Supplemental Figure S1: Distribution of sequence length (bp) of the individual reads that constitute the TOBEST database. These reads were assembled by the CAP3 software into overlapping clusters (contigs) and singlets.