Supplemental Figure 1. *ycf1* distribution based on HMM searches against a dataset of all protein data from the plastid and cyanobacterial genomes plus additional tBLASTn hits. To avoid missing any YCF1 homologs (false negatives) due to miss-annotation, a successive screen using tBLASTn and subsequent HMM analyses was performed on those genomes, for which in our first round of HMM analysis no YCF1 homologs were detected. They were then plotted in the same way as in Figure 1A. Note that the majority of additionally detected sequences are very fragmented and likely do not stem from functional proteins. Labeling and color code is identical to Figure 1A.