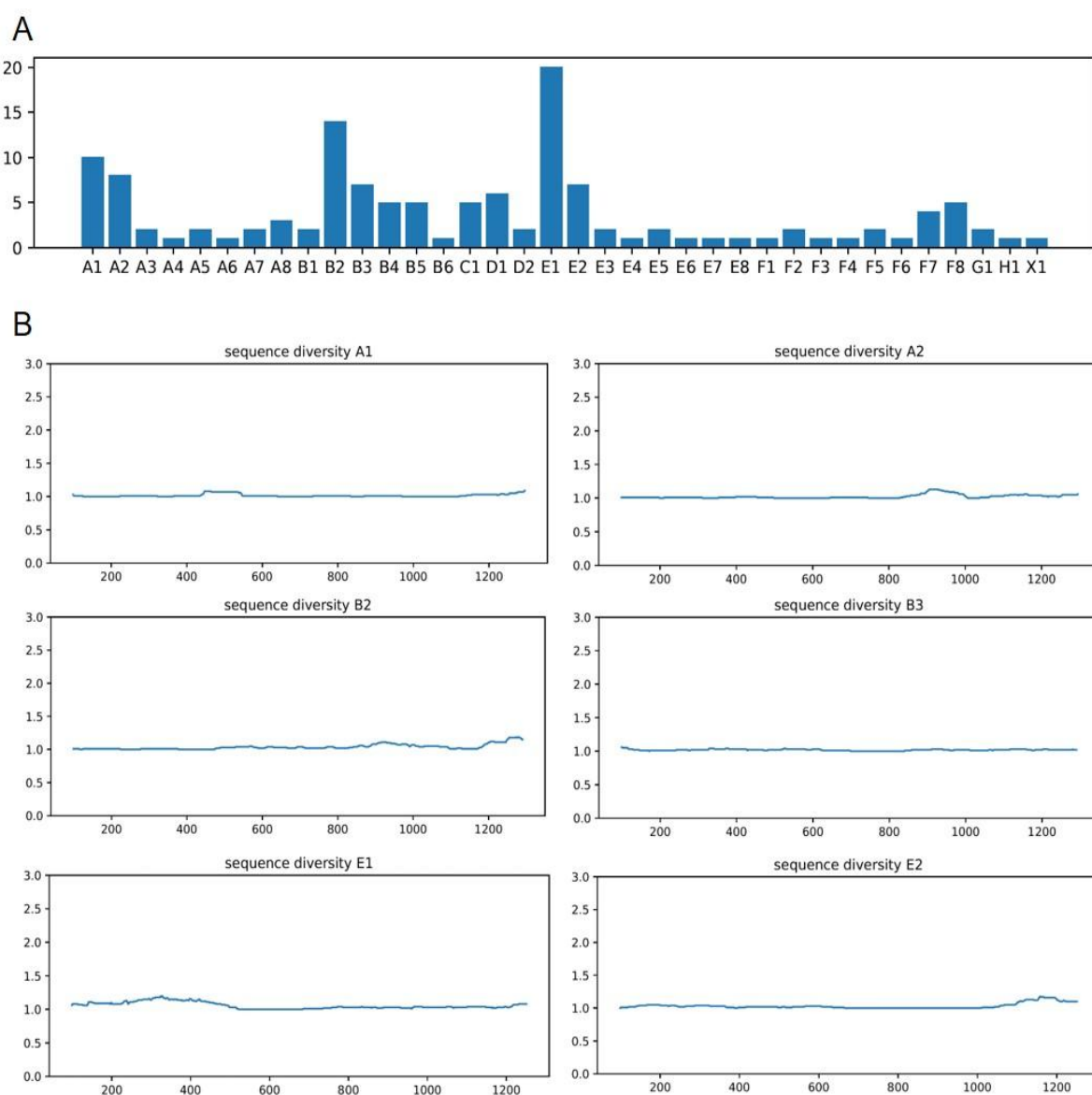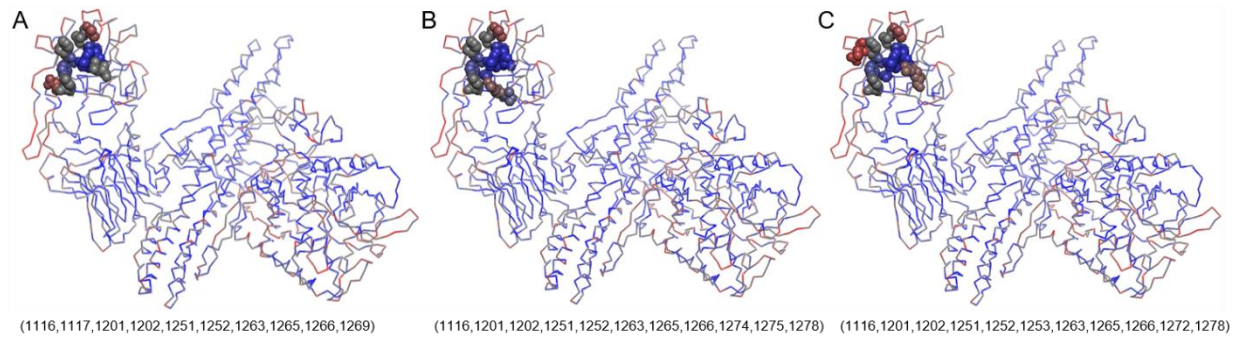**Supplementary Materials**

Figure S1. Number of *bont* genes collected for each type. The types C, D, G, H and X had < 7 sequences and were thus not considered for further comparison analysis.

Figure S2. The diversity index at subtype level. Subtypes were generated by protein sequence clustering at 97%. (A) shows the number of protein sequences of each subtype. (B) shows the diversity index, counting the

1

number of unique amino acids of each column in the multiple alignment, of each subtype. The moving average (n = 100) of the diversity index were shown for an overview of the sequence variation.

(1116,1117,1201,1202,1251,1252,1263,1265,1266,1269)   (1116,1201,1202,1251,1252,1263,1265,1266,1274,1275,1278)   (1116,1201,1202,1251,1252,1253,1263,1265,1266,1272,1278)

Figure S3. Interacting sites of BoNT with ganglioside receptors. The highlighted spheres corresponds to the interacting sites with ganglioside (A) GD1a (in 5PTC), (B) GT1B (in 2VU9), (C) GD1a (in 7QPT). The interacting sites were mapped onto the BoNT structure from PDB database (3BTA). The amino acid residue numbers were displayed.