

Supplementary Data

Supplementary Figure S1 – The statistical enrichment analyses of KEGG and GO terms for human cycling genes against the entire human proteome, using the hypergeometric test. Altogether, significantly over-represented KEGG pathways were detected (p -value < 0.001). For GO terms, the top most significantly enriched biological processes (p -value < 10^{-13}), molecular functions (p -value < 10^{-8}), and cellular components (p -value < 10^{-5}) were shown, respectively. The results were visualized using an online wordcloud generator (<http://www.wordclouds.com>), and the font size denotes the enrichment ratio of each term, while the color denotes the approximate p -value.

Supplementary Figure S2 – The statistical enrichment analyses of KEGG terms for mouse (p -value < 0.001), *Drosophila* (p -value < 0.01), *Neurospora* (p -value < 0.01) and *Arabidopsis* (p -value < 0.001) cycling genes against the whole proteomes, using the hypergeometric test. The results were visualized using an online wordcloud generator (<http://www.wordclouds.com>), and the font size denotes the enrichment ratio of each term, while the color denotes the approximate p -value.

Supplementary Table S1 – The detailed information of 27,964 experimentally identified cycling genes collected from both small-scale and high-throughput experiments. One gene may have multiple entries, since the phase information can be different under distinct external conditions or tissues/cells. This data set can also be downloaded at: <http://cgdb.biocuckoo.org/download.php>.

Supplementary Table S2 – The numbers of cycling genes of 148 eukaryotic organisms in CGDB. The 25 organisms from which the 27,964 experimentally validated cycling genes were collected were marked in gray with a number in the bracket indicating the number of experimentally validated genes for each species.