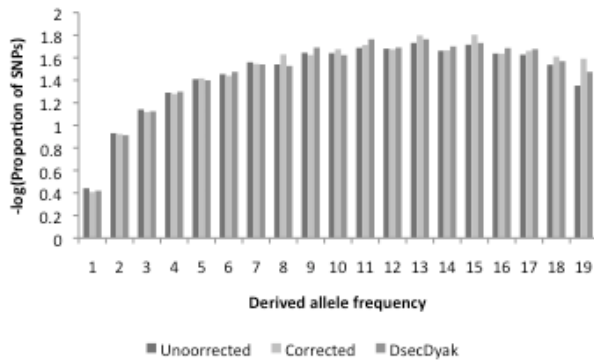
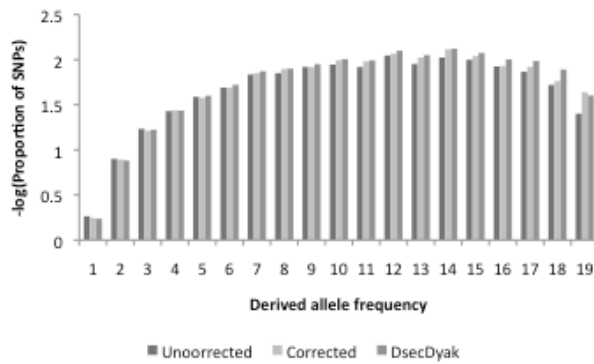


a)



b)



**Figure S4** Site frequency spectrum plotted on a negative log scale for a) long introns and b) third codon positions. The uncorrected site frequency spectrum ('uncorrected'), the corrected site frequency spectrum ('corrected', see Methods), and the site frequency spectrum obtained by polarizing SNPs to both *D. sechellia* and *D. yakuba* ('DsecDyak') are presented. Note that the corrected site frequency spectrum shows a lowest proportion of SNPs with high derived allele frequencies.