



**Figure S6.** The figures show how G-C content affects insertion index in different genomes in this study. 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.