Genetic Profiles of *Escherichia coli* Isolated from Indonesian Tempeh Based on Enterobacterial Repetitive Intergenic Consensus-Polymerase Chain Reaction (ERIC-PCR)

QURROTA A'YUN¹, ANTONIUS SUWANTO^{1,2}*, AND TATI BARUS²

¹Department of Biology, Faculty of Mathematics and Natural Sciences, Institut Pertanian Bogor, Darmaga Campus, Bogor 16680, Indonesia; ²Department of Biology, Faculty of Biotechnology, Universitas Katolik Atma Jaya, Jalan Jenderal Sudirman No 51, Jakarta 12930, Indonesia

Tempeh is a famous Indonesian fermented food derived from soybeans inoculated with *Rhizopus* sp. Tempeh production varies depend on the producers and often conducted in an uncontrolled condition. This condition could lead to the growth of *Escherichia coli* which is known as bacterial indicators of environmental hygiene. Some strains of *E. coli* could induce diarrhea, acute gastroenteritis or gastrointestinal tract infections. The aim of this study was to compare genetic diversity of *E. coli* isolates from tempeh with medical isolates employing ERIC-PCR method. In this study, 63f and 1387r primers were used to amplify 16S rRNA genes, and ERIC 1R and ERIC 2 primers were used for ERIC-PCR analysis. Tempeh samples were obtained from four producers in Bogor. Thirty-three isolates of *E. coli* isolates from the other two producers. In addition, the same tempeh samples could carry different genotypes of *E. coli*. On the other hand, the same genotypes could be found in different tempeh samples. Based on phylogenetic tree analysis, *E. coli* from tempeh could be separated from medical isolates. We showed that *E. coli* isolates derived from tempeh were genetically different from those of medical or pathogenic isolates.

Key words: ERIC-PCR, Escherichia coli, tempeh

Tempe merupakan makanan khas Indonesia yang berasal dari fermentasi kacang-kacangan yang diinokulasi menggunakan *Rhizopus* sp. Pembuatan tempe bervariasi tergantung pengrajin dan sering pada kondisi tidak terkontrol. Kondisi yang tidak higienis dapat mengakibatkan tumbuhnya *Escherichia coli* yang dikenal sebagai bakteri indikator kebersihan lingkungan. Beberapa strain *E. coli* dapat menimbulkan diare, gastroenteritis akut atau infeksi saluran pencernaan. Tujuan penelitian ini adalah untuk membandingkan keragaman genetik isolat *E. coli* dari tempe dengan isolat medis menggunakan metode ERIC-PCR. Pada penelitian ini, primer 63f dan 1387r digunakan untuk mengamplifikasi gen 16S rRNA, dan primer ERIC 1R dan ERIC 2 digunakan untuk analisis ERIC-PCR. Sampel tempe diambil dari empat pengrajin tempe di Bogor. Tiga puluh tiga isolat *E. coli* telah berhasil diisolasi dari sampel tempe yang diproduksi oleh dua pengrajin, isolat *E. coli* tidak ditemukan dari dua pengrajin lainnya. Lebih lanjut, sampel yang sama dapat membawa genotip *E. coli* yang berbeda. Disisi lain, genotip yang sama dapat ditemukan dalam sampel tempe yang berbeda. Berdasarkan analisis pohon filogenetik, *E. coli* dari tempe terpisah dengan isolat medis. Penelitian ini menunjukkan bahwa *E. coli* yang berasal dari tempe secara genetik berbeda dengan isolat medis atau *E. coli* patogen.

Kata kunci: ERIC-PCR, Escherichia coli, tempe

Tempeh is traditional Indonesian food derived from soybeans fermented by *Rhizopus* sp. Tempeh contains vitamins B_{12} , phytochemicals, antioxidative (Keuth and Bisping 1994; Astuti *et al.* 2000), essential fatty acid (Utari 2010), and isoflavon compounds (Haron *et al.* 2009). Some report also stated that tempeh have many health benefits such as can prevent the free radicals (Utari 2010), decreased risk of heart disease and strokes, osteoporosis, cancer, and menopause (Babu *et al.* 2009). Production of tempeh is done in different ways by different producers, and generally in uncontrolled conditions. The quality of tempeh is determined by three factors i.e raw material, microorganisms, and environment. The function of *Rhizopus* provides substrates for bacteria in synthesizing components of taste and other nutritional components (Seumahu 2012). Some of bacteria were known to contribute to increase quality of tempeh, such as *Citrobacter freundii* and *Klebsiella pneumoniae* (Keuth and Bisping 1994), *Lactobacillus reuteri* (Taranto *et al.* 2003), *Bacillus* sp. (Barus *et al.* 2008) and lactic acid bacteria (Efriwati *et al.* 2013). Besides that, in tempeh

^{*}Corresponding author; Phone: +62-21-5719060, Fax: +62-21-5727615; Email: antoniussuwanto@gmail.com

also found Enterobacterial group namely *Escherichia coli* (Barus *et al.* 2008). Unhygienic conditions led to the growth of *E. coli* in water provides an indication of fecal contamination.

Normally *E. coli* is found in the digestion of both humans and animals, but *E.coli* also have an important role in the spread of zoonotic diseases through food. Certain serotypes can cause disease such as O157:H7 which produces shiga toxin (Todar 2004). Some strains of *E. coli* could induce diarrhea, acute gastroenteritis or gastrointestinal tract infections. Strains of *E. coli* pathogen consists of five groups namely: enteropathogenic *E. coli* (EPEC), enterotoxigenic *E. coli* (ETEC), enterohemorrhagic *E. coli* (EHEC), enteroaggregative *E. coli* (EaggEC) and enteroinvasive *E. coli* (EIEC) (Nataro and Kaper 1998).

Genetic differences may affect the characteristics of E. coli, particularly in relation to the medical fields. Therefore, the study of genetic diversity is important for the identification, characterization and to study evolution and epidemiology of the pathogenicity of the bacteria (Rademarker and de Bruijn 1997). One of the many molecular techniques to study genetic diversity is Enterobacterial Repetitive Intergenic Consensus-Polymerase Chain Reaction (ERIC-PCR). ERIC sequences are short sequences, which is usually 126 bp long with replications area that conserve and in the non-coding area, sequences that are not encoded into proteins (Lupski and Weinstock 1992). This technique was used because it is fast, simple, discriminative (Meacham et al. 2003) and has been successfully used to analyze the diversity of some groups of bacteria such as Lactobacillus (Stephenson et al. 2009), V. cholerae (Waturangi et al. 2012), and Klebsiella spp. (Barus et al. 2013).

MATERIALS AND METHODS

Medical Isolates of *E. coli*. Three medical isolates were used for comparison with *E. coli* from tempeh i.e. *E. coli* ATCC 25922 (collection of Microbiology Laboratory, Faculty of Medicine, University of Indonesia, Jakarta), *E. coli* O157 (collection of Microbiology Laboratory, Faculty of Medicine, Atma Jaya Indonesia Catholic University, Jakarta) EPEC K.1.1 (collection of Microbiology Laboratory, Faculty of Mathematics and Natural Science, Bogor Agricultural University, Bogor) and *E. coli* DH5a was used as control positive (collection of Microbiology Laboratory, Faculty of Technobiology, Atma Jaya Indonesia Catholic University, Jakarta). *E. coli* Isolates from Tempeh. Fresh tempeh obtained from producer EMP, DRG, WJB and CLR in Bogor, West Java, Indonesia. Tempeh sampling were conducted in four month i.e. December 2013 until March 2014. A total of 10 g of each fresh tempeh was homogenized in 90 mL of sterile physiological salt 0.85% (w/v) (NaCl). Dilution was made from 10^{-1} until 10^{-5} . A total of 100 µL from dilution of 10^{-2} , 10^{-3} , 10^{-4} , and 10^{-5} were spread on Eosin Methylene Blue (EMB) Agar (Oxoid) and incubated at 37 °C for 24 h. Colonies with characteristic metallic green on EMB Agar were purified by quadrant method. These colonies were further analyzed by cultivating them on Simmon's Citrate Agar (SCA) (Difco), and incubated at 37 °C for 24 h.

Amplification of 16S rRNA Genes E. coli. Genom of each E. coli was isolated using Cetyl Trimethyl Ammonium Bromide (CTAB) method (Sambrook and Russell 2001). The 16S rRNA gene was amplified employed a PCR machine (Applied Biosystems, 2720 Thermal Cycler) using 63f (5'CAG GCC TAA CAC ATG CAA GTC-3') and 1387r primers (5'-CCC GGG AAC GTA TTC ACC GC-3') (Marchesi et al. 1998). PCR master mix (50 µL) consists of 25 µL GoTaq Green[®] Master Mix (Promega), 2 µL primer forward and reverse (25 pmol μ L⁻¹), 19 μ L nucleas free water and 1 µL DNA template. The PCR protocol was as follows: initial denaturation at 94 °C for 5 min, denaturation at 94 °C for 30 s, annealing at 55 °C for 30 s, elongation at 72 °C for 1 min, and post extention at 72 °C for 20 min. The cycle was repeated for 35 times. A total of 5 µL of PCR amplification products were further verified by electrophoresis in 1% agarose (Bioline) for 60 min, 80 V and visualized by using UV transluminator (UVB 36 ultralum, Carson, California). Sequencing of PCR product were performed in First Base Sequensing INT, Malaysia, and were analyzed using MEGA 5.2 program. Sequencing results were compared to the database with the Basic Local Alignment Search Tool Nucleotide (BLASTN) program which is provided by National Center for Biotechnology Information (NCBI).

Analysis of Genetic Profiling of *E. coli* Isolates by ERIC-PCR. ERIC sequence of *E. coli* was amplified using ERIC 1R (5'- ATGTAAGCTCCTGG GGATTCAC-3') and ERIC 2 primers (5'-AAGTAAGTGACTGGGGTGAGCG-3') (Versalovic *et al.* 1991). PCR master mix (25 μ L) consists of 12.5 μ L GoTaq Green[®] Master Mix (Promega), 1 μ L of each other primer (25 pmol μ L⁻¹), 9.5 μ L nuclease free water, and 1 μ L DNA template. The PCR protocol was as follows: initial denaturation at 95 °C for 7 min, denaturation at 95 °C for 30 s, annealing at 49 °C for 1 min, elongation at 65 °C for 8 min, and post extention at 65 °C for 16 min. The PCR cycle was repeated for 30 times. A total of 3 μ L of PCR product was verified by electrophoresis in 1.5% agarose (w/v) for 120 min and 113 V and visualized by using UV transluminator (UVB 36 ultralum, Carson, California). Band profiles were then compared as binary number and analyzed using MEGA 5.2 software. Phylogenetics tree construction was performed employing Unweighted Pair Groups Method Analysis (UPGMA).

RESULTS

E. coli in Tempeh. Eighty-one isolates of *E. coli* were isolated from EMP, DRG, and CLR tempeh (Table 1). The colonies showed metallic green on EMB medium. However, further analysis on Simmon's Citrate medium showed only 33 isolates showed positive characters of *E. coli*, i.e the *E. coli* colonies showed not be able to utilize citrate.

16S rRNA gene of each isolate have been successfully amplified using 63f and 1387r primers with a size of DNA fragments about 1.3 kb (Fig 1). Based on BLASTN, sequence of 16S rRNA gene showed similarity with *E. coli* with maximum identities for each isolate ranging from 96% to 100%, with E-value 0.0.

E. coli was not always found in tempeh samples. In this study, *E. coli* was found only in the EMP and CLR tempeh with different population (Table 1) of four different times for sample collections. *E. coli* was found in EMP tempeh from second until fourth sampling and only once was found in the CLR tempeh.

Genetic Profiling of *E. coli* **Isolates.** ERIC sequences of 33 *E. coli* isolates from tempeh and 3 medical isolates were successfully amplified using ERIC 1R and ERIC 2 primers (Fig 2). Visualization of the ERIC-PCR profiles shows that DNA banding patterns of *E. coli* from tempeh were different to that of medical isolates. Band profiles of *E. coli* isolates from tempeh showed consistencies of size 0.25 kb and 1.0 kb. However, we found inconsistencies in the DNA bands of *E. coli* from medical isolates.

ERIC-PCR profiles isolates of *E. coli* from CLR tempeh showed similar pattern while *E. coli* isolates from EMP tempeh showed more heterogenome patterns (Fig 2). Although genetic profiles of *E. coli* isolates in EMP tempeh more varied than CLR tempeh, we found no identical profiles when isolates from tempeh compared to *E. coli* Dh5 α .

Phylogenetic tree based on ERIC-PCR profiles showed the relation between *E. coli* from tempeh and medical isolates (Fig 3). It showed that *E. coli* from tempeh were genetically different from those of medical isolates. *E. coli* from tempeh formed a separate group (Group I-IV), while medical isolates formed Group VI and *E. coli* DH5a formed Group V.

Tempeh samples	Numbering of sampling	EMB Agar (Number of isolates)	Simmon's Citrate Agar (Number of isolates)
EMP	1	_	_
	2	20	17
	3	12	11
	4	4	3
DRG	1	23	-
	2	-	-
	3	-	-
	4	20	-
WJB	1	-	-
	2	-	-
	3	-	-
	4	-	-
CLR	1	-	-
	2	-	-
	3	3	3
	4	-	-
Total of isolates		81	33

Table 1 The frequency of E. coli presence in fresh tempeh

Note (-) = Colonies which are not green metallic on EMB agar or did not utilize citrate in Simmon's Citrate agar

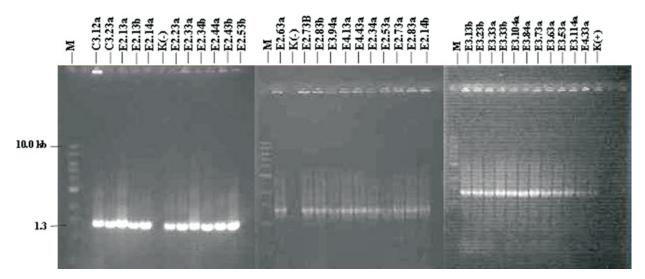


Fig 1 Amplification results of 16S rRNA genes *E. coli* from tempeh. Molecular markers (M), isolates from tempeh (C3.12a-E4.33a), positive control (K+), negative control (K-).

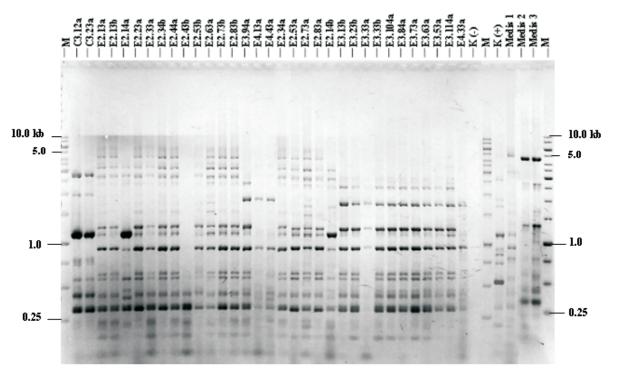


Fig 2 Genetic profiles ERIC-PCR isolates of *E. coli* from tempeh and medical isolates. Molecular markers (M), isolates from CLR tempeh (C3.12a, C3.23a), isolates from EMP tempeh (E2.13a-E4.33a), negative control (K-), *E. coli* DH5α (K+), *E. coli* ATCC 25922 (Medis 1), EPEC K.1.1 (Medis 2), and *E. coli* O157 (Medis 3).

DISCUSSION

The presence of *E. coli* in tempeh have been reported by Barus *et al.* (2008). *E. coli* from tempeh were varied in each sampling of the same producers (Table 1). *E. coli* was not always found in tempeh even the ones produced by the same producers despite of in the same tempeh. *E. coli* in tempeh may be affected by various sources such as raw materials, tools, people, and surrounding environment. The spread way of *E.*

coli in tempeh can be occurred due to cross contamination directly (by hand) and indirectly (through the water) during processing (Antara and Gunam 2002). The presence of *E. coli* in food indicated contamination due to non-hygienic food processing or handling. Therefore, *E. coli* is often used as an indicator of sanitary (Githiri *et al.* 2009).

Molecular approach can be used to detect the presence of *E. coli* by amplifying the DNA sequences which encoded 16S rRNA gene. 63f and 1387r primers

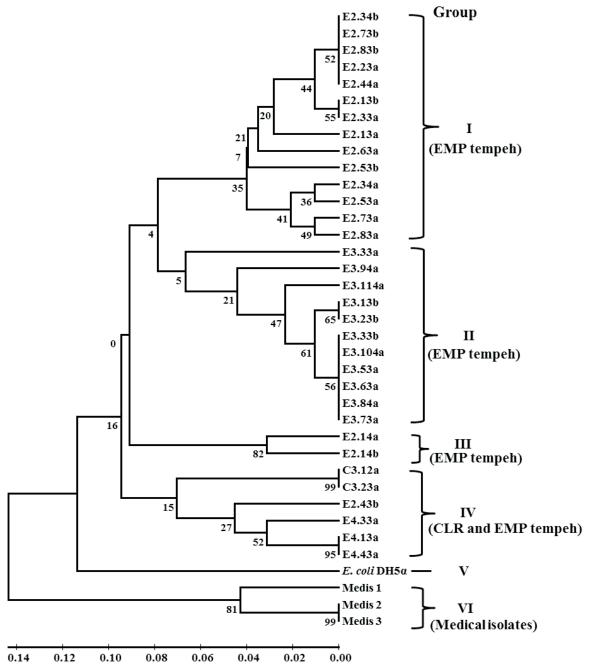


Fig 3 Phylogenetic tree constructed from of ERIC-PCR profiles using UPGMA method through MEGA 5.2 program with 1000x bootstrap. *E. coli* DH5α, *E. coli* ATCC 25922 (Medis 1), EPEC K.1.1 (Medis 2), *E. coli* O157 (Medis 3), and Group (I-VI).

were used to amplify 16S rRNA gene sequences of *E. coli*. During the amplification process, 63f primer starts to amplify the DNA template on the position of 43-63, meanwhile 1387r primer were amplified the DNA on the position of 1387–1370, so that the resulting product approximately 1.3 kb in size (Marchesi *et al.* 1998). The 16S rRNA sequences were used as moleculer markers because of their ubiquity and their identical function in all cellular-based organisms in addition to their conservative and varied sequence (Pangastuti 2006).

Analysis of bacterial virulence can be recognized

from the phenotype (Stenutz *et al.* 2006). Some serotypes of *E. coli* known to have virulence factors are O (somatic), H (flagellar), and K (capsular) surface antigen profiles (Nataro and Kaper 1998). Serotype of O157:H7 can induce the secretion of body fluids excessively and continuously that can lead to diarrhea or meningitis. *E. coli* O157 harbors VT1 and VT2 genes, generally found on pig (Suardana *et al.* 2007).

In this study, the multiple DNA fragments of all *E. coli* strains generated from ERIC primers were composed of 4-13 bands ranging in sizes between 0.25 kb and 5 kb (Fig 3A). The results indicated the extent of genetic diversity in *E. coli*. Meacham (2003) stated that the difference in the number and size of the ERIC-PCR profiles described genetic diversity among bacterial isolates. ERIC-PCR technique employs consensus primers during PCR to amplify DNA sequences located between successive repetitive sequence. It is often used for subtyping Gram-negative enteric bacteria (Hulton *et al.*1991).

Isolates of *E. coli* from tempeh showed similar characteristics in EMB media. However, the genomic profiles of *E. coli* from tempeh were different from those of medical isolates. UPGMA phylogenetic tree (Fig 3) showed that isolates of *E. coli* from tempeh formed separates group from medical isolates. Isolates of *E. coli* from CLR tempeh did not showed genetic diversity and closely related to some isolates of EMP tempeh which clustered in Group IV. In contrast to *E. coli* isolates from CLR tempeh, *E. coli* from EMP tempeh showed more genetic diversity. *E. coli* isolate from EMP tempeh are both closely and distantly related.

ERIC-PCR technique was successfully employed to determine genetic profiles of *E.coli* from tempeh and medical isolates and it showed more discriminatory power than 16S rRNA genes analysis. Previous study reported that ERIC-PCR was employed to obtain ERIC sequences of *Klebsiella* spp. from tempeh and it showed more discriminative results than 16S rRNA genes analysis (Barus *et al.* 2013). Ayu *et al.* (2014) reported that ERIC-PCR was able to differentiate genetic profiles of tempeh-derived *K. pneumoniae* isolates from those of medical isolates.

Our results showed that *E. coli* isolates from tempeh could be a distinctive non-pathogenic groups. Not all tempeh contain *E. coli* and the presence of *E. coli* in tempeh was not always found in every sample and every batch. Based on the analysis of genetic DNA profile using ERIC-PCR showed that *E. coli* from tempeh were different from *E. coli* pathogens or associated with medical isolates. However, there is no single report that consumption tempeh which contain *E. coli* will cause disease. Non aesthetic traditional tempeh production may contain some *E. coli* isolates. However, it does not mean necessarily contain dangerous bacteria.

ACKNOWLEDGMENT

This research was supported by thesis scholarship Lembaga Pengelola Dana Pendidikan (LPDP) and DANA DIPA IPB 2014 to Antonius Suwanto.

- Antara S, Gunam IBW. 2002. Dunia Mikroba (Bahaya Mikrobiologis pada Makanan). Pusat Kajian Keamanan Pangan. Denpasar (ID): Universitas Udayana.
- Astuti M, Meliala A, Dalais FS, Wahlqvist ML. 2000. Tempe, a nutritious healthy food from Indonesia. Asia Pac J Clin Nurt. 9(4):322-325.
- Ayu E, Suwanto A, Barus T. 2014. *Klebsiella pneumoniae* from Indonesian tempeh were genetically different from that of phathogenic isolates. J Microbiol Indones. 8(1):9-15. doi: 10.5454/mi.8.1.2.
- Babu PD, Bhakyaraj R, Vidhyalaksmi R. 2009. A low cost nutritious food "tempeh". World J Dairy food Sci. 4(1): 22-27.
- Barus T, Suwanto A, Wahyudi AT, Wijaya H. 2008. Role of bacteria in tempe bitter taste formation; microbiological and molecular biological analysis based on 16S rRNA gene. J Microbiol Indones. 2(1):17-21. doi:10.5454/mi.2.1.4.
- Barus T, Hanjaya I, Sadeli J, Lay BW, Suwanto A, Yulandi A.
 2013. Genetic diversity of *Klebsiella* spp. isolated from tempe based on enterobacterial repetitive intergenic consensus-polymerase chain reaction (ERIC-PCR). Hayati J Biosci. 20(4):171-176. doi:10.4308/hjb.20.4. 171.
- Efriwati, Suwanto A, Rahayu G, Nuraida L. 2013. Population dynamics of yeasts and lactic acid bacteria (LAB) during tempeh production. Hayati J Biosci. 20(2):57-64. doi:10.4308/hjb.20.2.57.
- Githiri M, Okemo P, Kimiywe J. 2009. Hygienic practices and occurrence of coliforms and *Staphylococcus* on food at a public hospital in Kenya. J Appl Biosci. 27: 1727-1731.
- Haron H, Ismail A, Azlan A, Shahar S, Loh Su Peng. 2009. Daidzein and genestein contents in tempeh and selected soy products. Food Chem. 115:1350-1356. doi: 10.1016/j.foodchem. 2009.01.053.
- Hulton CSJ, Higgins CF, Sharp PM. 1991. ERIC sequences: a novel family of repetitive elements in the genomes of *Escherichia coli*, *Salmonella typhimurium* and other enterobacteria. Mol Microbiol. 5: 825-834.
- Keuth S, Bisping B. 1994. Vitamin B_{12} production by *Citrobacter freundii* or *Klebsiella pneumoniae* during tempeh fermentation and proof of enterotoxin absence by PCR. Appl Environ Microbiol. 60(5):1495-1499.
- Lupski JR, Weinstock GM. 1992. Short, interspersed repetitive DNA sequences in prokaryotic genomes. J Bacteriol. 174(14):4525–4529.
- Marchesi JR, Sato T, Weihtman J, Martin TA, Fry JC, Hiom SJ, Wade WG. 1998. Design and evaluation of useful

bacterium-specific PCR primers that amplify genes coding bacterial 16S rRNA. Appl Envirol Microbiol. 64:795-799.

- Meacham KJ, Zhang L, Foxman B. Bauer RJ, Mars CF. 2003. Evaluation of genotyping large numbers of *Eschericia coli* isolates by enterobacterial repetitive intergenic consensus-PCR. J Clin Microbiol. 41(11):5224-5226. doi: 10.1128/JCM.41.11.5224-5226.2003.
- Nataro JP, Kaper JB. 1998. Diarrhegenic *Escherichia coli*. J Clin Microbiol. 11(2):142-201.
- Pangastuti A. 2006. Species definition of procaryotes based on 16S rRNA and protein coding genes sequences. Biodivers. 7(3):292-296.
- Rademaker JLW, de Bruijn FJ. 1997. Characterization and classification of microbes by REP-PCR genomic fingerprinting and computer assisted pattern analysis, DNA Markers: Protocols, Applications and Overviews. New York (US): Wiley and Sons. p151-171.
- Sambrook J, Russell DW. 2001. Molecular Cloning: A Laboratory manual, 3th Ed, New York (US): Cold Spring Harbor Laboratory Pr.
- Seumahu CA. 2012. Metagenome analysis to fingerprint bacterial and fungi communities in tempe [dissertation]. Bogor (ID): Institut Pertanian Bogor.
- Stephenson DP, Moore RJ, Allison GE. 2009. Comparison and utilization of repetitive-element PCR techniques for typing *Lactobacillus* isolates from the chicken gastrointestinal tract. Appl Environ Microbiol.

75(21):6764-6776. doi:10.1128/AEM.01150-09.

- Suardana IW, Sumiarto B, Lukman DW. 2007. Isolasi dan identifikasi *Escherichia coli* O157:H7 pada daging sapi di Kabupaten Badung Provinsi Bali [Isolation and identification *Escherichia coli* O157:H7 on beef in Kabupaten Propinsi Badung Bali]. J Vet. 8(1):16-23.
- Taranto MP, Vera JL, Hugenholtz J, De Valdez GF, Sesma F. 2003. Lactobacillus reuteri CRL1098 produces cobalamin. J Bacteriol. 185(18):5643-5647. doi: 10.1128/JB.185.18.5643-5647.2003.
- Todar K. 2004. The Enteric Bacteria. Online Textbook's Todar of Bacteriology/ University of Wisconsin-Madison, Dept. of Bacteriology aeruginosa Bakteriologi. University of Wisconsin-Madison, Department.
- Utari DM. 2010. Kandungan asam lemak, zink, dan copper pada tempe, bagaimana potensinya untuk mencegah penyakit degeneratif?.[Fatty acids, zinc, and copper in tempe, how its potential to prevent degenerative diseases?] Gizi Indon. 33(2):108-115.
- Versalovic J, Keuth T, Lupski JR. 1991. Distribution of repetitive DNA sequences in eubacteria and application to fingerprinting of bacterial genomes. Nucleic Acids Res. 19(24):6823-6831. doi:10.1093/nar/19.24.6823.
- Waturangi DE, Joanita I, Yogi Y, Thomas S. 2012. Use of REP- and ERIC-PCR to reveal genetic heterogeneity of *Vibrio cholerae* from edible ice in Jakarta, Indonesia. Gut Pathog. 4(2):1-9. doi:10.1186/1757-4749-4-2.