



**Concoct 1.** Single core gene (SCG) frequencies in the 76 CONCOCT clusters generated after binning contigs  $\geq 1$  kb. Each row corresponds to a cluster and columns are SCGs. The top 11 clusters, delineated by a thick black line, included  $\geq 75\%$  of the 36 SCGs. Gene copy number is indicated by the key.