



Figure S3. Rescaled error (eq. 30) as a function of the number of SNPs in the haplotype, for four quartiles of the sequence distance between the $(n-1)$ st SNP and the n th SNP, $n = 2, 5, 10, 30$. Quartiles (1, shortest distance; 4, longest distance) were determined separately for each n . The plot considers as the two populations the data from Europe and East Asia and uses SNP windows as the extended haplotypes. Each of the sixteen columns shows a distribution across 2,500 extended haplotypes. Whiskers extend to the largest and smallest observed rescaled error values no larger and smaller than 1.5 times the interquartile range above and below the median, respectively. The y-axis is truncated at 5. Kruskal-Wallis tests for equality of rescaled error across distance quartiles resulted in non-significant p -values for three of four haplotype lengths ($p = 0.5055$ for $n = 2$, $p = 0.1185$ for $n = 5$, $p = 0.2057$ for $n = 10$, and $p = 0.04991$ for $n = 30$).