



**Figure S5.** Scatter plot of the first two principal components of a PCA analysis based on the TA toxin genes in AIEC and non-AIEC *E. coli* strains. Data were analysed according to phylogroup (A) or type of strain (pathotypes) (B). Prediction were done by ClustVis and ellipses are such that with a probability 0.95. Percentages of the total variances described by the first two extracted factors are given on the axes.