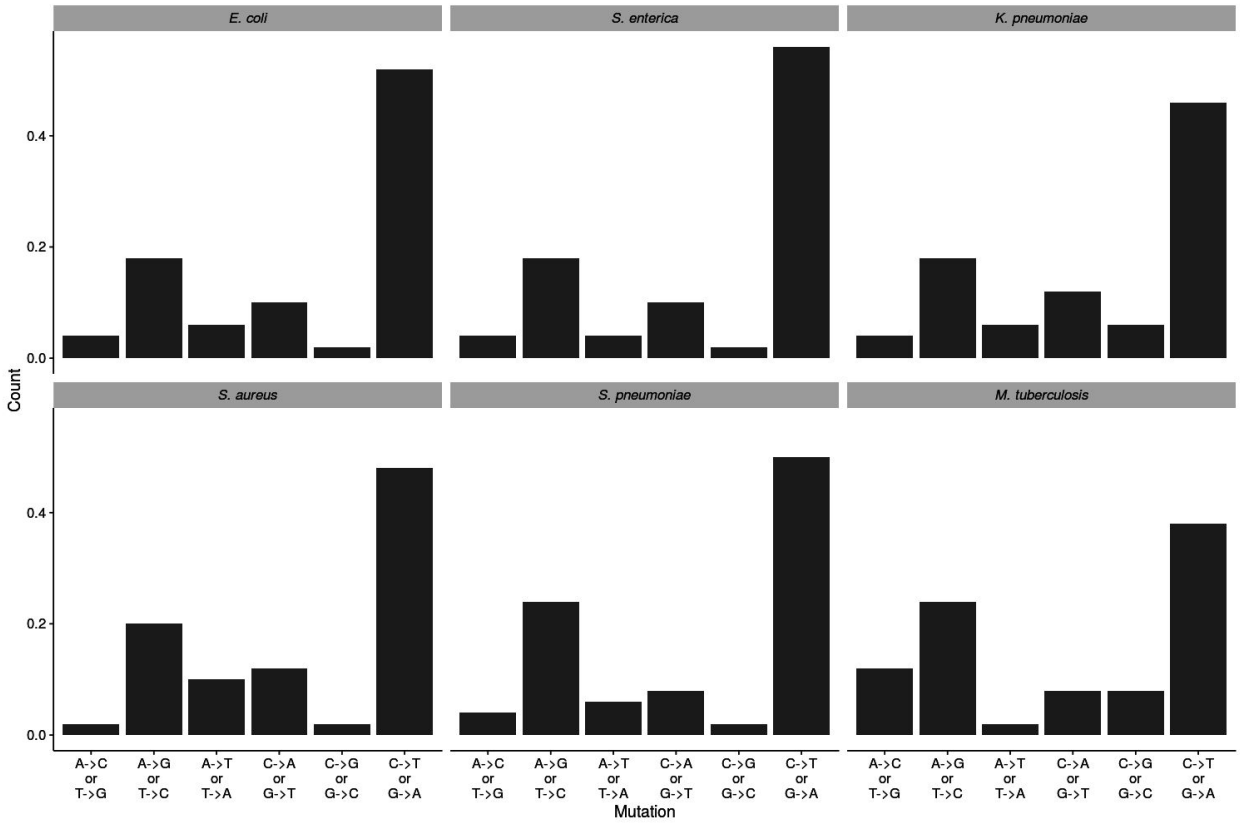

Fig S1: Per site mutation biases. The per site mutation bias was calculated for each species from the singleton SNPs. This was done by dividing the number of SNPs by the number of sites (e.g. the number of A->C and T->G mutations was divided by the number of A and T sites in the genome). These rates were then converted to a proportion so that the mutation types summed to 1 for each species.

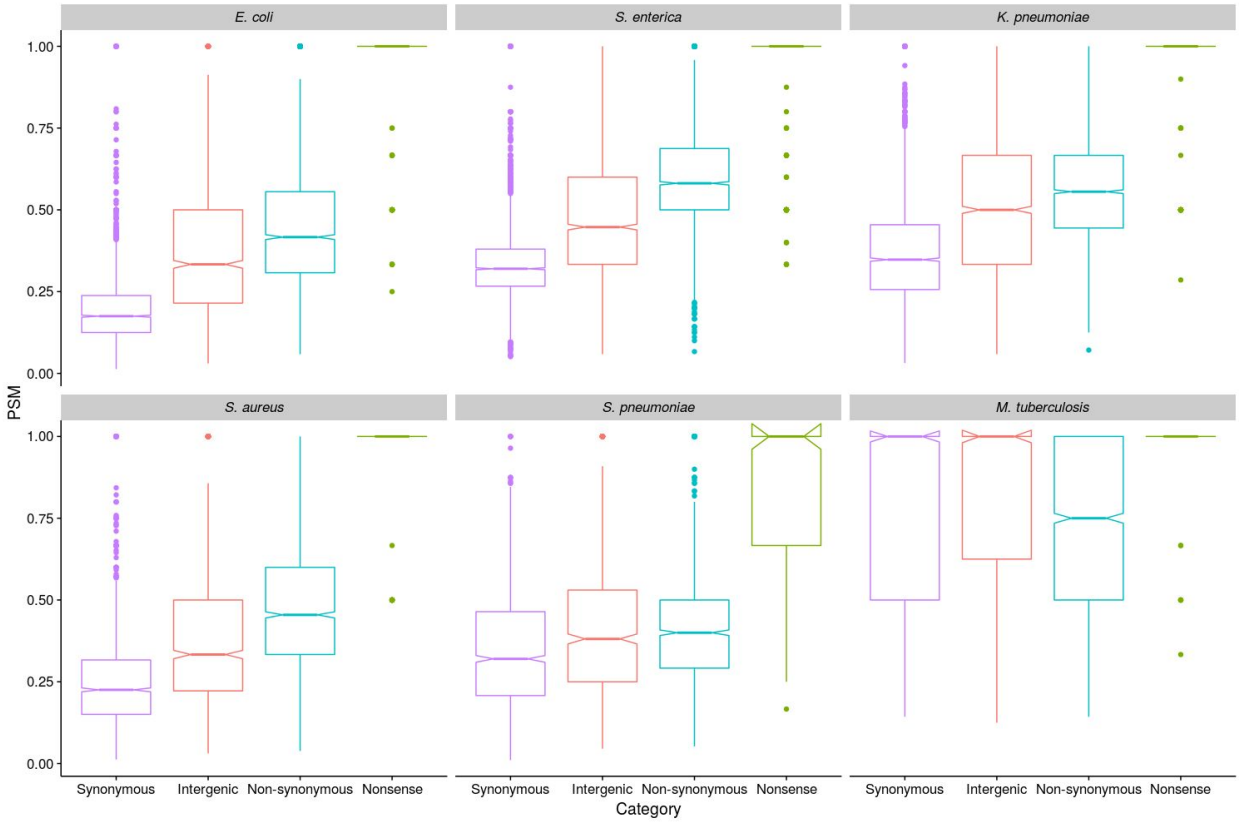

Fig S2: Analysis of selection on individual genes and IGRs. PSM values were calculated for each gene and IGR separately to check that our analysis was not confounded by a small number of highly conserved, unrepresentative IGRs. The notches in the box plots represent 95% confidence intervals around the median.

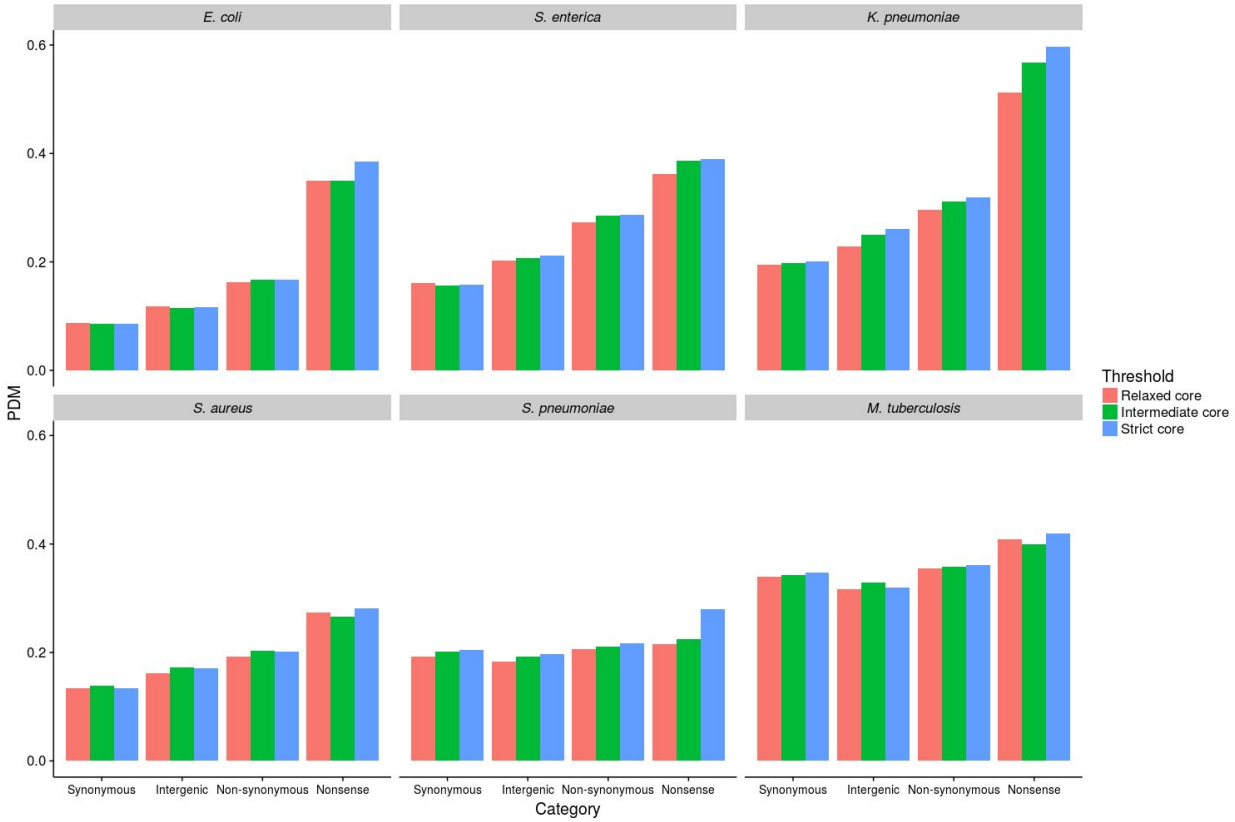

Fig S3: PDM (Proportion of Doubleton Mutations) analysis of selection on different mutation categories. PDM values were calculated by dividing the number of doubleton SNPs (those present in two genomes) by the total number of SNPs within that mutation category (excluding singletons).

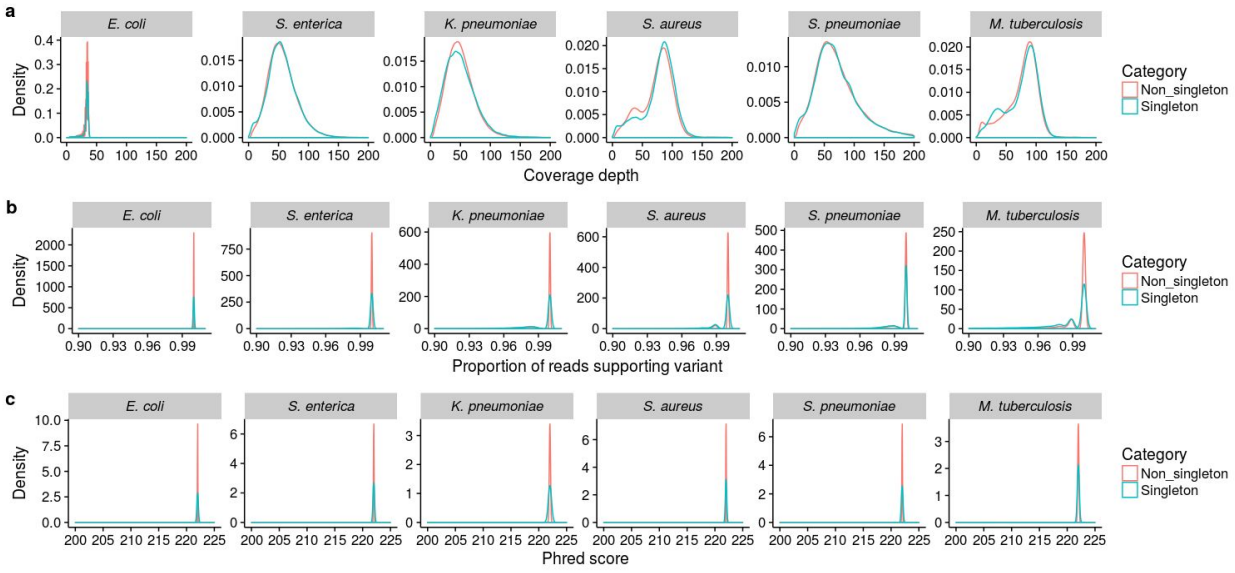

Fig S4: Validation of the quality of singleton SNPs. Singleton and non-singleton SNPs were analysed in order to validate the quality of the singleton SNPs. **a.** Depth of coverage of SNP positions. **b.** The proportion of reads supporting the SNP. **c.** The Phred Quality Q score of SNP positions.

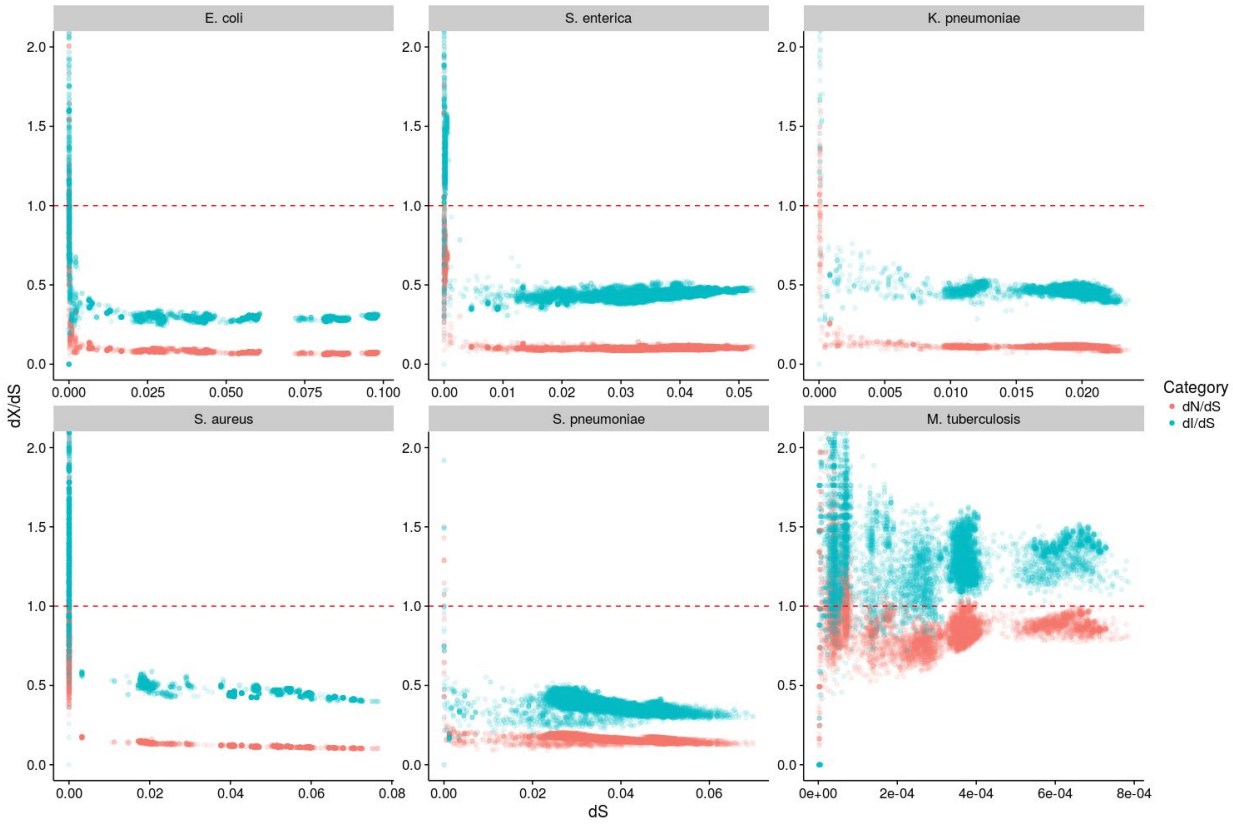

Fig S5: dN/dS and dI/dS analysis of selection. dN/dS and dI/dS were calculated between isolates in a pairwise manner, and these were both plotted against dS to explore the effect of divergence time on observed levels of selection. In order to control for the non-independence between the axes, we calculated dN/dS, dI/dS, and dS from different sites as described in Methods. The dashed red line shows where dN/dS and dI/dS = 1, and therefore indicates neutrality.

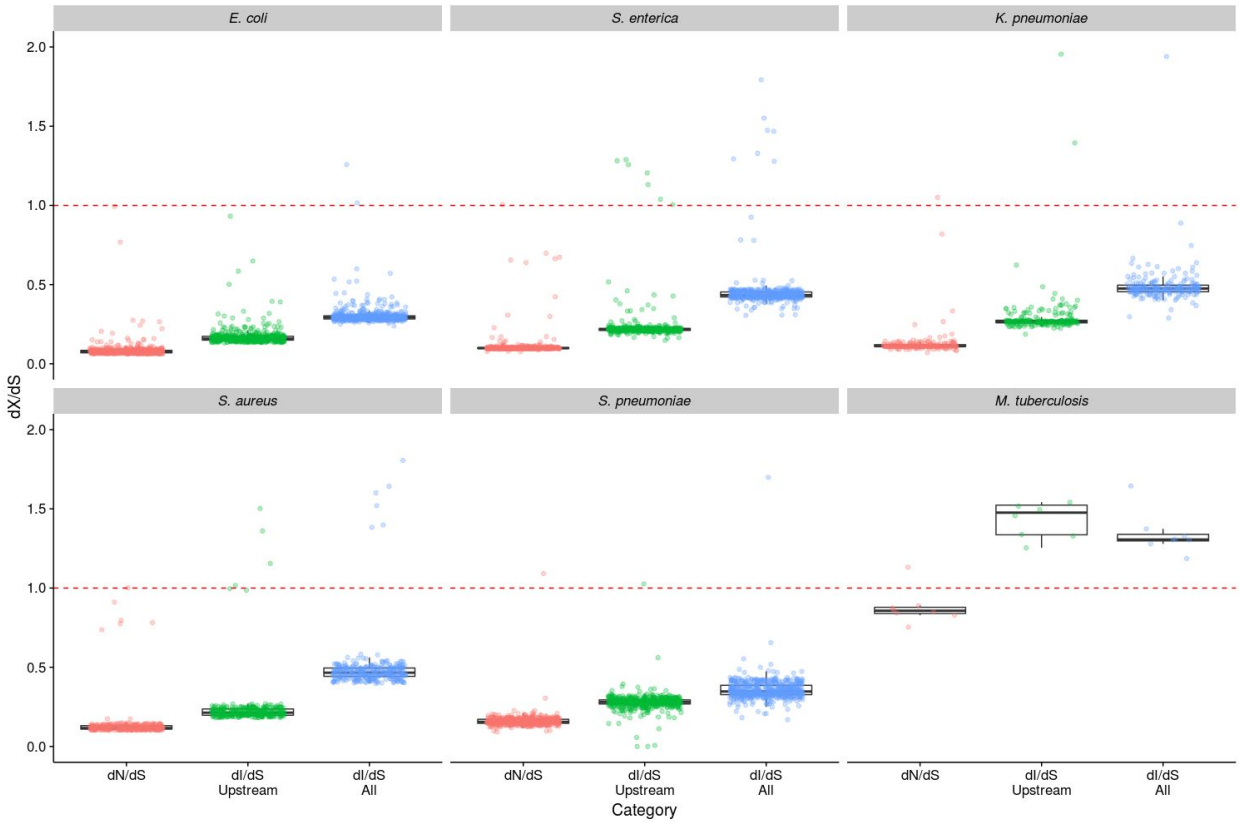

Fig S6: dI/dS analysis of IGRs upstream from gene starts. dI/dS was calculated from all intergenic sites, and intergenic sites 30 bp upstream from gene starts separately, by comparing isolates in a pairwise manner. The results were binned by dS (bin width = 0.0001) to control for oversampling of very closely related isolates (such as those belonging to the same CC). The genome-wide dN/dS values are included to enable comparisons to be made between non-synonymous sites and the intergenic sites. The dashed red line shows where dN/dS and dI/dS = 1, and therefore indicates neutrality.

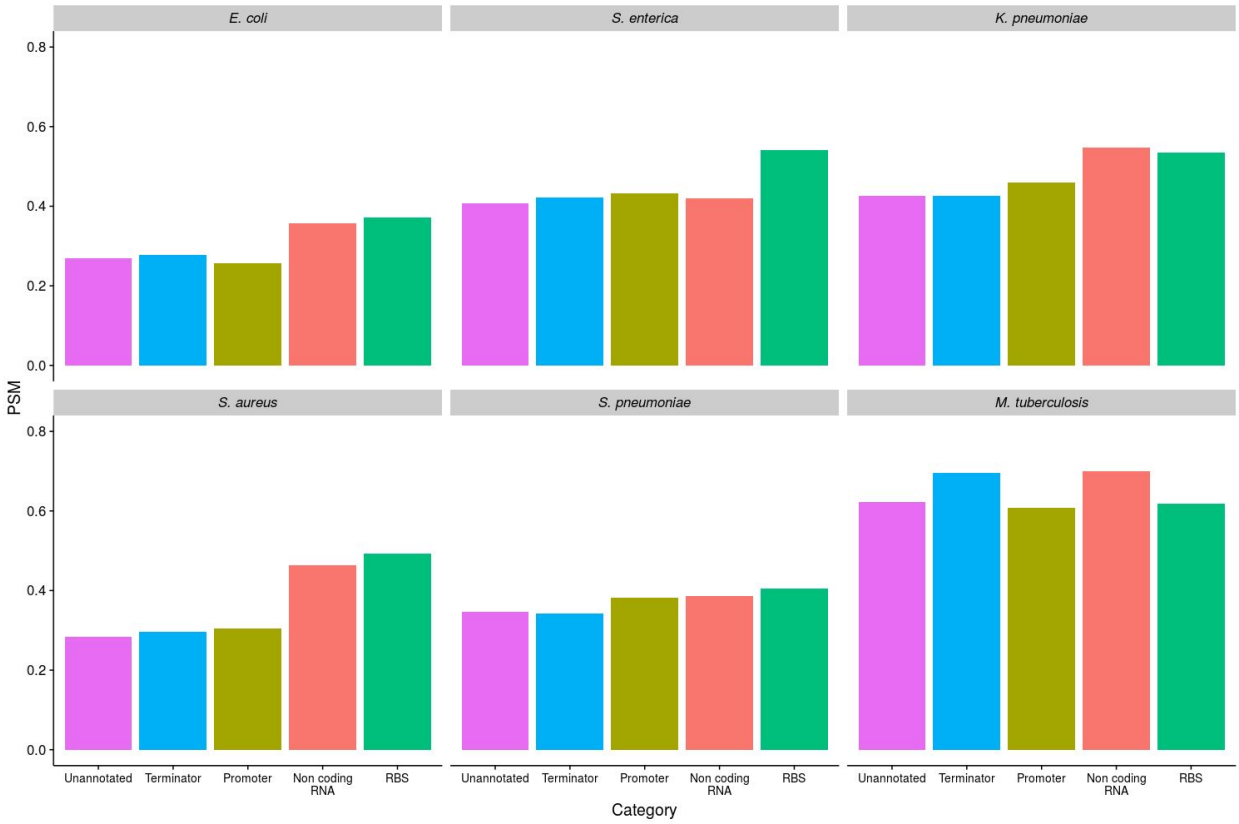

Fig S7: PSM analysis of selection on different regulatory elements within IGRs. PSM values were calculated by dividing the number of singleton SNPs by the total number of SNPs within that mutation category.

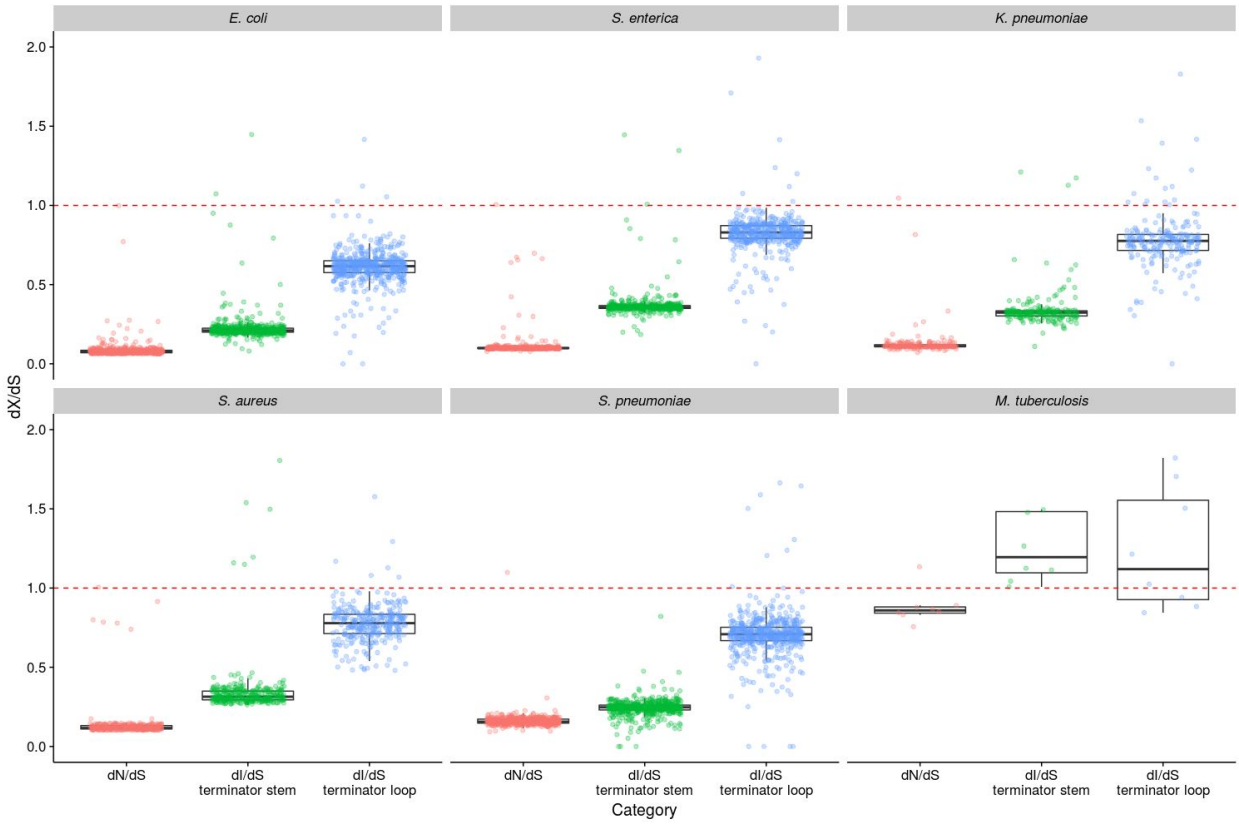

Fig S8: dI/dS analysis terminator stem and loop regions. dI/dS was calculated from terminator stem and loops separately, by comparing isolates in a pairwise manner. The results were binned by dS (bin width = 0.0001) to control for oversampling of very closely related isolates (such as those belonging to the same CC). The genome-wide dN/dS values are included to enable comparisons to be made between non-synonymous sites and the intergenic sites. The dashed red line shows where dN/dS and dI/dS = 1, and therefore indicates neutrality.

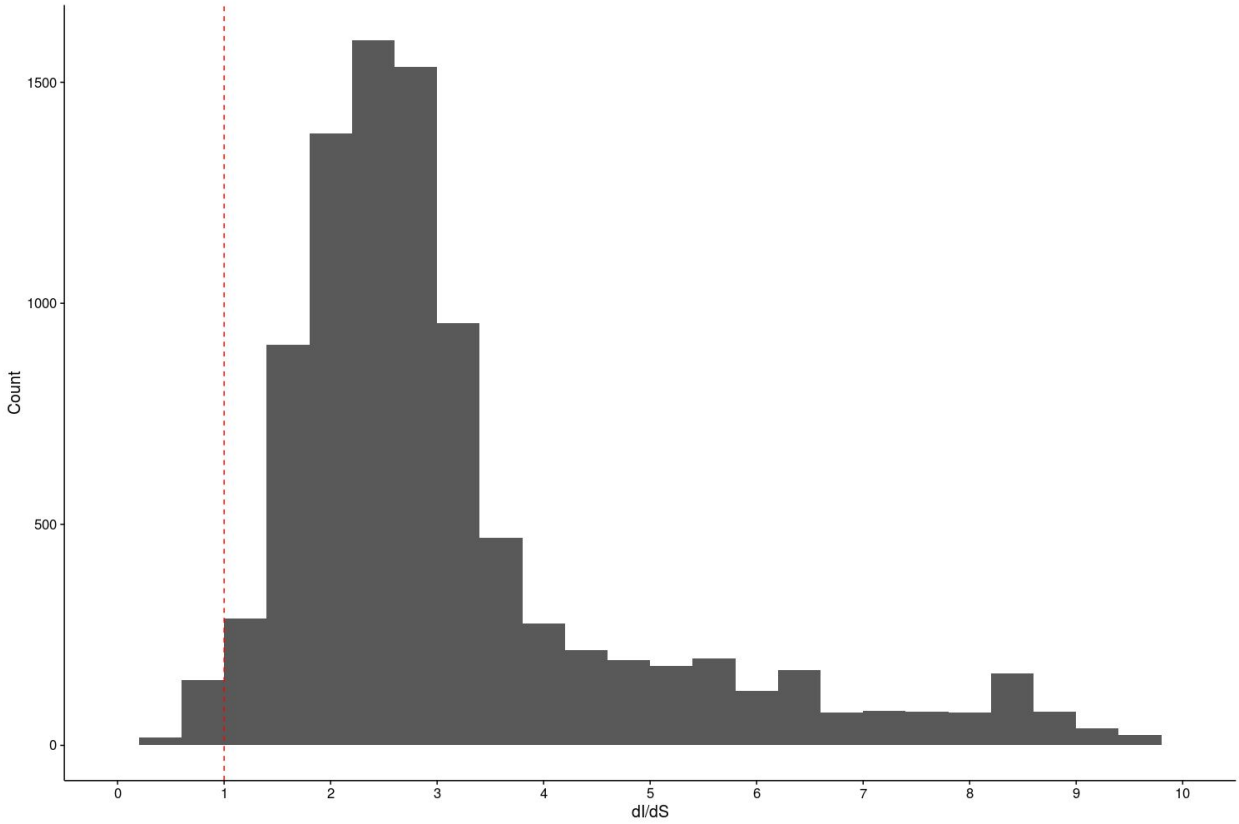

Fig S9: The distribution of promoter dI/dS values in *M. tuberculosis*. dI/dS was calculated from promoter sequences in a pairwise manner between *M. tuberculosis* isolates, and the histogram shows the distribution of these dI/dS values. 85% of the comparisons are > 1. The dashed red line shows where dI/dS = 1, and therefore indicates neutrality.

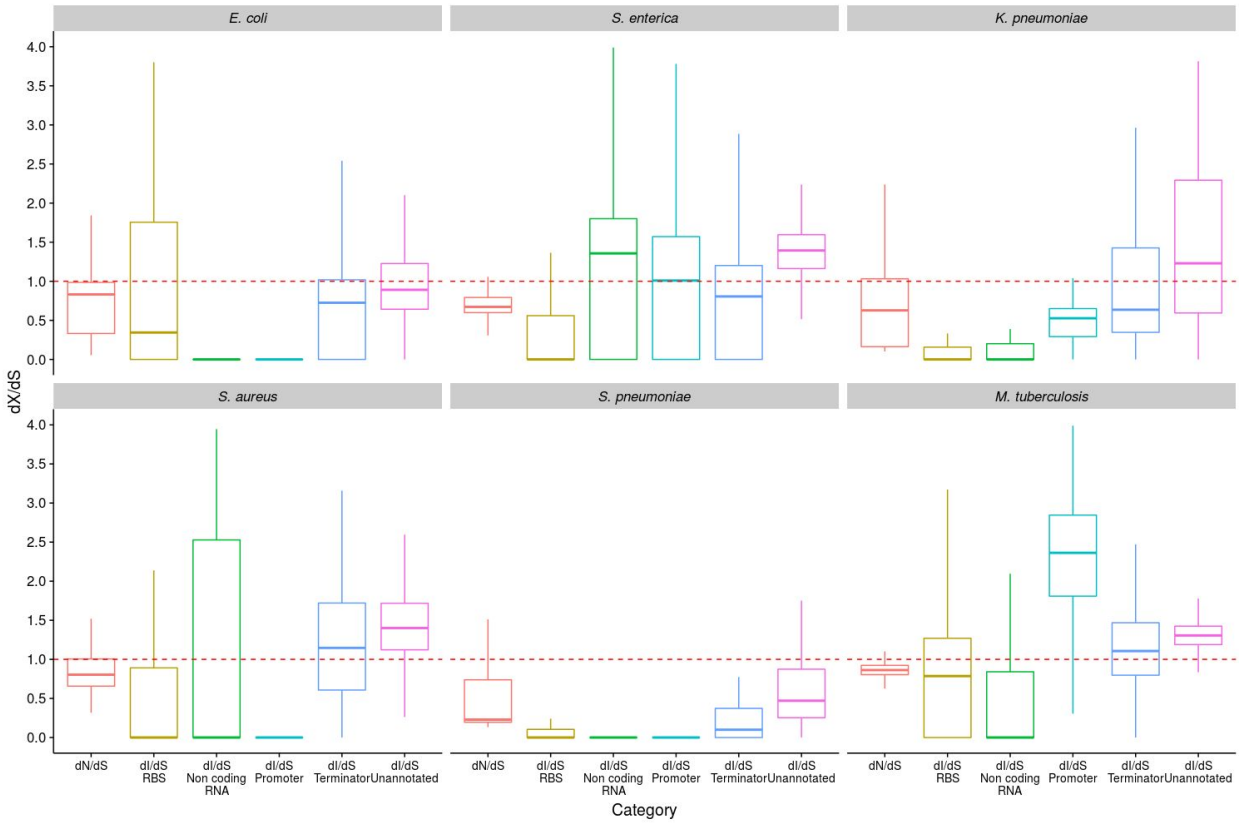

Fig S10: dI/dS analysis of selection on different regulatory elements in within-CC comparisons. dI/dS was calculated between isolates in a pairwise manner, and comparisons with dS < 0.003 were included to represent within-CC comparisons. The genome-wide dN/dS values are included to enable comparisons to be made between non-synonymous sites and the different regulatory intergenic sites. The dashed red line shows where dN/dS and dI/dS = 1, and therefore indicates neutrality.