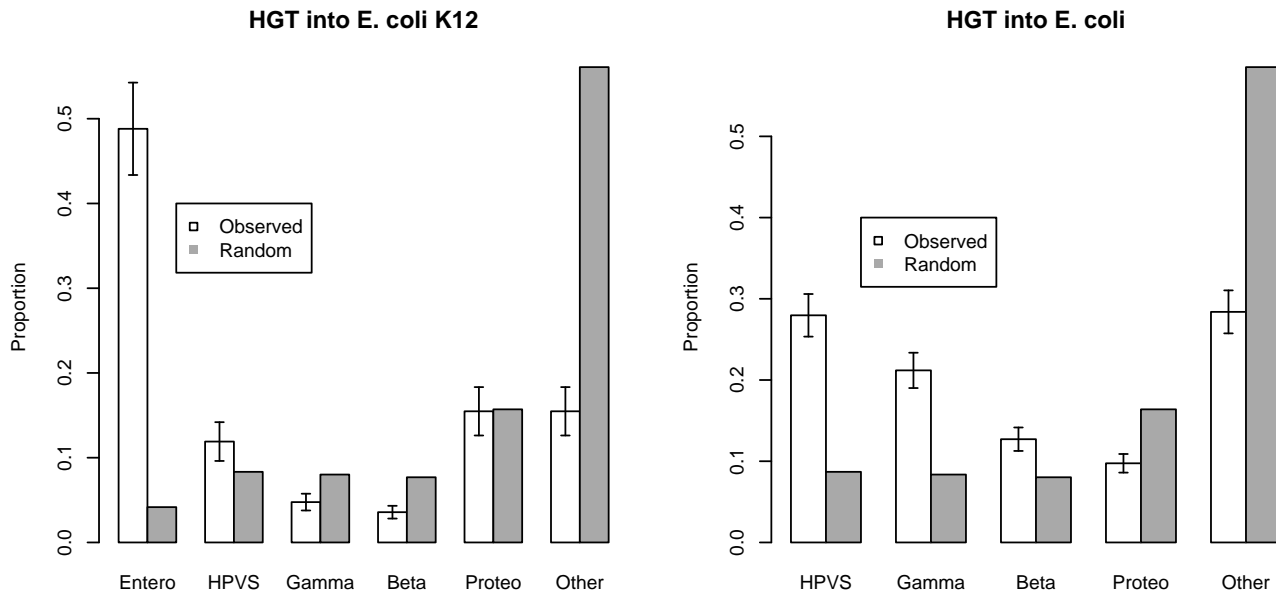


Figure S2: A preference for HGT between related genomes



We tabulate the potential sources (as identified by best hits) of genes that were acquired by *E. coli* K12 after its divergence from other strains of *E. coli* (left panel) and the sources of genes that were acquired after the divergence of *E. coli* from *Salmonella* (right panel). We compare the distribution to that if the sources of genes were chosen at random from completely sequenced genomes. Error bars show the mean \pm two standard deviations according to the binomial distribution. The most closely related groups of genomes are at the left, and the groups' names are abbreviations from Figure 9. Potential HGT events from other *E. coli* strains, from *Salmonella*, or (on the right) from other Enterobacteria are not considered because they are too close for HGT events to be identified by the automated method.