



**Figure S4.** Simulation of insertion density effects: The red and blue dots are obtained from real and simulated data, respectively. The black line shows the loess curve with 0.2 span and the gray regions show 95% confidence intervals. Insertion resolution is calculated by dividing genome length by the number of insertion sites. (a) Matthew's correlation coefficient remains almost constant after the insertion index reaches 0.04. (b) The percentage of genes with insertions rises as we increase the insertion density, however the rate slows than after 0.04 threshold.