Scientific Foundation SPIROSKI, Skopje, Republic of Macedonia
Open Access Macedonian Journal of Medical Sciences. 2022 Mar 30; 10(A):752-760.
https://doi.org/10.3889/oamjms.2022.8990
eISSN: 1857-9655
Category: *A - Basic Sciences*
Section: *Microbiology*

OPEN ACCESS

Escherichia coli Clonal Variability Based on Genetic Diversity Pattern with Enterobacterial Repetitive Intergenic Consensus-Polymerase Chain Reaction Methods for Traveler's Diarrhea Cases in Bali

I Dewa Made Sukrama[1]*, Komang Januartha Putra Pinatih[1], Made Agus Hendrayana[1], Burhannuddin Rasyid[2], Ni Luh Putu Harta Wedari[3]

[1]Department of Clinical Microbiology, Faculty of Medicine, Universitas Udayana, Bali, Indonesia; [2]Department of Medical Laboratory Technology, Polytechnic of Health Denpasar, Bali, Indonesia; [3]Clinical Microbiology Specialist Program, Faculty of Medicine, Universitas Udayana, Sanglah General Hospital, Denpasar, Bali, Indonesia

Edited by: Slavica Hristomanova-Mitkovska
Citation: Sukrama IDM, Pinatih KJP, Hendrayana MA, Rasyid B, Wedari NLPH. *Escherichia coli* Clonal Variability Based on Genetic Diversity Pattern with Enterobacterial Repetitive Intergenic Consensus-Polymerase Chain Reaction Methods for Traveler's Diarrhea Cases in Bali. Open Access Maced J Med Sci. 2022 Mar 30; 10(A):752-760. https://doi.org/10.3889/oamjms.2022.8990
Keywords: *Escherichia coli*; Genetic diversity; ERIC-PCR; Traveller's Diarrhea
***Correspondence:** I Dewa Made Sukrama, Department of Clinical Microbiology, Faculty of Medicine, Universitas Udayana, Bali, Indonesia.
E-mail: sukrama.idm@gmail.com
Received: 14-Feb-2022
Revised: 19-Mar-2022
Accepted: 20-Mar-2022

Funding: This research did not receive any financial support
Competing Interests: The authors have declared that no competing interests exist

Abstract

BACKGROUND: Bali is a favorite tourism destination in the world. As a major tourist destination, the incidence of illness that afflicts tourists greatly affects the image of tourism. Diarrhea is a health problem that is most often experienced and is a major obstacle for foreign tourists when traveling, especially to Bali. *Escherichia coli (E. coli)* bacteria cause diarrhea more often than viruses in some developing countries. Genetic differences can affect the characteristics of *E. coli*, especially in relation to the medical field.

AIM: We would like to assess the genetic diversity of the different pathogenic *E. coli* from various clinical isolates including those from traveler's diarrhea in Bali, Indonesia.

MATERIALS AND METHODS: One of the molecular techniques used in this study is to use enterobacterial repetitive intergenic consensus-polymerase chain reaction (ERIC-PCR). The sample in this study was the feces of foreign tourists with traveler's diarrhea in Bali. This study carried out research procedures in the form of Isolation of *E. coli* genome DNA from culture, amplification of *E. coli* 16S rRNA encoding genes, sequencing of *E. coli* 16S rRNA encoding genes, phylogenetic tree construction, and then analysis of *E. coli* genetic diversity with ERIC-PCR sequences.

RESULTS: The results showed that the ERIC-PCR method was more discriminatory than other methods to analyze the genetic diversity of *E. coli* from fecal samples of patients with traveler's diarrhea. It was found that clonal variability based on the genetic similarity of all sample *E. coli* isolates varied from 0% to 100%.

CONCLUSIONS: This shows that the source of transmission and the strains of *E. coli* that cause it come from diverse populations.

Introduction

Bali is a favorite tourism destination in the world. As a favorite tourist destination, Bali was selected based on Travelers Choice Awards from the best 25 places to visit around the world in the year of 2017 [1]. Balinese food is one of the most popular tourist attractions. This phenomenon shows that the increase in tourist visits to Bali is a challenge for the health sector in maintaining the image of tourism [2]. Based on data from the Bali Central Statistics Agency (*Badan Pusat Statistik*), in 2016, foreign tourist arrivals reached 4.92 million people during 2016. This number increased by 23.14% compared to the previous record (4.001 million people) [3]. Based on disease data from the *Kantor Kesehatan Pelabuhan* (harbor's health office) clinic at the International Departure Terminal of I Gusti Ngurah Rai Airport from 2015 to November 2017, diarrheal disease always occupies the top three cases of infectious diseases handled. In its development, there was an increase in cases of 20% from the previous year [4].

Diarrhea is a health problem that is most often experienced and is a major obstacle for foreign tourists when traveling, especially to developing countries. For people from developed countries, diarrhea in foreign tourists is a common disease when traveling to developing countries or countries with low economies. Bacteria and parasites cause diarrhea more often than viruses in some developing countries [5]. In several studies during an outbreak of food-related diarrhea in Bali, *Escherichia coli (E. coli)* bacteria were found in the food samples [6].

Research published in 2017 in several countries still states enterotoxigenic *E. coli* ETEC as the most common cause of diarrhea in foreign tourists, in addition to enteroaggregative *E. coli* (EAEC) [7]. Enteroinvasive *E. coli* (EIEC) and enterohemoragic *E. coli* bacteria currently have a low prevalence, so they are sometimes difficult to find in developing countries [8]. EAEC is the

highest pathotype identified as diarrheagenic *E. coli* (DEC) after ETEC in several developing countries [9]. The polymerase chain reaction (PCR) method can be used to detect pathogenic genes in isolated bacteria. PCR can be done to diagnose cases with DEC more quickly and effectively. PCR is needed to identify strains of DEC. PCR can be used to identify various subtypes that cause diarrhea, including enteropathogenic *E. coli*, EIEC, ETEC, EAEC, and Shiga toxin-producing *E. coli* [8].

Genetic diversity is a level of biodiversity that refers to the total amount of genetic variation in all species that inhabit part or all of the habitable surface of the earth. It differs from genetic variability, which describes the genetically controlled tendency of a character's ability to vary. Genetic differences can affect the characteristics of *E. coli*, especially in relation to the medical field. Therefore, molecular analysis based on genotypic traits is important to identify and characterize, study the evolution and epidemiology of the pathogenicity of a bacterium [10]. One of the molecular techniques used in this study is to use Enterobacterial Repetitive Intergenic Consensus-PCR (ERIC-PCR). ERIC-PCR is a DNA amplification method using ERIC sequences. The ERIC sequence is a short sequence (126 bp) with a conserved repeat area and a non-coding area, i.e., a sequence that is not encoded into protein [10] and is usually found in bacteria belonging to the family Enterobacteriaceae. This technique is used because it is simple, fast, and discriminatory [11]. The ERIC-PCR method has succeeded in analyzing *E. coli* bacteria isolated from tempeh [12].

Research related to tourist health has been widely carried out in various countries, associated with special and unique things of each country, one of which is culinary or special food. Meanwhile, in Indonesia itself, there are not many studies that raise this problem. Several studies related to diarrhea among foreign tourists (traveler's diarrhea) with tourism have been carried out but the number is still limited, so this study was conducted to determine the clonal variability of *E. coli* based on the Genetic Diversity Pattern of the bacterium *E. coli* that causes traveler's diarrhea in Bali based on the analysis of ERIC-PCR results and analyzing distribution patterns and epidemiological studies.

Materials and Methods

This research was done in the Clinical Microbiology Department, Faculty of Medicine, Universitas Udayana. Target population were foreign tourists visiting Bali, whereas the scope was foreign tourists with traveler's diarrhea who come for treatment at hospitals and clinics in Bali. Ethical approval granted by Faculty of Medicine, Universitas Udayana (Ethical Clearance number 1851/UN14.2.2.VII.14/LT/2020). Samples obtained from the hospital were then stored at 4°C for later culture with specific media for identification. After obtaining the stool sample from the patient, if the stool sample was not cultured directly, it should has been stored at 4°C for no more than 2 hours while waiting for preparation for culture. Stool samples that were stored did not use certain media, remain in their containers. The culture was performed no more than 2 hours from the time it was obtained from the patient. DNA isolation of *E. coli* bacteria using the DNA isolation kit DNeasy Blood and tissue kit from Qiagen by following the procedure according to the kit protocol. PCR examination using isolated DNA mixed with a mixture of KAPA2G Fast HotStart PCR^{TM}. The PCR process uses Biometra PCR Thermal cycler. The amplification of the gene encoding 16S rRNA used primers 63f (5'CAG GCC TAA CAC ATG CAA GTC-3') and 1387r primers (5'-CCC GGG AAC GTA TTC ACC GC-3') [11], [12], [13].

The PCR conditions were as follows: pre-denaturation at 95°C for 3 min one cycle, denaturation at 95°C for 30 s 35 cycles, primer annealing at 55°C for 30 s, elongation at 72°C for 1 min, and post-PCR final extension at 72°C for 5 s. The amplicon of the 16S RNA gene encoding *E. coli* was characterized by 1.5% agarose gel electrophoresis at a voltage of 80 V for 1 h. Furthermore, the isolated DNA bands were viewed under UV transilluminator. DNA sequencing was carried out by 1st base through *PT. Genetika Science Indonesia*. The DNA sequencing process was carried out using the Sanger dideoxy method. Furthermore, the analysis of the sequencing results was carried out by BLAST nucleotide sequences from the sequencing results with the database available on the website www.ncbi.nlm.nih.gov which was used to find the similarity of a nucleotide or protein sequence (query sequence) with a database sequence (subject sequence). A sequence alignment was carried out using the MEGA-X Program [12], [13].

The complete sequence of the gene encoding 16S rRNA is required to carry out the BLAST process. The complete sequence results are processed using the clone manager program. The compiling of the 16S rRNA encoding gene for *E. coli* was carried out after the alignment process to the 16S rRNA sequencing results. The results of the 16S rRNA coding gene sequences from the isolates were tracked for homology to the 16S rRNA sequences belonging to other *E. coli* bacteria samples through the BLAST program with the website address http://blast.ncbi.nlm.nih.gov/Blast.cgi. 16S rRNA sequences obtained from the BLAST program were stored in FASTA format and reprocessed using the MEGA-X program to obtain similarity figures. The profile of the fragments resulting from the ERIC-PCR amplification was interpreted into a binary data matrix and used as input for the creation of a phylogenetic tree using the MEGA-X program. The phylogenetic tree

construction used the Unweighted Pair Groups Method Analysis method. The topology of the phylogenetic tree construction was evaluated using bootstrap analysis with 1000x [11], [12], [13].

The phylogenetic tree was designed using the 16S rRNA encoding gene sequences of *E. coli* bacteria. The gene sequences encoding 16S rRNA are molecules that have conserved regions and constant function in each organism, are universally distributed and have well-conserved sequences among a wide range of phylogenetic members [14]. The phylogenetic tree was designed using the MEGA-X program. The method used to design the phylogenetic tree was the Neighbor-Joining Tree [13]. *E. coli* isolates were further analyzed using a phylogenetic tree construction and based on the geneBank database and then further analyzed to determine the clonal variability of *E. coli* based on genetic diversity pattern.

Results

Sample was the feces of foreign tourists with traveler's diarrhea in Bali. The number samples confirmed with *E. coli* infection were 30.

The results of the gel documentation system analysis after electrophoresis showed that the gene amplicon was 1.35 kilobases (kBp). Figure 1 shows one of the samples, namely sample 02 grown on different media to identify bacteria from the sample with some media used as follows:

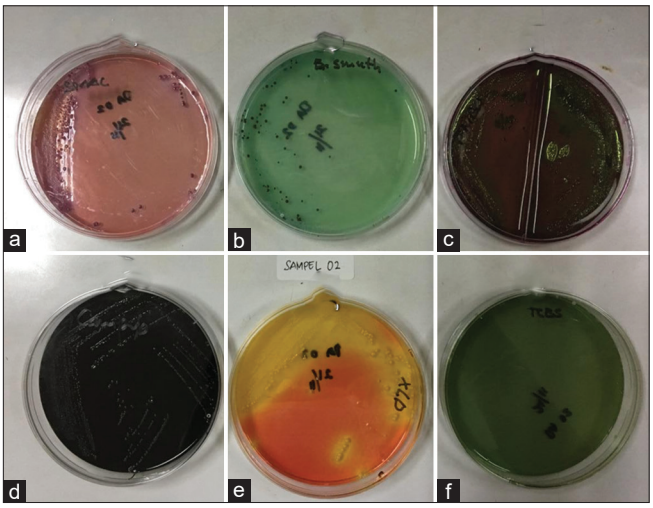

Figure 1: Sample code SAMPEL02 after being planted in specific media for identification of enterobacter bacteria that cause Traveler's Diarrhea

- A used SMAC media, the result looks colorless colonies
- B used bismuth media to find out if there was growth of *Salmonella* spp. bacteria, the result was that black colonies grown
- C used Emba media the results show metallic luster colonies which indicate *E. coli* bacteria colonies
- D used Campylobacter media, the results showed no colonies showing the growth of Campylobacter bacteria
- E used XLD media, to find out if there was any growth of *Shigella* spp. bacteria which were red with a black core, the results were not there.
- F used TCBS media to find out whether there was growth of *Vibrio cholera* bacterial colonies, eventually the results were not there.

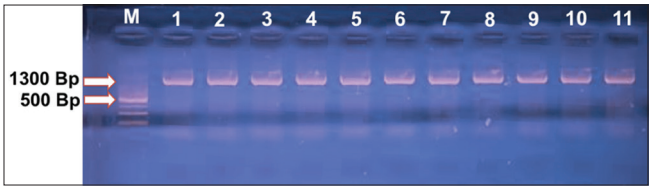

Figure 2: The results of purification of the 16S RNA encoding gene for Escherichia coli from some samples were characterized by agarose gel electrophoresis

PCR products from samples showed positive electrophoresis results, it can be seen that the amplified DNA fragment has a length of about 1350 bp. This DNA fragment is a region encoding 16S rRNA (Figure 2). This fragment has a variable region, wherein this region, the sequence of nitrogen bases in each bacterial species is different so that it can determine the species of the isolate. The results of the identification of DNA analysis of bacterial isolates that have been blasted are compared with several other types of bacteria, there are several comparison bacteria from the BLAST results with higher levels, such as the phylogenetic results of each isolate of *E. coli* bacteria there is a kinship between each isolate.

Homologous similarity analysis between Escherichia coli samples

The following table shows the similarity of the *E. coli* samples. The results of the sequence analysis of the blasted bacterial isolates were compared with several other bacteria between samples. There are several comparison bacteria from the BLAST results with higher levels and have a relationship with each isolate. The table of BLAST results shows the closeness of the *E. coli* bacteria from the sample to other samples.

In the BLAST homology similarity Table 1, it is known that *E. coli* from the samples had different homologous 16S rRNA sequences similarity. The range of similarity numbers varies from 0% to 100%. The closest similarity number is sample 16Emb with sample 17Emb, which is 100%. The most distant similarity is 0%, namely between sample number ECis and sample 34, and sample 15 and sample 21. Many of the samples also have high homology similarity numbers, which are above 75%, namely between sample 19 and sample 21 with 75% similarity, sample 20 and sample no.33

AQ6 **Table 1: Homology similarity number between *Escherichia coli* samples**

Strain Number	E. coli E Cis	E. coli 1 Emb	E. coli 4 Bis	E. coli 8 Bis	E. coli 16 Emb	E. coli 17 Emb	E. coli 19 Emb	E. coli 01 Camp	E. coli 02 Camp	E. coli 03 Camp	E. coli 04 Emb	E. coli 06 Emb	E. coli 02	E. coli 09	E. coli 10	E. coli 14	E. coli 15	E. coli 16	E. coli 17	E. coli 18	E. coli 19	E. coli 20	E. coli 21	E. coli 22	E. coli 25	E. coli 28	E. coli 32	E. coli 33	E. coli 34	E. coli 15F
E. coli E Cis	1	0.63	0.38	0.50	0.27	0.27	0.29	0.53	0.50	0.42	0.67	0.43	0.38	0.59	0.59	0.40	0.00	0.29	0.13	0.50	0.38	0.56	0.50	0.35	0.24	0.38	0.45	0.57	0.00	0.35
E. coli 1 Emb	0.63	1	0.88	0.63	0.40	0.53	0.14	0.27	0.75	0.32	0.67	0.14	0.25	0.35	0.59	0.40	0.17	0.14	0.13	0.50	0.50	0.67	0.38	0.59	0.47	0.38	0.45	0.57	0.43	0.35
E. coli 4 Bis	0.38	0.88	1	0.63	0.53	0.67	0.43	0.40	0.63	0.42	0.67	0.29	0.50	0.47	0.59	0.53	0.17	0.14	0.40	0.50	0.63	0.67	0.50	0.59	0.59	0.38	0.64	0.57	0.43	0.59
E. coli 8 Bis	0.50	0.63	0.63	1	0.67	0.67	0.43	0.53	0.50	0.32	0.53	0.43	0.38	0.47	0.59	0.53	0.33	0.29	0.40	0.38	0.63	0.56	0.50	0.47	0.71	0.25	0.36	0.57	0.57	0.47
E. coli 16 Emb	0.27	0.40	0.53	0.67	1	1.00	0.77	0.71	0.53	0.33	0.71	0.15	0.40	0.38	0.38	0.57	0.00	0.46	0.43	0.67	0.53	0.35	0.40	0.50	0.38	0.40	0.48	0.30	0.31	0.50
E. coli 17 Emb	0.27	0.53	0.67	0.67	1.00	1	0.62	0.43	0.67	0.56	0.43	0.31	0.67	0.38	0.50	0.43	0.18	0.15	0.43	0.53	0.53	0.59	0.40	0.38	0.38	0.53	0.48	0.50	0.31	0.50
E. coli 19 Emb	0.29	0.14	0.43	0.43	0.77	0.62	1	0.92	0.57	0.24	0.62	0.33	0.57	0.53	0.53	0.62	0.00	0.50	0.15	0.57	0.43	0.38	0.43	0.40	0.53	0.29	0.50	0.42	0.50	0.40
E. coli 01 Camp	0.53	0.27	0.40	0.53	0.71	0.43	0.92	1	0.67	0.22	0.57	0.31	0.53	0.63	0.63	0.71	0.00	0.46	0.14	0.53	0.40	0.47	0.53	0.38	0.50	0.27	0.48	0.50	0.15	0.38
E. coli 02 Camp	0.50	0.75	0.63	0.50	0.53	0.67	0.57	0.67	1	0.42	0.40	0.14	0.38	0.47	0.59	0.67	0.17	0.29	0.40	0.50	0.50	0.67	0.50	0.35	0.47	0.50	0.55	0.57	0.29	0.35
E. coli 03 Camp	0.42	0.32	0.42	0.32	0.33	0.56	0.24	0.22	0.42	1	0.33	0.71	0.74	0.50	0.50	0.33	0.27	0.24	0.56	0.32	0.32	0.38	0.32	0.30	0.40	0.53	0.40	0.50	0.35	0.60
E. coli 04 Emb	0.67	0.67	0.67	0.53	0.71	0.43	0.62	0.57	0.40	0.33	1	0.15	0.53	0.63	0.50	0.57	0.36	0.15	0.43	0.67	0.67	0.47	0.53	0.75	0.50	0.40	0.38	0.50	0.31	0.50
E. coli 06 Emb	0.43	0.14	0.29	0.43	0.15	0.31	0.33	0.31	0.14	0.71	0.15	1	0.86	0.40	0.67	0.46	0.40	0.50	0.46	0.29	0.43	0.38	0.29	0.27	0.27	0.57	0.30	0.32	0.33	0.53
E. coli 02	0.38	0.25	0.50	0.38	0.40	0.67	0.57	0.53	0.38	0.74	0.53	0.86	1	0.47	0.71	0.67	0.33	0.43	0.53	0.50	0.50	0.33	0.50	0.47	0.35	0.63	0.55	0.57	0.57	0.59
E. coli 09	0.59	0.35	0.47	0.47	0.38	0.38	0.53	0.63	0.47	0.50	0.63	0.40	0.47	1	0.56	0.75	0.15	0.67	0.38	0.71	0.59	0.53	0.59	0.33	0.56	0.35	0.52	0.55	0.40	0.56
E. coli 10	0.59	0.59	0.59	0.59	0.38	0.50	0.53	0.63	0.59	0.50	0.50	0.67	0.71	0.56	1	0.88	0.31	0.53	0.63	0.35	0.47	0.63	0.59	0.56	0.56	0.59	0.52	0.64	0.40	0.44
E. coli 14	0.40	0.40	0.53	0.53	0.57	0.43	0.62	0.71	0.67	0.33	0.57	0.46	0.67	0.75	0.88	1	0.18	0.62	0.43	0.53	0.40	0.59	0.53	0.38	0.50	0.40	0.57	0.60	0.31	0.38
E. coli 15	0.00	0.17	0.17	0.33	0.00	0.18	0.00	0.00	0.17	0.27	0.36	0.40	0.33	0.15	0.31	0.18	1	0.20	0.36	0.17	0.17	0.14	0.00	0.15	0.31	0.50	0.11	0.24	0.40	0.31
E. coli 16	0.29	0.14	0.14	0.29	0.46	0.15	0.50	0.46	0.29	0.24	0.15	0.50	0.43	0.67	0.53	0.62	0.20	1	0.46	0.43	0.14	0.25	0.29	0.27	0.40	0.29	0.40	0.32	0.33	0.13
E. coli 17	0.13	0.13	0.40	0.40	0.43	0.43	0.15	0.14	0.40	0.56	0.43	0.46	0.53	0.38	0.63	0.43	0.36	0.46	1	0.27	0.40	0.47	0.40	0.38	0.38	0.53	0.57	0.50	0.46	0.38
E. coli 18	0.50	0.50	0.50	0.38	0.67	0.53	0.57	0.53	0.50	0.32	0.67	0.29	0.50	0.71	0.35	0.53	0.17	0.43	0.27	1	0.63	0.56	0.63	0.59	0.47	0.50	0.64	0.38	0.14	0.47
E. coli 19	0.38	0.50	0.63	0.63	0.53	0.53	0.43	0.40	0.50	0.32	0.67	0.43	0.50	0.59	0.47	0.40	0.17	0.14	0.40	0.63	1	0.78	0.75	0.71	0.71	0.63	0.55	0.57	0.57	0.71
E. coli 20	0.56	0.67	0.67	0.56	0.35	0.59	0.38	0.47	0.67	0.38	0.47	0.38	0.33	0.53	0.63	0.59	0.14	0.25	0.47	0.56	0.78	1	0.78	0.63	0.74	0.56	0.67	0.78	0.50	0.63
E. coli 21	0.50	0.38	0.50	0.50	0.40	0.40	0.43	0.53	0.50	0.32	0.53	0.29	0.50	0.59	0.59	0.53	0.00	0.29	0.40	0.63	0.75	0.78	1	0.59	0.71	0.50	0.55	0.57	0.43	0.71
E. coli 22	0.35	0.59	0.59	0.47	0.50	0.38	0.40	0.38	0.35	0.30	0.75	0.27	0.47	0.33	0.56	0.38	0.15	0.27	0.38	0.59	0.71	0.63	0.59	1	0.67	0.47	0.61	0.55	0.40	0.56
E. coli 25	0.24	0.47	0.59	0.71	0.38	0.38	0.53	0.50	0.47	0.40	0.50	0.27	0.35	0.56	0.56	0.50	0.31	0.40	0.38	0.47	0.71	0.74	0.71	0.67	1	0.47	0.43	0.73	0.67	0.56
E. coli 28	0.38	0.38	0.38	0.25	0.40	0.53	0.29	0.27	0.50	0.53	0.40	0.57	0.63	0.35	0.59	0.40	0.50	0.29	0.53	0.50	0.63	0.56	0.50	0.47	0.47	1	0.45	0.48	0.29	0.59
E. coli 32	0.45	0.45	0.64	0.36	0.48	0.48	0.50	0.48	0.55	0.40	0.38	0.30	0.55	0.52	0.52	0.57	0.11	0.40	0.57	0.64	0.55	0.67	0.55	0.61	0.43	0.45	1	0.74	0.40	0.52
E. coli 33	0.57	0.57	0.57	0.57	0.30	0.50	0.42	0.50	0.57	0.50	0.50	0.32	0.57	0.55	0.64	0.60	0.24	0.32	0.50	0.38	0.57	0.78	0.57	0.55	0.73	0.48	0.74	1	0.63	0.64
E. coli 34	0.00	0.43	0.43	0.57	0.31	0.31	0.50	0.15	0.29	0.35	0.31	0.33	0.57	0.40	0.40	0.31	0.40	0.33	0.46	0.14	0.57	0.50	0.43	0.40	0.67	0.29	0.40	0.63	1	0.67
E. coli 15F	0.35	0.35	0.59	0.47	0.50	0.50	0.40	0.38	0.35	0.60	0.50	0.53	0.59	0.56	0.44	0.38	0.31	0.13	0.38	0.47	0.71	0.63	0.71	0.56	0.56	0.59	0.52	0.64	0.67	1

with 78% similarity, sample 20 with sample 21 with 78% similarity, sample 19 with sample 20 with 78% similarity, sample 06Emb with sample 02 with 86% similarity, sample 10 with sample 14 with 88% similarity, sample no. 1Emb with sample 4Bis with 88% similarity, sample 19Emb with sample 01Camp with 92% similarity. This shows that phylogenetically, *E. coli* bacteria from the sample have various species similarity similarities.

Analysis of genetic profile of Escherichia coli samples and construction of phylogenetic trees

The results of the phylogenetic tree design of *E.coli* bacteria are shown in the following figure:

Phylogenetic tree on Figure 3 shows isolate 01CAMP was in the same strain as isolate 09. Isolate 1EMB was in the same strain as *E. coli* strain KC819123.1. Isolate number 02CAMP was in the same strain as *E. coli* strain MW026027.1. Isolate 17EMB was in the same strain as isolate 04EMB. Isolate no.01 was in the same strain as isolate no.28. Isolate no.17 was in the same strain as *E. coli* strain JQ907530.1. Isolate number ECIS was in the same strain as *E. coli* strain KU156697.1. Isolate 02 was in the same strain as isolate 10. Isolate 22 was in the same strain as isolate 33. Isolate 14 was in the same strain as isolate 5F. Isolate 29 was in the same strain as isolate 32. Isolate 4BIS was in the same strain as isolate 16. Figure 3 shows that many of the *E. coli* bacteria from the sample were not in the same branch or node as bacteria from other samples in the phylogenetic tree. *E. coli* bacteria from the sample have branches that vary between bacterial strains of other samples in the phylogenetic tree.

Figure 4 shows that the *E. coli* bacteria from several samples are in the same branch. Isolate 17EMB was in the same strain as isolate 04EMB. Isolate 19EMB was in the same strain as isolate 09. Isolate no.21 was in the same strain as isolate 28. Isolate 02 was in the same strain as isolate 10. Isolates that are on the same branch indicate the similarity of species [15].

Enterobacterial repetitive intergenic consensus-polymerase chain reaction genetic profile analysis and phylogenetic tree construction

ERIC sequences from 30 *E. coli* isolates from travel diarrhea samples were successfully amplified using ERIC 1R and ERIC 2 primers. Visualization of ERIC-PCR profiles showed that the banding pattern of *E. coli* isolates from samples was identical between one isolate and another isolate, and some were not. The ERIC-PCR profiles of sample isolates have varied patterns.

A total of 30 *E. coli* isolates from traveller's diarrhea samples were analyzed using ERIC-PCR to determine their genetic diversity (Figure 5). The results of the analysis using the ERIC-PCR phylogenetic tree showed that the isolates from the sample had varied genetic diversity. The ERIC-PCR profile phylogenetic tree of the sample isolates in Figure 6 shows that there is still a relationship between the sample *E. coli* isolates. Based on this phylogenetic tree, it appears that *E. coli* isolates from traveler's diarrhea samples were genetically diverse so they were in separate groups. The group can be divided into five major groups, namely groups I, II, III, IV and V. Based on the results of the reconstruction of the phylogenetic tree and the results of ERIC-PCR (Figures 3, 4, and 6), it was known that there were types of *E.coli* bacteria from the traveler's diarrhea samples that have the same phenotype and have more dominant numbers than those with different phonotypes. Identification based on the gene encoding 16S rRNA stated that 30 isolates from patients with traveler's diarrhea were in the same branch and some were not in the same phylogenetic branch. The isolates of *E. coli* bacteria had similarities

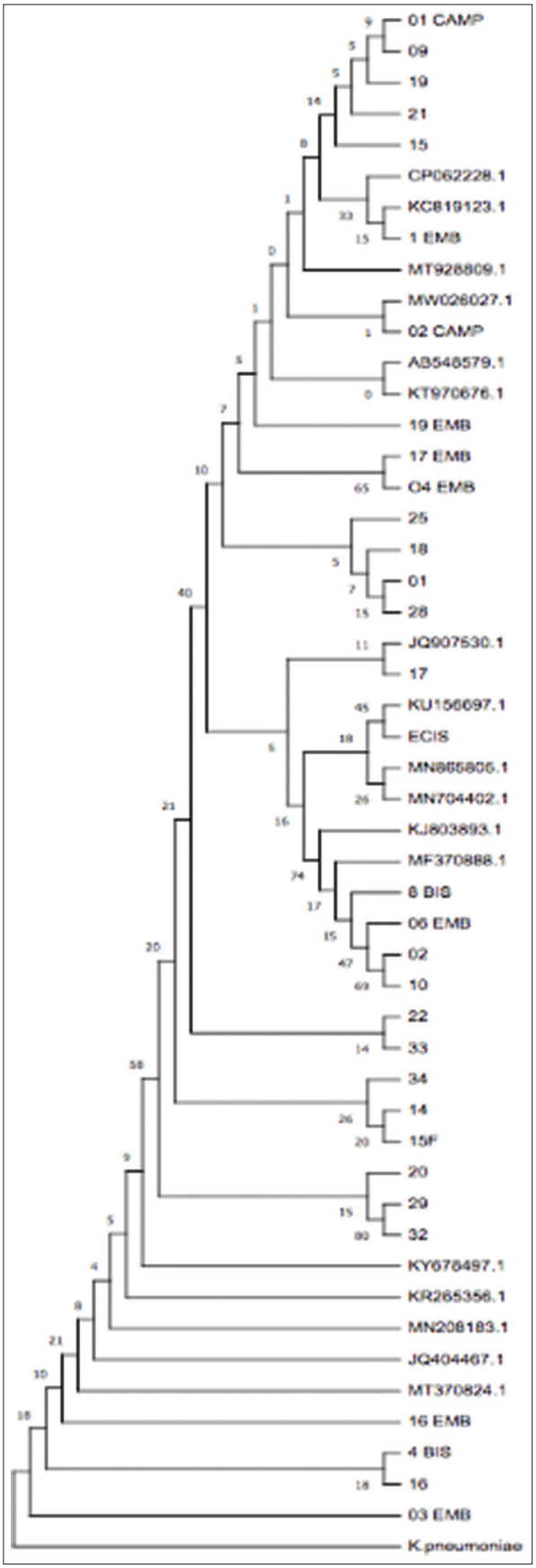

Figure 3: Phylogenetic tree of Escherichia coli isolates from 31 samples and using several isolates taken from GenBank for later comparison

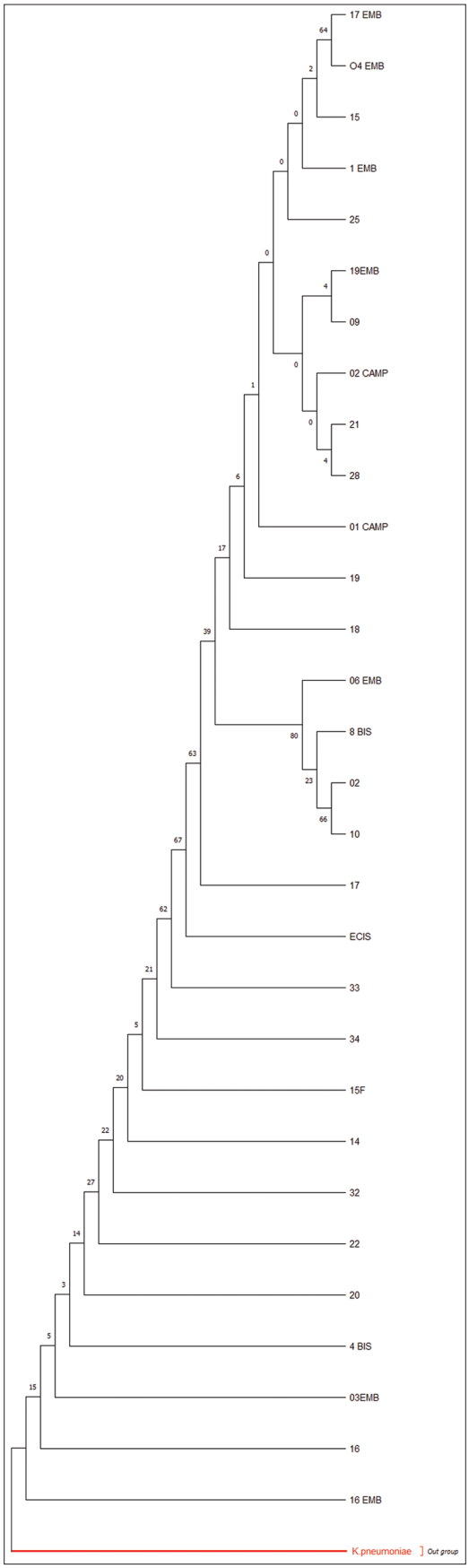

Figure 4: Phylogenetic tree of Escherichia coli isolates from 30 samples and 1 as control using K. pneumoniae isolates

that varied from 0% to 100%. The results of genetic DNA profile analysis using ERIC-PCR showed that *E. coli* isolates from traveler's diarrhea specimens had genetic clones (clonal variability) that varied compared

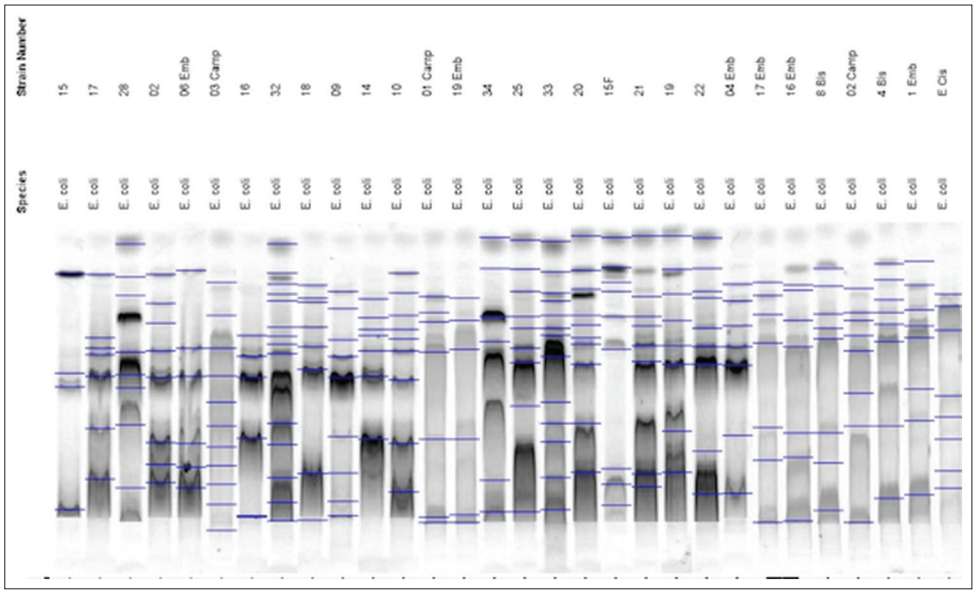

Figure 5: Enterobacterial repetitive intergenic consensus-polymerase chain reaction profile of Escherichia coli from isolates of traveller's diarrhea samples

to *E. coli* isolates between samples and other *E. coli* isolates from GenBank.

E. coli isolates from the sample formed five groups, namely group I consisting of 1 isolate that showed different diversity, namely isolate code 15. Group II showed isolates that had similar genetic diversity as indicated by 5 isolates, namely isolate code 17, 28, 02, 06Emb, 03Camp. Group III showed isolates that had similar genetic diversity as shown by 8 isolates, namely isolates coded 16, 32, 18, 09, 14, 10, 01Camp and 19Emb. Group IV showed isolates that had similar genetic diversity as shown by 15 isolates, namely isolates coded 34, 25, 33, 20, 15F, 21, 19, 22,

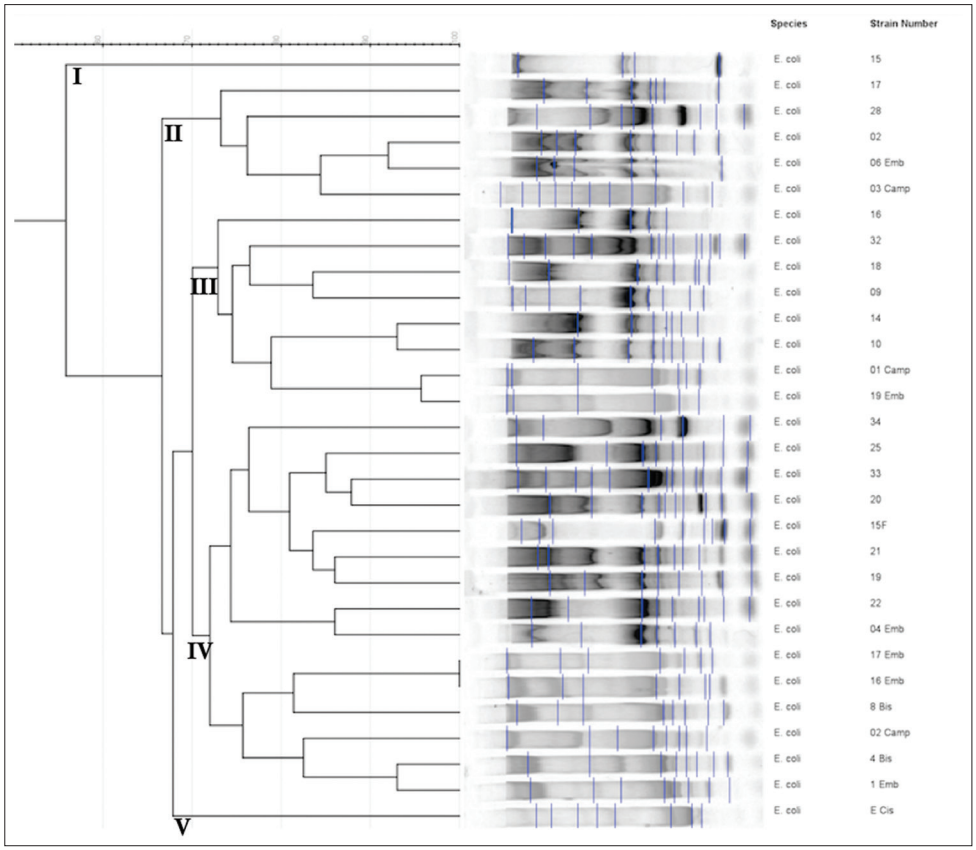

Figure 6: Enterobacterial repetitive intergenic consensus-polymerase chain reaction phylogenetic tree of Escherichia coli isolates which divided into 5 large groups, namely groups I, II, III, IV and V

04Emb, 17Emb, 16Emb, 8Bis, 02Camp, 4Bus and 01Emb. Group V consisted of 1 isolate that showed different diversity, namely isolates coded ECis.

Discussion

In this research, many cases of traveler's diarrhea are not recorded or reported because tourists do not come to health services to check themselves and seek treatment because tourists usually try to treat their illness themselves by taking anti-diarrheal drugs that they have brought from their country or bought at pharmacies or a drug store in Bali assisted by a tour guide. Another reason is because they choose to rest in hotels and not go to health services. The pattern of tourists in terms of looking for a place of health services in an effort to help for the treatment of diarrhea also has variations. The health services they can look for are: hospitals, health centers, port health offices, 24-h clinics and private doctor services. For hospital health services, we found in this research there are also variations, namely more tourists visit private hospitals compared to government central hospitals because of considerations of locations close to tourist areas. When we compare the incidence of diarrhea to tourists who come to the hospital compared to 24-hour clinics or private doctors, both at the practice and visiting tourists' residences (on call) also tend to be higher because of easy access and simpler administration. The incidence appears to be lower outside the hospital compared to other reports of coming to the hospital. Possibly because it is not only hospitals that tourists visit when they suffer from diarrhea and also because they are constrained by poor documentation.

Currently, a faster molecular-based identification method with a high level of sensitivity and specificity has been developed, namely by sequencing analysis of the 16S rRNA gene. 16S rRNA gene become the most commonly used molecular marker in microbial ecology [16]. The 16S rRNA gene is also often referred to as 16S rDNA, but according to the consensus of the American Society for Microbiology, the term 16S rRNA is considered more appropriate. The gene coding for ribosomal RNA (rRNA) is the most conserved gene [17]. The proportion of rDNA sequences of each genetically correlated organism is generally the same. Thus each organism that has a certain kinship distance can be aligned so that it is easier to determine the differences in the sequences that characterize these organisms. The gene coding for rRNA is used to determine taxonomy, phylogeny (evolutionary relationships) and estimate the distance of diversity between species (rates of species divergence) of bacteria. A comparison of rDNA sequences can show evolutionary relationships between organisms [16].

The analysis of 16S rRNA uses the PCR principle, which involves several cycles, each of which consists of three sequential steps, namely the separation (denaturation) of the DNA template chain, annealing of the primer pair on the target DNA and primer extension or polymerization reaction catalyzed by DNA polymerase. This conserved region also causes this gene to be used as a universal primer used in PCR, and its nucleotide sequence can be determined through sequencing [13].

Figure 3 depicts the result of phylogenetic tree of same strain E. coli from GenBank and other samples.

Figure 4 shows the relationship between one bacterium and another. From the design of the phylogenetic tree, it can be concluded that the *E. coli* isolates from the sample were not all in the same phylogenetic branch but were still in the same clade (species), and one node (genus). This indicates that some *E. coli* isolates from the sample have similar nucleotide base sequences and phylogenetic proximity to each other, and some do not. These isolates had the same root (ancestor) but underwent different changes from each other when they evolved. In addition, the isolates of *E. coli* bacteria from the sample were not new bacterial species because the homology values of the two isolates of *E. coli* bacteria from the samples varied widely between 0 and 100%. Bacteria can be said to be a new species if they have a nucleotide base homology <70%. A species can be said to have a relationship with one of the existing species groups if it has a gene homology value >70% when compared to all genes undergoing DNA-DNA hybridization. The total value of genes undergoing DNA-DNA hybridization is the main key in determining and limiting the relationship between the new species and existing species [18].

Based on the sequence data of the gene encoding 16S rRNA, a newly discovered bacterial isolate can be said to be in the same genus group as bacteria that already exist in the GenBank data if it has a 16S rRNA gene sequence homology with a value between 97 - 99%. If the homology value of the 16S rRNA gene sequence is less than 97%, then the bacteria cannot be referred to as new bacteria or classified as bacteria of a different genus [19]. However, this needs to be explored through phylogenetic analysis by looking at the branching formed by isolates through observing the position occupied among other species or comparison species.

The phylogenetic tree in this study was constructed using the neighbor-joining tree method. The neighbor joining tree method selects sequences which, when combined, will provide the best estimate of the branch length that most closely reflects the actual

distance between the sequences. A group of organisms whose members share many characteristics or traits are considered to have a very close relationship and are thought to be descended from a common ancestor. Ancestors and members are assumed to carry the same genetic and biochemical traits or patterns [20].

To determine the exact taxonomic position of the new bacteria, it is necessary to carry out several evaluations. The evaluation includes the evaluation of the phylogenetic position of bacteria with the entire phylogenetic group in GenBank, evaluation of chemotaxonomy, and evaluation of the phenotype of bacteria that have the closest phylogenetic relationship. If the results of the phenotypic and chemotaxonomic evaluation support the results of the phylogenetic evaluation, then the group of bacteria can be determined by taxa and named according to the genus that has been evaluated [21].

Genetic clone diversity is the variation in genetic clones within a species both between geographically separated populations and between individuals within a population. An individual in a population has genetic differences from one another. Variations in genetic clones arise because each individual has unique gene forms. Genetic clonal variation increases when offspring receive unique combinations of genes and chromosomes from their parents through gene recombination. This process increases the potential for random genetic variation, resulting in different combinations. Genetic clone diversity is a variation within a population that occurs due to the diversity among individuals who are members of the population [22].

Intraspecies variation and subspecies genotype variant is a concept of microheterogeneity in a species. This microheterogeneity usually shows a difference of <0.5% or only a few base pairs per 16S rRNA gene sequence. The significance of the microheterogeneity of a strain makes it possible to distinguish important phenotype, pathogenicity, and genetic differences between strains. Microheterogeneity has also been used for strain tracking and epidemiological studies [17].

Conclusions

Not all sample isolates contained *E. coli* bacteria and the presence of *E. coli* in traveler's diarrhea patients was not always consistently found in every health facility. Identification based on the 16S rRNA encoding gene stated that 30 isolates from patients with traveler's diarrhea were in the same branch, and some were not in the same phylogenetic branch. Isolates of *E. coli* bacteria had similarities that varied from 0% to 100%. The results of genetic DNA profile analysis using ERIC-PCR showed that *E. coli* isolates from traveler's diarrhea specimens had genetic clones (clonal variability) that varied compared to *E. coli* isolates between samples and other *E. coli* isolates from GenBank.

Acknowledgments

We would like to acknowledge medical doctors and laboratory analysts in hospitals and clinics in Bali, Indonesia, who participated in this study. We also would like to thank staffs in Port Health Office I Denpasar, Bali, Indonesia.

References

1. Pratiwi A. Popular Destinations in Tourism: Case Study of TripAdvisor. 3rd International Conference on Vocational Higher Education (ICVHE 2018). Advances in Social Science, Education and Humanities Research; 2020. p. 426.
2. Velyniawati P, Dewantari NM, Suarjana IM. Tingkat penerimaan wisatawan asing terhadap makanan tradisional Bali. Jurnal Ilmu Gizi. 2015;6:58-65.
3. BPS Provinsi Bali. Jumlah wisatawan asing ke Bali menurut bulan 1982-20182019. Available from: https://bali.bps.go.id/statictable/2018/02/09/21/jumlah-wisatawan-asing-ke- bali-menurut-bulan-1982-2018.html. [Last accessed on 2022 Feb 07].
4. Wahyuni NW, Wirawan IM, Hendrayana MA. Risks factors for diarrhea among travellers visiting Bali. Public Health Prev Med Arch. 2019;2:121-6. https://doi.org/10.15562/phpma.v7i2.222
5. Steffen R. Epidemiology of travellers'diarrhea. J Travel Med. 2017;24:2016-9. https://doi.org/10.1093/jtm/taw072
6. Trisdayanti NP, Sawitri AA, Sujaya IN. Hygiene, sanitation and potential existence of virulent genes of *E. coli* in lawar Bali in Kuta. Public Health Prev Med Arch. 2015;3(2):124-32.
7. Jiang, ZD, DuPont, HL. Etiology of travellers' diarrhea. J Travel Med. 2017;24(1):S13-6. https://doi.org/10.1093/jtm/tax003
8. Toma C, Yan L, Higa N, Nakasone N, Chinen I, Baschkier A, *et al*. Multiplex PCR assay for identification of human diarrheagenic *Escherichia coli*. J Clin Microbiol. 2003;41(6):2669-71. https://doi.org/10.1128/JCM.41.6.2669-2671.2003
PMid:12791900
9. Jafari A, Aslani M, Bouzari S. *Escherichia coli*: A brief review of diarrheagenic pathotypes and their role in diarrheal disease in Iran. Iran J Microbiol. 2012;4(3):102-17.
PMid:23066484
10. Akondi KB, Lakshmi VV. Emerging trends in genomic approaches for microbial bioprospecting. OMICS. 2013;17(2):61-70. http://dx.doi.org/10.1089/omi.2012.0082
PMid:23308385
11. Park YS, Lee S, Sharma A. Molecular typing tools for identifying and characterizing lactic acid bacteria: A review. Food Sci Biotechnol. 2020;29:1301-18. https://doi.org/10.1007/s10068-020-00802-x
PMid:32995049
12. A'yun Q. Profil genetik *Escherichia coli* yang diisolasi dari tempe

berdasarkan Enterobacterial Repetitive Intergenic Consensus-Polymerase Chain Reaction (ERIC-PCR). Tesis Magister Sains Pada Program Studi Mikrobiologi Sekolah Pascasarjana Institut Pertanian Bogor; 2015.

13. Srinivasan R, Karaoz U, Volegova M, MacKichan J, Maeda MK, Miler S, *et al*. Use of 16S rRNA gene for identification of a broad range of clinically relevant bacterial pathogens. PLoS One. 2015;10(2):e0117617. https://doi.org/10.1371/journal.pone.0117617

    PMid:25658760

14. Madigan MT, Martinko JM, Parker J. Brock, the Biology of Microorganisms. 9th ed. Upper Saddle River, New Jersey: Pearson Prentice Hall; 2000.

15. Ludwig, W, Klenk HP. Overview: A phylogenetic backbone and taxonomic framework for procaryotic systematics. In: Boone DR, Castenholz RW, editors. The *Archae*a and the Deeply Branching and Phototrophic Bacteria. In: Garrity GM, editor. Bergey's Manual of Systematic Bacteriology. 2nd ed. New York: Springer-Verlag; 2001. p. 49-65.

16. Case RJ, Boucher Y, Dahllof I, Holmstrom C, Doolittle WF, Kjelleberg S. Use of 16S rRNA and rpoB genes as molecular markers for microbial ecology studies. Appl Environ Microbiol. 2007;73(1):278-88. https://doi.org/10.1128/AEM.01177-06

    PMid:17071787

17. Clarridge JE 3rd. Impact of 16S rRNA gene sequence analysis for identification of bacteria on clinical microbiology and infectious diseases. J Clin Microbiol. 2004;17(4):840-62. https://doi.org/10.1128/CMR.17.4.840-862.2004

    PMid:15489351

18. Moore ER, Mihaylova SA, Vandamme P, Kriscevsky M. Microbial systematics and taxonomy: Relevance for a microbial commons. Res Microbiol. 2010;161(6):430-8. http://doi.org/10.1016/j.resmic.2010.05.007

    PMid:20670913

19. Drancourt M, Bollet C, Carlioz A, Martelin R, Gayral JP, Raoult D. 16S ribosomal DNA sequence analysis of a large collection of environmental and clinical unidentifiable bacterial isolates. J Clin Microbiol. 2000;38:3623-30. https://doi.org/10.1128/JCM.38.10.3623-3630.2000

    PMid:11015374

20. Mahapatro G, Mishra D, Shaw K, Mishra S, Jena T. Phylogenetic tree construction for DNA sequences using clustering methods. Procedia Eng. 2012;38:1362-6. https://doi.org/10.1016/j.proeng.2012.06.169

21. Johnson JS, Spakowicz DJ, Hong BY, Petersen LM, Demkowicz P, Chen L, *et al*. Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. Nat Commun. 2019;10(1):5029. https://doi.org/10.1038/s41467-019-13036-1

    PMid:31695033

22. Roberfroid S, Vanderleyden J, Steenackers H. Gene expression variability in clonal populations: causes and consequences. Crit Rev Microbiol. 2016;42(6):969-84. https://doi.org/10.3109/1040841X.2015.1122571

    PMid:26731119