https://dx.doi.org/10.4314/ijs.v21i3.2

Ife Journal of Science vol. 21, no. 3 (2019)

16S rRNA GENE SEQUENCING: A PRACTICAL APPROACH TO CONFIRMING THE IDENTITY OF FOOD BORNE BACTERIA

[*1]Chukwu, E.E., [2]Nwaokorie, F.O., [3]Coker, A.O., [4]Avila-Campos, M.J., [5]Ogunsola, F.T.

[1]Diarrhoea, Immunology and Parasitology Division, Microbiology Department, Nigerian Institute of Medical Research Yaba, Lagos, Nigeria;

[2]Department of Medical Laboratory Sciences, College of Medicine, University of Lagos, Nigeria;

[3]Department of Medical Microbiology and Parasitology, Babcock University Teaching Hospital, Ilishan, Remo, Ogun state, Nigeria.

[4]Anaerobe Laboratory, Department of Microbiology, Institute of Biomedical Sciences, University of São Paulo-USP, Av. Professor Lineu Prestes, 05508-900, São Paulo, SP, Brazil.

[5]Department of Medical Microbiology and Parasitology, College of Medicine, University of Lagos, Nigeria.

*Corresponding Author: emeldachukwu123@gmail.com

(Received: 27th May, 2019; Accepted: 12th August, 2019)

ABSTRACT

One of the main challenges for clinical practice and public health surveillance is rapid and accurate identification of food borne pathogens. In Nigeria, the use of sophisticated molecular tools for foodborne pathogens detection has not been extensively adopted. Bacterial strains isolated from food products sold in Lagos, Nigeria, were subjected to phenotypic techniques and 16S rRNA gene sequencing for specie level identification. Bacterial identification was performed using colonial morphology, Gram staining, conventional biochemical tests, and 16S rRNA gene sequencing. The pairwise sequence similarities between all the bacterial species with official names were analysed. Out of the 30 bacterial strains isolated and sequenced, 29 (96.7%) strains had nucleotide identity to known bacterial species in the GenBank. One (3.3%) sequence contained a large number of undetermined nucleotides and was not associated with specific strain identification. This technique was efficient showing 96.6% of the isolates being identified to the genus level and 93.1% being identified to the species level. The strains that were not identified due to low similarity levels were assigned phylogenetic positions, suggesting that they may belong to new taxa. This study demonstrates that 16S rDNA sequence analysis is more accurate and objective for identification of foodborne pathogens and also offers the possibility of rapidly recognizing yet undescribed taxa.

Keywords: Foodborne pathogens; Microbial detection; Phylogenetic relationship; Phenotypic identification; 16S rRNA gene sequencing

INTRODUCTION

Foodborne pathogens constitute a global menace and are the leading causes of illness in developing as well as developed countries (Joshy *et al*., 2008). In developed countries, food borne pathogens are responsible for millions of cases of infectious gastrointestinal diseases each year, costing billions of dollars in medical care and social costs (Fratamico *et al*., 2005, Moffatt *et al*., 2011, Scallan *et al*., 2011,). These pathogens consist of bacteria, viruses, parasites and fungi (Bamaiyi, 2011). However, bacteria foodborne pathogens are the more common cause of illness, and the implicated pathogens are *Campylobacter jejuni* (Humphrey *et al*., 2007), Clostridial species especially *C. perfringens* and *C. botulinum* (Moffatt *et al*., 2011), *Salmonella* spp. (Tribe *et al*., 2002, Bell *et al*., 2010) and *Escherichia coli* O157:H7 and 0104: H4 (Bamaiyi, 2011). The surveillance of food borne illnesses outbreak is fairly established in developed countries but in spite of that only less than 10% are recorded in official statistics and this does not represent the true burden of the disease (Osterholm, 2011). In the United States, the Centers for Disease Control and Prevention (CDC) estimated there were 47.8 million food borne illnesses per year (Scallan *et al*., 2011, CDCP, 2017). Unfortunately, the problem appears more crucial in developing countries, with less than 1% of the cases reported (WHO, 2006). There is a dearth of information concerning food borne illnesses in various developing countries, including Nigeria. It is therefore important that rapid and accurate technique should be used in the identification of these pathogens in these countries to promote public health. Presently, one of the major challenges for clinical practice and

public health surveillance is rapid and accurate identification of food borne pathogens. These clinical challenges have resulted in misdiagnosis and inappropriate antibiotic administration which have contributed to a large extent to the development of resistant microorganisms (Larson, 2007; Hawkey, 2008). Several studies have demonstrated that rapid and appropriate diagnosis followed by adequate antibiotic treatment significantly improves patient outcomes, particularly in the case of an outbreak (Iregui *et al*., 2002; Shorr *et al*., 2011). Despite advances in technology, identification and profiling of microbial species by the majority of public health and hospital microbiology laboratories in Nigeria is largely reliant upon phenotypic and biochemical techniques. Such approaches are time-consuming often considerably more in the case of anaerobes and other fastidious organisms such as *Clostridium* species. Hence conventional method of identification is rigorous and frequently fails to produce relevant data within the crucial timeline to permit rapid and appropriate therapeutic decisions to be made (Bouvet *et al*., 2014; Rossi-Tamisier *et al*., 2015). In view of the fact that many diverse species can share the same biochemical phenotypes; there has been debate on whether conventional approaches can correctly distinguish between isolates (Srinivasan *et al*., 2015). On the other hand, molecular testing allows for highly specific and sensitive identification of a large number of pathogens from clinical isolates as well as directly from clinical specimens (Lehmann *et al*., 2008; Barbut *et al*., 2011).

The 16S rRNA gene which is ubiquitous to the members of the prokaryotes is a 1500 base pair gene coding for an RNA that is part of the 30S ribosomal subunit. It has conserved and variable sequence regions that is important for determination of both close and distant phylogenetic relationships. These characteristics allow the use of 16S rRNA in the assignment of genetic relationship at the genus and in many cases at the species level (Conlan *et al*., 2012; Fettweis *et al*., 2012). Thanks to ever-expanding databases of sequence information, 16S rRNA gene sequencing has become a useful tool for bacterial identification (Srinivasan *et al*., 2015).

Several studies have evaluated the usefulness of 16S rRNA gene sequencing for clinical microbiology (Woo *et al*., 2005; Werner *et al*., 2012; Srinivasan *et al*., 2015). In an era where DNA sequencing is rapidly becoming a commodity, Nigeria should not be left behind in this evolving science. Unfortunately, in Nigeria, bacteria identification is still largely dependent on routine phenotypic identification. The goal of this study therefore, was to ascertain accuracy of routine microbiological technique for identification of a range of foodborne pathogens as compared to 16S rRNA gene sequencing.

MATERIALS AND METHODS

Bacterial Strains

The food products used for this study included raw meat and meat products, raw milk (*fura*), yoghurts, vegetables and canned foods randomly purchased from open markets, cafeterias and supermarkets in Lagos. Food samples were collected into sterile plastic bags and transported to the laboratory for processing within 2 h of collection. One gram of the food samples were inoculated into duplicate Robertson cooked meat broth for enrichment and plated after 24 h onto either Trypticase soy agar, 5% sheep blood agar or chocolate agar as the case maybe and incubated aerobically at 37 °C. The broth was simultaneously inoculated into fastidious anaerobic agar and incubated anaerobically at 37 °C. Thirty bacterial strains isolated from food products sold in Lagos state, Nigeria were used to methodically evaluate the ability of routine microbiological techniques as compared with 16S rRNA gene sequencing for identification at species level of foodborne pathogens.

Phenotypic Identification

Phenotypic identification was performed using morphological appearance, Gram staining, pigment production and biochemical tests. The number of biochemical tests performed depended upon the ease of characterization and biochemical activity of the organism isolated. Biochemical testing done included but were not limited to the following: lecithinase, urease, motility, indole, lipase, nitrate reduction, oxidase, catalase, iron uptake, tolerance to 5% sodium chloride, hydrolysis of esculin, acid phosphatase,

citrate test, polymyxin B and ciprofloxacin inhibition, and sugar fermentation (glucose, maltose, sucrose, lactose, fructose, mannitol, inositol, and sorbitol). Growth at temperatures of 25, 28, 31, 37, and 42 °C; pigmentation; growth rates, spore formation, hemolysis on blood agar and colony characteristics were also assessed.

The isolates were stored, in triplicate, in skimmed milk, thioglycollate broth and brain heart infusion broth at – 80 °C for further analysis.

DNA Extraction/ PCR Amplification

Bacterial DNA was extracted using GeneJET Genomic DNA Purification kit (Thermo Scientific #K0722 Lot 00232256). A 1,400-bp 16S rRNA gene, was amplified from all 30 bacterial DNA as described by Wilson *et al*., (1990), using the primers 27F and 1492R (F 5`-AGAGTTTGATCCTGGCTCAG-3` and R 5`-ACGGCTACCTTGTTACGACTT-3`; (Invitrogen, Sao Paulo, SP, Brazil). The PCR mixture (25 µl) contained bacterial DNA, PCR buffer, 2 mM $MgCl_2$, 200 mM of each dNTP, and 1.0 U *Taq* polymerase (Invitrogen, Sao Paulo, SP, Brazil). The mixtures were amplified by initial denaturation at (94 °C) for 5 min, followed by 35 cycles at 94 °C for 30 sec, 56 °C for 30 sec, and 72 °C for 2 min, with a final extension at 72 °C for 5 min in an automated 0.5 ml GeneAmp PCR system 9700 (Applied Biosystems, Foster City, California, USA). DNase I treated distilled water was used as the negative control. A 10 µl aliquot of each amplified product was electrophoresed in 1.0% (wt/vol) agarose gel, with a molecular size marker 1kb in parallel. Electrophoresis in Tris/borate/EDTA buffer was performed at 90 V for 1.5 h. The gel was stained with ethidium bromide (0.5 mg/ml) for 15 min, rinsed, and photographed under ultraviolet light illumination. Bands were observed at the predicted 1.4kb size. A reference strain *Clostridium perfringens* ATCC 13124 was included in this study.

Purification of PCR Product

The PCR products were purified using the PCR DNA and gel band purification kit (Illustra, GE healthcare, Sao Paulo, SP, Brazil). Capture buffer was added to 25 µl PCR products and mixed thoroughly. The mixture was then loaded onto the GFX Microspin column and collection tube and centrifuged at 16,000 × g for 30 seconds to bind the product to the spin column. The column was subsequently washed by adding 500 µl of wash buffer and spinning at 16000 × g for 30 seconds. Finally, the DNA was eluted by adding 50 µl elution buffer into the GFX Microspin column in collection tube (DNase-free 1.5 ml eppendorf tube), incubated for 1 minute at room temperature and spinned at 16,000 × g for 1 minute. The purified DNA was stored at -20 °C and used for sequencing.

16S rRNA Gene Sequencing Analysis

The 16S rRNA gene sequence was determined using an automated sequencer ABI 3730 DNA Analyser (Life Technologies- Applied Biosystem) using BigDye(R) Terminator V3.1 cycle sequencing kit. Sequences were converted to FASTA format and compared with sequences available in the GenBank database by using the multisequence advanced BLAST comparison software from the National Centre for Biotechnology Information (NCBI-http://www.ncbi.nlm.nih.gov). Identification at the species level was defined with a similarity of ⩾ 98.7% with that of the prototype sequence in GenBank while identification at the genus level was defined as a 16S rDNA sequence similarity of ≥ 95% with that of the prototype in GenBank. A failure to identify was defined as a similarity score of lower than 95% with those deposited in GenBank at the time of analysis. The sequence data were submitted to GenBank under accession nos. MN064635–MN064658. After molecular Identification was completed, re-assessment of phenotypic data was done in order to determine what caused the conventional identification to fail. These faults were categorized under biochemical determination failures, growth requirement determination failures, oxidase and catalase activity determination failures and morphology and Gram stain determination failures.

RESULTS

Using 16S rRNA gene sequencing, twenty-nine (96.7%) strains that have names with standing in nomenclature, classified within 5 families, 11 genera and 16 species were included in our study (Table 1). One (3.3%) strain containing a large

number of undetermined nucleotides, for which 16S rRNA gene sequence accession numbers was not available on the GenBank, was not included in the study.

Phenotypic identification technique using morphological appearances and biochemical tests showed comparable genus level identification with 16S rRNA sequencing technique (86.2% and 93.1%, respectively). However, the former showed much lower species level identification 55.2% than the later 89.7%. 16S rRNA sequence-based identification confirmed phenotypic identification for 16 isolates. However, it failed to Identify 13 isolates accurately to species levels. Failures were inferred to be due to inappropriate and/or insufficient biochemical profile determination in 7 (24.1%) isolates, Gram staining in 2 isolates (6.9%), oxidase and catalase activity determination in 3 isolates (10.3%), and growth requirement determination in one isolate (3.4%) see figure 2. Failure to accurately identify *Clostridium botulinum* by conventional method was as a result of insufficient biochemical profile determination as well as mismatched Gram determination. The two un-identified isolates were due to inappropriate biochemical profile which led to in-determinate results

Table 2 is a comparison of the phenotypic method of identification and 16S rRNA gene sequencing technique. The turnaround times for phenotypic identification ranged from 7 to 14 days, depending on the biochemical activity of the organism isolated while the turnaround times for 16S rRNA sequencing ranged from 2 to 3 days, regardless of growth characteristics or biochemical activity.

Table 1: Bacterial Classification of Strains Isolated from Food

Family	Genus	Species (No. of strains)
Clostridiaceae	***Clostridium***	*C. perfringens* (5)
		C. botulinum (2)
		C. novyi (1)
		C. carboxidivorans (1)
		C. ramosum (1)
Enterobacteriaceae	***Enterobacter***	*Enterobacter cloacae subsp. Cloacae* (1)
		Enterobacter aerogenes (1)
	Proteus	*Proteus mirabilis* (1)
	Klebsiella	*Klebsiella pneumoniae subsp. Pneumonia* (2)
	Escherichia	*Escherichia coli* (7)
	Citrobacter	*Citrobacter freundii* (1)
	Salmonella	*Salmonella enterica subsp. enterica serovar Typhi* (1)
Lactobacillaceae	***Pediococcus***	*Pediococcus acidilactici* (1)
Streptococcaceae	***Lactococcus***	*Lactococcus garvieae* (2)
Enterococcaceae	***Enterococcus***	*Enterococcus faecium* (1)
	Vagococcus	*Vagococcus lutrae* (1)

Table 2: Comparison of Identification Techniques

Identification techniques	Genus level (%)	Specie level (%)	Turnaround time (days)	Accuracy (%)
Conventional tests	25/29 (86.2)	16/29 (55.2)	7-14	55.2
16S rRNA gene sequencing	27/29 (93.1)	26/29 (89.7)	2-3	89.7

Figure 1 show(s) bands observed at the predicted 1.4 kb size following amplification of 16S rRNA gene. The mean length of the sequences was ≈ 933 nucleotides. Ten isolates had sequencing with percentage nucleotide identity to that of known *Clostridium* species while 14 had nucleotide identity to known members of the *Enterobacteriaceae* family. The remaining five isolates were identified as Gram positive lactic acid bacteria (LAB). Of these 29 isolates, 26 were associated with clinically relevant infections; two (*Lactococcus garvieae*) are rare human pathogens with low virulence, whereas the remaining one (*Pediococcus acidilactici*) is a known organism associated with fermented dairy products. Ten of the 26 clinically relevant isolates were identified as known food borne pathogens including five *C. perfringens*, two *C. botulinum*, two *Escherichia coli* O104:H4 and one *Salmonella enterica*. (Table 3). 16S rRNA gene sequencing failed to identify two members of the *Clostridiaceae* to genus level and one to specie level. These isolates had sequence similarity of ≥ 95% with that of the prototype in GenBank which maybe indicative of new species or new genera.

Twenty-seven isolates (93.1%) had at least one sequence in GenBank that yielded a similarity score of ≥ 95% and twenty-six (89.7%) had at least one sequence in GenBank that yielded a similarity score of ≥ 98.7%. These similarity score values were used to define identification at the genus and species levels, respectively. Among the 29 studied bacterial species, the lowest and highest inter-species pairwise 16S rRNA gene sequencing similarities were 88% and 99.9%, respectively. Hence, for the 29 food strains included in this study, phenotypic characterization identified 9 of 29 (31%) isolates at the genus level; 16 of 29 (55.2) isolates at specie level; 2 of 29 (6.9%) isolates were mis-assigned to different taxa while 2 of 29 (6.9%) could not be discriminated at any taxonomic level. In comparison, 16S rDNA sequencing identified 26 of 29 (89.7) isolates at the species level, 1 of 29 (3.4%) isolates at the genus level, and 2 of 29 (6.9%) isolates at the family level (Figure 2).

For those isolates which were not identified by 16S rDNA sequence analysis, taxonomic relationships were inferred from 16S rDNA sequence comparison. Plotting a 16S rDNA based phylogenetic tree assigned phylogenetic positions to the isolates. Species of *Clostridium botulinum* and *Clostridium novyi* which could not be identified to species level using 16S rDNA sequence clustered separately from other members of the genus (Figure 3).

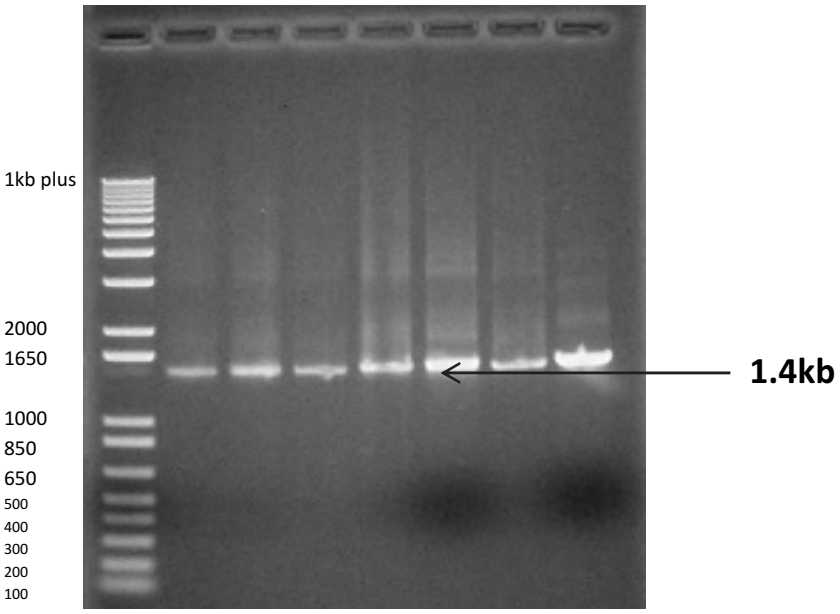

Figure 1: Agarose Gel Containing Representative Amplicon using 16S Rrna Universal Primers. Lane M, molecular marker (1kb DNA ladder). Lane 1: reference strain *Clostridium perfringens* ATCC 13124; and lines 2 to 7 are bacterial strains showing bands at 1,400 bp

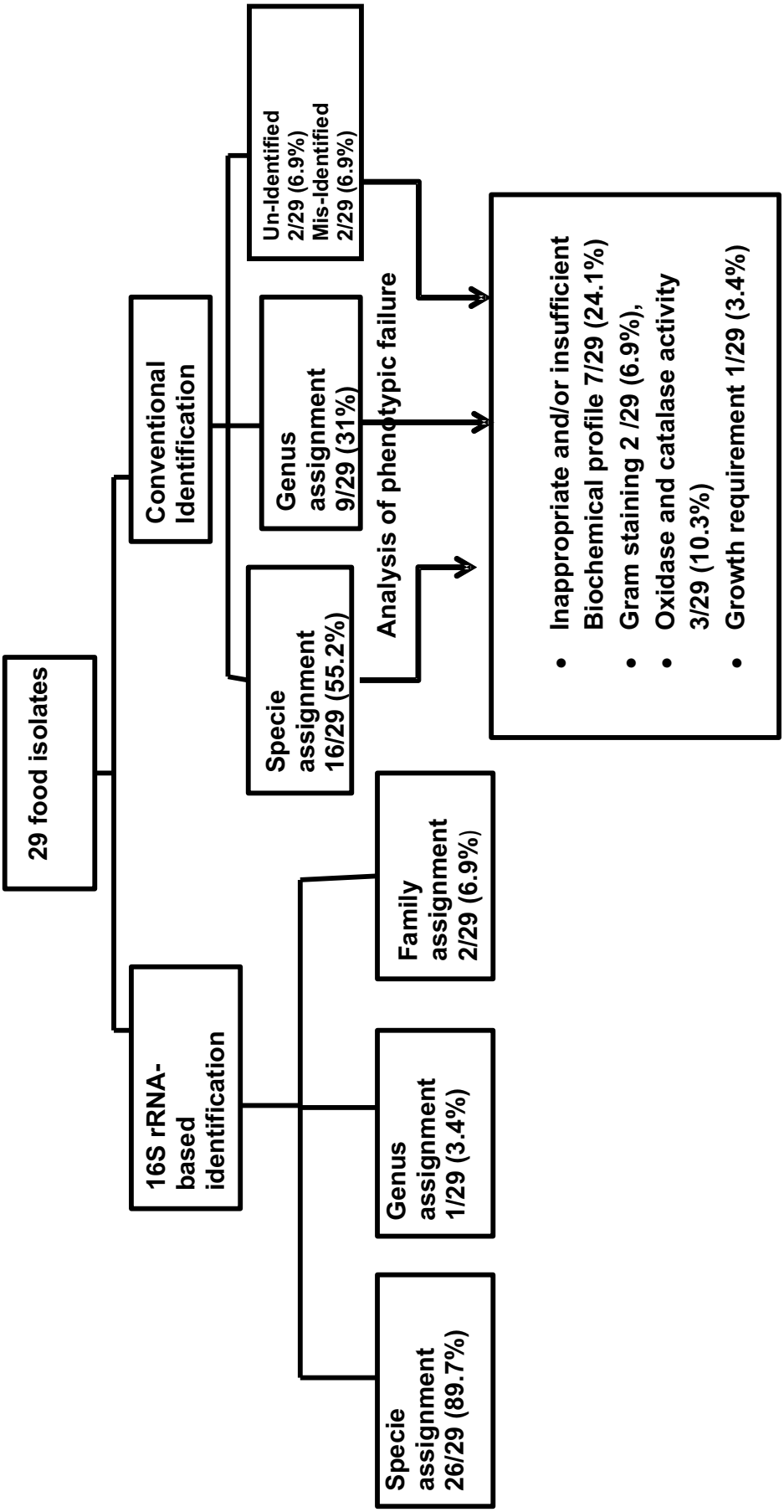

Figure 2: Schematic Representation of the Identification Process for Food Borne Bacteria

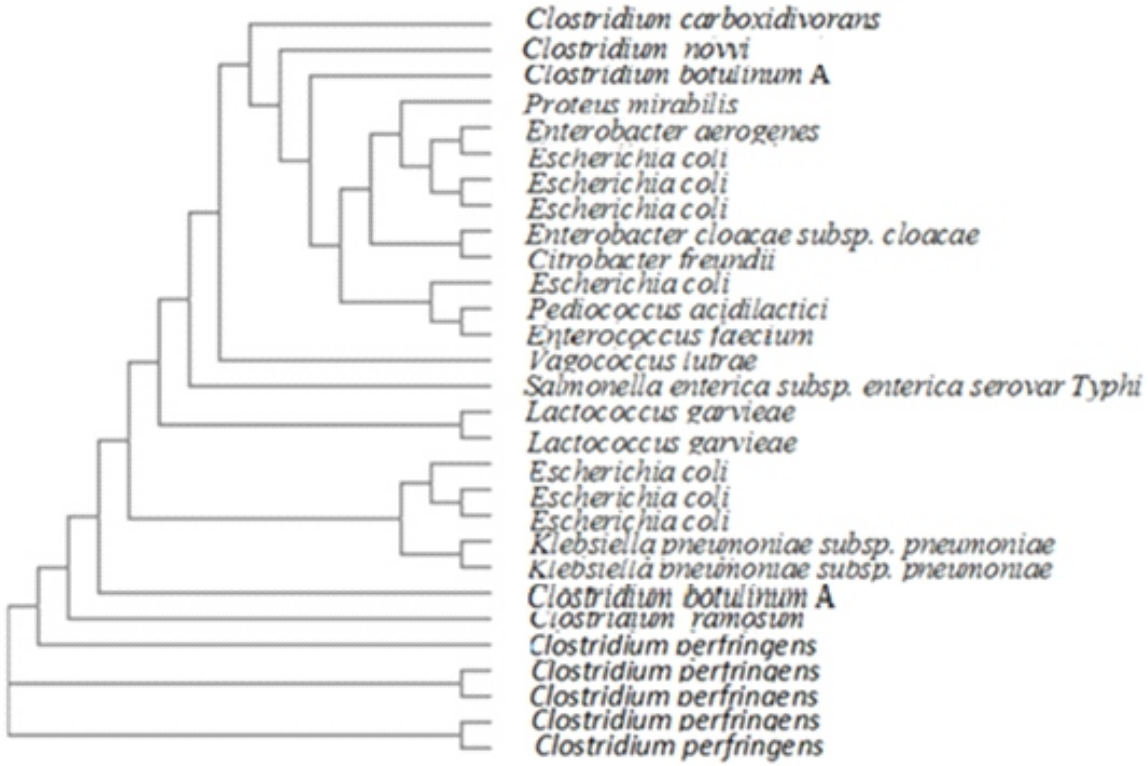

Figure 3: Dendrogram showing the Phylogenetic Positions of Bacterial Species based on 16S rRNA Gene Sequencing.
The tree was constructed using the neighbour-joining method and is based on a comparison of a continuous stretch of approximately 1400 bases with Mega 5.02

Table 3: Identification of Food Isolates by Conventional and 16S Ribosomal RNA-Based Technique

S/N	Conventional Identification	16S rRNA-based Identification	Reference strains	Corresponding Strain number	GenBank AC number/Version	% Nucleotide identity*
1	*Proteus mirabilis*	*Proteus mirabilis*	*P. mirabilis*	*strain HI4320*	NC_010554.1	99.0
2	*Un-identified*	*Vagococcus lutrae*	*V. lutrae*	*LBD1 adhBo-supercont1.12*	NZ_AYSH01000012.1	99.0
3	*Pediococcus spp.*	*Pediococcus acidilactici*	*P. acidilactici*	*DSM 20284*	NZ_GL397069.1	99.0
4	*Klebsiella pneumonia*	*Klebsiella pneumoniae subsp. Pneumonia*	*K. pneumoniae subsp. Pneumonia*	*HS11286*	NC_016845.1	98.7
5	*Klebsiella pneumonia*	*Klebsiella pneumoniae subsp. Pneumonia*	*K. pneumoniae subsp. Pneumonia*	*HS11286*	NC_016845.1	99.0
6	*Streptococcus minor*	*Lactococcus garvieae*	*L. garvieae*	*ATCC 49156*	NC_015930.1	98.7
7	*Streptococcus varani*	*Lactococcus garvieae*	*L. garvieae*	*ATCC 49156*	NC_015930.1	99.9
8	*Enterobacter spp.*	*Enterobacter cloacae subsp. Cloacae*	*E. cloacae subsp. Cloacae*	*ATCC 13047*	NC_014121.1	99.0
9	*Enterobacter spp.*	*Enterobacter aerogenes*	*E. aerogenes*	*KCTC 2190*	NC_015663.1	98.7
10	*Enterococcus faecium*	*Enterococcus faecium*	*E. faecium*	*DO*	NC_017960.1	98.7
11	*Citrobacter spp.*	*Citrobacter freundii*	*C. freundii*	*CFNIH1*	NZ_CP007557.1	99.0
12	*Escherichia coli*	*Escherichia coli*	*E. coli*	IAI39	NC_011750.1	99.0
13	*Escherichia coli*	*Escherichia coli*	*E. coli*	*O83:H1 str. NRG 857C*	NC_017634.1	98.7
14	*Escherichia coli*	*Escherichia coli*	*E. coli*	*O104:H4 str. 2011C-3493*	NC_018658.1	98.7
15	*Escherichia coli*	*Escherichia coli*	*E. coli*	*IAI39*	NC_011750.1	98.7
16	*Escherichia coli*	*Escherichia coli*	*E. coli*	*O104:H4 str. 2011C-3493*	NC_018658.1	99.0
17	*Escherichia coli*	*Escherichia coli*	*E. coli*	*IAI39 chromosome*	NC_011750.1	98.7
18	*Escherichia coli*	*Escherichia coli*	*E. coli*	*IAI39* chromosome	NC_011750.1	99.0
19	*Salmonella spp.*	*Salmonella enterica subsp. enterica serovar Typhi*	*S. enterica subsp. enterica serovar Typhi*	*str. CT18*,	NC_003198.1	98.7
20	*Clostridium perfringens*	*C. perfringens*	*C. perfringens*	Str 13	NC_003366.1	99.9
21	*Clostridium perfringens*	*C. perfringens*	*C. perfringens*	Str 13	NC_003366.1	99.0
22	*Clostridium perfringens*	*C. perfringens*	*C. perfringens*	ATCC 13124	NC_008261.1	98.7
23	*Clostridium perfringens*	*C. perfringens*	*C. perfringens*	ATCC 13124	NC_008261.1	98.8
24	*Clostridium perfringens*	*C. perfringens*	*C. perfringens*	ATCC 13124	NC_008261.1	99.0
25	*Clostridium spp*	*C. botulinum* A	*C. botulinum* A	ATCC 3502	NC_009495.1	98.7
26	*Clostridium spp.*	*Clostridium spp*	*C. botulinum* A	ATCC 3502	NC_009495.1	97.5
27	*Clostridium spp.*	*Clostridiaceae*	*C. novyi*	NCTC 538	NC_008593.1	88.0
28	*Un-identified*	*Clostridiaceae*	*C. carboxidivorans*	P7	NZ_GG770705.1	93.8
29	*Clostridium spp.*	*Clostridium ramosum*	*C. ramosum*	M23731	M23731.1	99.9

S/N= Serial number
AC= Accession number
*% Nucleotide identity of 16S rRNA gene sequence to the closest match in gene bank

DISCUSSION

Accurate identification of bacteria is important for patient care and public health surveillance so as to understand the epidemiology of foodborne pathogens for better infection control strategies and patient management. We used 16S rRNA gene sequencing technique to compare phenotypic identification of isolates in other to ascertain accuracy of routine microbiological identification of a range of food isolates. Our collection of isolates for this study included ten *Clostridium* species, fourteen members of the *Enterobacteriaceae* and five members of the *Lactobacillaceae* (Table 1). Although, this set represents a limited subset of possible food borne pathogens, we demonstrated very high confidence genus level identification and good species level identification using 16S rRNA gene sequencing. Mixed cultures or inappropriate DNA extraction led to one (3.3%) strain containing a large number of undetermined nucleotides which was excluded from the study. This may either be because the

wrong colony was selected on subculture or because more than one bacterial species was accidentally included in the amplification, resulting in ambiguous 16S rDNA data.

Applying the 95% for genus and the 98.7% to 99% limit classically proposed for species delineation within a genus (Figueras *et al*., 2011), we were able to achieve 93.1% genus level and 89.7% species level identification using 16S rRNA gene sequencing. Our findings are in keeping with previous reports (Bosshard *et al*. 2003; Caporaso *et al.,* 2012; Srinivasan *et al*., 2015). Drancourt *et al*. (2000) compared phenotypic and 16S based identification for a collection of 177 isolates of which 81 were from clinical samples. The authors concluded that 16S rDNA sequencing provides unambiguous data even for rare isolates, which are reproducible in and between laboratories.

Bacterial phenotypic characterization using the biochemical identification of the isolates included in this study showed that some biochemical differences exist among the different species but in some cases, they were not sufficient to allow for a differential identification. This is evident in the low species level identification recorded with phenotypic technique (55.2%). Discrepant results were found in 2 of 29 strains (Table 3) in which conventional identification resulted in *Streptococcus minor* and *Streptococcus varani* whereas sequence comparison with public databases resulted in 98.7% and 99.0% similarity with *Lactococcus garvieae.* Upon re-analysis, it was found that conventional identification was incorrect due to inappropriate biochemical profile determination. This simply reflects the known limitations of phenotypic identification namely; intra-species variability, interspecies similarity and bias towards previously identified bacterial species. Phenotypic identification is known to be more effective for well-known established bacterial species, with less known or rare species posing difficulty in species level identification. This constraint can be resolved to a great extent with the use of 16S rRNA gene sequencing. Unfortunately, 16S rRNA gene sequencing is not routinely used for strain identification in Nigeria. This maybe as a result of several factors including limited resources and lack of trained personnel. Consequently, phenotypic identification is still done routinely while 16S rDNA sequencing is typically performed on special request and for research purposes.

We identified twenty-six clinically relevant bacterial species. The larger number (14) of members of the *Enterobacteriaceae* isolated maybe suggestive of contamination of food with human and/or animal faecal material. In this study, the use of 16S rDNA sequencing allowed for the identification of rare organisms that are difficult to identify using conventional techniques. *Lactococcus garvieae* which was mis-identified by phenotypic technique is an important pathogen in aquaculture. It is a rare pathogen with a low virulence in human infection. This organism, though not an established food borne pathogen, has been reported in patients with gastrointestinal disorders who admitted to consuming raw fish (Wang *et al*., 2007). *Pediococcus acidilactici* is a homofermentative bacterium commonly found in fermented vegetables, fermented dairy products, and meat. A study has also suggested possible probiotic properties of *P. acidilactici* in the treatment of multiple sclerosis (Takata *et al*., 2011). Another rare organism which was identified in this study with the help of 16S rDNA sequencing is *Vagococcus lutrae.* This organism was initially isolated from the common otter (*Lutra lutra*) but has recently been reported in a patient hospitalized with extensive skin lesions in Marseille, France (Garcia *et al*., 2016). The authors speculate that the infection originated from a food-mediated acquisition of the pathogen, particularly through fish and other seafood. The skin lesions were adduced to be as a result of excretion of the bacterium via faeces directly onto the skin, as a result of the patient's poor hygiene. Ten of the 26 clinically relevant isolates were identified as known food borne pathogens including *C. perfringens*, *C. botulinum*, *Escherichia coli* and *Salmonella enterica*. The presence of these organisms in foods sold in Lagos implies that food hygiene is an existing problem in Lagos state and this has been indicated in our previous study (Chukwu *et al*., 2016).

Going further to plot a 16S rDNA phylogenetic tree using the neighbour-joining method which is

based on a comparison of a continuous stretch of approximately 1,400 bases, we noted that even those isolates which could not be identified at the genus and/or species level were assigned a phylogenetic position. This is in contrast to phenotypic identification, which is biased by errors and the variability of character expression. Identification of one *Clostridium botulinum* and *C. novyi* isolated in this study to the species level proved difficult because of low similarity levels, suggesting that, may be, too few *Clostridium* sequences have been deposited in GenBank. These organisms clustered separately from other members of the genus. Previous studies have also reported cases of different isolates of the same species being assigned to different clusters in the phylogenic tree after 16S rDNA sequencing analysis (Tang *et al.*, 1998; Drancourt *et al*., 2000). We also noted that the two *Enterobacter* isolates (*Enterobacter cloacae* and *Enterobacter aerogenes*) were unidentifiable at the species level using conventional method. This trend has also been reported in a previous study (Drancourt *et al*., 2000). In a study by Tang *et al*. (1998), six isolates identified as *Enterobacter cloacae* by the conventional method fell into three different clusters after 16S rDNA sequence analysis, with a 1.52% divergence rate among them. In the present study, the enterobacter species clustered with *Escherichia coli* suggesting that current *Enterobacter* taxonomy may not be sufficient or may be in-appropriate.

The usefulness of 16S rRNA gene sequencing for clinical microbiology has been evaluated by several researchers. Bosshard *et al*. (2003) evaluated the suitability of 16S rRNA for the identification of clinical strains of aerobic gram-positive rods. They concluded that rDNA sequencing is an effective means for the identification of aerobic gram-positive rods which are difficult to identify by biochemical techniques. Subsequently, Srinivasan *et al.* (2015) compared the accuracy of 16S rRNA gene based identification to that of non-16S based clinical identification for a broad range of clinically relevant bacterial species. The authors pointed to multiple cases of probable clinical misidentification with traditional culture based identification across a wide range of Gram-negative rods and Gram-positive cocci as well as common Gram negative cocci.

Presently, the quality of DNA sequences generated in laboratories has been greatly improved by the introduction of automated sequencing systems and DNA alignment software. Some authors have argued that the 95% (for genus) and 98.7% (for species) sequence similarity thresholds that are currently recommended to classify bacterial isolates were defined by comparison of a limited number of bacterial species, and may not apply to many genera that contain human-associated species (Drancourt *et al.*, 2000; Rossi-Tamisier *et al*., 2015). Certain factors like the purity of the DNA template and number of overlapping nucleotide fragment in the alignment, may affect the reliability of the 16S rRNA gene sequencing. This notwithstanding, the use of the 16S rRNA gene biomarker still offers a more practical method of identification in the situation of outbreak of food borne illness. More so for difficult to identify organisms. It not only identifies the organism but can also be used to determine the genetic relatedness of the organisms by constructing the phylogenetic tree.

CONCLUSION

This study demonstrates the value of 16S rRNA gene sequence as a tool for rapid and accurate initial screening of food borne pathogens and for precise assessment of phylogenetic relatedness. The results of this investigation provide insight into the strengths of molecular identification of common foodborne bacterial pathogens using 16S rRNA gene sequencing. We conclude that high confidence identification of food borne pathogens to the species-level can be achieved by 16S rRNA gene sequencing despite poor taxonomic boundaries for certain organisms. Acceptance of 16S rRNA techniques is based on demonstrated greater accuracy and rapid laboratory turn-around times, which ultimately improves clinical decision-making and overall patient management during an outbreak.

ACKNOWLEDGEMENTS

The authors are grateful to Liliana Rosa Solis of the Anaerobe Laboratory, Department of Microbiology, Institute of Biomedical Sciences,

University of Sao Paulo-USP for technical assistance.

REFERENCES

Bamaiyi, P. H. 2011. Emerging food pathogens: A case study of *E. coli* 0104:H4. *Continental Journal of Animal and Veterinary Research*, 3(2): 21-27.

Barbut, F., Monot, M., Rousseau, A., Cavelot, S., Simon, T., Burghoffer, B., Lalande, V., Tankovic, J., Petit, J.C., Dupuy, B., Eckert, C. 2011. Rapid diagnosis of *Clostridium difficile* infection by multiplex real-time PCR. *European Journal of Clinical Microbiology & Infectious Diseases*, 30: 1279–1285.

Bell, R., Bright, A., Butow, B., Combs, B., Franklin, N., Fullerton, K., Gibbs, R., Gradie, D., Gregory, J., Gunn, J., Harlock, M., Heilbronn, C., Hogg, G., Kardamanidas, K., Kirk, M., Knope, K., Lalor, K., Leader, R., McCallum, L., McKercher, C., Miller, M., Moffatt, C., Munnoch, S., Pingault, N., Raupach, J., Roper, K., Shadbolt, C., Stafford, R., Stephens, N. 2010. Monitoring the incidence and causes of diseases potentially transmitted by food in Australia: annual report of the OzFoodNet Network, 2009. *Communicable Diseases Intelligence*, 234(4):396-426.

Bosshard, P.P., Abels, S., Zbinden, R., Bottger, E.C. and Altwegg, M. 2003. Ribosomal DNA sequencing for identification of aerobic gram-positive rods in the clinical laboratory (an 18-month evaluation). *Journal of Clinical Microbiology*, 41: 4134–4140.

Bouvet, P., Ferraris, L., Dauphin, B., Popoff, M., Butel, M. J. and Aires, J. 2014. 16S rRNA gene sequencing, multilocus sequence analysis, and mass spectrometry identification of the proposed new species "*Clostridium neonatale*". *Journal of Clinical Microbiology*, 52(12): 4129–4136.

Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons, D., Huntley, J., Fierer, N., Owens, S. M., Betley, J., Fraser, L., Bauer, M., Gormley, N., Gilbert, J. A., Smith, G. and Knight, R. 2012. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME Journal*, 6: 1621–1624.

CDCP. 2017. Centers for Disease Control and Prevention. Surveillance for Foodborne Disease Outbreaks, United States, 2015, Annual Report. Atlanta, Georgia: US Department of Health and Human Services. http://www.cdc.gov/foodsafety/fdoss. Accessed 28th November 2018.

Chukwu, E.E., Nwaokorie, F.O., Coker, A.O., Avila-Campos, M.J., Solis, R.L., Llanco, L.A. and Ogunsola, F.T. 2016. Detection of toxigenic *Clostridium perfringens* and *Clostridium botulinum* from food sold in Lagos, Nigeria, *Anaerobe*, 42: 176e181, http://dx.doi.org/10.1016/j.anaerobe.2016.10.009.

Conlan, S., Kong, H.H. and Segre, J.A. 2012. Species-level analysis of DNA sequence data from the NIH human microbiome project. *PLoS One*, 7: e47075. doi: 10.1371/journal.pone.0047075.

Drancourt, M., Bollet, C., Carlioz, A., Martelin, R., Gayral, J. and Raoult, D. 2000. 16S Ribosomal DNA sequence analysis of a large collection of environmental and clinical unidentifiable bacterial isolates. *Journal of Clinical Microbiology*, 38(10): 3623–3630.

Fettweis, J.M., Serrano, M.G., Sheth, N.U., Mayer, C.M., Glascock, A.L., Brooks, J.P., Jefferson, K.K., Vaginal Microbiome Consortium (additional members) and Buck G.A. 2012. Species-level classification of the vaginal microbiome. *BMC Genomics*, 13 (8): S17.

Figueras, M.J., Beaz-Higalgo, R., Collado, L. and Martinez-Murcia, A. 2011. Recommendations for new bacterial species description based on analyses of the unrelated genera *Aeromonas* and *Arcobacter*. *Bull. BISMiS*, 2:1–16.

Fratamico, P. M., Bhunia A. K. and Smith, J. L. 2005. Foodborne pathogens: Microbiology and Molecular biology. *Caister Academic Press*, ISBN: 1-978-1-904455-00-4.

Garcia, V., Abat, C. and Rolain, J.M. 2016. Report of the first Vagococcus lutrae human infection, Marseille, France. *New Microbes and New Infections*, 9: 56–57.

Hawkey, P.M. 2008. The growing burden of antimicrobial resistance. *Journal of Antimicrobial Chemotherapy*, 62 (1): i1–9.

Humphrey, T., O'Brien, S. and Madsen, M. 2007. Campylobacters as zoonotic pathogens: A food production perspective. *International Journal of Food Microbiology*, 117 (3): 237–257.

Iregui, M., Ward, S., Sherman, G., Fraser, V.J. and Kollef, M.H. 2002. Clinical importance of delays in the initiation of appropriate antibiotic treatment for ventilator-associated pneumonia. *Chest*; 122: 262–268.

Joshy, L., Chaudhry, R. and Chandel, D. S. 2008. Multiplex PCR for the detection of *Clostridium botulinum* and *C. perfringens* toxin genes. *Indian Journal of Medical Research*, 128: 206-208.

Larson, E. 2007. Community factors in the development of antibiotic resistance. *Annual Review of Public Health*, 28: 435–447.

Lehmann, L.E., Hunfeld, K.P., Emrich, T., Haberhausen, G., Wissing, H., Hoeft, A. and Stüber, F. 2008. A multiplex real-time PCR assay for rapid detection and differentiation of 25 bacterial and fungal pathogens from whole blood samples. *Medical Microbiology and Immunology*, 197: 313–324.

Moffatt, C. R., Howard, P. J. and Burns, T. 2011. A mild outbreak of gastroenteritis in long-term care facility residents due to *Clostridium perfringens. Foodborne Pathogens and Disease*, 8(7):791-796.

Osterholm, M. T. 2011. Foodborne Disease in 2011- The rest of the story. *New England Journal of Medicine*, 364 : 889-891.

Rossi-Tamisier, M., Benamar, S., Raoult, D. and Fournier, P. 2015. Cautionary tale of using 16S rRNA gene sequence similarity values in identification of human-associated bacterial species. *International Journal of Systematic and Evolutionary Microbiology*, 65, 1929–1934.

Scallan, E., Griffin, P. M., Angulo, F. J., Tauxe, R. V. and Hoekstra, R. M. 2011. Foodborne illness acquired in the United States-unspecified agents. *Emerging Infectious Diseases*, 17 (1): 16–22.

Shorr, A.F., Micek, S.T., Welch, E.C., Doherty, J.A., Reichley, R.M. and Kollef, M.H. 2011. Inappropriate antibiotic therapy in Gram-negative sepsis increases hospital length of stay. *Critical Care Medicine*, 39: 46–51.

Srinivasan, R., Karaoz, U., Volegova, M., MacKichan, J., Kato-Maeda, M., Miller, S., Nadarajan, R., Brodie, E. L. and Lynch, S. V. 2015. Use of 16S rRNA gene for identification of a broad range of clinically relevant bacterial pathogens. *PLOS ONE* DOI:10.1371/journal.pone.0117617.

Takata, K, Kinoshita, M, Okuno, T, Moriya, M, Kohda, T, Honorat, J. A., Sugimoto, T., Kumanogoh, A., Kayama, H., Takeda, K., Sakoda, S. and Nakatsuji, Y. 2011. The lactic acid bacterium *Pediococcus acidilactici* suppresses autoimmune encephalomyelitis by inducing IL-10-producing regulatory T cells. *PLoS One*. 2011, 6(11): e27644.

Tang, Y.W., Ellis, N. M., Hopkins, M. K., Smith, D. H., Dodge, D. E. and Persing, D. H. 1998. Comparison of phenotypic and genotypic techniques for identification of unusual aerobic pathogenic gram-negative bacilli. *Journal. Clinical. Microbiology*. 36**:**3674–3679.

Tribe, I. G., Cowell, D., Cameron, P. and Cameron, S. 2002. An outbreak of *Salmonella typhimurium* phage type 135 infection linked to the consumption of raw shell eggs in an aged care facility. *Communicable Disease Intelligence*, **26**(1): 38-39.

Wang, C.Y.1., Shie, H.S., Chen, S.C., Huang, J.P., Hsieh, I.C., Wen, M.S., Lin, F.C. and Wu, D. 2007. *Lactococcus garvieae* infections in humans: possible association with aquaculture outbreaks. *International Journal of Clinical Practice*, 61(1):68-73.

Werner, J.J., Koren, O., Hugenholtz, P., DeSantis, T.Z., Walters, W.A., Caporaso, J.G., Angenent, L.T., Knight, R. and Ley, R.E. 2012. Impact of training sets on classification of high-throughput bacterial 16s rRNA gene surveys. *ISME Journal* 6: 94–103.

Wilson, K.H., Blitchington, R.B. and Greene, R.C. 1990. Amplification of bacterial 16S ribosomal DNA with polymerase chain reaction, *Journal of Clinical Microbiology*, 28 1942-1946.

Woo, P. C. Y., Lau, S. K. P., Chan, K., Fung, A. M. Y., Tang, B. S. F. and Yuen, K. 2005. *Clostridium* bacteraemia characterised by 16S ribosomal RNA gene sequencing. *Journal of Clinical Pathology*, 58:301–307.

WHO. 2007. World Health Organization.. Food safety and foodborne illnesses. Fact Sheet No. 237. https://foodhygiene2010.files.wordpress.com/2010/06/who-food_safety_fact-sheet.pdf. Accessed 28 November 2018