



**S1 Fig. Distance dependence of spurious reassortment counts in non reassorting sequence.** (a) Histograms of the number of events found in a HA tree as a function of  $\delta$ , for sequences of total length  $L$ . Error bars represent the standard deviation obtained from 5 different random choices of the sites for each  $\delta$ . (b) The decay exponent  $\gamma$  is shown as a function of  $L$  (cf. Materials and Methods, equation 6). The inferred values are stable for large values of  $L$ , allowing extrapolation to  $L = L_{\text{HA}} + L_{\text{NA}}$ .