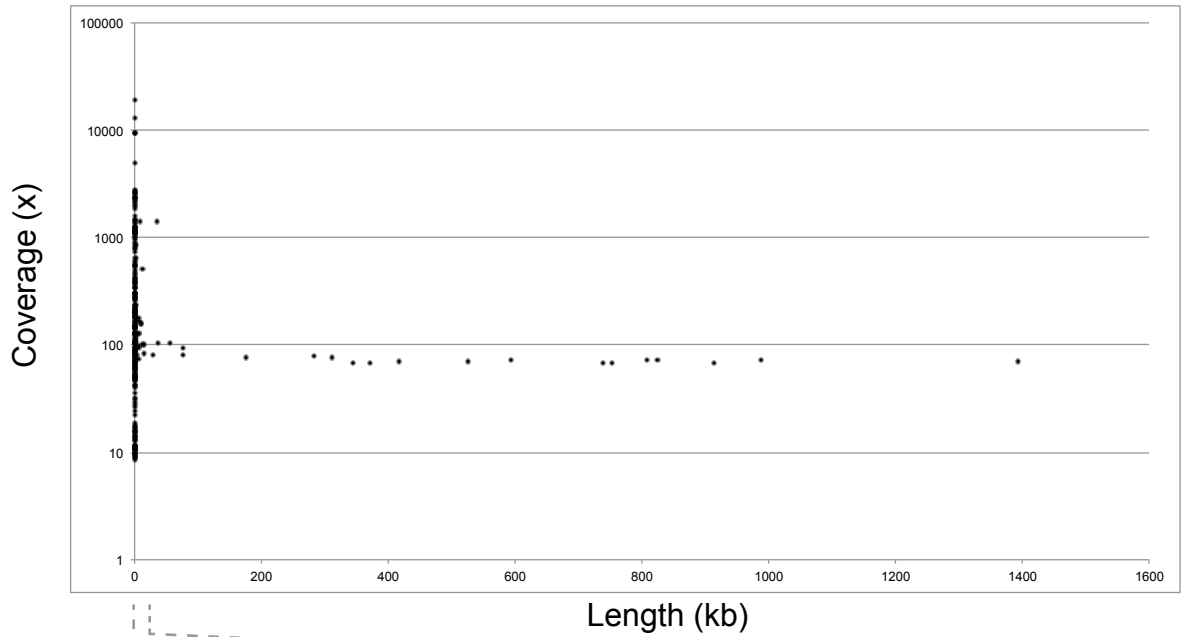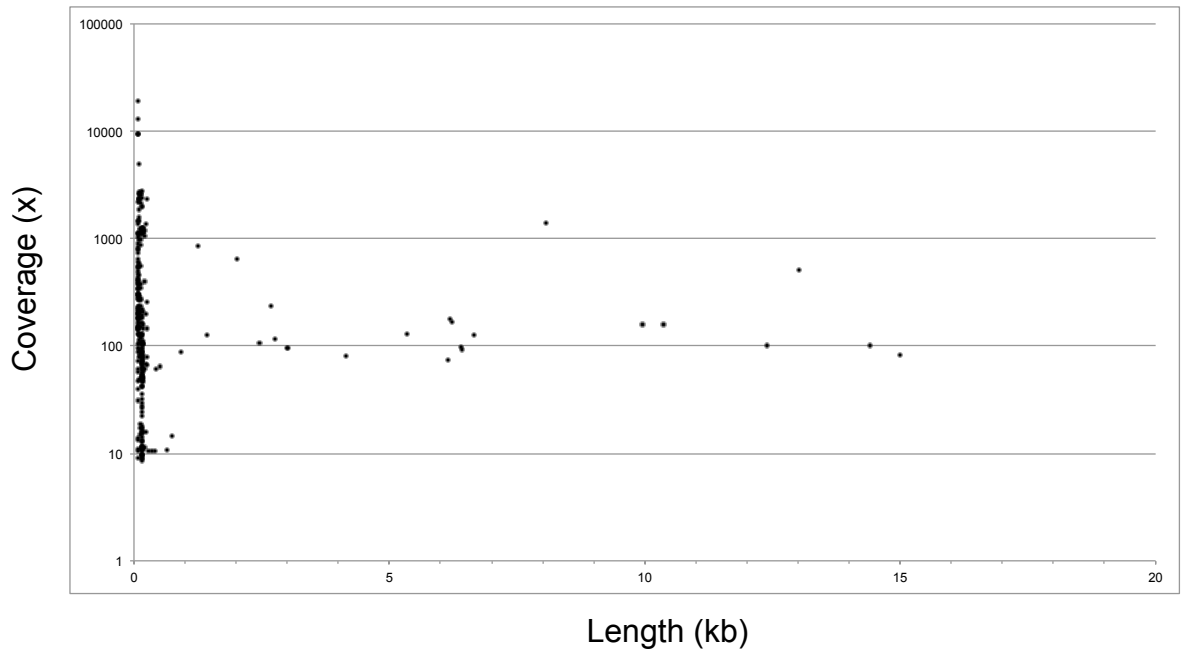**Figure S1.** CVL plot of an ABySS v2.0 assembly of *K. phaffii* strain CBS7435 showing (A) all scaffolds, and (B) scaffolds shorter than 20 kb. The Illumina raw data files were the same as those used to make the SPAdes assembly shown in Figure 2. In an ABySS assembly, each scaffold's header contains three fields: SEQ_ID, SEQ_LEN and KMERS, where SEQ_ID is a unique identifier, SEQ_LEN is the length of the scaffold, and KMERS is the total number of *k*-mers that mapped to the scaffold during assembly. We used this information to calculate *k*-mer coverage for each scaffold as C_k = KMERS / (SEQ_LEN - *k* + 1), where *k* is the *k*-mer size used for assembly (*k* = 77 in this example).