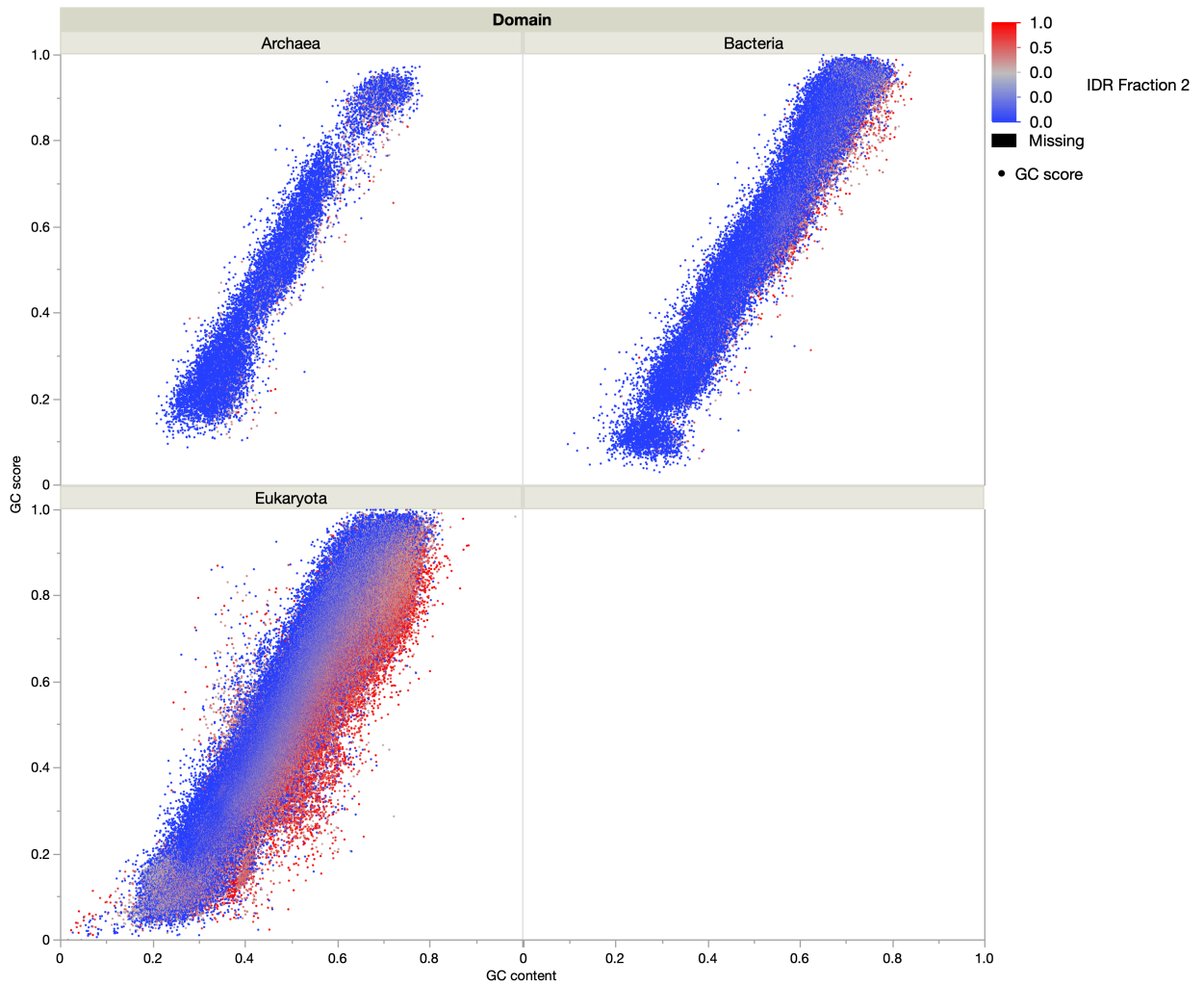


Supplementary Figure 3.



This figure shows the scatter plots of the GC scores and GC contents of all target protein genes by the domains to which they belong. In this figure, the plots are colored according to their fraction (proportion) of intrinsically disordered regions in the protein sequence. Proteins without intrinsically disordered regions are colored blue, while proteins with intrinsically disordered regions are colored gray to red, depending on their fraction of the protein. These fractions were calculated using information from the UniProt protein database [5].

The deviation from the correlation between GC score and GC content is greatest in eukaryotes, and these deviating proteins were found to consist predominantly of proteins with larger fractions of intrinsically disordered regions.