



**Figure S5** The patterns of the slope (a & c) and the coefficient of determination  $R^2$  values (b & d) for the usages of amino acids between two different GC-groups of sequences. CDSs (a & b) and 50-codon motifs (c & d) were separately used as analysis units, with a GC-content difference less than 0.25 plotted.