



Figure S4 Comparison of site-frequency spectra on human chromosome 6 estimated by different methods. The site-frequency spectrum estimated using the method by Nielsen *et al.* (2012) (ANGSD) and the Samtools or GATK genotype likelihood model is compared to that estimated using the proposed method. The theoretical prediction is in a population with constant size, and is given by $[1/j + 1/(162-j)]/[\sum_{k=1}^{161} 1/k]$ for $1 \leq j < 81$ and $(1/j)/[\sum_{k=1}^{161} 1/k]$ for $j = 81$, where j is the minor-allele count (Ewens 2004).