



**Supplemental Figure S6.** Viterbi Decoding and LRT attack on proxy panels. **(a)** Allele decoding accuracy. Each plot shows the allele error fraction with increasing window length used for the decoding. In each plot, bar colors indicate the error for Viterbi (Blue) and error rate for the randomly matched haplotypes (Red). Between different plots, the hashing allele hashing window length is increased from 7 to 17-mers indicated on the horizontal arrow at the bottom. Note that the vicinity variant size for hashing corresponding to these parameters varies from 3 to 8. **(b)** Distribution of Sankaraman's LRT statistic using matching panel (Red) and non-matching panel (Blue) as the query individuals and the decoded panel as the pool. **(c)** Distribution of LRT statistics using matching panel (Red) and non-matching panel (Blue) as the query individuals and the original matching panel as the pool panel (No proxy panel). This is used as the baseline re-identification scenario when pool panel's allele frequencies are exactly known. **(d)** Distribution of LRT statistics matching panel (Red) and non-matching panel (Blue) as the query individuals and the resampled matching panel as the pool without hashing of variants.