



**Figure S1** Phred quality scores of individual base calls belonging to the first six classes of the site frequency spectrum in the North American population (calculated from the DPGP1 raw fastq files). The middle bands indicate the median values, the boxes the upper and lower quartiles and the whiskers the minimum and maximum values.