



Metagenomic analysis of bacterial diversity in pigeon pea after soaking in water

Yuni Sine¹, Donny Widiyanto^{1,2}, Yekti Asih Purwestri^{1,3}, Byong Hoon Lee⁴, Widodo Widodo^{1,5,*}

¹The Graduate School of Biotechnology, Universitas Gadjah Mada, Jl Teknika Utara Berek, Yogyakarta 55281, Indonesia

²Department of Microbiology, Faculty of Agriculture, Universitas Gadjah Mada, Jl Flora Bulaksumur, Yogyakarta 55281, Indonesia

³Laboratory of Biochemistry, Faculty of Biology, Universitas Gadjah Mada, Jl Teknika Selatan, Sekip Utara, Yogyakarta 55281, Indonesia

⁴Departments of Microbiology/Immunology and Food Science, McGill University, Montreal, QC, Canada H3A 2B4

⁵Faculty of Animal Science, Universitas Gadjah Mada, Jl Fauna 3 Bulaksumur, Yogyakarta 55281, Indonesia

*Corresponding author: widodohs@ugm.ac.id

SUBMITTED 22 February 2024 REVISED 21 May 2024 ACCEPTED 22 May 2024

ABSTRACT This study investigated the diversity of bacterial community in the samples of pigeon pea (*Cajanus cajan* L. Millsp.) soaked in water for 12 h and 24 h. The detection of certain bacterial species in the samples that can be isolated and potentially be used as starter cultures in the development of pigeon pea-based functional foods is the importance of this study. For bacterial identification, the V1–V9 regions on the 16S ribosomal RNA gene were amplified using 27F and 1492R primers under specific polymerase chain reaction conditions. Genomic DNA (130 ng) was sequenced on the R9.4 flow cell by Oxford Nanopore Technologies using a GridION sequencer. Library preparations were conducted using a Native Barcoding Kit 24 V14 (SQK-NBD114.24). Primary data were acquired using MinKNOW version 22.05.7. A total of 13 bacterial families and 89 genera were identified in the pigeon pea sample soaked for 12 h, and 26 families and 90 genera were identified in the pigeon pea soaked for 24 h. The values of five diversity indices showed that the sample soaked in water for 24 h had richer bacterial abundance and diversity than for 12 h. Shannon and Simpson values revealed the higher bacterial diversity in the samples collected at 24 h than in those collected at 12 h. Species observation and abundance-based coverage estimators (ACE) values demonstrated that the samples collected at 24 h harbored higher bacterial richness than those collected at 12 h. Bacterial communities during soaking of the pigeon pea were dominated by the family Enterobacteriaceae and genus *Enterobacter*. The presence of bacterial genera like *Lacticaseibacillus*, *Lentilactobacillus*, and *Secundilactobacillus* is interesting because of their importance as starter cultures for fermented plant-based milk products.

KEYWORDS Bacterial diversity; Lactic acid bacteria; Metagenomic analysis; Pigeon pea; Soaking time

1. Introduction

Functional foods have positive effects on human health. Its physiological properties are determined by its bioactive components, such as dietary fiber, antioxidants, phytochemicals, polyunsaturated fatty acids, prebiotics, probiotics, postbiotics, and synbiotics. In Indonesia, the utilization of local food sources for functional foods is essential to meet the protein needs of the rural community. One example of functional foods produced using local food sources is tempeh, a traditional fermented soybean dish prepared using *Rhizopus* and that has been consumed in various countries around the world. Qiao et al. (2022) reported the health benefits of tempeh and other fermented soy foods due to their antioxidative, antihypertensive, anti-inflammatory, anticancer, and neuroprotective properties. During tempeh fermentation, certain microorganisms degrade proteins to amino acids that determine the flavor and

aroma of tempeh. These native microorganisms in soybeans affect the chemical composition and aroma of final product (Nurdini et al. 2015).

Apart from soybean, pigeon pea (*Cajanus cajan* L. Millsp., local name: *kacang gude*) is another local legume that is widely used as the raw ingredient for tempeh fermentation, which belongs to the family Fabaceae (Valenzuela and Smith 2002). Indonesian pigeon pea has several advantages over other legume crops, such as being drought- and infertile soils-tolerant (Varshney et al. 2012) due to its deep roots. Therefore, pigeon pea can be cultivated in dry areas where soybean plants do not grow well. Its deep rooting does not interfere with nutrition absorption of other plants, thus pigeon pea can be intercropped (Sheahan 2012).

Pigeon pea is cultivated in several region in Indonesia due to its edible seeds. The fruit of pigeon pea is a pod that is 4–10 centimeter long, hairy, flat, and green in

color. Pigeon pea seeds are round and small, with the number of seeds per pod ranging from four to nine (Valenzuela and Smith 2002) and the pods are straight and crescent in shape, and the seed coat color varies from being grayish white, cream, yellow, purplish brown, or black. In addition, the seed coats are smooth and shiny, and the seed weight varies between 4 and 26 g per 100 grains (Maaesen 1985). Pigeon pea seeds are composed of the seed coat (14%), embryo (1%), and cotyledons (85%). The nutritional content of pigeon pea seeds per 100 g is 21.7 g of protein, 1.5 g of fat, 62.8 g of carbohydrates, 12.7 g of water content, and 336 kcal total energy provided (Abebe 2022). Pigeon pea has been traditionally used as a medicinal plant (Akande et al. 2010); however, it contains cyanide and antinutritional compounds. Hence, further processing is warranted to reduce such compounds. Pigeon pea fermentation increases the availability of nutrients and health benefits of pigeon pea as a functional food (Nwosu et al. 2013).

Soybean and many other pulses are an excellent food source due to their high amounts of either dietary fibers, proteins, or micronutrients and phytochemicals. However, the consumption of pulse products has been somewhat limited because of intestinal disturbances such as flatulence, which is caused by the presence of oligosaccharides, such as raffinose and stachyose, which are nondigestible and not assimilated in the small intestine by the human GI enzymes, are fermented by the microbiota and thus responsible for flatulence (Liu et al. 2022).

Fermentation is a food biotransformation process that provides positive nutritional and sensory properties to the products depending on the microorganisms being used as starter cultures (Gan et al. 2017; Tiwari et al. 2020). Certain microorganisms can synthesize vitamins, such as folic acid, riboflavin, niacin, thiamin (Yang et al. 2015), and B12 (Gu et al. 2015) from the fermentation of grain legumes as a substrate. Other microorganisms can metabolize n-hexanal and pentanal that give a beany flavor to the products (Desai et al. 2002). Furthermore, fermentation can reduce the level of oligosaccharides, which can cause postprandial flatulence from legumes (Sandberg 2011), and decrease antinutritional components, such as tannins, phytic acid (Fredrikson et al. 2002), and trypsin inhibitors (Fredrikson et al. 2002). Different strains of the same species can have completely distinct metabolic patterns, which consequently affect the taste and texture of the product (Hickisch et al. 2016).

The process of making tempeh from pigeon pea is carried out in two stages. The first stage is soaking the pigeon pea for approximately 24 h. Soaking is carried out for pigeon pea acidification which occurs through bacterial activity. At this stage, the pH drops from 7 to 4, which is important for the growth of *Rhizopus oryzae*. This species is inoculated with pigeon pea in the second stage, which lasts approximately 48 h. However, information regarding the types of bacteria that play a role in the acidification of pigeon pea tempeh is limited. Our interest is to determine the bacteria responsible for pigeon pea acidification

and select those that can be used as starter cultures for the preparation of pigeon pea-based functional food products. In this study, we use metagenomic analysis by sequencing 16S ribosomal RNA (16S rRNA) genes that have widely been used for monitoring microbial population. This study aims to detect and identify all bacteria that play a role in the acidification of pigeon pea soaked for 12 and 24 h and measure their abundance in water-immersed pigeon pea. Soaking pigeon pea for 12 and 24 h has previously been reported to increase hydration of grain legumes and accelerate the growth of lactic acid bacteria, and increase the acidity which provides a suitable pH for the growth of *Rhizopus oryzae* (Prativi et al. 2023).

2. Materials and Methods

2.1. Pigeon pea collection and sample preparation

Pigeon pea samples were collected from Nusa Tenggara Timur Province in Indonesia (-10.110251059792311, 123.81441076672097) and were sorted to obtain peas with good or intact pods. In brief, 500 g of pigeon pea were soaked in sterile distilled water at room temperature (± 25 °C) for 12 and 24 h. Afterward, 10 ml of the soaking water was collected to measure the bacterial diversity using metagenomic analysis (Demarinis et al. 2022; Yarlina et al. 2022).

2.2. Genomic DNA extraction

In brief, 10 mL of each soaking water sample was centrifuged at $2,500 \times g$ for 10 min. The obtained pellets were washed first with saline NaCl 1.5 M and then with sterile distilled water and used for genomic DNA extraction with ZymoBIOMICS DNA Miniprep Kit D4300 (Zymo Research, Cambridge, UK). DNA concentration was determined using NanoDrop spectrophotometers and Qubit fluorometer (Thermo Fisher, Waltham, Massachusetts, USA), and DNA quality was assessed via 1% agarose gel with electrophoresis 100 V, followed by visualization using Gel-Doc EZ imager (Bio-Rad, California, USA).

2.3. Amplification of the 16s rRNA V1-V9 region

The V1-V9 regions of the 16S rRNA gene for bacterial species identification were amplified using 27F and 1492R primers (final concentration 10 μ M) with MyTaq™ HS Red Mix 2X (Bioline, Essex, UK). The following PCR conditions were applied: preliminary denaturation at 95 °C for 3 min; followed by 5 cycles of 95 °C for 15 s, 55 °C for 15 s, and 72 °C for 30 s; 30 cycles of 95 °C for 15 s, 62 °C for 15 s, and 72 °C for 30 s; and a final extension at 72 °C for 1 min.

2.4. Library preparation and sequencing

A library was prepared using a Native Barcoding Kit 24 V14 (SQK-NBD114.24) (Oxford Nanopore Technologies, Oxford, UK). The amplicons from the 16S rRNA amplification (130 ng) was sequenced using Oxford Nanopore Technology (ONT, United Kingdom), which provides the

long-read sequencing that covers the full-length sequence of the 16S rRNA gene. Sequencing was conducted on the R9.4 flow cell by ONT using a GridION sequencer (ONT, United Kingdom). Primary data were acquired using MinKNOW, version 22.05.7, the operating software that operates nanopore sequencing devices.

2.5. Bioinformatics analysis

Base calling was performed using Guppy version 6.1.5 with high-accuracy model (Wick et al. 2019). The quality of FASTQ files was visualized using NanoPlot 1.40.0 (De Coster et al. 2018). Reads were classified using Centrifuge classifier (Kim et al. 2016). An index for bacteria and archaea was built using NCBI 16S RefSeq database (<https://ftp.ncbi.nlm.nih.gov/refseq/TargetedLoci/>). Downstream analysis and visualizations were performed with Pavian (<https://github.com/fbreitwieser/pavian>), Krona Tools (<https://github.com/marbl/Krona>), and RStudio using R version 4.2.0 (<https://www.R-project.org/>) to obtain diversity index (observed species index, Chao1 index, ACE index, Shannon index, and Simpson index), and bacterial composition.

3. Results and Discussion

3.1. Metagenomic data

The bacterial diversity in the soaking water of the pigeon pea samples soaked for different times (12 and 24 h) was examined by sequencing the hypervariable regions (V1–V9) of the 16S rRNA gene. This 16S rRNA gene is well preserved and uniquely found in all bacteria and archaea, suggesting that it specifically targets and identifies bacteria and archaea present in the samples. Given that all the informative sites of 16S rRNA genes are considered, the full-length 16S rRNA sequences can provide a high level of taxonomic and phylogenetic resolution for bacte-

rial identification (Bahram et al. 2019). Using this technology, we identified a total of 13 bacterial families, 89 genera, and 495 species in the pigeon pea sample soaked for 12h; 26 families, 90 genera, and 533 species in the pigeon pea sample soaked for 24 h.

Shannon and Simpson diversity indices were applied to measure the bacterial diversity in the samples. Both consider the number of species living in a habitat and their relative abundance (Fang et al. 2015). Meanwhile, the observed species, Chao1, and abundance-based coverage estimators (ACE) indices reflect sample richness. The values of these five diversity indices are shown in Figure 1. The sample collected at 24 h had richer genera abundance and diversity than that collected at 12 h. The Shannon and Simpson values showed that the samples collected at 24 h harbored higher bacterial diversity than those collected at 12 h. Species observation and ACE values revealed that the samples collected at 24 h harbored higher sample richness than those collected at 12 h. All these findings indicated that the bacterial diversity in pigeon pea increases with the soaking time.

After being soaked in water for 12 and 24 h, pigeon pea's dominant phylum was Proteobacteria (Figure 2a and b). Among Proteobacteria, Enterobacteriaceae was the predominant family, and *Enterobacter* and *Klebsiella* were the most abundant genera in the samples (Figure 2a and b). At the species level, *Enterobacter ludwigii* (15%), *Enterobacter asburiae* (11%), and *Enterobacter cloacae* (8%) were the most abundant species in the samples soaked in water for 12 h. These three *Enterobacter* species were still the most dominant species after soaking pigeon pea in water for 24 h. Other phyla were detected, such as Firmicutes and Actinobacter; however, their relative abundance remained below 1.00%. Nur et al. (2020) reported that Firmicutes and Actinobacteria were consistently found in tempeh fermentation, with Firmicutes be-

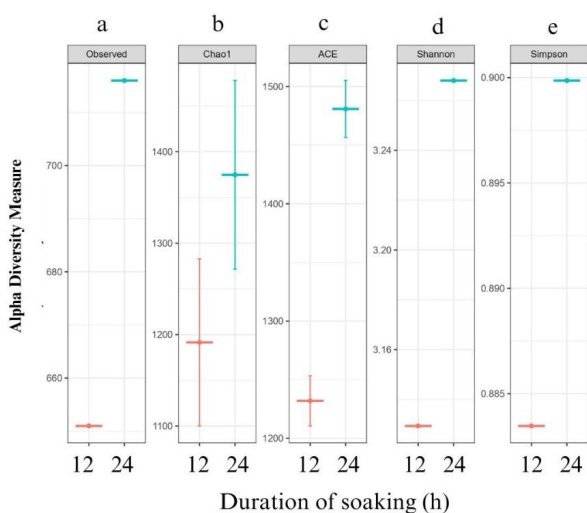
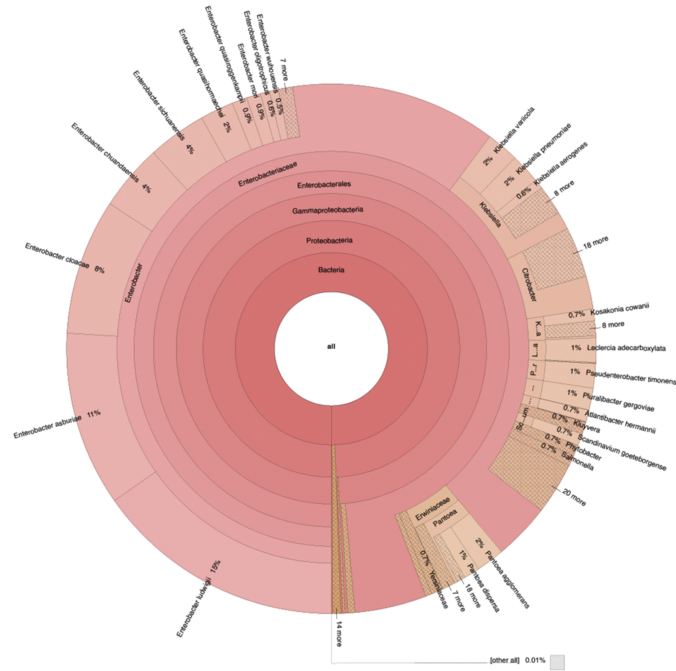


FIGURE 1 Diversity index box figure of bacteria in pigeon pea at different soaking times. a: Observed species index; b: Chao1 index; c: ACE index; d: Shannon index; e: Simpson index.

TABLE 1 Classification of the bacterial family of the pigeon pea sample after being soaked in water for 12 h using Centrifuge's nt database

Family name	taxID	taxRank	Number of uniquely classified reads
Enterobacteriaceae	543	family	2607
Erwiniaceae	1903409	family	22
Pectobacteriaceae	1903410	family	10
Calotrichaceae	2661849	family	4
Halomonadaceae	28256	family	2
Morganellaceae	1903414	family	2
Vibrionaceae	641	family	2
Alteromonadaceae	72275	family	1
Balneolaceae	1813606	family	1
Flavobacteriaceae	49546	family	1
Intrasporangiaceae	85021	family	1
Kofleriaceae	224464	family	1
Micrococcaceae	1268	family	1



(a)



(b)

FIGURE 2 Bacterial composition of pigeon pea after being soaked in water. a. 12 h, b. 24 h.

ing the most dominant bacteria.

Metagenomic analysis was previously used to reveal the bacterial diversity during fermentation process of farmhouse sauce from Northeast China (Hao and Sun 2020). By sequencing the 16S rRNA, they identified 16 phyla, 37 classes, 56 orders, 96 families, 139 genera, and 81 species. *Tetragenococcus*, *Weissella*, *Lactobacillus*, and *Leuconostoc* were all detected during the fermentation process, but *Lactobacillus* was the most im-

portant microflora. Metagenomic analysis has also been applied to reveal microbial community in Toddy, a popular fermented palm beverage of India (Das and Tamang 2023). Using shotgun-based metagenomic, they identified 54 phyla, 363 families, 1087 genera and 1885 species. The most abundant bacterial phylum was *Bacillota* (49.3%) and several bacterial species were identified, including *Leuconostoc mesenteroides*, *Leuconostoc citreum*, *Lactobacillus helveticus*, *Lactiplantibacillus plantarum*, *Lac-*

TABLE 2 Classification of the bacterial genus of the pigeon pea sample after being soaked in water for 12 h using Centrifuge's nt database.

Genus name	taxID	taxRank	Number of uniquely classified reads
<i>Enterobacter</i>	547	genus	9484
<i>Citrobacter</i>	544	genus	1516
<i>Klebsiella</i>	570	genus	819
<i>Pantoea</i>	53335	genus	138
<i>Kluyvera</i>	579	genus	100
<i>Kosakonia</i>	1330547	genus	98
<i>Serratia</i>	613	genus	85
<i>Tatumella</i>	82986	genus	54
<i>Erwinia</i>	551	genus	46
<i>Cronobacter</i>	413496	genus	30
<i>Vibrio</i>	662	genus	23
<i>Pectobacterium</i>	122277	genus	21
<i>Bacillus</i>	1386	genus	18
<i>Dickeya</i>	204037	genus	14
<i>Rahnella</i>	34037	genus	14
<i>Buttiauxella</i>	82976	genus	13
<i>Mangrovibacter</i>	451512	genus	12
<i>Planctopirus</i>	1649480	genus	12
<i>Pseudocitrobacter</i>	1504576	genus	11
<i>Cedecea</i>	158483	genus	10
<i>Lentilactobacillus</i>	2767893	genus	10
<i>Mixta</i>	2100764	genus	9
<i>Acinetobacter</i>	469	genus	9
<i>Franconibacter</i>	1649295	genus	8
<i>Chimaeribacter</i>	2716544	genus	7
<i>Lelliottia</i>	1330545	genus	7
<i>Edwardsiella</i>	635	genus	6
<i>Actinobacillus</i>	713	genus	5
<i>Escherichia</i>	561	genus	5
<i>Lactocaseibacillus</i>	2759736	genus	5
<i>Pseudomonas</i>	286	genus	4
<i>Yersinia</i>	629	genus	4
<i>Izhakiella</i>	1780190	genus	3
<i>Leclercia</i>	83654	genus	3
<i>Shigella</i>	620	genus	3
<i>Acetomicrobium</i>	49894	genus	2
<i>Bradyrhizobium</i>	374	genus	2
<i>Brenneria</i>	71655	genus	2
<i>Brevitalea</i>	2048911	genus	2
<i>Gibbsiella</i>	929812	genus	2
<i>Haemophilus</i>	724	genus	2
<i>Lonsdalea</i>	1082702	genus	2
<i>Paraburkholderia</i>	1822464	genus	2
<i>Psychromonas</i>	67572	genus	2
<i>Raoultella</i>	160674	genus	2
<i>Trabulsiella</i>	158851	genus	2
<i>Xenorhabdus</i>	626	genus	2

TABLE 2 (continued).

Genus name	taxID	taxRank	Number of uniquely classified reads
<i>Actinoplanes</i>	1865	genus	1
<i>Aeribacillus</i>	1055323	genus	1
<i>Aeromonas</i>	642	genus	1
<i>Atlantibacter</i>	1903434	genus	1
<i>Brucella</i>	234	genus	1
<i>Burkholderia</i>	32008	genus	1
<i>Caldimonas</i>	196013	genus	1
<i>Catenulispora</i>	414878	genus	1
<i>Crinalium</i>	241421	genus	1
<i>Croceicoccus</i>	1295327	genus	1
<i>Flavisolibacter</i>	398041	genus	1
<i>Gloeobacter</i>	33071	genus	1
<i>Halomonas</i>	2745	genus	1
<i>Humidesulfovibrio</i>	356	genus	1
<i>Hyphomicrobium</i>	81	genus	1
<i>Iningainema</i>	1932705	genus	1
<i>Ktedonobacter</i>	363276	genus	1
<i>Leucothrix</i>	45247	genus	1
<i>Marichromatium</i>	85076	genus	1
<i>Massilia</i>	149698	genus	1
<i>Mathylobacillus</i>	404	genus	1
<i>Micromonospora</i>	1873	genus	1
<i>Microvirga</i>	186650	genus	1
<i>Morganella</i>	581	genus	1
<i>Motilimonas</i>	1914248	genus	1
<i>Niastella</i>	354354	genus	1
<i>Novosphingobium</i>	165696	genus	1
<i>Oceanisphaera</i>	225143	genus	1
<i>Paraglaciecola</i>	1621534	genus	1
<i>Pedomicrobium</i>	47494	genus	1
<i>Permianibacter</i>	1649479	genus	1
<i>Phenylobacterium</i>	20	genus	1
<i>Providencia</i>	586	genus	1
<i>Pseudoxanthomonas</i>	83618	genus	1
<i>Rhabdothermincola</i>	2820403	genus	1
<i>Rhodoplanes</i>	29407	genus	1
<i>Salmonella</i>	590	genus	1
<i>Sphingomonas</i>	13687	genus	1
<i>Thermithiobacillus</i>	119979	genus	1
<i>Thermosynthropha</i>	54293	genus	1
<i>Virgisporangium</i>	65504	genus	1

tococcus lactis, *Acetobacter malorum*, *Gluconobacter japonicus*, *Gluconacetobacter liquefaciens*, *Fructobacillus durionis*, *Zymomonas mobilis*. Meanwhile, Erhardt et al. (2023) identified 135 lactic acid bacteria species of 7 genera from artisanal cheeses produced in Brazil by amplifying the V3/V4 region of the 16S rRNA with *Lactococcus lactis* as the most abundance species in all samples.

The identified lactic acid bacteria were *Bavariicoccus*, *Enterococcus*, *Lactobacillus*, *Lactococcus*, *Leuconostoc*, *Marinilactibacillus*, and *Pediococcus*. The bacterial diversity reported by Das and Tamang (2023) was higher than that reported by Hao and Sun (2020), and Erhardt et al. (2023), and by our current study. The findings certainly prove that shotgun metagenome sequencing captures more microbial diversity than 16S rRNA gene sequencing.

3.2. Bacterial diversity

Bacterial 16S rRNA gene sequences were classified at the family and genus levels to find out what makes up the bacterial community in pigeon pea soaked in water. A total of 13 families (Table 1) and 89 genera (Table 2) were identified in the pigeon pea sample soaked in water for 12 h. Meanwhile, 26 bacterial families (Table 3) and 90 genera were identified in the pigeon pea soaked in water for 24 h (Table 4).

The relative abundances of the different bacterial communities in all the samples at the family and genus lev-

TABLE 3 Classification of the bacterial family of the pigeon pea sample after being soaked in water for 24 h using Centrifuge's nt database.

Family name	taxID	taxRank	Number of uniquely classified reads
Enterobacteriaceae	543	family	2836
Yersiniaceae	1903411	family	34
Erwiniaceae	1903409	family	23
Vibrionaceae	641	family	6
Calotrichaceae	2661849	family	5
Pectobacteriaceae	1903410	family	5
Bacillaceae	186817	family	4
Burkholderiaceae	119060	family	2
Methylococcaceae	403	family	2
Acidobacteriaceae	204434	family	1
Aeromonadaceae	84642	family	1
Anaerolineaceae	292628	family	1
Bruguierivoracaceae	2812006	family	1
Flavobacteriaceae	49546	family	1
Gomontiellaceae	1892255	family	1
Halanaerobiaceae	972	family	1
Halomonadaceae	28256	family	1
Heliobacteriaceae	31984	family	1
Hyphomicrobiaceae	45401	family	1
Intrasporangiaceae	85021	family	1
Microbacteriaceae	85023	family	1
Morganellaceae	1903414	family	1
Pasteurellaceae	712	family	1
Pseudomonadaceae	135621	family	1
Rhodospirillaceae	41295	family	1
Succinivibrionaceae	83763	family	1

TABLE 4 Classification of the bacterial genus of the pigeon pea sample after being soaked in water for 24 h using Centrifuge's nt database.

Genus name	taxID	taxRank	Number of uniquely classified reads
<i>Enterobacter</i>	547	genus	11233
<i>Klebsiella</i>	570	genus	1061
<i>Citrobacter</i>	544	genus	842
<i>Kluyvera</i>	579	genus	150
<i>Kosakonia</i>	1330547	genus	144
<i>Pantoea</i>	53335	genus	98
<i>Serratia</i>	613	genus	68
<i>Tatumella</i>	82986	genus	58
<i>Erwinia</i>	551	genus	45
<i>Buttiauxella</i>	82976	genus	33
<i>Cronobacter</i>	413496	genus	31
<i>Dickeya</i>	204037	genus	23
<i>Mangrovibacter</i>	451512	genus	21
<i>Pectobacterium</i>	122277	genus	18
<i>Vibrio</i>	662	genus	17
<i>Planctopirus</i>	1649480	genus	16
<i>Edwardsiella</i>	635	genus	13
<i>Rahnella</i>	34037	genus	12
<i>Shigella</i>	620	genus	12
<i>Pseudocitrobacter</i>	1504576	genus	11
<i>Lactacaseibacillus</i>	2759736	genus	10
<i>Lentilactobacillus</i>	2767893	genus	9
<i>Acetomicrobium</i>	49894	genus	8
<i>Lelliottia</i>	1330545	genus	8
<i>Escherichia</i>	561	genus	7
<i>Trabulsiella</i>	158851	genus	6
<i>Chimaeribacter</i>	2716544	genus	5
<i>Brevitalea</i>	2048911	genus	4
<i>Rosenbergiella</i>	1356488	genus	4
<i>Yersinia</i>	629	genus	4
<i>Gibbsiella</i>	929812	genus	3
<i>Raoultella</i>	160674	genus	3
<i>Salmonella</i>	590	genus	3
<i>Xenorhabdus</i>	626	genus	3
<i>Bradyrhizobium</i>	374	genus	2
<i>Burkholderia</i>	32008	genus	2
<i>Crinalium</i>	241421	genus	2
<i>Franconibacter</i>	1649295	genus	2
<i>Gloeobacter</i>	33071	genus	2
<i>Legionella</i>	445	genus	2
<i>Microbacterium</i>	33882	genus	2
<i>Miltoncostaea</i>	2843200	genus	2
<i>Providencia</i>	586	genus	2
<i>Pseudaeromonas</i>	1929090	genus	2
<i>Reyranella</i>	445219	genus	2
<i>Salinicola</i>	404432	genus	2
<i>Shewanella</i>	22	genus	2

TABLE 4 (continued).

Genus name	taxID	taxRank	Number of uniquely classified reads
<i>Thalassotalea</i>	1518149	genus	2
<i>Achromobacter</i>	222	genus	1
<i>Aeromonas</i>	642	genus	1
<i>Alkalimonas</i>	265980	genus	1
<i>Ammonifex</i>	42837	genus	1
<i>Bosea</i>	85413	genus	1
<i>Caballeronia</i>	1827195	genus	1
<i>Cephalothrix</i>	1844514	genus	1
<i>Chloroflexus</i>	1107	genus	1
<i>Clostridium</i>	1485	genus	1
<i>Conexibacter</i>	191494	genus	1
<i>Defluviitalea</i>	1185408	genus	1
<i>Dongia</i>	1146845	genus	1
<i>Euryhalinema</i>	2661529	genus	1
<i>Halomonas</i>	2745	genus	1
<i>Iningainema</i>	1932705	genus	1
<i>Laceyella</i>	292635	genus	1
<i>Leclercia</i>	83654	genus	1
<i>Lonsdalea</i>	1082702	genus	1
<i>Lysobacter</i>	68	genus	1
<i>Marinicella</i>	863253	genus	1
<i>Massilia</i>	149698	genus	1
<i>Motilibacter</i>	1434021	genus	1
<i>Motilimonas</i>	1914248	genus	1
<i>Neobacillus</i>	2675232	genus	1
<i>Nitrospira</i>	1234	genus	1
<i>Oceanisphaera</i>	225143	genus	1
<i>Paraburkholderia</i>	1822464	genus	1
<i>Pelagicoccus</i>	455433	genus	1
<i>Phenylbacterium</i>	20	genus	1
<i>Photobacterium</i>	657	genus	1
<i>Pseudoalteromonas</i>	53246	genus	1
<i>Pseudomonas</i>	286	genus	1
<i>Psychromonas</i>	67572	genus	1
<i>Rhizobium</i>	379	genus	1
<i>Secundilactobacillus</i>	2767892	genus	1
<i>Shimwellia</i>	1335483	genus	1
<i>Solirubrobacter</i>	207599	genus	1
<i>Tepidimonas</i>	114248	genus	1
<i>Thermoanaerobacter</i>	1754	genus	1
<i>Thermodesulfobivrio</i>	28261	genus	1
<i>Trinickia</i>	2571160	genus	1
<i>Tumebacillus</i>	432330	genus	1

els are shown in Figure 3. Among the bacterial families observed and identified from the pigeon pea sample soaked for at 12 and 24 h, five were predominant, namely, Enterobacteriaceae, Erwiniaceae, Yersiniaceae, Peptobacteriaceae, and Lactobacillaceae (Figure 3a). Meanwhile,

according to the relative abundance of the observed and identified bacterial genera, nine of them were predominant (Figure 3b), namely, *Enterobacter*, *Klebsiella*, *Citrobacter*, *Pantoea*, *Kosakonia*, *Pseudoenterobacter*, *Pluralibacter*, *Leclercia*, and *Kluyvera* (Figure 3b). At the genus level, a slight increase in the abundance of *Klebsiella*, *Kosakonia*, and *Pluralibacter* and a slight decrease in the abundance of *Citrobacter* were observed after prolonged incubation from 12 h to 24 h (Figure 3b).

The bacterial communities in the pigeon pea's soaking water were dominated by bacterial genus *Enterobacter*. *Enterobacter* species reported in this study included *E. ludwigii*, *E. cloacae*, *E. asburiae*, *E. chuandaensis*, *E. sichuanensis*, *E. quasiroggkampii*, *E. oligotrophicus*, *E. wuhouensis*, *E. bugandensis*, *E. cancerogenus*, *E. mori*, and *E. kobei*. Apart from the genus *Enterobacter*, some other genera of the family Enterobacteriaceae were identified from the pigeon pea's samples, including *Salmonella* and *Shigella* that are considered as pathogenic microorganisms causing foodborne diseases (Tables 2 and 4). The family Enterobacteriaceae is the major causative agent of foodborne diseases and is widely dispersed in nature. The

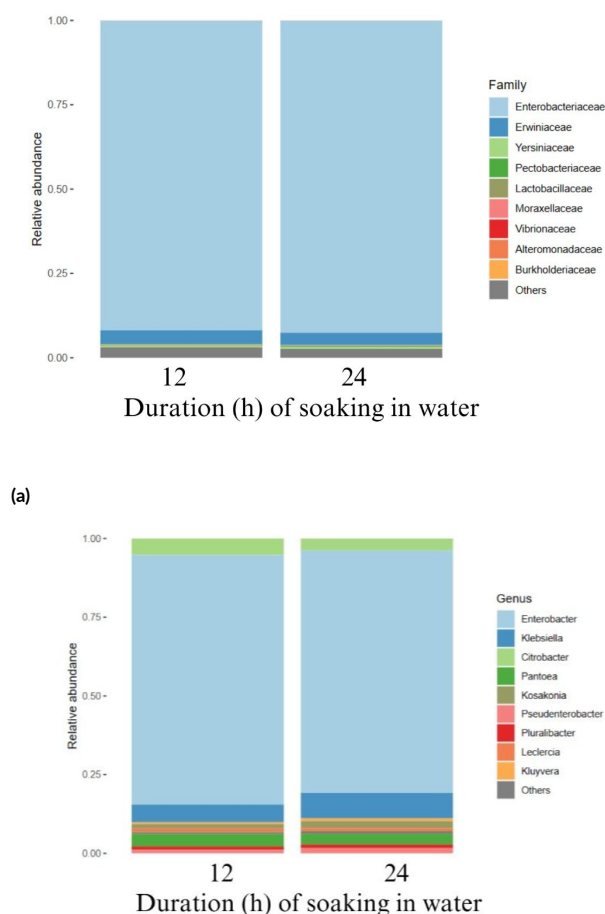


FIGURE 3 Relative abundance of bacterial (a) family and (b) genera in pigeon pea after being soaked for 12 and 24 h.

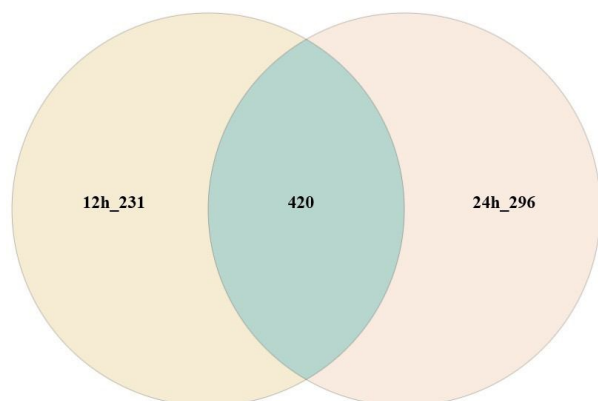


FIGURE 4 Venn diagram of pigeon pea soaked in water for 12 hours (left) and 24 hours (right)

high abundance of this family in the pigeon pea samples is not surprising because they are present and have been detected in natural ecosystems, including in the gastrointestinal tract of vertebrates, vegetation, and aquatic habitats (Janda and Abbott 2021).

Small differences in bacterial communities were observed among the pigeon pea samples under different soaking times. Figure 4 shows that 231 and 296 bacteria were identified in the pigeon pea samples soaked in water for 12 and 24 h, respectively. From the total number of observed and identified bacteria, those detected in the pigeon pea samples soaked in water for 12 h accounted for 55%, and those detected in the samples soaked for 24 h accounted for 70.4%. This difference increased with the soaking time and was consistent with the abovementioned results for diversity indices.

The presence of some members of the family Lactobacillaceae in the pigeon pea's soaking water is interesting; some bacterial species of this family are industrially important as starter cultures for dairy fermentation (Figure 5). The bacterial species belonging to family Lactobacillaceae observed and identified in this study were *Lacticaseibacillus paracasei* (32%), *Lentilactobacillus parabuchneri* (14%), *Lentilactobacillus hilgardii* (7%), *Liquorilactobacillus vini* (9%), *Liquorilactobacillus nagelii* (7%), and *Liquorilactobacillus sicerae* (2%) (Figure 5). The occurrence of some lactic acid bacteria species during soybean fermentation has been reported (Ma et al. 2022; Sirilun et al. 2017). Lactic acid bacteria belong to the phylum Firmicutes. This group is the dominant bacteria in tempeh because of their role in the acidification of pigeon pea during soaking (Nurdini et al. 2015). The lactic acid bacteria genera *Lacticaseibacillus*, *Lentilactobacillus*, and *Liquorilactobacillus* observed in this study might have important roles for pigeon pea acidification. Radita et al. (2018) reported that during fermentation, the low pH of pigeon pea inhibited the growth of spoilage microorganisms.

The lactic acid bacteria species isolated from the soaking water of pigeon pea samples can be of interest as po-

tential starter fermentation to improve nutritional value, flavor characteristic, and health-promoting effects of soy beverages. The use of *Lacticaseibacillus rhamnosus* GG and *L. rhamnosus* INIA P344 as starters for soy beverage fermentation increased antioxidant capacity in the fermented products (Ruiz de la Bastida et al. 2023). Luo et al. (2023) reported flavor improvement of fermented soybean foods due to co-culture of *Bacillus velezensis* and *Lactiplantibacillus plantarum* which produced acetoin and pyrazines. Promoting health benefit of the fermented pigeon pea products by lactic acid bacteria species as starters was reported by Yogeswara et al. (2023). They concluded the increased γ -aminobutyric acid in pigeon pea fermented by *Lactiplantibacillus plantarum* Dad-13 as starters.

Phylum Actinobacteria is another subdominant group in tempeh but is present in lesser amounts compared with Firmicutes and Proteobacteria. All three bacterial phyla make up the lipolytic bacteria in tempeh that play an important role in flavor production (Nur et al. 2020). Some Proteobacteria species such as *Klebsiella pneumoniae* and *Citrobacter freundii* play a major role in the production of vitamin B12 in tempeh. Different food products exhibit certain differences in bacterial communities. Li et al. (2017) reported that Firmicutes and Proteobacteria were the predominant phyla in a Chinese traditional fermented broad bean (*Vicia faba* L.) paste. Peng et al. (2018) reported the unique microbial diversity of fermented vegetables obtained from different regions of Hainan, China and found that *Lactobacillus* was the most dominant genus. Within this genus, *Lactobacillus plantarum* was the most abundant species, followed by *L. fermentum* and *L. pentosaceus*. Similarly, an analysis of the microbial composition of kimchi showed that main bacteria were lactic acid bacteria, including *Lactobacillus*, *Leuconostoc*, and *Weissella* (Kim and Chun 2005).

Previous studies on pigeon pea grain's microflora frequently detected *Lentilactobacillus* and *Lactococcus* (Balogun et al. 2021; Demarinis et al. 2022). Demarinis et al. (2022) reported the isolation and identification of lactic acid bacteria and acetic acid bacteria from pigeon pea grains. Among the identified lactic acid bacteria species, *Lentilactobacillus casei*, *Leuconostoc mesenteroides*, and *Gluconobacter hansenii* were the most abundant. Balogun et al. (2021) reported the presence of *Lactococcus lactis* and *Enterococcus faecalis* in fermented pigeon pea grains. *Lentilactobacillus* genus accounted for approximately 30% of the microorganisms found in the fermented pigeon pea samples examined in North America. One *Lentilactobacillus* species, *Lentilactobacillus hilgardii*, is known to produce exopolysaccharide, which affects the physicochemical quality of products, during fermentation. Marsh et al. (2013) reported that *Leuconostoc* species are rarely found in pigeon pea grains. Other bacterial genera found at low levels in pigeon pea grains included *Acetobacter* and *Gluconoacetobacter* (Gulitz et al. 2011). *Gluconoacetobacter* was initially present in the grain but disappeared after fermentation. Meanwhile, Acetobacteria was detected at low levels in pigeon pea grains, but its role

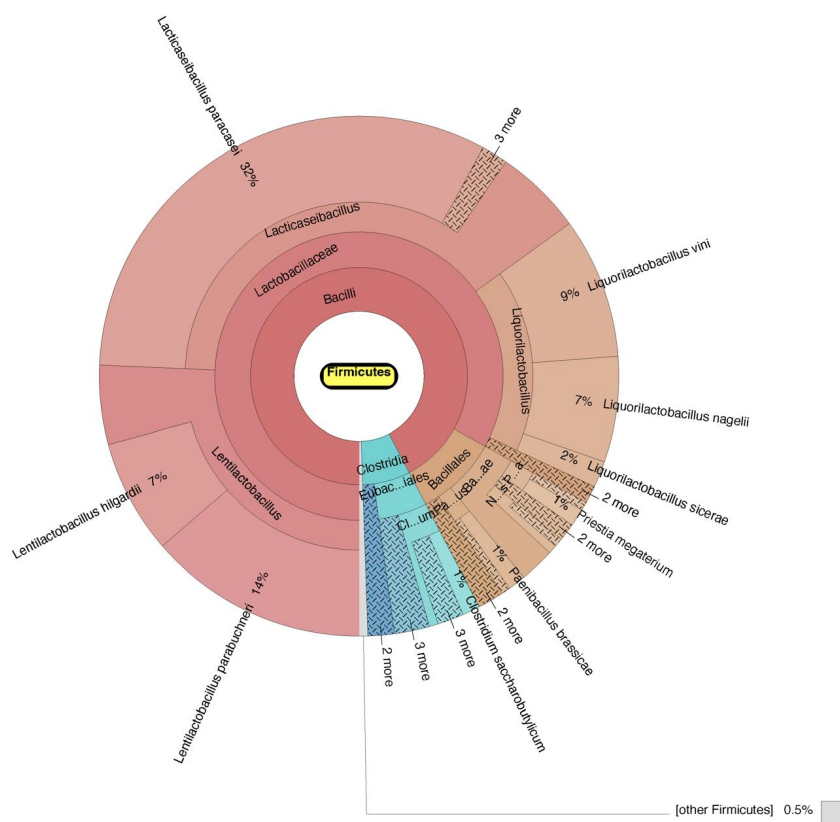


FIGURE 5 Bacterial composition of family Lactobacillaceae in pigeon pea soaked for 24 h.

remains unclear (Gulitz et al. 2011).

Metagenomic analysis results showed the presence of diverse bacterial families and genera in pigeon pea soaked in water. During fermentation, certain microorganisms with metabolic activity were able to synthesize vitamins, increase the nutritional content, and produce metabolites that affect the taste and texture of the products. Some lactic acid bacteria were reported in this study, suggesting their importance for pigeon pea fermentation. These species can be isolated to assess their capability as starter cultures for developing pigeon pea-based functional foods or drinks.

4. Conclusions

The metagenomic analysis of pigeon pea soaked in water for 12 and 24 h shows the diversity of bacterial families and genera, mainly from phylum Proteobacteria. The main bacterial families observed and detected were Enterobacteriaceae, Erwiniaceae, Yersiniaceae, Peptobacteriaceae, and Lactobacillaceae. The pigeon pea samples collected after 24 h of soaking showed higher bacterial diversity and richness than those collected after 12 h. The presence of *Lactocaseibacillus*, *Lentilactobacillus*, and *Secundilactobacillus* is interesting because of their importance as starter cultures for the development of fermented pigeon pea beverages for lactose-intolerant individuals. These three lactic acid bacteria genera are potential for improving nutritional content, flavor characteristics, and health-promoting ef-

fects of soy beverage products. These products may prevent malnutrition (low lysine, etc. in pulses), and for reducing antinutritional and flatulence factors.

Acknowledgments

Yuni Sine is a recipient of LPDP scholarship provided by the Indonesian Government that enable her to pursue a doctoral degree at the Graduate School of Biotechnology, Universitas Gadjah Mada in Yogyakarta Indonesia.

Authors' contributions

Conceptualization, WW, DW, YAP., and YS; methodology, YS, WW, DW, YAP, and BHL; validation, YS, WW, DW, YAP, and BHL; formal analysis, YS, WW, DW, and YAP; investigation, YS, WW, DW, and YAP; resources, YS, and WW; data curation, YS, WW, YAP, DW, and BHL; writing—original draft preparation, YS, DW, YAP, DW, HRP, and BHL; supervision, WW, DW, and YAP; funding acquisition, YS, and WW. All authors have read and agreed to the published version of the manuscript.

Competing interests

The authors declare no competing interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- Abebe B. 2022. The dietary use of pigeon pea for human and animal diets. *Sci. World J.* 2022:4873008. doi:10.1155/2022/4873008.
- Akande KE, Abubakar MM, Adegbola TA, Bogoro SE, Doma UD. 2010. Chemical evaluation of the nutritive quality of pigeon pea [*Cajanus cajan* (L.) Millsp.]. *Int. J. Poult. Sci.* 9(1):63–65. doi:10.3923/ijps.2010.63.65.
- Bahram M, Anslan S, Hildebrand F, Bork P, Tedersoo L. 2019. Newly designed 16S rRNA metabarcoding primers amplify diverse and novel archaeal taxa from the environment. *Environ. Microbiol. Rep.* 11(4):487–494. doi:10.1111/1758-2229.12684.
- Balogun MA, Ahmed RN, Akintayo OA, Aruna TE, Omovbude MO, Shittu T. 2021. Microbial, chemical and sensory evaluation of pigeon pea condiment from wild and controlled fermentation. *Ceylon J. Sci.* 50(3):269–277. doi:10.4038/cjs.v50i3.7908.
- Das S, Tamang JP. 2023. Metagenomics and metabolomics of Toddy, an Indian fermented date palm beverage. *Food Res. Int.* 172:113205. doi:10.1016/j.foodres.2023.113205.
- De Coster W, D’Hert S, Schultz DT, Cruts M, Van Broeckhoven C. 2018. NanoPack: Visualizing and processing long-read sequencing data. *Bioinformatics* 34(15):2666–2669. doi:10.1093/bioinformatics/bty149.
- Demarinis C, Verni M, Pinto L, Rizzello CG, Baruzzi F. 2022. Use of selected lactic acid bacteria for the fermentation of legume-based water extracts. *Foods* 11(21):3346. doi:10.3390/foods11213346.
- Desai A, Small D, Mc Gill AE, Shah NP. 2002. Metabolism of *Raffinose* and *Stachyose* in reconstituted skim milk and of n-hexanal and pentanal in soymilk by *Bifidobacteria*. *Biosci. Microflora* 21(4):245–250. doi:10.12938/bifidus1996.21.245.
- Erhardt MM, Oliveira WdC, Fröder H, Marques PH, Oliveira MBPP, Richards NSPdS. 2023. Lactic bacteria in artisanal cheese: Characterization through metagenomics. *Fermentation* 9(1):41. doi:10.3390/fermentation9010041.
- Fang RS, Dong YC, Chen F, Chen QH. 2015. Bacterial diversity analysis during the fermentation processing of traditional Chinese yellow rice wine revealed by 16S rDNA 454 pyrosequencing. *J. Food Sci.* 80(10):2265–2271. doi:10.1111/1750-3841.13018.
- Fredrikson M, Andlid T, Haikara A, Sandberg AS. 2002. Phytate degradation by micro-organisms in synthetic media and pea flour. *J. Appl. Microbiol.* 93(2):197–204. doi:10.1046/j.1365-2672.2002.01676.x.
- Gan RY, Li HB, Gunaratne A, Sui ZQ, Corke H. 2017. Effects of fermented edible seeds and their products on human health: Bioactive components and bioactivities. *Compr. Rev. Food Sci. Food Saf.* 16(3):489–531. doi:10.1111/1541-4337.12257.
- Gu Q, Zhang C, Song D, Li P, Zhu X. 2015. Enhancing vitamin B12 content in soy-yogurt by *Lactobacillus reuteri*. *Int. J. Food Microbiol.* 206:56–59. doi:10.1016/j.ijfoodmicro.2015.04.033.
- Gulitz A, Stadie J, Wenning M, Ehrmann MA, Vogel RF. 2011. The microbial diversity of water kefir. *Int. J. Food Microbiol.* 151(3):284–288. doi:10.1016/j.ijfoodmicro.2011.09.016.
- Hao Y, Sun B. 2020. Analysis of bacterial diversity and biogenic amines content during fermentation of farmhouse sauce from Northeast China. *Food Control* 108:106861. doi:10.1016/j.foodcont.2019.106861.
- Hickisch A, Beer R, Vogel RF, Toelstede S. 2016. Influence of lupin-based milk alternative heat treatment and exopolysaccharide-producing lactic acid bacteria on the physical characteristics of lupin-based yogurt alternatives. *Food Res. Int.* 84:180–188. doi:10.1016/j.foodres.2016.03.037.
- Janda JM, Abbott SL. 2021. The changing face of the family enterobacteriaceae (Order: Enterobacterales): New members, taxonomic issues, geographic expansion, and new diseases and disease syndromes. *Clin. Microbiol. Rev.* 34(2):e00174–20. doi:10.1128/CMR.00174-20.
- Kim D, Song L, Breitwieser FP, Salzberg SL. 2016. Centrifuge: Rapid and sensitive classification of metagenomic sequences. *Genome Res.* 26(12):1721–1729. doi:10.1101/gr.210641.116.
- Kim M, Chun J. 2005. Bacterial community structure in kimchi, a Korean fermented vegetable food, as revealed by 16S rRNA gene analysis. *Int. J. Food Microbiol.* 103(1):91–96. doi:10.1016/j.ijfoodmicro.2004.11.030.
- Li Z, Rui J, Li X, Li J, Dong L, Huang Q, Huang C, Wang Z, Li L, Xuan P, Tang Y, Chen F. 2017. Bacterial community succession and metabolite changes during doubanjiang-meju fermentation, a Chinese traditional fermented broad bean (*Vicia faba* L.) paste. *Food Chem.* 218:534–542. doi:10.1016/j.foodchem.2016.09.104.
- Liu L, Chen X, Hao L, Zhang G, Jin Z, Li C, Yang Y, Rao J, Chen B. 2022. Traditional fermented soybean products: processing, flavor formation, nutritional and biological activities. *Crit. Rev. Food Sci. Nutr.* 62(7):1971–1989. doi:10.1080/10408398.2020.1848792.
- Luo YY, Guo Y, Hu XY, Liu WH, Liu BQ, Yang J, Tu ZC, Huang YH. 2023. Flavor improvement of fermented soybean foods by co-fermentation with *Bacillus velezensis* and *Lactiplantibacillus plantarum*. *LWT* 186:115257. doi:10.1016/j.lwt.2023.115257.
- Ma H, Wang L, Yu H, Wang W, Wu G, Qin G, Tan Z, Wang Y, Pang H. 2022. Protease-producing lactic acid bacteria with antibacterial properties and their potential use in soybean meal fermentation. *Chem. Biol. Technol. Agric.* 9(1):40. doi:10.1186/s40538-022-00303-5.
- Maesen LVD. 1985. *Cajanus* DC. and *Atylosia* W. & A. (Leguminosae), volume 85-4. Wageningen: Wageningen Centre for Agricultural Publishing.

- geningen Agricultural University.
- Marsh AJ, O'Sullivan O, Hill C, Ross RP, Cotter PD. 2013. Sequencing-based analysis of the bacterial and fungal composition of kefir grains and milks from multiple sources. *PLoS One* 8(7):69371. doi:10.1371/journal.pone.0069371.
- Nur N, Meryandini A, Suhartono MT, Suwanto A. 2020. Lipolytic bacteria and the dynamics of flavor production in Indonesian tempeh. *Biodiversitas* 21(8):3818–3825. doi:10.13057/biodiv/d210850.
- Nurdini AL, Nuraida L, Suwanto A, Suliantari. 2015. Microbial growth dynamics during tempe fermentation in two different home industries. *Int. Food Res. J.* 22(4):1668–1674.
- Nwosu JN, Ojukwu M, Ogueke CC, Ahaotu I, Owuamanam CI. 2013. The antinutritional properties and ease of dehulling on the proximate composition of pigeon pea (*Cajanus cajan*) as affected by malting. *Int. J. Life Sci.* 2(2):60–67.
- Peng Q, Jiang S, Chen J, Ma C, Huo D, Shao Y, Zhang J. 2018. Unique microbial diversity and metabolic pathway features of fermented vegetables from Hainan, China. *Front. Microbiol.* 9:399. doi:10.3389/fmicb.2018.00399.
- Prativi MBN, Astuti DI, Putri SP, Laviña WA, Fukusaki E, Aditiawati P. 2023. Metabolite changes in Indonesian tempe production from raw soybeans to over-fermented tempe. *Metabolites* 13(2):300. doi:10.3390/metabo13020300.
- Qiao Y, Zhang K, Zhang Z, Zhang C, Sun Y, Feng Z. 2022. Fermented soybean foods: A review of their functional components, mechanism of action and factors influencing their health benefits. *Food Res. Int.* 158:111575. doi:10.1016/j.foodres.2022.111575.
- Radita R, Suwanto A, Kurosawa N, Wahyudi AT, Rusmana I. 2018. Firmicutes is the predominant bacteria in tempeh. *Int. Food Res. J.* 25(6):2313–2320.
- Ruiz de la Bastida A, Peirotén Á, Langa S, Rodríguez-Mínguez E, Curiel JA, Arqués JL, Landete JM. 2023. Fermented soy beverages as vehicle of probiotic lactobacilli strains and source of bioactive isoflavones: A potential double functional effect. *Heliyon* 9(4):14991. doi:10.1016/j.heliyon.2023.e14991.
- Sandberg AS. 2011. Developing functional ingredients: a case study of pea protein. Woodhead Publishing. p. 358–382. doi:10.1533/9780857092557.3.358.
- Sheahan CM. 2012. Plant guide for sunn hemp (*Crotalaria juncea*). Soil Quality Agronomy Technical Note 10:8210.
- Sirilun S, Sivamaruthi BS, Kesika P, Peerajan S, Chaiyasut C. 2017. Lactic acid bacteria mediated fermented soybean as a potent nutraceutical candidate. *Asian Pac. J. Trop. Biomed.* 7(10):930–936. doi:10.1016/j.apjtb.2017.09.007.
- Tiwari BK, Gowen A, McKenna B. 2020. Pulse foods: Processing, quality and nutraceutical applications. Academic Press. p. 363. doi:10.1016/C2009-0-61252-6.
- Valenzuela H, Smith J. 2002. Cowpea. Honolulu: College of Tropical Agriculture and Human Resources, University of Hawaii at Manoa.
- Varshney RK, Chen W, Li Y, Bharti AK, Saxena RK, Schlueter JA, Donoghue MT, Azam S, Fan G, Whalley AM, Farmer AD, Sheridan J, Iwata A, Tuteja R, Penmetsa RV, Wu W, Upadhyaya HD, Yang SP, Shah T, Saxena KB, Michael T, McCombie WR, Yang B, Zhang G, Yang H, Wang J, Spillane C, Cook DR, May GD, Xu X, Jackson SA. 2012. Draft genome sequence of pigeon pea (*Cajanus cajan*), an orphan legume crop of resource-poor farmers. *Nat. Biotechnol.* 30(1):83. doi:10.1038/nbt.2022.
- Wick RR, Judd LM, Holt KE. 2019. Performance of neural network basecalling tools for Oxford Nanopore sequencing. *Genome Biol.* 20(1):129. doi:10.1186/s13059-019-1727-y.
- Yang H, Zhang L, Xiao G, Feng J, Zhou H, Huang F. 2015. Changes in some nutritional components of soymilk during fermentation by the culinary and medicinal mushroom *Grifola frondosa*. *LWT* 62(1):468–473. doi:10.1016/j.lwt.2014.05.027.
- Yarlina VP, Andoyo R, Djali M, Lani MN. 2022. Metagenomic analysis for indigenous microbial diversity in soaking process of making tempeh jack beans (*Canavalia ensiformis*). *Curr. Res. Nutr. Food Sci.* 10(2):620–632. doi:10.12944/CRNFSJ.10.2.18.
- Yogeswara IBA, Kusumawati IGAW, Nursini NW, Mariyatun M, Rahayu ES, Haltrich D. 2023. Health-promoting role of fermented pigeon pea (*Cajanus cajan* L (Mill)) milk enriched with γ -aminobutyric acid (GABA) using probiotic *Lactiplantibacillus plantarum* Dad-13. *Fermentation* 9(7):587. doi:10.3390/fermentation9070587.