

Figure X: Gene loss from sporulation protein set correlates with evolutionary distance for predicted sporulators. Evolutionary distance calculated for 796 genomes from the Parks et al “Firmicutes” negatively correlated with genome classified as “LIKELY\_SPORULATOR” (grey dots) or not (black dots). **(A)** R2 = 0.15 for a linear model based on the full reference, while **(C)** R2 = 0.59 when non-spourlators when excluded from the model. **(B)** Segmented regression predicted a breakpoint at 26.25 weighted genes. Solid red lines indicate the model predictions based on the reference set, while dashed lines represent the 95% CI.

**C**

**B**

**A**