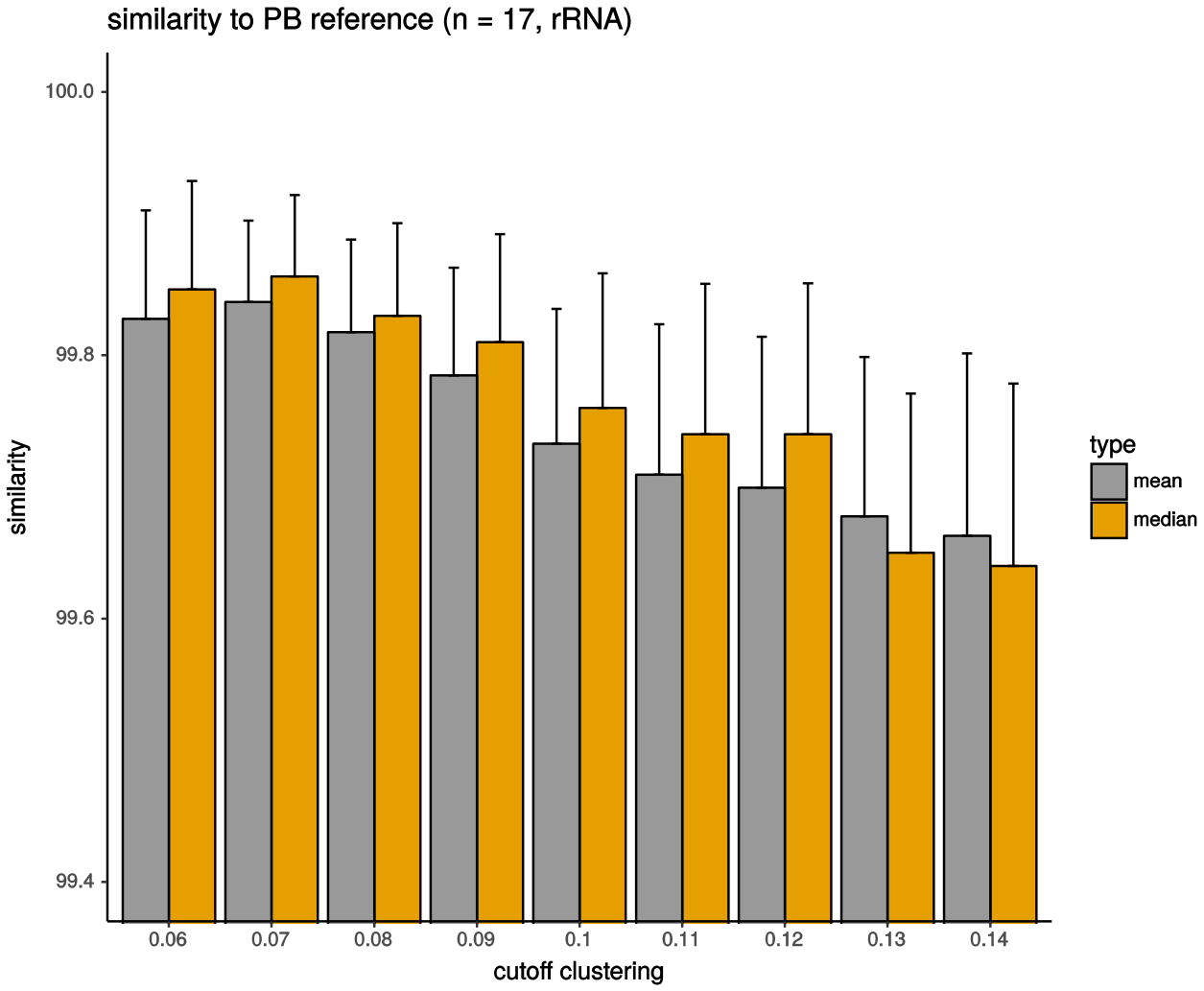
Supplemental Material S3

Figure S3.1. To test the effect of clustering in Mothur (with Opticlust) with different clustering hard-cutoffs ranging between 0.05 and 0.14 on the consensus accuracy (default parameters), we took the Pacbio generated consensus sequences as control (PB reference) and used local blast to assess the similarity values for 17 specimens. For each cutoff, we calculated the mean and the median similarity, plus the upper standard deviation. The test was done with operon sequences obtained from the rDNA-PCR. We had to remove the 0.05 cutoff, due to the inability to generate OTUs with sufficient sequences.



The clustering at 0.07 seems to be most successful and generated the highest accuracy. In the longer TR fragments 0.08 perfomed best (not included in the graph). In general we recommend to use a clustering cutoff below 0.08 for the following reason: In one specimen we encountered the phenomenon of a polymorphic intron in the SSU region (specimen 17080 had a long intron in the SSU). When we clustered at 0.09 or above, the intron got more and more integrated into the consensus sequences, i.e. it got clustered into the dominant OTU (see alignment Figure S3.2 below). So as a precaution against introns, we'd recommend to cluster at 0.08 or below.

Figure S3.2. Comparison of the effect of the clustering threshold on the consensus sequence of the Nanopore generated sequence data (Nano008-Nano012, stands for clustering thresholds of 0.08 to 0.12, respectively) in an alignment view. We compared these to the PacBio reference OTU, that had the intron, and three closely related species from the same genus (Inocybe) taken from the SILVA

reference database (SSUref, v132; Inoc1SILVA to Inoc3_SILVA) to confirm this irregularity. The intron is 374 bp long and ranges from alignment position 1175 to 1549.