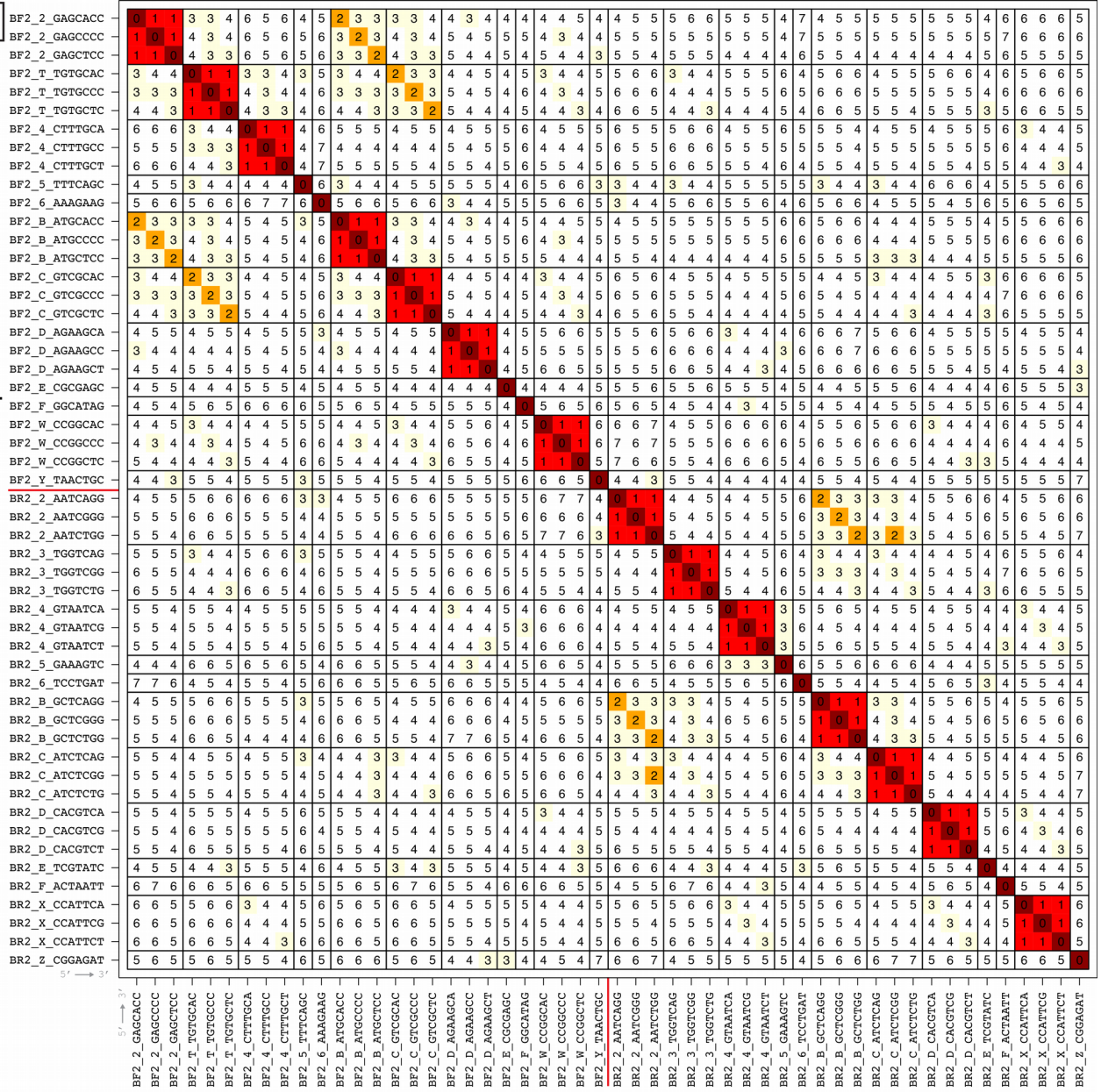


A

illumina P5 adapter



Levenshtein distance between tags

Illumina P7 adapter

B

- x Substitution
- + Insertion
- Deletion

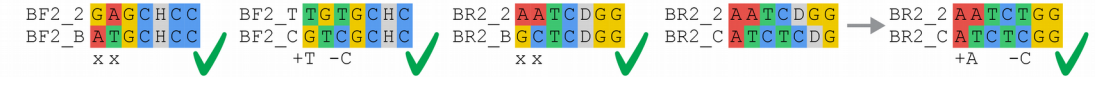


Figure S4: Plot showing the Levenshtein distance (counting substitutions, insertions and deletions) between all fusion primer combinations using a 7 bp in-line tag (A). All tags show very high dissimilarity, except for four combinations detailed in (B). To verify that a Levenshtein distance of 2 is actually 2 when considering the full primers sequences, the 4 cases were verified manually. The 7 bp tags are a combination of in-line inserts and the 5' tail of the primer sequence. Single insertion or deletion differences of a nucleotide will lead to a sequence shift that requires the insertion of another nucleotide on the tag it is compared to (increasing the distance to 2, see Figure 1 in Faircloth & Glenn 2012). Consequently, all distances of 2 should be verified, as they also might indicate a single nucleotide difference, if the full primer sequence is considered (e.g. in demultiplexing). However, as shown in B, the distance here is always 2, thus the tagging sets should be reliable in demultiplexing.