Principal Component Analysis and Hierarchical Clustering The data from the NPDIB were downloaded as described above. The data matrix contains a total of 35,753 rows, each of which corresponds to an isolate sample of L. monocytogenes entered from January 2010 through December 2021. These samples were further analyzed to obtain a matrix in which each row represents one gene while each column represents one year with the detection occurrence of the gene in the corresponding year recorded in the matrix cell. The data contained a total of 65 AMR genes for 2010 to 2021 resulting in a matrix that contains 65 rows and 12 columns. Due to the number of dimensions, these data were analyzed using principal component analysis (PCA) and hierarchical clustering (HC) to identify the highly occurring AMR genes by region and setting. PCA allows the visualization of multi-dimensional data in two dimensions. The data are expressed in terms of new variables that are linear combinations of the existing variables. The principal components PC1 and PC2 are those that retain the most variation from the original data [23]. In this work, PCA is used to project AMR genes into the PC1~PC2 two-dimensional space so that the outlier genes, which typically show higher occurrence over years, are identified for further investigation. While AMR genes can be visualized in PCA, certain genes are lumped together. PCA does not directly provide the correlation relationship between individual AMR genes. Therefore, hierarchical clustering is further used to group the genes projected onto the PC1~PC2 space into clusters that are similar to each other in the format of dendrograms. PCA and HC were performed, and graphs were generated using the free statistical software package R, version 1.4.1106, implemented in RStudio (Boston, MA, USA) [24].