



Additional Figure 1. The quality of the 196 budding yeast genomes sequenced using Illumina paired-end technology is comparable to the quality of the 136 publicly available genomes. (a) N50 value (i.e., the contig or scaffold value above which 50% of the total length of the sequence assembly can be found) for newly sequenced Y1000+ Project genomes and publicly available genomes (including the unpublished RIKEN genomes). (b) GC content (i.e., the percentage of G+C in a given genome) for newly sequenced and publicly available genomes. (c), Genomic completeness between newly sequenced and publicly available genomes. Genomic completeness was assessed as the percentage of the full-length BUSCO genes present in each genome (see “Assessment of genome assemblies” section of the Supplementary Methods). (d) Number of annotated genes in a given genome for newly sequenced and publicly available genomes. Each violin plot describes the distribution of each of the four quality measures and their densities, with dots denoting median values. Values for all four measures for each of the 332 genomes are provided in Table S1.