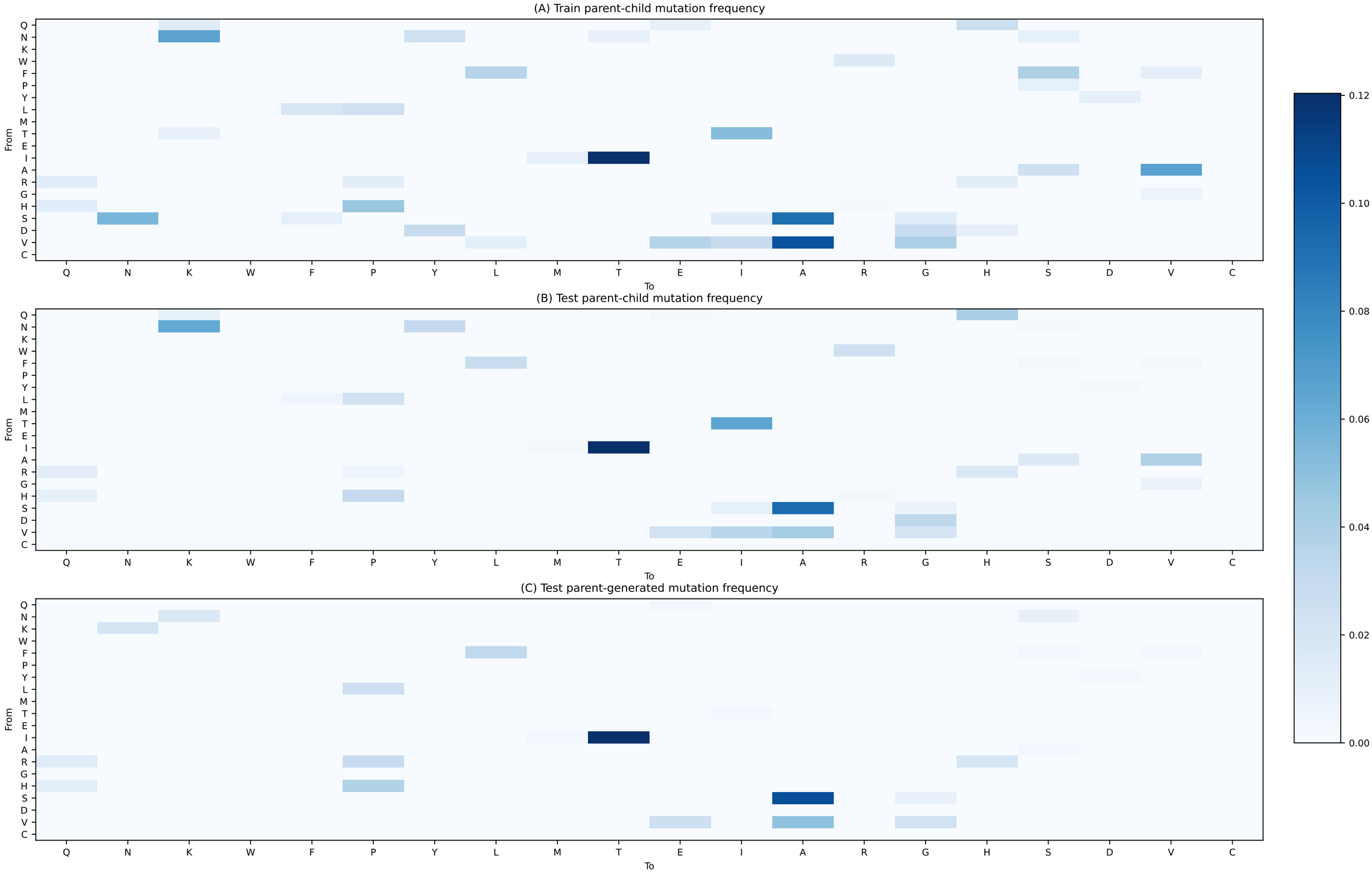Mutation frequency in test, train and generated datasets. Pearson correlation of A & B: 0.89, A & C: 0.78, B & C: 0.82