

Supplementary Text

Distribution of restriction-modification (R-M) systems in sequence genomes

As methyltransferases (MTases) show highly conserved features unlike restriction endonucleases (REases), genome analyses of MTases were carried out to understand the distribution of R-M systems. Analysis of the sequenced genomes for the number of MTase genes was taken from the restriction enzyme database, REBASE (<http://rebase.neb.com/rebase/rebase.html>). A total of 1028 sequenced genomes were grouped into different bins with a defined genome size range. The scatter plot (Fig. S1), depicting the distribution with genome size interval on x-axis and number of MTase genes on y-axis, was generated using GraphPad prism V4. The median of the distribution was determined and plotted in Fig. 2.

For the analysis of the distribution of R-M systems in RecBC- organisms, the genomes of RecBC- organisms were taken from Montague *et.al.* and Rocha *et.al.* (1, 2). The genomes were analysed for the number of genes encoding MTases as described in the above section. The mean value obtained for the RecBC- and the total organisms are plotted in Fig. 6. The distribution of the R-M systems the RecBC- organisms with median value is given in Fig. S3.

References

1. **Montague M, Barnes C, Smith HO, Chuang RY, Vashee S.** 2009. The evolution of RecD outside of the RecBCD complex. *J. Mol. Evol.* **69**:360-371.
2. **Rocha EP, Cornet E, Michel B.** 2005. Comparative and evolutionary analysis of the bacterial homologous recombination systems. *PLoS Genet.* **1**:e15.