



**Figure S5.** The figures show how the distance from the origin of replication affects the insertion index in different genomes. The origin of replication in these figures is assumed to be the dnaA gene. Each dot represents a gene and the red line shows the GAM curve when the formula is  $y \sim s(x)$ . The pink area shows the 95% confidence interval.