



Article

Investigation of Incidents and Trends of Antimicrobial Resistance in Foodborne Pathogens in Eight Countries from Historical Sample Data

Katherine Yang ¹, Annie Wang ¹, Matthew Fu ¹, Aaron Wang ¹, Kevin Chen ¹, Qian Jia ² and Zuyi Huang ^{1,*}

¹ Department of Chemical Engineering, Villanova University, Villanova, PA 19085, USA; katherineyang2002@gmail.com (K.Y.); agani.first@gmail.com (A.W.); zmfu2002@gmail.com (M.F.); aaronwangca@gmail.com (A.W.); maykevin.chen@gmail.com (K.C.)

² Department of Health, Nutrition & Exercise Sciences, Immaculata University, Immaculata, PA 19345, USA; qjia@immaculata.edu

* Correspondence: zuyi.huang@villanova.edu; Tel.: +1-610-519-4848

Received: 8 December 2019; Accepted: 27 December 2019; Published: 10 January 2020



Abstract: Antimicrobial resistance (AMR) causes millions of illnesses every year, threatening the success of lifesaving antibiotic therapy and, thus, public health. To examine the rise and spread of antimicrobial resistance around the world, our study performs a multivariate statistical analysis of antimicrobial resistance gene data from eight different countries: the US, the UK, China, Brazil, Mexico, Canada, Australia, and South Africa. Multi-dimensional data points were projected onto a two-dimensional plane using principal component analysis and organized into a dendrogram utilizing hierarchical clustering to identify significant AMR genes and pathogens. Outlier genes/pathogens were typically involved in high occurrences of antimicrobial resistance, and they were able to indicate the trend of antimicrobial resistance in the future. Statistical analysis of the data identified: (1) *tet(A)*, *aph(3'')-Ib*, *aph(6)-Id*, *bla_{EC}*, *bla_{TEM-1}*, *qacEdelta1*, *sul1*, *sul2*, and *aadA1* as the nine most common AMR genes among the studied countries; (2) *Salmonella enterica* and *E. coli* and *Shigella* as the most common AMR foodborne pathogens; and (3) chicken as the most prevalent meat carrier of antimicrobial resistance. Our study shows that the overall number of reported antimicrobial resistance cases in foodborne pathogens is generally rising. One potential contributing factor for this is the increasing antimicrobial usage in the growing livestock industry.

Keywords: foodborne pathogens; antimicrobial resistance; principal component analysis; hierarchical clustering

1. Introduction

Antimicrobial resistance is increasingly a threat to global public health [1]. AMR-related infections limit the efficacy of lifesaving antibiotic therapy necessary for the treatment of infectious disease and limit the success of advanced surgical procedures like organ transplants [2,3]. This results in millions of illnesses and deaths worldwide. For example, the US Centers for Disease Control and Prevention (CDC) estimated that AMR bacteria account for over 2 million illnesses and at least 23,000 deaths in the US every year [4]. Similarly, it is estimated that 426,000 infections and 33,000 deaths from antimicrobial resistant microorganisms occur in the European Union every year [5]. Furthermore, AMR places a considerable medical and financial burden on global societies and health care systems as a whole [1,3,6–8]. It is therefore essential to develop a better understanding of the cause and spread of AMR around the world.

As a result of selection pressures provided by the usage of antimicrobials, resistance genes utilize various mechanisms to inhibit the antimicrobial's effects—enzymatic inactivation, antimicrobial target modification or protection, drug permeability reduction, and drug efflux [2]. These genes are becoming increasingly prevalent. In response to the need for further research in the spread of resistant foodborne pathogens and illness, the NCBI Pathogen Detection Isolates Browser (NPDIB) was created to integrate bacterial pathogen genomic sequences originating in food, patients, and environmental sources [9]. It analyzes samples from around the world and compares their genomic sequences to others already in the database. When AMR foodborne pathogens access humans, the AMR genes may be passed from the pathogens to human cells via horizontal gene transfer (HGT) through mechanisms of transformation, transduction, or conjugation [10]. This is implied by a study in which 13,514 high confidence HGT genes had been found in 308 human microbes [11]. The development of antibiotic resistance in foodborne bacteria and in HGT can significantly endanger human health. NPDIB clusters and identifies related sequences to uncover potential food contamination outbreaks and resistance genes. Through the database, people can search for foodborne pathogen isolates and identify pathogens with particular resistance genes. In previous literature, the database was typically only used to compare the DNA from detected pathogens to those in the database or provide data for a historical analysis of resistance in specific pathogens [12,13]. Little research, however, has been done to examine specific AMR genes over large geographic areas.

Currently, only one previous study used the database to study specific resistance genes [14]. That 2019 study examined the spread of resistance genes across six different US states. It found common antimicrobial-resistance genes in both geographically near and geographically far states in the US. However, it did not study the historical occurrence trends of either antimicrobial resistance genes or pathogens. Therefore, the study was unable to indicate how the antimicrobial resistance issue evolved over time. In order to address these gaps, our study aims to expand the previous research by extending the scope of its work and identifying genes with increasing occurrence over time. A multivariate statistical method like principal component analysis [15,16] was used in this work to visualize the data into comparable points on a two-dimensional plane. On the basis of this, the hierarchical clustering approach [17–19] was further used to identify the genes, foodborne pathogens, and meats that were mostly involved in antimicrobial resistance. The occurrence trends of specific resistant genes and foodborne pathogens over time were examined over eight different countries in geographically different areas around the world—the US, the UK, China, Brazil, Mexico, Canada, Australia, and South Africa—which have the most amount sample data in NPDIB for different continents.

2. Materials and Methods

2.1. Data Source: NCBI Pathogen Isolates Browsers

The NPDIB database was created to monitor patterns and trends of illness-causing pathogens and genes. These pathogens were sampled from food, the environment, and human patients by health agencies in various countries, such as the Centers for Disease Control and Prevention, the U.S. Food and Drug Administration, the U.S. Department of Agriculture, and Public Health England. The database contains information on geographic location, sample time, isolation type, food source, and host. In this study, data for hundreds of samples collected from 2010 to 2019 were extracted from the NCBI database for eight countries: the US, the UK, South Africa, China, Australia, Canada, Mexico, and Brazil. These countries have the most data in the NPDIB database for the continents where they are located, making them the most suitable countries for analysis. In addition, the countries were selected because of their geographical, cultural, and developmental diversity. The chosen countries include highly developed countries, such as the US and the UK, along with less industrialized countries like South Africa and Brazil. Some countries have very high populations while others have fewer people. Using MATLAB software programs, the data for each country was organized into matrices in which each row represented one sample and each column represented one of five categories of information:

(1) the scientific name of the foodborne pathogen, (2) the collection date, (3) the sample source (e.g., turkey or chicken), (4) the number of AMR genes identified in the sample, and (5) individual AMR genes (with a value of zero/one that indicated the absence/presence of the corresponding gene in the sample). Figure 1 shows the general procedure of the multivariate statistical analysis implemented in this work. Figure 1A,B illustrate the steps of extracting data from the NPDIB database to the matrix format for further statistical analysis.

2.2. Statistical Analysis

Given each sample's large number of dimensions, principal component analysis (PCA) was applied to each country's data to visualize the complex data matrices into two-dimensional plots, providing a clear image of data and their relationships. Utilizing dimensionality reduction, PCA created new individual dimensions that can retain the most information in the data matrix. These new dimensional variables, called principal components, were orthogonal vectors that represented different levels of variance in the data. In order to offer a strong representation of the data, a sizable amount of variation needed to be accounted for by the principal components used in the visual illustration of the data. By looking at the proportion of variance explained (PVE), PC1 (the vector with the greatest PVE and, thus, the highest variance in its projections), and PC2 (the vector with the second highest PVE and variance in its projections) were chosen to create the two dimensional graph of the data. When plotting the sample data points, PC1 was used as the *x*-axis and PC2 as the *y*-axis. The graph created revealed clusters of similar data points along with significant outliers (Figure 1C).

On the basis of the PCA plots, hierarchical clustering was used to separate the studied objects (e.g., AMR genes, foodborne pathogens, and meats) so that the objects showing similar involvement with antimicrobial resistance were put into the same group. Initially treating each observation as a unique object, hierarchical clustering merged the two objects that were in the closest proximity to each other and, thus, had the highest similarity to form a new cluster. It then continued to merge close clusters until only one remained. The hierarchical clustering output a dendrogram, a tree diagram that represents different clusters and how they merge together. The vertical height of the link between two objects (i.e., foodborne pathogens, AMR genes, or meats) revealed their proximity and similarity, in that the smaller the height at which the two objects were merged, the more similar their involvement in antimicrobial resistance. In Figure 1C, for example, Genes E and F were the closest in proximity in the PCA plot and were detected in similar samples. They, thus, had the lowest linkage in the dendrogram. The dendrograms produced by the antimicrobial resistance data were analyzed, and significant outlier genes and foodborne pathogens were identified. Those outlier genes and foodborne pathogens were further confirmed from the PCA plot. Special attention should be paid to their indications of antimicrobial resistance for the following reasons: (1) they were of high occurring frequencies in the isolate samples; (2) the genes they indicated were involved in the resistance to various antimicrobials; (3) the pathogens they indicated carried various AMR genes. These genes and foodborne pathogens were used to generate the historical occurrence profiles of frequently occurring genes and foodborne pathogens over time. The profiles of absolute frequency were plotted for both AMR genes and pathogens to indicate the number of samples containing the selected gene or pathogen that was collected each year. In addition, the profiles of the relative frequency, which was calculated by the number of samples for a gene/pathogen in each year divided by the total number of samples obtained in that year, were also provided in this work so as to reduce bias resulting from the variation of sample numbers across different years. These graphs were then used to quantify the movement of AMR (upwards or downwards) by country based on the historical data from 2010 to 2019.

R, a programming language capable of performing statistical analysis and generating graphs [20], was used to create the analytical data set and perform the statistical analysis in this project. R offers many statistical and graphical packages and easily produces plots and calculations on arrays. For these reasons, it was well-suited for multivariate statistical analysis in this work.

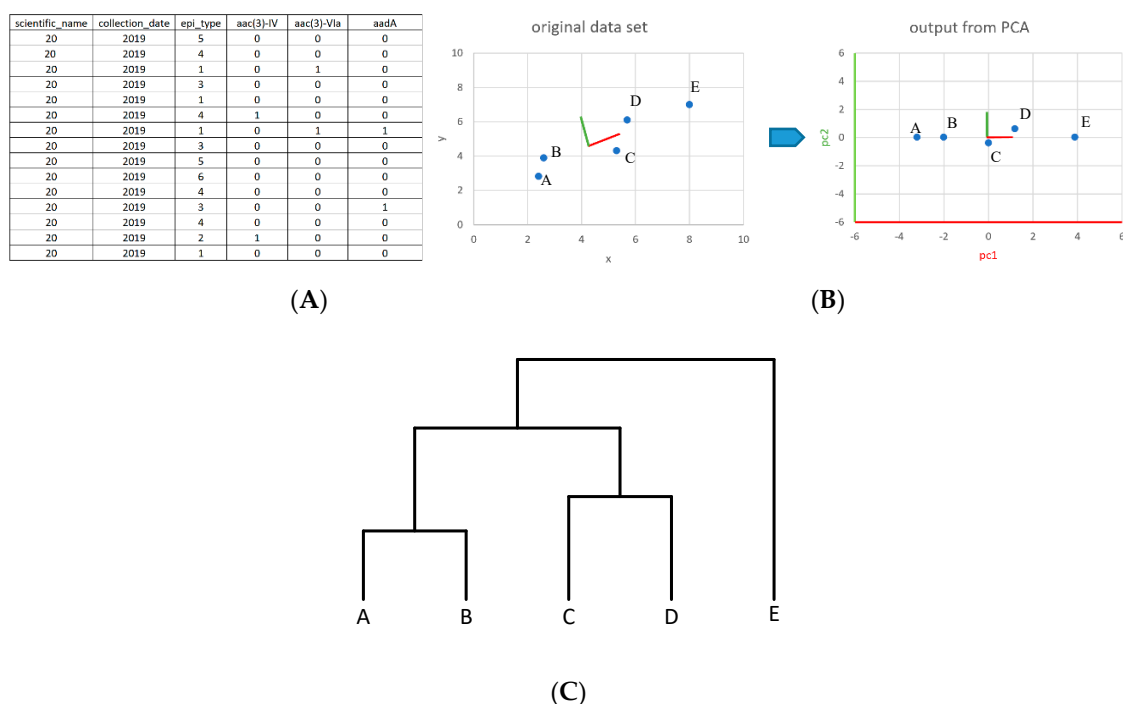


Figure 1. The general procedure of multivariate statistical analysis implemented in this study: (A) A portion of the data matrix extracted from NCBI Pathogen Detection Isolates Browser Database for the US; (B) principal component analysis was used to reduce the data matrix dimensions so that the data points (e.g., genes/pathogens/meats) can be visualized in a two dimensional space (e.g., the x - y space can be reduced to the PC1 space); (C) the hierarchical clustering approach was used to separate the data points on the basis of a two-dimensional space produced by principal component analysis. Simple diagrams are used in (A–C) to illustrate the general steps for principal component analysis and hierarchical clustering.

3. Results

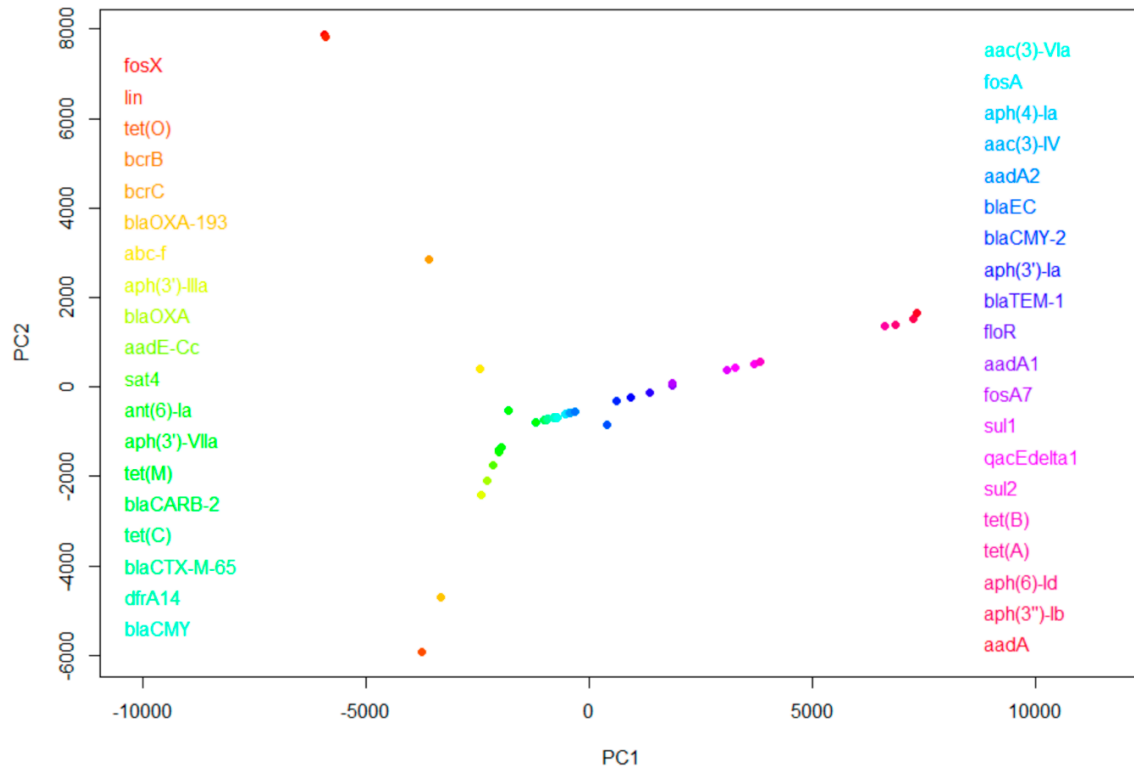
Data from the NPDIB was analyzed using PCA and hierarchical clustering in this section to identify: (1) the common highly-occurring antimicrobial-resistance genes in the studied countries: the US, the UK, South Africa, China, Australia, Canada, and Mexico; (2) the common highly-occurring foodborne pathogens carrying antimicrobial-resistance genes in each of the countries studied; (3) and the parallelism in antimicrobial resistance trends among the eight countries; (4) the similarities in the occurrence of AMR genes and foodborne pathogens in the selected eight countries.

3.1. Identification of Common Highly-Occurring AMR Genes in the Studied Countries

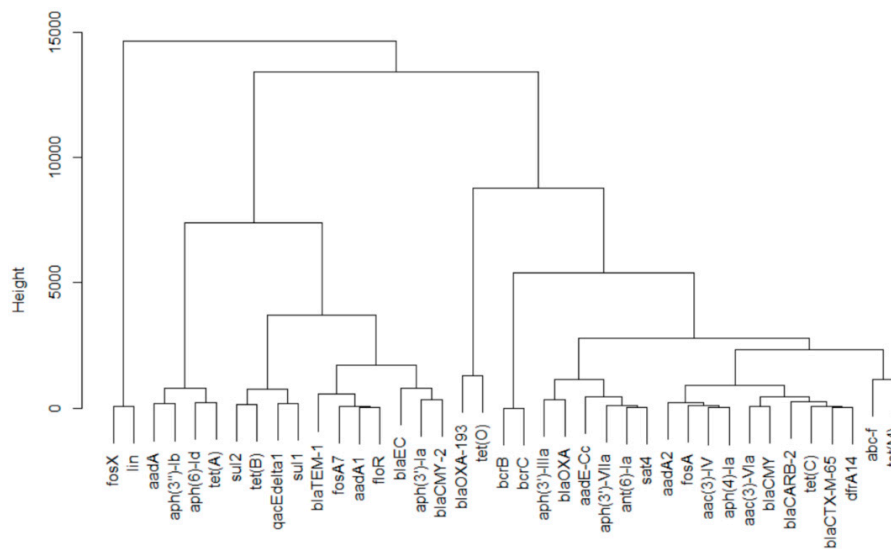
3.1.1. Common Frequently-Occurring Antimicrobial-Resistance Genes for Individual Countries for All Years

Highly-occurring AMR genes in each of the eight countries were identified using PCA and hierarchical clustering. For each country, the antimicrobial resistance gene data for different pathogens was organized into a matrix, each row representing one gene and each column representing one pathogen. The elements in the matrix recorded the number of samples in which the gene in the row was detected in the pathogen shown in the column. As an example, Figure 2 shows the results from such analysis for the United States' Data: Figure 2A projects the AMR genes onto the PC1~PC2 space and Figure 2B shows the hierarchical clustering constructed from the PCA projection. Genes in close proximity to each other in the principal component analysis plots had similar trends of occurrence and, thus, were grouped closely in the hierarchical clusters. As seen in the center of Figure 2A, the majority of genes—including *tet(C)*, *blaCMY*, and *fosA*—were grouped together in the principal component

analysis, and were correspondingly clustered in the same branch of the hierarchical cluster, highlighting their similar trends of occurrence. However, genes with unique trends of higher levels of occurrence were separated from the majority of cluster of genes in PCA. For example, *fosX* and *lin* stand out from the rest of the genes, revealing their significance. Through detailed hierarchical clustering, Figure 2B illustrates this significance with a clear visual: it groups *fosX* and *lin*, along with the other highly occurring genes, into their respective clusters.



(A)



(B)

Figure 2. An illustrative example of PCA and hierarchical analysis: (A) projections of antimicrobial genes onto the PC1~PC2 space based on the occurrence of these genes detected in foodborne pathogens; (B) clustering of the genes based on their projections in (A).

The AMR genes that stand out as the most frequently-occurring from the analysis of each country's data are listed in Table 1. In Table 1, the genes are sorted by their frequency, and it can be seen that (1) *tet(A)*, *aph(3'')-Ib*, *aph(6)-Id*, *blaEC*, *blaTEM-1*, *qacEdelta1*, *sul1*, *sul2*, and *aadA1* are the nine most common AMR genes in the majority of the studied countries; (2) South Africa's set of highly-occurring AMR genes are relatively unique; and (3) China harbors the greatest number of genes uncommon to the other seven countries. Supplementary Table S1 shows the mechanisms of the antimicrobial resistance associated with those frequently-occurring genes listed in Table 1.

Table 1. Highest occurring AMR genes sorted and color coded by their presence in each country (indicated by PC and hierarchical clustering for countries: the US, the UK, South Africa, China, Australia, Canada, Mexico, and Brazil). The genes are listed from the highest occurring frequency (commonly present in all eight countries) to the lowest frequency (commonly present in only one country). Different occurring frequencies are marked in different colors, e.g., dark green for AMR genes commonly present in eight countries, and light green for AMR genes commonly present in seven countries.

All Genes	US	UK	South Africa	China	Australia	Canada	Mexico	Brazil
<i>tet(A)</i>	<i>tet(A)</i>	<i>tet(A)</i>	<i>tet(A)</i>	<i>tet(A)</i>	<i>tet(A)</i>	<i>tet(A)</i>	<i>tet(A)</i>	<i>tet(A)</i>
<i>aph(3'')-Ib</i>	<i>aph(3'')-Ib</i>	<i>aph(3'')-Ib</i>		<i>aph(3'')-Ib</i>	<i>aph(3'')-Ib</i>	<i>aph(3'')-Ib</i>	<i>aph(3'')-Ib</i>	<i>aph(3'')-Ib</i>
<i>aph(6)-Id</i>	<i>aph(6)-Id</i>	<i>aph(6)-Id</i>		<i>aph(6)-Id</i>	<i>aph(6)-Id</i>	<i>aph(6)-Id</i>	<i>aph(6)-Id</i>	<i>aph(6)-Id</i>
<i>blaEC</i>	<i>blaEC</i>	<i>blaEC</i>	<i>blaEC</i>	<i>blaEC</i>	<i>blaEC</i>	<i>blaEC</i>		<i>blaEC</i>
<i>blaTEM-1</i>	<i>blaTEM-1</i>	<i>blaTEM-1</i>		<i>blaTEM-1</i>	<i>blaTEM-1</i>	<i>blaTEM-1</i>	<i>blaTEM-1</i>	<i>blaTEM-1</i>
<i>qacEdelta1</i>	<i>qacEdelta1</i>	<i>qacEdelta1</i>		<i>qacEdelta1</i>	<i>qacEdelta1</i>	<i>qacEdelta1</i>	<i>qacEdelta1</i>	<i>qacEdelta1</i>
<i>sul1</i>	<i>sul1</i>	<i>sul1</i>		<i>sul1</i>	<i>sul1</i>	<i>sul1</i>	<i>sul1</i>	<i>sul1</i>
<i>sul2</i>	<i>sul2</i>	<i>sul2</i>		<i>sul2</i>	<i>sul2</i>	<i>sul2</i>	<i>sul2</i>	<i>sul2</i>
<i>aadA1</i>	<i>aadA1</i>	<i>aadA1</i>		<i>aadA1</i>	<i>aadA1</i>	<i>aadA1</i>		<i>aadA1</i>
<i>tet(B)</i>	<i>tet(B)</i>	<i>tet(B)</i>	<i>tet(B)</i>			<i>tet(B)</i>		<i>tet(B)</i>
<i>aadA2</i>		<i>aadA2</i>		<i>aadA2</i>	<i>aadA2</i>		<i>aadA2</i>	<i>aadA2</i>
<i>fosX</i>	<i>fosX</i>	<i>fosX</i>			<i>fosX</i>	<i>fosX</i>	<i>fosX</i>	
<i>lin</i>	<i>lin</i>	<i>lin</i>			<i>lin</i>	<i>lin</i>	<i>lin</i>	
<i>aadA</i>	<i>aadA</i>	<i>aadA</i>	<i>aadA</i>					<i>aadA</i>
<i>floR</i>	<i>floR</i>			<i>floR</i>			<i>floR</i>	<i>floR</i>
<i>fosA7</i>	<i>fosA7</i>	<i>fosA7</i>					<i>fosA7</i>	<i>fosA7</i>
<i>aph(3')-Ia</i>	<i>aph(3')-Ia</i>			<i>aph(3')-Ia</i>	<i>aph(3')-Ia</i>			
<i>dfrA12</i>				<i>dfrA12</i>	<i>dfrA12</i>		<i>dfrA12</i>	
<i>blaCMY-2</i>	<i>blaCMY-2</i>	<i>blaCMY-2</i>						<i>blaCMY-2</i>
<i>bcrB</i>	<i>bcrB</i>					<i>bcrB</i>		
<i>bcrC</i>	<i>bcrC</i>					<i>bcrC</i>		
<i>abc-f</i>		<i>abc-f</i>				<i>abc-f</i>		
<i>sul3</i>				<i>sul3</i>	<i>sul3</i>			
<i>qnrB19</i>							<i>qnrB19</i>	<i>qnrB19</i>
<i>blaCTX-M-55</i>				<i>blaCTX-M-55</i>				
<i>blaOXA-193</i>	<i>blaOXA-193</i>							
<i>catA1</i>			<i>catA1</i>					
<i>mph(A)</i>				<i>mph(A)</i>				
<i>tet(O)</i>	<i>tet(O)</i>							
<i>aadA5</i>				<i>aadA5</i>				
<i>ble</i>				<i>ble</i>				
<i>bleO</i>				<i>bleO</i>				
<i>fosA3</i>				<i>fosA3</i>				
<i>mcr-1.1</i>				<i>mcr-1.1</i>				
<i>oqxA</i>				<i>oqxA</i>				
<i>oqxB</i>				<i>oqxB</i>				
<i>aac(2')-Ic</i>			<i>aac(2')-Ic</i>					
<i>blaA</i>			<i>blaA</i>					
<i>erm(37)</i>			<i>erm(37)</i>					
<i>cmlA1</i>					<i>cmlA1</i>			
<i>dfrA5</i>					<i>dfrA5</i>			
<i>qacL</i>					<i>qacL</i>			
<i>fosA</i>							<i>fosA</i>	

3.1.2. Historical Profiles of Highly-Occurring AMR Gene Occurrence from 2010 to 2019

One of the crucial aspects of the NPDIB data was the relative frequency of occurrence. Since there are a large number of AMR genes detected in each country, this section only focuses on the highly-occurring AMR genes that are listed in Table 1. The historical graphs of these genes should reflect the occurrence trend of the AMR genes in those countries. Genes that increased in occurrence frequency were identified and listed in Table 2 for all the eight countries studied.

Table 2. AMR genes with an increasing occurrence from 2010 to 2019 in the eight countries

US	UK	South Africa	China	Australia	Canada	Mexico	Brazil
<i>aadA1</i>	<i>aadA1</i>	<i>aph(3'')-Ib</i>	<i>aadA2</i>	<i>blaTEM-1</i>	<i>aadA1</i>	<i>aadA2</i>	<i>aph(3'')-Ib</i>
<i>blaOXA-193</i>	<i>aadA2</i>	<i>aph(6)-Id</i>	<i>blaTEM-1</i>	<i>dfrA5</i>	<i>tet(B)</i>	<i>blaTEM-1</i>	<i>aph(6)-Id</i>
<i>blaTEM-1</i>	<i>abc.f</i>	<i>blaEC</i>	<i>ble</i>	<i>sul2</i>		<i>dfrA12</i>	<i>blaCMY-2</i>
<i>floR</i>	<i>aph(3'')-Ib</i>	<i>blaTEM-1</i>	<i>floR</i>			<i>fosA</i>	<i>blaTEM-1</i>
<i>qacEdelta1</i>	<i>aph(6)-Id</i>	<i>dfrA5</i>	<i>mph(A)</i>			<i>qacEdelta1</i>	<i>floR</i>
<i>sul1</i>	<i>blaCMY.2</i>	<i>sul2</i>	<i>oqxA</i>			<i>sul1</i>	<i>fosA7</i>
<i>tet(A)</i>	<i>blaTEM.1</i>	<i>tet(A)</i>	<i>oqxB</i>			<i>sul2</i>	<i>sul2</i>
<i>tet(O)</i>	<i>fosA7</i>	<i>tet(B)</i>	<i>qacEdelta1</i>			<i>tet(A)</i>	<i>tet(A)</i>
	<i>qacEdelta1</i>		<i>sul1</i>				<i>tet(B)</i>
	<i>sul1</i>		<i>sul2</i>				
	<i>sul2</i>						
	<i>tet(A)</i>						

3.2. Identification of Common, Highly Occurring Pathogens in the Studied Countries

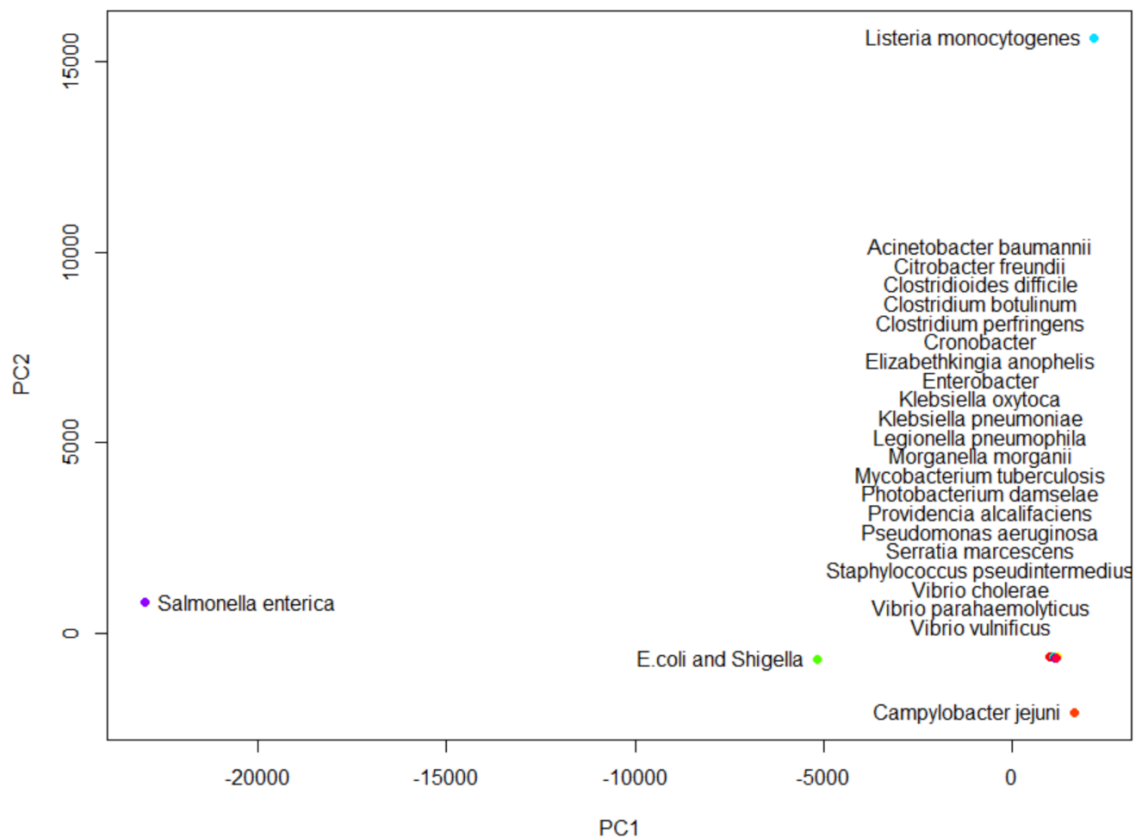
As in the previous section, the antimicrobial resistant gene data for each country was organized into a matrix. To isolate highly occurring antimicrobial resistant pathogens, the data was reorganized so that each row represented one pathogen and each column represented one gene. PCA plots and hierarchical clusters were then created for each country grouping pathogens with similar trends of resistance.

3.2.1. Common, Highly Occurring Pathogens of Individual Countries for All Years

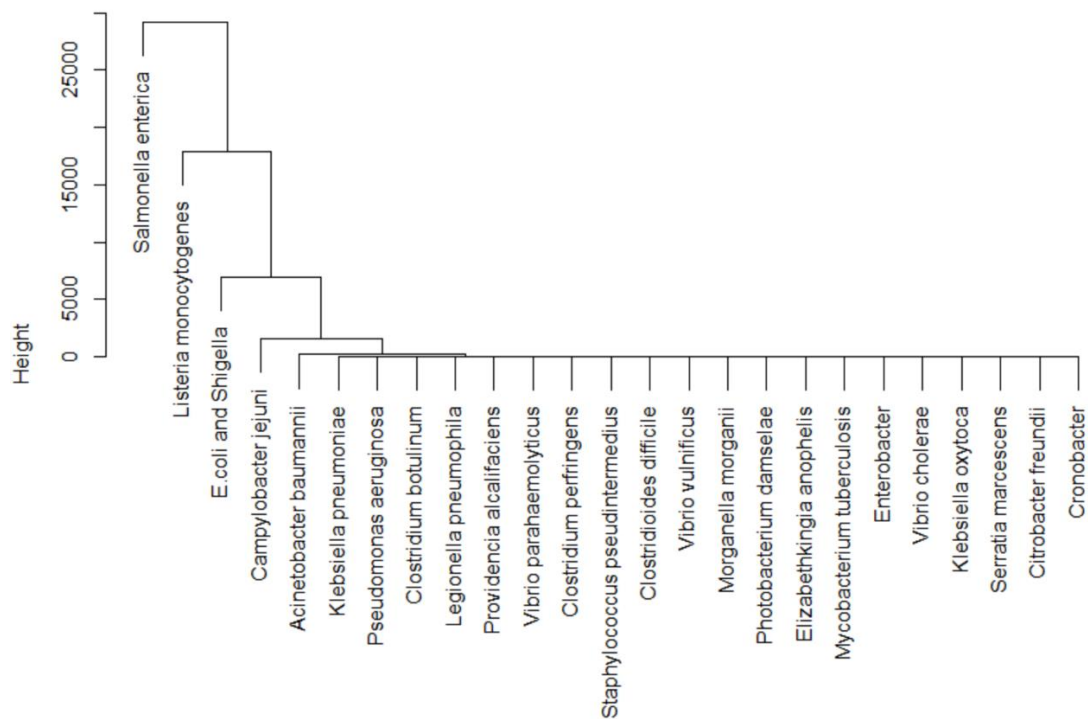
The key highly occurring pathogens in each of the eight countries were identified using PCA and hierarchical clustering. As an example, Figure 3 shows the results from such analysis for the United States' Data: Figure 3A projects the pathogens onto the PC1~PC2 space and Figure 3B shows the hierarchical clustering constructed from the PCA projection. As seen in Figure 3A, pathogens such as *Listeria monocytogenes* and *Salmonella enterica* stand out from the rest of the pathogens, revealing their notably high levels of occurrence. Figure 3B, through a detailed hierarchical clustering, illustrates this significance through a clear visual. It groups *Listeria monocytogenes* and *Salmonella enterica*, along with the other highly-occurring pathogens, into their respective clusters.

Through this process, it can be determined that the most highly-occurring pathogens in the US are: *Listeria monocytogenes*, *Salmonella enterica*, *E. coli* and *Shigella*, and *Campylobacter jejuni*. Each of these pathogens forms its own cluster alone and is extremely variant from the rest of the pathogens which form one big cluster apart from these four.

All the other countries were analyzed with PCA and hierarchical clustering in a manner similar to the analysis of the US. Highest occurring AMR pathogens are sorted and color-coded by their presence in each country (the US, the UK, South Africa, China, Australia, Canada, Mexico, and Brazil) in Table 3. *Salmonella enterica* and *E. coli* and *Shigella* were the most common highly-occurring pathogens. *Listeria monocytogenes* and *Klebsiella pneumoniae* were each highly-occurring in at least two of the countries studied, suggesting similarities and relations between specific partner countries. Similarly, as in the case for genes, China has the most unique AMR pathogens among the studied countries, *Enterobacter* and *Cronobacter*.



(A)



(B)

Figure 3. An illustrative example of PCA and hierarchical clustering: (A) projections of pathogens onto the PC1-PC2 space based on the resistance of antimicrobials; (B) clustering the pathogens based on their projections in (A).

Table 3. Highest occurring AMR pathogens sorted and color coded by their presence in each country (the US, the UK, South Africa, China, Australia, Canada, Mexico, and Brazil). Note: red indicates AMR pathogen presence in 7–8 out of the 8 countries; orange indicates AMR pathogen presence in 2–5 out of the 8 countries; and yellow indicates the AMR pathogen presence in 1 out of the 8 countries.

Pathogens	Countries							
	US	UK	South Africa	China	Australia	Canada	Mexico	Brazil
<i>E. coli and Shigella</i>	x	x	x	x	x	x	x	x
<i>Salmonella enterica</i>	x	x	x	x		x	x	x
<i>Listeria monocytogenes</i>	x	x			x	x	x	
<i>Klebsiella pneumoniae</i>				x				x
<i>Mycobacterium tuberculosis</i>			x					
<i>Campylobacter jejuni</i>	x							
<i>Pseudomonas aeruginosa</i>								x
<i>Enterobacter</i>				x				
<i>Cronobacter</i>				x				

3.2.2. Historical Profiles of Highly-Occurring Pathogen Occurrence between 2010 and 2018

The relative frequency graphs of highly-occurring foodborne pathogens present in the US were plotted together in Figure 4 for an example. It can be seen that the occurrence of *Listeria monocytogenes* and *Salmonella enterica* generally decreased in recent years in the US. The occurrence of *Campylobacter jejuni* decreased first, but then increased since 2014. The historical profile of *E. coli* and *Shigella* looks similar to the one of *Campylobacter jejuni* but with a generally decreasing trend in 2018 and 2019. The occurrence trend of *E. coli* and *Shigella* is thus not regarded as an increasing one. A similar approach was implemented to the other selected countries to find the pathogens with a generally increasing occurrence. The result is shown in Table 4 where most countries have at least a foodborne pathogen increasingly detected in the samples. This implies that certain foodborne pathogens will keep carrying those AMR genes over time and that there is still an urgent need to combat antimicrobial resistance issues throughout the world.

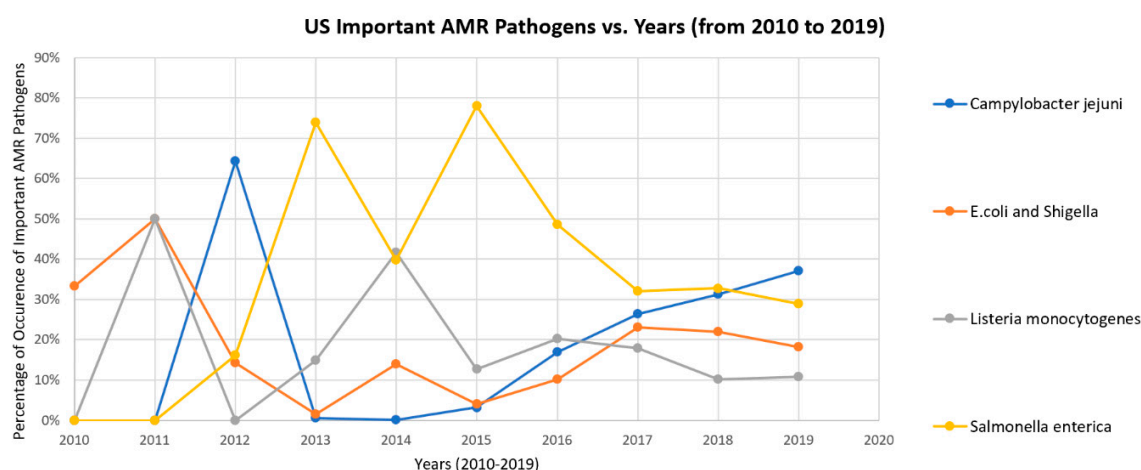


Figure 4. Frequency of highly-occurring pathogens over the years 2010–2019 in the US.

3.3. Major Meat Carriers Found in the Studied Countries

All the four major meats (beef, chicken, pork, and turkey) were studied through PCA and hierarchical clustering analysis of the AMR genes they carried for each of the countries studied. The hierarchical cluster diagrams for the US are provided in Figure 5 as an example. This diagram was used to rank the meat carriers in terms of their relative occurrence in carrying AMR genes in the range

of 1 to 4, 1 represents the most and 4 represents the least relative occurrence. For example, in the US, chicken was ranked as 1, beef as 2, and turkey and pork were ranked equally as 4. The individual ranking for each country and the total ranking for all the studied countries are provided in Table 5.

Table 4. Foodborne pathogens with an increasing occurrence in the eight selected countries

US	UK	South Africa	China	Australia	Canada	Mexico	Brazil
<i>Campylobacter jejuni</i>	<i>Listeria monocytogenes</i>	<i>E. coli</i> and <i>Shigella</i>	<i>Klebsiella pneumoniae</i>	<i>E. coli</i> and <i>Shigella</i>	<i>E. coli</i> and <i>Shigella</i>	<i>E. coli</i> and <i>Shigella</i>	<i>Salmonella enterica</i>
	<i>Salmonella enterica</i>		<i>Salmonella enterica</i>				

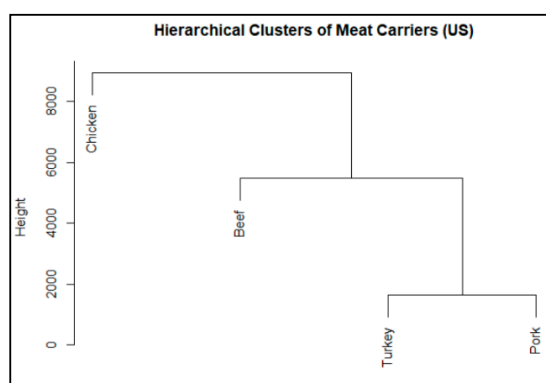


Figure 5. Hierarchical clusters of meat carriers are displayed for the US.

Table 5. Ranking of beef, chicken, pork, and turkey based on their prevalence in each country. The ranking is on a scale of 1 to 4, 1 represents the most prevalent carrier while 4 represents the least prevalent carrier.

Meats \ Countries	US	UK	South Africa	China	Australia	Canada	Mexico	Brazil	Total Rank
	Beef	2	4	No Data	4	1	1	1	4
Chicken	1	1	No Data	1	4	2	2	1	12
Pork	4	2	No Data	2	4	4	4	2	22
Turkey	4	4	No Data	4	4	4	4	4	28

From Table 5, a couple of trends can be identified: (1) overall rank of prevalent carriers are chicken, beef, pork, and turkey (with chicken as the most prevalent), which correlates well with the world per capita consumption of the different meats [21] and (2) beef and chicken are dominant in the US, Canada, and Mexico (three north American countries); pork and chicken are dominant in the UK, China, and Brazil; and beef is dominant in Australia with the other three meats being equal.

3.4. Identification of Similarity in Antimicrobial Resistance Occurrence in Studied Countries

Analysis was performed on the data to study the similarity in antimicrobial resistance in the eight selected countries. The data was first organized into a matrix in which each row represented one country and each column represented each highly-occurring AMR gene from Table 1. PCA (Figure 6A) and hierarchical clustering (Figure 6B) revealed the similarities in highly-occurring AMR gene frequencies and occurrences in the eight countries. In hierarchical clustering, countries connected together at a lower vertical level exhibited a greater similarity in the occurrences of AMR genes. Therefore, a few clusters are notable in Figure 6B: (1) the US and the UK, the most developed countries, are clustered together; (2) China and Brazil, high developing countries, are clustered together; (3) Canada and Mexico’s cluster is vertically very close to the US and the UK; and (4) South Africa is the most vertically unique and varying country.

Furthermore, PCA and hierarchical clustering analysis highlighted the similarity of the highly-occurring resistant pathogen occurrence across the eight countries studied. Figure 7A shows the projections of countries onto the PC1–PC2 space based on the occurrence of the foodborne pathogens shown in Table 3. Figure 7B displays the clustering of the countries based on their projections in Figure 7A. South Africa is again the most vertically unique, as its height in the hierarchical cluster is significantly higher than those of the other seven countries. It therefore exhibits the most unique and different patterns of resistance in its pathogens. Among the other seven countries, Brazil and Mexico were coupled together, and Canada and China were clustered together. As found previously, the US and UK also formed a clustered pair. Australia was located close to Canada and China, but was not in a cluster pair. Figure 7 is different from Figure 6, as the AMR genes can be carried by multiple pathogens. In other words, AMR genes and specific pathogens were not necessarily correlated in their occurrence. For example, the AMR genes can be carried by the pathogens not shown in Table 3, on the basis of which Figure 7 was generated.

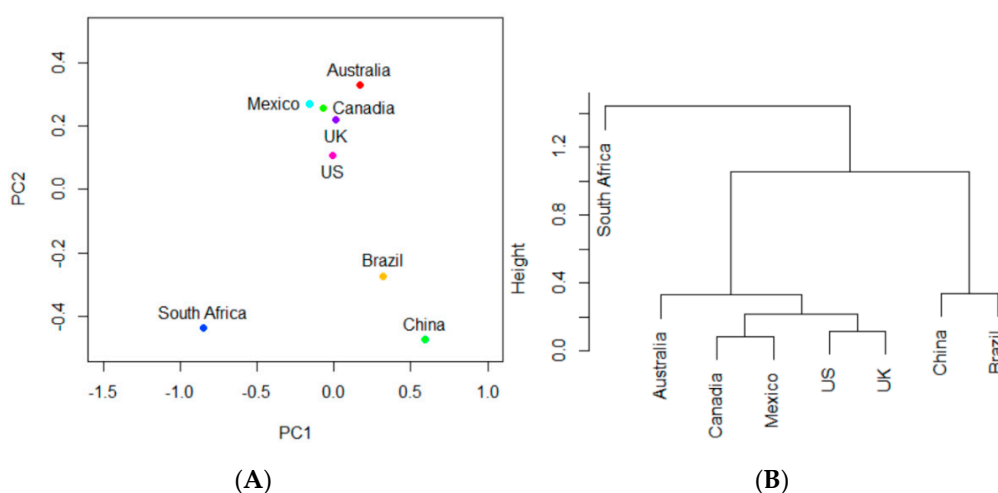


Figure 6. (A) Projections of South Africa, Mexico, Canada, the UK, the US, Brazil, China, and Australia onto the PC1-PC2 space based on the occurrence of AMR genes shown in Table 1; (B) clustering the countries based on their projections in (A).

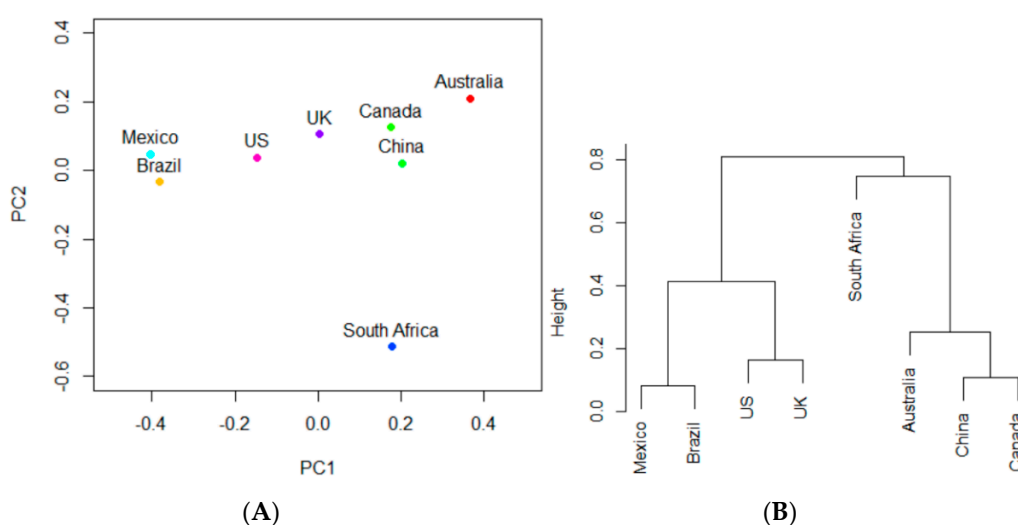
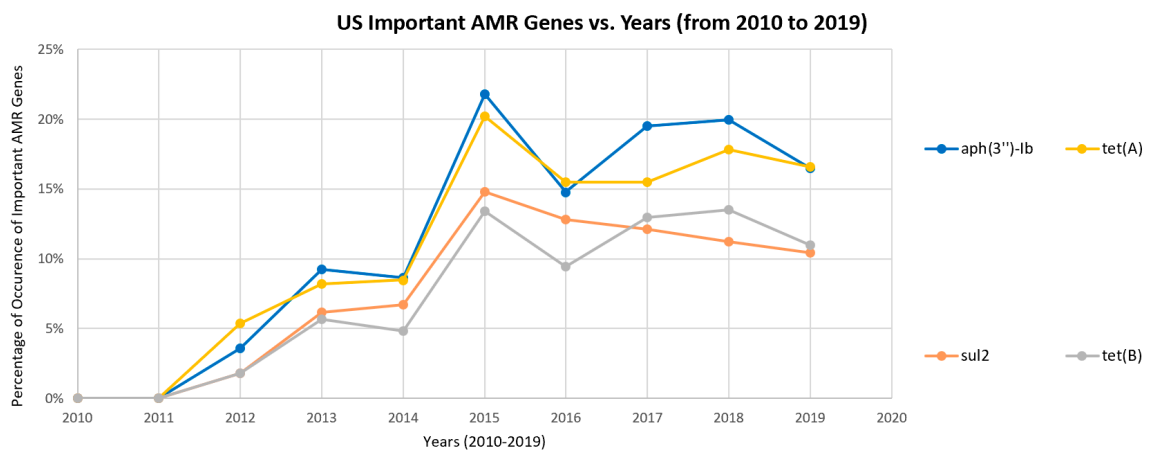


Figure 7. (A) projections of South Africa, Mexico, Canada, the UK, the US, Brazil, China, and Australia onto the PC1-PC2 space based on the occurrence of foodborne pathogens shown in Table 3; (B) clustering the countries based on their projections in (A).

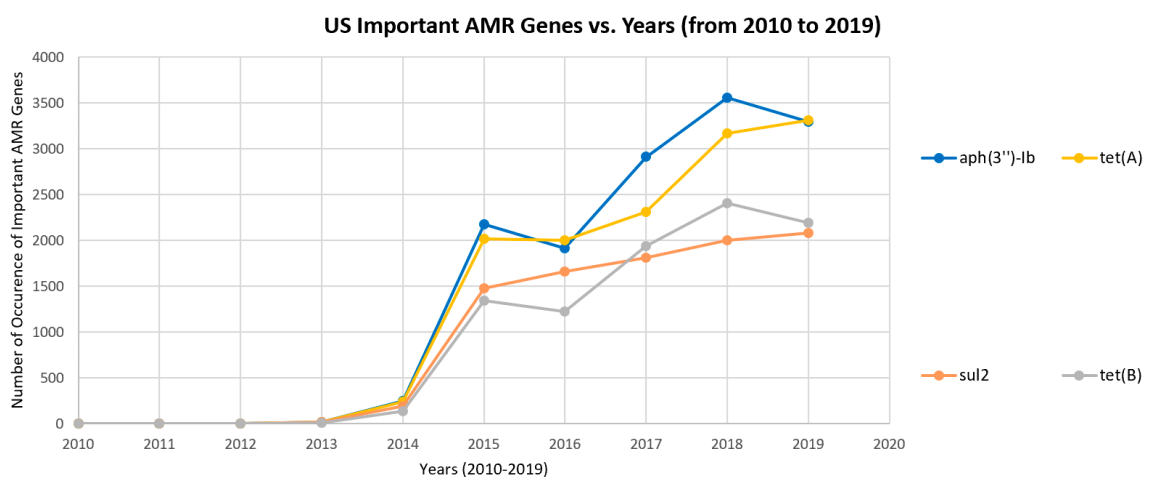
4. Discussion

4.1. Absolute Occurrence versus the Relative Occurrence for Studying the Time Trend of Antimicrobial Resistance

Figure 8 shows the occurrence profiles of several representative genes occurring with increasing frequency. The relative occurrence, instead of the absolute occurrence, was preferred for studying the trend of antimicrobial resistance. This is because the number of available samples in early years (e.g., from 2010 to 2013) is relatively low compared to later years. The rise magnitudes shown in the absolute frequency profiles in Figure 8B are much higher than those shown in Figure 8A, even with only seven months of data available for 2019. Figure 8B indicates a strong increase in the absolute frequencies of AMR genes over time. However, this may be due to the fact that fewer samples were collected in the early years because of the limited interest from researchers and because fewer facilities were available for sample collection. The modern understanding of the rise of AMR may be inflated by the skewed absolute occurrences of resistance. However, the number of AMR genes showing a general increase in their relative occurrences from 2011 still implies the existence of a potential rise in the relative occurrence of AMR genes, regardless of the influence of the increased frequency of measurement.



(A)



(B)

Figure 8. Frequencies of occurrence among the important AMR genes *aph(3'')-Ib*, *tet(A)*, *sul2*, and *tet(B)* between 2010 and 2019 in the US with (A) normalized and (B) absolute numbers of samples that contain the AMR genes.

4.2. Antimicrobial Usage and the Rising Overall Trend of AMR Genes

Figure 8 indicates an absolute frequency profile that shows a stronger increasing trend than the relative one from 2010 to 2019. In the eight studied countries, 43 different antimicrobial resistant genes were deemed highly-occurring. This upward trend shown in the absolute frequency profile occurs because of the following potential factors: (1) the increase in reports of antimicrobial resistance isolates because of developments in microbial measurement technology; (2) the population increase in those countries; (3) the increased populations with lower immune systems (e.g., people with obesity or nosocomial infections); and (4) the growth of antimicrobial usage in the livestock industry.

In particular, many countries are globalizing, joining the international meat trade web. Propelled by the rapid rise in participating countries, the meat industry is expected to grow to \$1.5 T by 2022. This amount doubles the number from 2016, \$714 B. The expansion of the livestock industry has led to a greater exchange of animal products—and, thus, their AMR genes and resistant pathogens—and an increased administration of excessive antibiotics in livestock. The rise in specific antimicrobial resistant genes in various countries may be correlated with the excessive administration of antibiotics to livestock. Such a relationship is most strongly seen in the UK, in which nearly all resistant genes were resistant to either aminoglycoside antibiotics, beta-lactam related antibiotics (cephalosporin, penam), sulfonamide antibiotics, or tetracycline antibiotics. Interestingly, this corresponds with the top four antibiotics used in livestock in the UK. While the use of antibiotics as growth promoters was banned in the UK in the early 2000s, antibiotics were still used in other sectors of the livestock industry, such as for prophylaxis. Of the antibiotics used on livestock in the UK, 37% was tetracycline, 26% penicillin (which utilizes beta lactam resistance), 11% sulfonamide, and 8% aminoglycosides [22]. These four antibiotics add up to 82% of UK antibiotic usage in livestock, implying a potential correlation between antibiotic usage in livestock and the rising trend of AMR genes [22]. On the other hand, the relative frequency profile (Figure 8A) does not show as strong an increasing trend as the absolute frequency profile (Figure 8B). This indicates that the aforementioned factors other than the growth of antimicrobial usage may be contributing to the increase in the absolute frequency profile.

Of the 43 highly-occurring antimicrobial resistant genes, only one had significant occurrences in all 8 of the studied countries—*tet(A)*. As the most common AMR gene, *tet(A)*, encoding an antibiotic efflux, shows resistance to at least 21 different antibiotics, including some of the critically important antibiotics shown in [23]. In addition to *tet(A)*, eight other highly-occurring antimicrobial resistant genes were common in the studied countries, showing significant appearances in all but one or two countries—*aph(3'')-Ib*, *aph(6)-Id*, *blaEC*, *blaTEM-1*, *qacEdelta1*, *sul1*, *sul2*, and *aadA1*. All of the nine common highly occurring genes are found in *E. coli*, *Shigella*, and *Salmonella Enterica*. Our data revealed that *Salmonella enterica*, *E. coli*, and *Shigella* were the most common highly-occurring foodborne pathogens, showing significant appearances in most of the eight countries. This suggests that these pathogens may be prevalent because of their access to and expression of these AMR genes. According to the FDA, *Salmonella* and *E. coli* represent some of the most common causes of foodborne disease, and cause millions of illnesses every year [24].

4.3. Geographic and Trade Relations between AMR Genes of Different Countries

Of the eight countries studied, South Africa hosted the least number of common highly-occurring AMR genes. While all of the other countries had significant appearances of at least seven of the nine highly-occurring genes, South Africa had only two—*tet(A)* and *blaEC*. Rather, it had the most unique set of highly-occurring resistant genes with the least increase in AMR gene occurrence over time. Similarly, in both of the hierarchical clusters of the countries based on their foodborne pathogens and AMR genes (Figures 7 and 8), South Africa was consistently separated from the other seven countries. This uniqueness may in part relate to a lack of access to technology and limited means for reporting gene occurrences. These limitations are further discussed in the section below, Limitations and Future Work. However, it may also be correlated to the limited involvement of South Africa in the global livestock market. Of the eight countries studied, South Africa has the lowest rate of importing and exporting of

animal products and, thus, is the most isolated from world trade and the subsequent exchange of AMR genes and foodborne pathogens. South Africa, thus, had quite different foodborne pathogens and genes compared to the other seven countries. While similarly isolated geographically, Australia is one of the most developed countries in the world. Moreover, it is a key player in the global meat industry as the third largest exporter of beef and the 55th most populated country in the world [25]. Although it implements the most conservative antimicrobial policy, it exhibits trends in abundance and time of occurrence of AMR genes similar to the trends in other countries that are also heavily involved in the global meat industry. This suggests that the occurrence of AMR genes (and thus antimicrobial resistance) is most likely related to the involvement of a country in the global livestock trade market and its trade relations with other countries, rather than its geographic location or its antimicrobial policy. Another example of this relationship between a country and its trade patterns is the clustering of the US and the UK together in both Figures 7 and 8. This may be associated with their close trade relations. For example, the US and the UK both underwent outbreaks of *Salmonella enterica* in recent years, specifically in 2013 and 2015 [26].

4.4. Limitations

The NPDIB database contains data from the last 30 years. During this time, this data has been affected by increasing technological accessibility; thus, an increase in the number of gene occurrences could possibly be attributed to a growing means of reporting it. In many countries, specifically developing ones, the lack of technology or uniform data collection policy may mean that less AMR genes and foodborne pathogens are reported. Moreover, since not every instance of every AMR gene was reported in each country, it is important to note that the database only contains a limited number of samples that may not be able to completely reflect the antimicrobial resistance status of each selected country. This work addressed these limitations by selecting the countries with the largest amounts of samples in various continents. It is hoped that economic and technological development can overcome these limitations and that more AMR samples can be obtained for countries all over the world in the future.

5. Conclusions

The NPDIB is a database containing genomic data for AMR genes and foodborne pathogen samples. The data has been provided from countries all over the world beginning in 2010. Data from this database during the period of 2010 to 2019 was analyzed to determine the occurrence of AMR genes over eight countries from various continents. Programs developed through the R-program performed principal component analysis and hierarchical clustering of high-dimensional data to identify genes and foodborne pathogens most involved in antimicrobial resistance. The results indicated that (1) *tet(A)*, *aph(3'')-Ib*, *aph(6)-Id*, *blaEC*, *blaTEM-1*, *qacEdelta1*, *sul1*, *sul2*, and *aadA1* were the 9 most common of 43 identified highly-occurring AMR genes; (2) *Salmonella enterica* and *E. coli* and *Shigella* were the most common highly-occurring foodborne pathogens; (3) chicken was the most prevalent meat carrier of AMR; and (4) South Africa had the most statistically unique resistance pattern. The large number of highly occurring AMR genes and the wide dissemination of many of these genes implies a potential rise of antimicrobial resistance. The spread of many highly occurring genes to numerous geographically distant countries with varying antimicrobial policies highlights the role of the global livestock trade in the transfer of AMR genes and foodborne pathogens. The identification of the nine globally significant highly-occurring genes, highly-occurring pathogens, and most prevalent meat carrier offer valuable insight for future research and development of new solutions for antimicrobial resistance.

Supplementary Materials: The following are available online at <http://www.mdpi.com/1660-4601/17/2/472/s1>, Table S1: Resistance Mechanisms and Drug Class of Highly-occurring AMR genes.

Author Contributions: Conceptualization, Q.J. and Z.H.; Methodology, K.Y., A.W., M.F., A.W., K.C., Q.J. and Z.H.; Formal analysis, K.Y., A.W., M.F., A.W. and K.C.; Data curation, Q.J. and Z.H.; Writing—original draft

preparation, K.Y., A.W., M.F., A.W. and K.C.; Writing—Review and editing, Q.J. and Z.H.; Supervision, Q.J. and Z.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: We appreciate Emily Carson from the Writing Center of Villanova University in polishing our manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Velez, R.; Sloand, E. Combating antibiotic resistance, mitigating future threats and ongoing initiatives. *J. Clin. Nurs.* **2016**, *25*, 1886–1889. [[CrossRef](#)] [[PubMed](#)]
2. Michael, G.B.; Freitag, C.; Wendlandt, S.; Eidam, C.; Feßler, A.T.; Lopes, G.V.; Kadlec, K.; Schwarz, S. Emerging issues in antimicrobial resistance of bacteria from food-producing animals. *Futur. Microbiol.* **2015**, *10*, 427–443. [[CrossRef](#)] [[PubMed](#)]
3. Davies, J.; Davies, D. Origins and Evolution of Antibiotic Resistance. *Microbiol. Mol. Boil. Rev.* **2010**, *74*, 417–433. [[CrossRef](#)] [[PubMed](#)]
4. Solomon, S.L.; Oliver, K.B. Antibiotic resistance threats in the United States: Stepping back from the brink. *Am. Fam. Phys.* **2014**, *89*, 938–941.
5. Friedrich, A.W. Control of hospital acquired infections and antimicrobial resistance in Europe: The way to go. *Wien. Med. Wochenschr.* **2019**, *169* (Suppl. 1), 25–30. [[CrossRef](#)] [[PubMed](#)]
6. Roberts, R.R.; Hota, B.; Ahmad, I.; Ii, R.D.S.; Foster, S.D.; Abbasi, F.; Schabowski, S.; Kampe, L.M.; Ciavarella, G.G.; Supino, M.; et al. Hospital and Societal Costs of Antimicrobial-Resistant Infections in a Chicago Teaching Hospital: Implications for Antibiotic Stewardship. *Clin. Infect. Dis.* **2009**, *49*, 1175–1184. [[CrossRef](#)] [[PubMed](#)]
7. Lekshmi, M.; Ammini, P.; Kumar, S.; Varela, M.F. The Food Production Environment and the Development of Antimicrobial Resistance in Human Pathogens of Animal Origin. *Microorganisms* **2017**, *5*, 11. [[CrossRef](#)] [[PubMed](#)]
8. Van Boeckel, T.P.; Brower, C.; Gilbert, M.; Grenfell, B.T.; Levin, S.A.; Robinson, T.P.; Teillant, A.; Laxminarayan, R. Global trends in antimicrobial use in food animals. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 5649–5654. [[CrossRef](#)] [[PubMed](#)]
9. Delgado-Suárez, E.J.; Selem-Mójica, N.; Ortiz-Lopez, R.; Gebreyes, W.A.; Allard, M.W.; Barona-Gómez, F.; Rubio-Lozano, M.S. Whole genome sequencing reveals widespread distribution of typhoidal toxin genes and VirB/D4 plasmids in bovine-associated nontyphoidal Salmonella. *Sci. Rep.* **2018**, *8*, 9864. [[CrossRef](#)] [[PubMed](#)]
10. Davison, J. Genetic Exchange between Bacteria in the Environment. *Plasmid* **1999**, *42*, 73–91. [[CrossRef](#)] [[PubMed](#)]
11. Liu, L.; Chen, X.; Skogerbø, G.; Zhang, P.; Chen, R.; He, S.; Huang, D.-W. The human microbiome: A hot spot of microbial horizontal gene transfer. *Genomics* **2012**, *100*, 265–270. [[CrossRef](#)] [[PubMed](#)]
12. Ceric, O.; Tyson, G.H.; Goodman, L.B.; Mitchell, P.K.; Zhang, Y.; Prarat, M.; Cui, J.; Peak, L.; Scaria, J.; Antony, L.; et al. Enhancing the one health initiative by using whole genome sequencing to monitor antimicrobial resistance of animal pathogens: Vet-LIRN collaborative project with veterinary diagnostic laboratories in United States and Canada. *BMC Veter-Res.* **2019**, *15*, 130. [[CrossRef](#)] [[PubMed](#)]
13. McMillan, E.A.; Gupta, S.K.; Williams, L.E.; Jové, T.; Hiott, L.M.; Woodley, T.A.; Barrett, J.B.; Jackson, C.R.; Wasilenko, J.L.; Simmons, M.; et al. Antimicrobial Resistance Genes, Cassettes, and Plasmids Present in Salmonella enterica Associated with United States Food Animals. *Front. Microbiol.* **2019**, *10*, 832. [[CrossRef](#)] [[PubMed](#)]
14. Zhang, N.; Liu, E.; Tang, A.; Ye, M.C.; Wang, K.; Jia, Q.; Huang, Z. Data-Driven Analysis of Antimicrobial Resistance in Foodborne Pathogens from Six States within the US. *Int. J. Environ. Res. Public Heal.* **2019**, *16*, 1811. [[CrossRef](#)] [[PubMed](#)]
15. Qin, S.J. Statistical process monitoring: Basics and beyond. *J. Chemom.* **2003**, *17*, 480–502. [[CrossRef](#)]
16. Kourti, T. Application of latent variable methods to process control and multivariate statistical process control in industry. *Int. J. Adapt. Control. Signal Process.* **2005**, *19*, 213–246. [[CrossRef](#)]

17. Arnau, V.; Mars, S.; Marin, I. Iterative cluster analysis of protein interaction data. *Bioinformatics* **2005**, *21*, 364–378. [[CrossRef](#)] [[PubMed](#)]
18. Wang, X.; Smith, K.; Hyndman, R.; Smith-Miles, K. Characteristic-Based Clustering for Time Series Data. *Data Min. Knowl. Discov.* **2006**, *13*, 335–364. [[CrossRef](#)]
19. Bar-Joseph, Z.; Demaine, E.D.; Gifford, D.K.; Srebro, N.; Hamel, A.M.; Jaakkola, T.S. K-ary clustering with optimal leaf ordering for gene expression data. *Bioinformatics* **2003**, *19*, 1070–1078. [[CrossRef](#)] [[PubMed](#)]
20. Ihaka, R.; Gentleman, R. R: A Language for Data Analysis and Graphics. *J. Comput. Graph. Stat.* **1996**, *5*, 299.
21. USDA. Per Capita Consumption of Poultry and Livestock, 1960 to Forecast 2020. 2019. Available online: <https://www.nationalchickencouncil.org/about-the-industry/statistics/per-capita-consumption-of-poultry-and-livestock-1965-to-estimated-2012-in-pounds/> (accessed on 16 September 2019).
22. Farm Antibiotics. Available online: <https://www.farmantibiotics.org/science-facts/antibiotic-infographs/> (accessed on 16 September 2019).
23. WHO. Highest Priority Critically Important Antimicrobials. 2019. Available online: <https://www.who.int/foodsafety/cia/en/> (accessed on 16 September 2019).
24. FDA. Most Common Foodborne Illness. 2019. Available online: <https://www.fda.gov/files/food/published/Most-Common-Foodborne-Illnesses-%28PDF%29.pdf> (accessed on 16 September 2019).
25. Australia, M.A.L. Know the Numbers: State of the Industry Report 2018. 2018. Available online: <https://www.mla.com.au/prices-markets/market-news/know-the-numbers-state-of-the-industry-report-2018/> (accessed on 16 September 2019).
26. CDC. Reports of Salmonella Outbreak Investigations from 2015. 2017. Available online: <https://www.cdc.gov/salmonella/outbreaks-2015.html> (accessed on 16 September 2019).



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).