DETECTION OF PATHOGENIC BACTERIA IN PIGEON FAECAL SAMPLES

CHAN WEN JUN

BACHELOR OF SCIENCE (HONS)

BIOTECHNOLOGY

FACULTY OF SCIENCE

UNIVERSITI TUNKU ABDUL RAHMAN

MAY 2023

DETECTION OF PATHOGENIC BACTERIA IN PIGEON FAECAL SAMPLES

By

CHAN WEN JUN

A project report submitted to the Department of Biological Science
Faculty of Science
Universiti Tunku Abdul Rahman
In partial fulfilment of the requirements for the degree of
Bachelor of Science (Hons) Biotechnology

MAY 2023

ABSTRACT

DETECTION OF PATHOGENIC BACTERIA IN

PIGEON FAECAL SAMPLES

Chan Wen Jun

Pigeons flock in public area especially dining area due to food availability. Pigeon droppings are found in abundance at Block C cafeteria of Universiti Tunku Abdul Rahman (UTAR) Kampar campus, which can lead to health issues due to the presence of pathogenic bacteria such as *Campylobacter jejuni*, *C. coli*, *Chlamydia psittaci*, *C. avium*, *Enterococcus* spp., *Escherichia coli*, *Salmonella* spp., *Shigella* spp., and *Staphylococcus aureus* that can lead to certain diseases such as campylobacteriosis, psittacosis and salmonellosis. Therefore, this project aims to detect the possible pathogenic bacteria in pigeon faecal samples at UTAR Kampar campus Block C Cafeteria. Furthermore, to study the bacterial diversity and lastly raise the awareness of UTAR community towards the potential health risks caused by pigeon faeces. Culture independent methods were used in this project. Genomic DNA was extracted from pigeon faecal sample that collected from Block C cafeteria, followed by amplification of *16S rRNA* gene using polymerase chain reaction (PCR) method. The *16S rRNA* gene with approximately band size 1.5 kb was cloned into pTG19-T

cloning vector. The plasmids were then transformed into TOP10 *E. coli* competent cells using heat-shock transformation method and plated on Luria-Bertani (LB) agar containing proper concentration of Ampicillin antibiotic for screening and selection process. Then, colony PCR was conducted on positive transformants to verify the presence of inserted *16S rRNA* gene and select the positive clones. Plasmid extraction was done for the twelve positive clones and *16S rRNA* gene PCR were done to verify the presence of *16S rRNA* gene in the plasmid. DNA sequencing was done for ten plasmid samples and analysis were conducted using BLASTn. Results from the BLASTn showed six *Lactobacillus crispatus*, one *Lactobacillus ingluviei*, one *Ligilactobacillus salivarius*, one *Candidatus* Mycoplasma liparidae and one uncultured bacterium. Among the identified bacterial species, all of them are non-pathogenic. Larger sample size is to be included and more robust method such as next-generation sequencing could be performed to enhance the screening efficiency.

ACKNOWLEDGEMENTS

First, I would like to express my most sincere gratitude and appreciation to my supervisor, Dr. Yap Moh Lan, from the Department of Biological Science, Faculty of Science, for her patience, teaching and supervision throughout this final year project. This project could not be completed without her expert and patience guidance and mentorship, especially when I was facing difficulties. Special thanks to my postgraduate senior, Fong Gui Ying for her guidance and mentorship. Special thanks to lab assistants who helped to provide materials, instructions and equipment when I required.

Besides, special thanks to my parents who gave me support and mentorship when I faced difficulties. Furthermore, special thanks to my lab mates, Wong Yong Han, Kevin Tan Yi Jian, Low Fu Song, Low Zhan Yi, Lo Zheng Xuan, Nur Armanda binti Mohamad Hafiz Ikhwan, Lo Zheng Xuan, Lee Jia Hui, Ho Yong Quan, Sabrina Khoo Pei Pei and Shantinii Priya a/p Pushpanathan who offered helps and supports. Finally, I would like to appreciate the support and encouragement from all my friends and family. Because of all of you, I was able to complete the research successfully.

DECLARATION

I hereby declare that the project report is based on my original work except for quotations and citations which have been duly acknowledged. I also declare that it has not been previously or concurrently submitted for any other degree at UTAR or other institutions.

CHAN WEN JUN

APPROVAL SHEET

This final year project entitled "**DETECTION OF PATHOGENIC BACTERIA IN PIGEON FAECAL SAMPLES**" was prepared by Chan Wen Jun and submitted as partial fulfilment of the requirements for the degree of Bachelor of Science (Hons) in Biotechnology at Universiti Tunku Abdul Rahman.

Approved by:

Date: 10 April 2023

(Dr. YAP MOH LAN)

Supervisor

Department of Biological Science

Faculty of Science

Universiti Tunku Abdul Rahman

FACULTY OF SCIENCE

UNIVERSITI TUNKU ABDUL RAHMAN

Date: 9 April 2023

PERMISSION SHEET

It is hereby certified that CHAN WEN JUN (ID No: 1901517) has completed this final year project report entitled "DETECTION OF PATHOGENIC BACTERIA IN PIGEON FAECAL SAMPLES" under the supervision of DR. YAP MOH LAN from the Department of Biological Science, Faculty of Science.

I hereby give permission to the University to upload the softcopy of my final year project dissertation in pdf format into the UTAR Institutional Repository, which may be made accessible to the UTAR community and public.

I hereby give permission to my supervisor to write and prepare manuscript of these research findings for publishing in any form, if I did not prepare it within six (6) months time from this date provided that my name is included as one of the authors for this article. Arrangement of the name depends on my supervisor.

Yours truly,

(CHAN WEN JUN)

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

× g	Acceleration of gravity
A	Absorbance
ATP	Adenosine triphosphate
Bp	Base pair
Ca^{2+}	Calcium ions
$CaCl_2$	Calcium chloride
DNA	Deoxyribonucleic acid
Dntp	Deoxyribonucleoside triphosphate
Kb	Kilobase pair
LB	Luria-Bertani
LPS	Lipopolysaccharide
$MgCl_2$	Magnesium chloride
OD	Optical density
PCR	Polymerase chain reaction
Rpm	Revolutions per minute
rRNA	Ribosomal ribonucleic acid
TAE buffer	Tris-acetate-EDTA buffer

CHAPTER 1

INTRODUCTION

Pigeon, with common name like “rock pigeons”, “homing pigeon” and “rock doves”, scientifically known as *Columba livia* which means bird with blue-grey or dull grey colour. The bird order of pigeons belongs to Columbiformes which consists of Columbidae family that has 40 genera and 315 species of pigeons. Smaller species is known as “dove” while larger species is known as “pigeon”. The subspecies of pigeons including *Columba livia atlantis, Columbia livia dakhlae, Columbia livia nigricans* and so on. The origins of pigeons are mainly from Asia, Europe, and North Africa where the most common subspecies of pigeon, feral pigeon is found in Europe with total number of 10 to 15 million out of 28 million bird population, due to domestication. In addition, feral pigeon, also known as *Columbia livia domestica*, is merely found in human’s habitation area while wild type pigeon usually can be found near coastal region (Pigeon Control Resource Centre, 2009) and sea cliff (Animals Network Team, 2017). Pigeons and doves such as tropical fruit dove usually dwell on trees, but there are also rock doves that dwell and nest on man-made buildings, bridges, roof, building beams, pipes and air vent in urban area nowadays (Gyimesi, 2015).

Pigeons are monogamous, which means mate to one male or female only at a time, pigeons’ mate for life (Britannica, 2020). The diet of pigeon depends

on the differ in species, but mainly in seeds form (Pigeon Control Resource Centre, 2009). Other common foods included vegetables, fruits, worms, and grains. Feral pigeons usually consume foods like seeds from human food wastes. Human refuse usually acts as the main food resource for domesticated pigeons like feral pigeons, due to various food types are available (Animals Network Team, 2017).

Pigeons usually remain in flocks for protection from predator and their helpless characteristics. Some pigeons in the flock will act as food searcher by searching food resources while others will feed on the located food source (Animals Network Team, 2017). However, the inhabit of pigeons in the form of flocks in urban area and cities may lead to droppings everywhere especially human's dining area due to high food availability. These pigeon faeces may pose health risks to human. For instance, the dried pigeon faeces or pigeon guano favour the growth of fungus that can lead to several diseases like cryptococcosis, candidiasis and histoplasmosis (Conventry Pest Control, n.d.). There are many types of bacteria, viruses and fungi can be found in pigeon faeces and some are pathogenic. Commonly found bacteria are *Chlamydia psittaci*, *Salmonella* species, *Escherichia coli*, *Campylobacter jejuni*, *Staphylococcus aureus* and *Enterococcus* species. These bacteria can transmit diseases like psittacosis and diarrhea to human via aerosol transmission, direct and indirect contact of contamination like food and water sources especially at dining area (Biomed Central, 2010; Executive, n.d.). Furthermore, the overuse of antibiotics can lead to antibiotic resistance microbes which are harmful towards people with lower immune system response.

The objectives of this project are to study the bacterial diversity and to detect the possible pathogenic bacteria in pigeon faeces that may pose a health risk to users of dining area at Block C of UTAR Kampar campus. This project also aims to raise the awareness of UTAR community towards the impact of pigeon faeces found at Block C of UTAR Kampar campus.

CHAPTER 2

LITERATURE REVIEW

2.1 Impact of pigeon domestication

Domestication of wild pigeons lead to appearance of high pigeon population in urban area, thus increasing interaction between human and pigeons. Pigeons can look for food, shelter and mates in urban area, with rare natural predators (Johnston and Janiga, 1995). However, pigeons are always reported as the carrier of zoonotic pathogens that pose health risk to human, especially at hospital and dining area. The exposure of pigeon faeces to patients with weak immune system can lead to longer hospitalization duration due to microbial infection. Pathogenic bacteria can be transmitted to human both direct and indirectly such as food, water and dust particles in air. A previous study revealed that pigeons are the carrier of more than 60 pathogenic microorganisms that pose health risk to human (Santos, et al., 2020). Out of 60 pathogen species, 45 are fungi, 9 are bacteria and remaining are viruses and protozoa (Vasconcelos, et al., 2018). *Campylobacter jejuni* that carried by pigeons had led to numerous infections and diseases like campylobacteriosis especially in England, Australia and Canada (BioMed Central, 2010).

2.2 Drug resistant bacteria

The appearance of drug-resistant bacteria in pigeon faeces is due to the unhygienic eating habits and abuse of antibiotics where pigeons contacted to antimicrobial agents. A study was conducted in Costa Rica to determine the presence of antimicrobial or antibiotic resistance genes such as *bla* $_{TEM}$, *cat*I, *mec*A, *qnr*S, *sul* II, *sul* I, *tet* (A), *tet* (Q) and *van* (A) in 141 pigeon faecal samples. The study finally concluded that pigeons that inhabit in urban area of Costa Rica are potentially carrying antibiotic resistant bacteria. Pigeons rely on contaminated foods and waters from the urban park which near to hospitals, clinics and contaminated rivers in the study as the travel distance of pigeons are short, approximately 5.29 km. Therefore, the contaminated foods and waters may contain antibiotics and faecal that lead to antimicrobial resistance genes (Blanco-Pena, et al., 2017). In addition, another study found that the presence of antibiotic resistance genes is highly associated with the contact between domestic animals and antimicrobial agents or human medicines (Luca, et al., 2004). Furthermore, a previous study in Brazil showed that around 38% of antibiotic-resistant *E. coli* were found in the fresh pigeon faeces samples. Among the diarrheagenic *E. coli* strains, there are around 12.1% are pathogenic, which includes Enteroinvasive *E. coli,* Enteropathogenic *E. coli,* Shiga toxin-producing *E. coli* and Enterotoxigenic *E. coli* (Silva, et al., 2009). Another study showed 85 out of total 100 *E. coli* strains from feral pigeon faeces are antibiotic-resistant, especially the strain PGB01 that can survive in both human 37°C and pigeon 42°C body temperature, is highly resistant to trimethoprim (Kumar, et al., 2015). The thermal adaptation of PGB01 strain may pose health risk when exposed to human, as it can lyse humans' blood and is serum resistant. Another

study focused on the *Staphylococcus aureus* and *Enterococcus* spp. from pigeon faeces near hospital area, due to high prevalence in epidemiological studies, especially those with antibiotic-resistant characteristics like vancomycin-resistant *Enterococcus* spp. and methicillin-resistant *S. aureus*. These antibiotic-resistant bacteria can potentially lead to health problems towards patients at hospital area, for instance, *Enterococcus faecium* strain ST1054 and *S. aureus* strain ST188 (Vasconcellos, et al., 2022).

Besides, there are also multidrug-resistant bacteria that have been isolated from pigeons in Bangladesh, such as MDR *E. coli* and *Salmonella* spp. These bacteria that found in pigeon faeces are resistant to certain antibiotics like ampicillin, amoxicillin, erythromycin, azithromycin, nalidixic acid, levofloxacin and tetracycline (Karim, et al., 2020). A study that previously conducted in Brno, Czech Republic in 2006, found that there are 203 antibiotic-resistant *E. coli* out of 247 pigeon samples, and around 31% *Enterococcus* isolates resist to one to four antibiotics, out of 143 isolated enterococci species (Radimersky, et al., 2010). High level of antibiotic-resistant characteristics was also found in *E. coli*, *Streptococcus gallolyticus* and *Salmonella enterica* serotype Typhimurium in Belgium (Kimpe, et al., 2002). Another study in Japan also provides the evidence of antimicrobial characteristics of *E. coli* towards chloramphenicol (Sato, et al., 1978). Bacteria like *Campylobacter* species like *Campylobacter jejuni* have found to be ciprofloxacin-resistant since the approval of the use of fluoroquinolones antibiotics in poultry flocks by the U.S. Food and Drug Administration (FDA) in 1995 and 1996. Research data from National Antimicrobial Resistance Monitoring System (NARMS) show the

increase of the antibiotic-resistant of *C. jejuni* from 17 % to 27% in 2017 (Centers for Disease Control and Prevention, 2019). In short, pigeons that are carrying drug-resistant pathogens are the key contributor of spreading of infectious agents to humans.

2.3 Pathogenic bacteria and diseases

Pathogenic bacteria like *Chlamydia psittaci* and *C. avium* can be found in pigeon faeces (Burt, et al., 2018). These bacteria can lead to psittacosis when human contact with contaminated food, water and even through aerosols transmission, therefore lead to pneumonia which will be fatal if not treated properly (Centers for Disease Control and Prevention, 2022). Many surveys were previously conducted in Europe to investigate the chlamydial infections of pigeons and its adverse effects towards public health (Magino, et al., 2009).

Thermophilic bacteria like *Campylobacter jejuni* and *Campylobacter coli* can lead to campylobacteriosis that further lead to inflammation and acute diarrhea. People infected may suffer from bloody diarrhea, dysentery syndrome, paralysis, even lead to some complications like irritable bowel syndrome (IBS) and arthritis. It can be fatal to patients with weak immune system like AIDS (Centers for Disease Control and Prevention, 2019). Previous study in Madrid showed high prevalence of *C. jejuni* which potentially pose health risk to human population (Vasquez, et al., 2010).

Salmonellosis that caused by *Salmonella* species bacteria can be transmitted to human via contaminated food. It can lead to food poisoning, inflammation, gastrointestinal problems with symptoms like fever, diarrhea, headaches, and cramps. It could be fatal to elder, child and people with low immune response (Food and Drug Administration, 2019).

Pathogenic bacteria like *Escherichia coli* strain O157:H7 and *Shigella sonnei* were found in pigeons' faeces in Block C cafeteria of UTAR Kampar campus (Singh, 2021). The *E. coli* strain O157:H7 can lead to diseases like haemolytic uremic syndrome and haemorrhagic colitis, due to production of Shiga Toxin 1 (Stx 1) and Shiga toxin 2 (Stx 2). The *E. coli* strain O157:H7 lead to infection by adhering on the epithelium or mucosa layer of intestine, releasing toxins into bloodstream which causes inflammation (Paton and Paton, 1999; Caprioli, et al., 2005). On the other hand, *Shigella sonnei* has low infective dose (ID_{50}) of range 10 to 200 to cause shigellosis (Pathogen Regulation Directorate, Public Health Agency of Canada, 2010: Koh, et al., 2012). Shigellosis can lead to symptoms like bloody diarrhea and fever. It is usually transmitted via contaminated food, water and contact with birds and animals, possibly lead to large outbreak (Centers for Disease Control and Prevention, 2020).

Apart from pathogenic bacteria, some beneficial, non-pathogenic bacteria can be found in pigeon faeces, especially *Lactobacillus* genus bacteria. Lactic acid bacteria (LAB) is important in lactic acid production and balancing

of normal microflora in gastrointestinal (GI) tract. *Ligilactobacillus salivarius* was found in domestic pigeon faeces, where the biodiversity and characteristic of different strains was studied (Dec, et al., 2021). Another study also showed isolated of *L. ingluviei* where the information regarding this LAB is still limited (Margo, et al., 2003). Previous study done in UTAR Kampar Block C cafeteria showed some identified lactic acid bacteria such as *Lactobacillus johnsonii*, *L. crispatus*, *L. reuteri* and *L. panis* (Singh, 2021).

CHAPTER 3

MATERIALS AND METHODS

3.1 Materials Preparation

3.1.1 Ampicillin stock solution (100 mg/mL)

To prepare 100 mg/mL ampicillin stock solution, a total of 3 g of ampicillin powder was dissolved in 30 mL of distilled water and filter sterilised through 0.22 µm membrane. The solution was kept at -20 °C when not in use.

3.1.2 Competent cells chemical reagents recipes

0.2 M calcium chloride ($CaCl_2$) and 30% glycerol stock (v/v)

A total of 2.22 g of $CaCl_2$ powder was dissolved in 100 mL distilled water and filter sterilised through 0.22 µm membrane. The solution was kept at 4 °C for storage. To prepare 30% glycerol stock, a total of 30 mL of glycerol was mixed with 70 mL distilled water and sterilised by autoclaving at 121°C for 15 minutes. The solution was kept at room temperature.

0.1 M $CaCl_2$ and 0.1 M $CaCl_2$/ 15% glycerol

To prepare 0.1 M $CaCl_2$, dilution of 0.2 M $CaCl_2$ was carried out by adding sterile distilled water in a ratio of 1:1. To prepare 0.1 M $CaCl_2$/ 15% glycerol, equal volume of 0.2 M $CaCl_2$ and 30% glycerol were mixed. The mixture was kept at room temperature.

3.1.3 Agarose gel buffer

10X Tris-Acetate-EDTA (TAE) buffer stock and 1X TAE buffer

To prepare 1 L of 10X TAE buffer stock, a total of 48.4 g of Tris powder, 3.7 g of EDTA powder and 11.4 mL of glacial acetic acid were mixed evenly and topped up to 1 L with distilled water. The buffer was kept at room temperature. To make 1X TAE buffer, 10X TAE buffer was diluted with distilled water in a ratio of 1:9. The buffer was kept at room temperature.

3.2 Methods

3.2.1 Sample collection

Fresh pigeon faecal sample was collected from the Block C Cafeteria of Universiti Tunku Abdul Rahman. A large paper box covered with black plastic bag was placed overnight under the common excretion area of pigeons to collect

the sample. The fresh faecal sample was collected in the early morning using flat side of spatula and stored into petri dish.

3.2.2 Genomic DNA Extraction

The collected faecal sample was subjected to genomic DNA extraction using GF-1 Soil Sample DNA Extraction Kit (Manufacturer: Vivantis). Buffer SL1 and Buffer SB were warmed in 60 °C water bath to prevent salt precipitation in these buffers due to cold temperature. Approximately 0.5 g of glass beads were aliquot into a 2 mL microcentrifuge tube. Approximately 250 mg of faecal samples were weighed and added into the 2 mL microcentrifuge tube containing glass beads. Then, 1 mL of Buffer SL1 was added and mixed well by inverting the tube. The tube was vortexed at maximum speed for 5 minutes and incubated at 70 °C pre-heated water bath for 10 minutes. Sample was vortexed to mix twice during incubation period. The sample was centrifuged at 10,000 x *g* for 5 minutes to pellet the contaminant. Then, 650 µL of supernatant was transferred into a new microcentrifuge tube carefully to avoid sucking up the pellet which can affect the efficiency of DNA extraction. A volume of 65 µL Buffer SL2 was added into the sample and mixed thoroughly by inverting the tube, and then incubated on ice for 5 minutes. Sample was centrifuged at maximum speed 16,000 x *g* for 5 minutes to pellet down contaminants. A volume of 600 µL supernatant was then transferred into new 2 mL microcentrifuge tube without sucking up any debris. A volume of 600 µL Buffer SB was added into the sample and mixed thoroughly using pipette mixing technique.

After that, 600 µL of sample was transferred into a provided column assembled in a clean collection tube and centrifuged at 10,000 x *g* for 1 minute. The flow through was discarded. This step was repeated for the remaining sample. The column was washed with 650 µL SPW Wash Buffer 1, stood for 1 minute, and centrifuged at 10,000 x *g* for 1 minutes. Flow through was discarded. Next, the column was washed with 650 µL SPW Wash Buffer 2 and centrifuged at 10,000 x *g* for 1 minute. The flow through was discarded and column was washed again with 650 µL SPW Wash Buffer 2 with flow through discarded. After washing process, the column was centrifuged at maximum speed with 16,000 x *g* for 2 minutes to remove all ethanol traces to ensure no interference of ethanol towards the downstream applications. The column was placed into a clean 1.5 mL microcentrifuge tube, with 50 µL of preheated Elution Buffer was added directly onto the centre of membrane and stood for 2 minutes for complete elution. Column was centrifuged at 10,000 x *g* for 1 minute to elute DNA. The eluted DNA was marked as 1st elution. For 2nd elution, the elution steps of adding 50 µL Elution Buffer with centrifugation were repeated. Lastly, both eluted DNA stocks were stored in -20 °C.

3.2.3 Verification of size and concentration of extracted DNA

The extracted DNA was verified using agarose gel electrophoresis method. Approximately 0.25 g of Agarose powder (Manufacturer: 1st BASE) was mixed with 25 mL of 1 × TAE buffer and melted completely in microwave to cast a

1% agarose gel. The solution was then poured into gel casting tray with placed small comb and left for 30 minutes to solidify. After that, the comb was removed, and gel was placed into gel tank. 1 × TAE buffer was used to fill up the tank until fully covered the gel. On a parafilm, a total volume of 6 µL containing 2 µL 1st eluted extracted genomic DNA sample, 1 µL 6 × loading dye and 3 µL water was loaded. This step was repeated for 2nd eluted extracted genomic DNA sample. A volume of 6 µL mixture containing 0.5 µL DNA ladder, 4.5 µL water and 1 µL 6 × loading dye was loaded. The gel was run at 90 V for 45 minutes and viewed under gel imager (Manufacturer: BIO-RAD, US). The concentrations of both eluted DNA samples were then measured using NanoDrop™ 2000 UV-Vis Spectrophotometer (Manufacturer: Thermo Fisher Scientific).

3.2.4 Amplification of *16S rRNA* gene using Polymerase Chain Reaction (PCR)

The extracted genomic DNA was undergone amplification of *16S rRNA* gene by PCR. Forward primer (27F) and reverse primer (1492R) were used in this PCR to amplify the *16S rRNA* gene. PCR master-mix including *Taq* polymerase, dNTPs and buffer with added magnesium chloride ($MgCl_2$) were added. Table 3.1 below shows the PCR reagents with respective volume. Firstly, the 1st eluted genomic DNA was selected as template to undergo *16S rRNA* gene PCR. A volume of 1 µL DNA template was mixed with 5 µL of PCR Master Mix and 2 µL for both Forward and Reverse primers, totally 10 µL mixture. For

negative control, total 9 µL mixture without adding DNA template. The mixtures were sent to PCR machine or thermocycler and undergone PCR based on the cycling condition below in Figure 3.1. These steps were repeated for diluted DNA where the 1st eluted DNA was diluted by adding 8 µL sterile distilled water to 2 µL DNA. After completing PCR, the PCR products were stored in -20°C freezer.

Table 3.1: PCR reagents and volume used in *16S rRNA* gene PCR.

	Volume prepared (µL)	
Components	**Negative control**	**Sample**
Forward Primer 27F (1 µM)	2	2
Reverse Primer 1492R (1 µM)	2	2
PCR Master Mix	5	5
DNA Template	-	1
Total	9	10

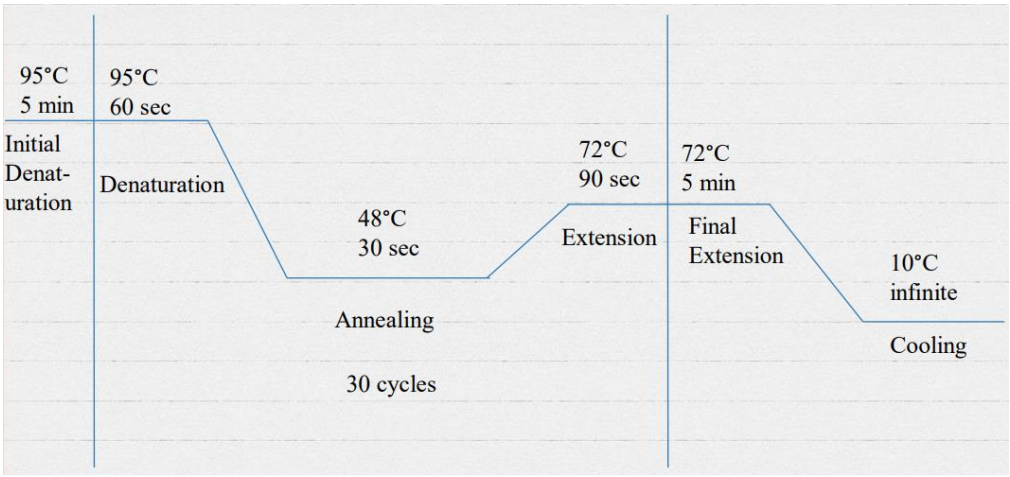

Figure 3.1: *16S rRNA* gene PCR cycling condition, showing the time, temperature and number of cycles, respectively.

3.2.5 Verification of *16S rRNA* Gene PCR using Agarose Gel Electrophoresis

The PCR products were analysed by 1% agarose gel electrophoresis to verify the successful of *16S rRNA* gene PCR. The concentrations of both PCR product containing 1st eluted DNA and diluted DNA were measured using NanoDrop™ 2000 UV-Vis Spectrophotometer (Manufacturer: Thermo Fisher Scientific).

3.2.6 Preparation of *Escherichia coli* competent cells

Overnight TOP10 *E. coli* culture (1 mL) was inoculated into 100 mL Luria-Bertani (LB) broth in 250 mL sterile conical flask. The culture was incubated at 37°C with constant agitation of 220 rpm until the optical density at wavelength 600 nm (OD_{600}= 0.4-0.6) is reached. The following steps were carried out and maintained at cold condition (4°C). The culture that reached the desired OD600 was incubated on ice for 10 minutes. Cold culture was distributed into 50 mL centrifuge tubes and centrifuged at 5,000 rpm, 4°C for 5 minutes and the supernatant was discarded. The pellet was resuspended in 10 mL ice cold 0.1 M $CaCl_2$. The mixture was incubated one ice for 20 minutes and centrifuged at 5,000 rpm, 4°C for 5 minutes. The supernatant was removed, and the pellet was resuspended on 1 mL 0.1 M $CaCl_2$/15% glycerol. Lastly, a volume of 100 µL of the resuspended cells was distributed into cooled 1.5 mL microcentrifuge tube. Ethanol bath that was stored in -80 °C one day before, was used to freeze the cells quickly to be stored in -80 °C.

3.2.7 Cloning into pTG19-T vector

Cloning process was done by using the provided pTG19-T PCR cloning vector kit (Manufacturer: Vivantis). The remaining 6 µL PCR product was mixed with 2 µL of 25 ng/µL pTG19-T vector, 1 µL of 10 × Buffer Ligase and 1 µL of T4 DNA Ligase, totally 10 µL of mixture. The mixture was then incubated in a PCR machine at 16 °C for 4 hours. The incubated mixture was then stored in 4 °C chiller overnight. Table 3.2 below shows the reagents and volume used for cloning process. Figure 3.2 below shows the map information of pTG19-T vector.

Table 3.2: Reagents and volume used in cloning process.

Reagents	Volume used (µL)
pTG19-T vector (25 ng/ µL)	2
10 × Buffer Ligase	1
T4 DNA Ligase	1
PCR product	6
Total	10

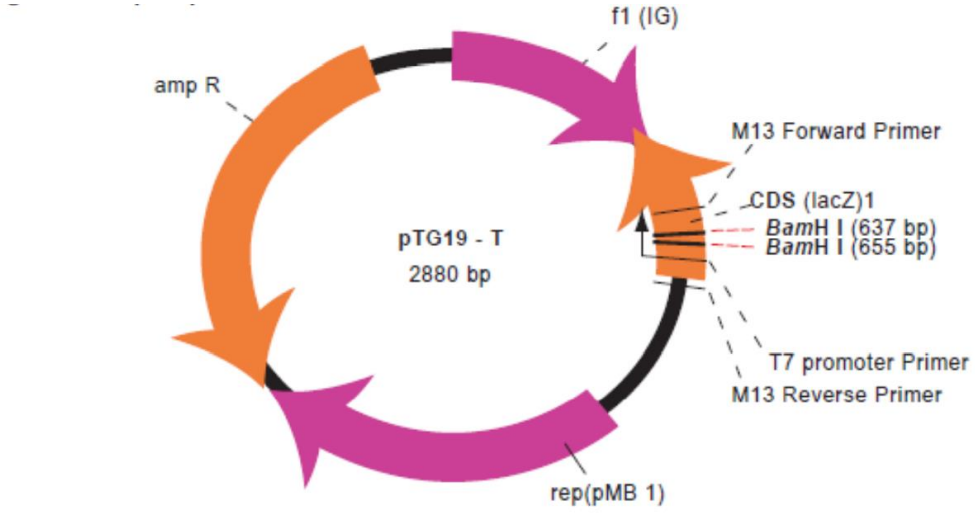

Figure 3.2: Map information of pTG19-T vector.

3.2.8 *E. coli* competent cells transformation

A total of 10 µL of ligation mix was mixed with 100 µL of competent *E. coli* cells and incubated on ice for 30 minutes. The mixture was heat-shocked for 30 seconds at 42°C and immediately incubated on ice for 5 minutes. Next, the mixture was supplemented with 900 µL of LB broth and incubated at 37°C with constant agitation of 220 rpm for 60 minutes. A total volume of 100 µL of the culture was plated on LB agar plate supplemented with 100µg/mL ampicillin antibiotics. Remaining 900 µL mixture was centrifuged at 6000 rpm for 2 minutes. Approximately 800 µL of supernatant was discarded while the remaining 100 µL pellet cells were resuspended and plated on LB agar plate supplemented with ampicillin antibiotics.

3.2.9 Colony PCR

From the transformed colonies that formed on agar plate supplemented with ampicillin, 30 colonies were randomly picked and subcultured carefully without picking up the satellite colonies, on LB agar plate supplemented with ampicillin for overnight growth. The 30 colonies were subcultured on 3 agar plates with maximum 10 colonies for each plate and labelled accordingly. After that, 20 colonies were undergone colony PCR to screen whether the presence of plasmid with *16S rRNA* gene insert in the transformed *E.* coli. During colony PCR, M13 Forward primer and M13 Reverse primer were used. For 20 colonies samples, a total volume of 10 µL with 2 µL of both M13 Forward and M13 Reverse primers, with 5 µL of PCR Master Mix and the colony. Total 20 reactions and 1 negative control for colony PCR. The preparation of PCR was done on ice to prevent degradation of heat sensitive reagents, thus ensuring the efficiency of PCR. Table 3.3 below shows the concentration and volume of each reagents used. Figure 3.3 below shows the PCR cycling conditions.

Table 3.3: PCR reagents and volume to be used for each reaction in colony PCR.

Components	Volume prepared (µL)	
	Negative control	Sample
M13 Forward Primer (1 mM)	2	2
M13 Reverse Primer (1 mM)	2	2
PCR Master Mix	5	5
Total	9	9

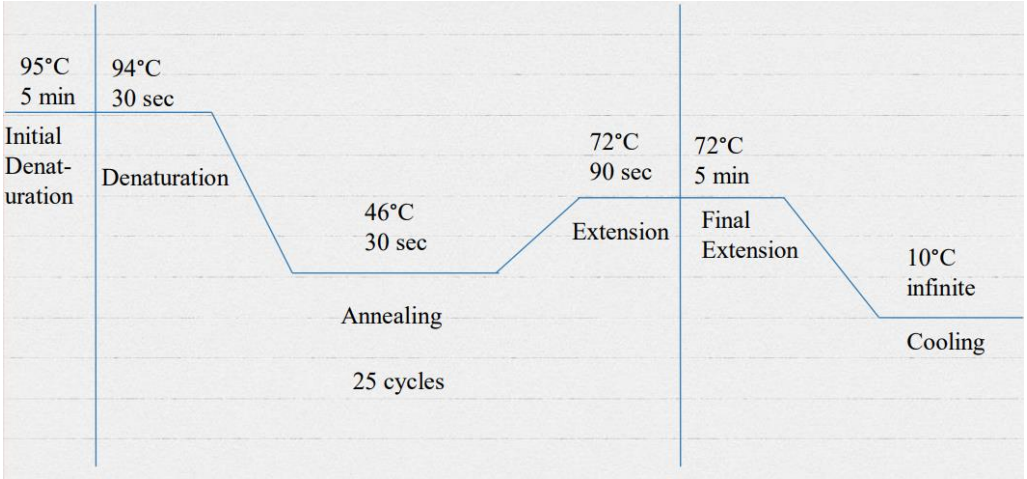

Figure 3.3: Colony PCR Cycling condition, showing the time, temperature and number of cycles, respectively.

After PCR, the PCR products were analysed by 1% agarose gel electrophoresis to verify and select for 10 positive clones to be proceeded to plasmid extraction. During agarose gel electrophoresis, for each PCR products, total 6 µL containing 3 µL PCR products, 1 µL loading dye and 2 µL of water were inserted into the well. For DNA ladder, 0.5 µL DNA ladder was mixed

with 1 µL loading dye and 4.5 µL water. Gel was run at 90 V for 45 minutes and viewed under gel imager (Manufacturer: BIO-RAD, US) to verify the DNA band size compared to the DNA ladder.

3.2.10 Plasmid Extraction and Verification

The plasmid extraction process was done using FavorPrep^{TM} Plasmid Extraction Mini Kit (Manufacturer: Favorgen Biotech Corp.), towards the samples labelled "1, 2, 3, 4, 6, 8, 11, 12, 13, 14" as 10 samples will be sent out for DNA sequencing. Additional samples labelled "16" and "18" were undergone plasmid extraction for additional backup.

For each sample, firstly, 5 mL well-grown overnight bacterial culture was transferred to a centrifuge tube and centrifuged at 11,000 x *g* for 1 minute to pellet the cells while supernatant was discarded completely. A volume of 250 µL of RNase A added FAPD1 Buffer was added to the pellet cells and resuspended completely by pipetting mixing, until no cell pellet was visible. Then, 250 µL of FAPD2 Buffer was added and the tube was immediately inverted gently for 5 to 10 times to ensure lysate become clear. The sample mixture was incubated at room temperature for 2 to 5 minutes to lyse the cells but the incubation period would not exceed 5 minutes. Then, 350 µL of FAPD3 Buffer was added and the tube was inverted 5 to 10 times immediately to neutralize the lysate and avoid asymmetric precipitation. The sample mixture was then centrifuged at maximum speed, approximately 18,000 x *g* for 10

minutes to clarify the lysate. After that, the supernatant was transferred carefully without picking up any white pellet into FAPD Column that placed in a Collection Tube. Centrifugation at 11,000 x *g* was carried out for 30 seconds. The flow through was discarded and column was placed back to the collection tube. A volume of 400 µL W1 Buffer was added to FAPD Column and centrifuged at 11,000 x *g* for 30 seconds. Flow through was discarded and colum was placed back again to the collection tube. Then, 700 µL Wash Buffer was added to the column, centrifuged at 11,000 x *g* for 30 seconds, flow through discarded and column placed back to collection tube. Centrifugation at maximum speed, approximately 18,000 x *g* was carried out for additional 3 minutes to dry the FAPD Column and removed the residual liquid thoroughly. FAPD Column was then placed to a new 1.5 mL microcentrifuge tube. A volume of 50 µL ddH_2O was added to membrane centre of FAPD Column and stood for 1 minute for complete absorption. Centrifugation at maximum speed, approximately 18,000 x *g* for 1 minute for 1st plasmid DNA elution. The elution step was repeated for 2nd elution. All DNA were stored in -20 °C.

The concentration of eluted plasmid DNA was measured using NanoDrop^{TM} 2000 UV-Vis Spectrophotometer (Manufacturer: Thermo Fisher Scientific). ddH_2O was used as blank. For size verification of eluted plasmid DNA, 1% agarose gel electrophoresis was carried out. Total volume of 6 µL containing 3 µL of eluted plasmid, 1 µL 6 × loading dye and 2 µL of water was loaded into well, this step was repeated for eluted plasmid samples “1”, “2”, “3”, “4”, “6”, “8”, “11”, “12”, “13”, “14”, “16” and “18”. Gel was run at 90 V

for 45 minutes and viewed under gel imager (Manufacturer: BIO-RAD, US) to verify the DNA band size compared to the DNA ladder.

3.2.11 DNA Sequencing

Samples "1, 2, 3, 4, 6, 11, 12, 14, 16, 18", which achieved the required concentration and volume of DNA sequencing were selected to be shipped out and labelled "C1, C2, C3, C4, C5, C6, C7, C8, C9, C10" accordingly in ascending order compared to the pre-labelled samples. These samples were prepared in sterile 1.5 mL microcentrifuge tubes with recommended volumes and were sealed using parafilm. Then, the samples were shipped out to First BASE Laboratories Sdn Bhd (604944-X) for sequencing.

When the sequencing results were received back from the vendor, nucleotide BLAST (BLASTn) program was used to analyze the result. From the chromatogram received, the targetted unknown bacterial sequence of each sample was searched and highlighted using M13 Forward primer as a key guide. The unknown bacterial sequence was subjected to BLASTn program of National Center for Biotechnology Information (NCBI) website to blast and look for the sequence ID or accession numbers, bacterial identity and identities percentage. MEGA X software was used to construct a phylogenetic tree for these 10 identified bacteria, using pre-setted Maximum Likelihood method and Tamua-Nei model.

CHAPTER 4

RESULTS

4.1. Genomic DNA Extraction

The extracted DNA concentration was measured using NanoDrop™ 2000 UV-Vis Spectrophotometer (Manufacturer: Thermo Fisher Scientific). For 1st DNA elution, the DNA concentration measured at 9.8 ng/µL with A_{260}/A_{280} ratio of 1.76 and A_{260}/A_{230} ratio of 1.60. For second DNA elution, the DNA concentration measured at 2.5 ng/µl with A_{260}/A_{280} ratio of 1.69 and A_{260}/A_{230} ratio of -8.31. Figure 4.1 shows 1st and 2nd DNA elution in 1% agarose gel.

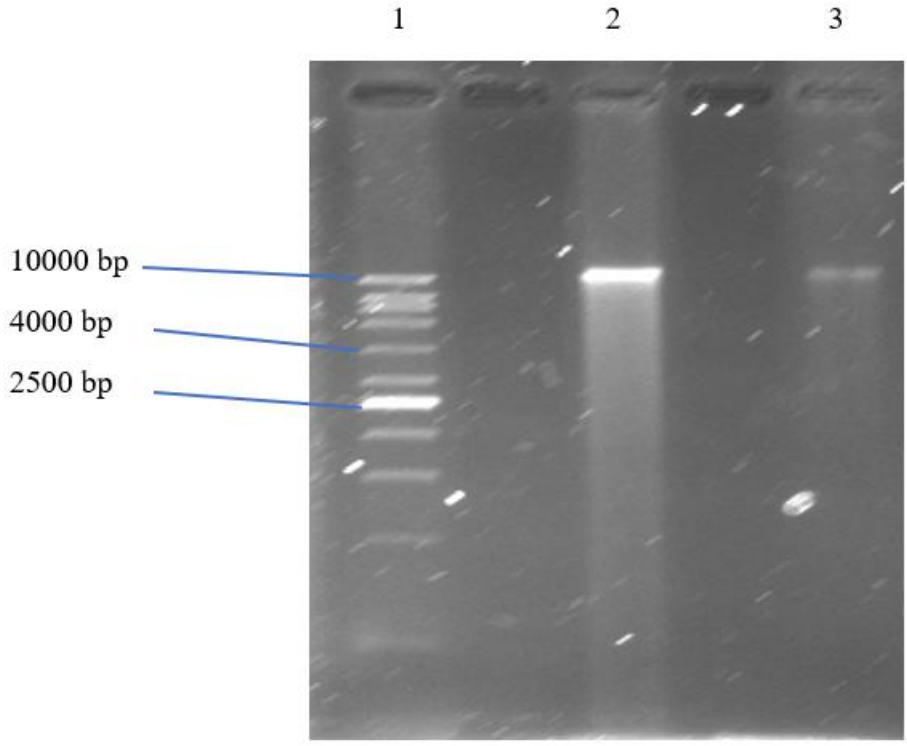

Figure 4.1: 1% agarose gel analysis of extracted genomic DNA. Lane 1: 1 kb DNA ladder; lane 2: 1st eluted DNA; lane 3: 2nd eluted DNA.

4.2 *16S rRNA* gene amplification

The *16S rRNA* gene was successfully amplified. Below is the Figure 4.2 showing the gene size is approximately 1.5 kb when compared to the DNA ladder on 1% agarose gel. Figure 4.3 shows the band size comparison between PCR products containing 1st eluted DNA and diluted DNA.

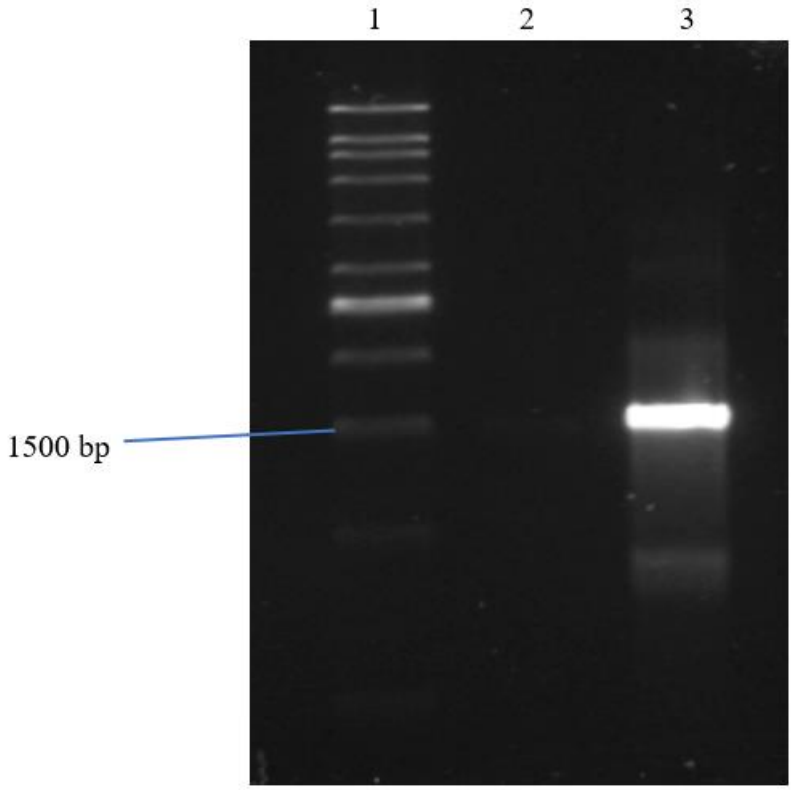

Figure 4.2: 1% agarose gel analysis of PCR products from *16S rRNA* gene amplification. Lane 1: 1 kb DNA ladder; lane 2: negative control; lane 3: amplified *16S rRNA* gene.

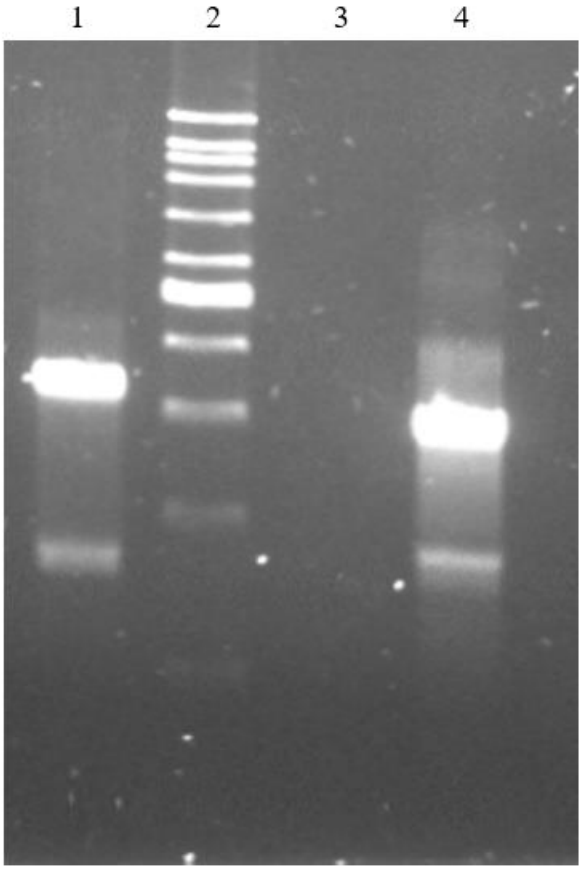

Figure 4.3: 1% agarose gel analysis of PCR products from *16S rRNA* gene amplification. Lane 1: 1st eluted DNA; lane 2: 1 kb DNA ladder; lane 3: negative control; lane 4: diluted DNA.

4.3 Cloning and transformation

The amplified 1*6S rRNA* gene was successfully cloned into pTG19-T vector and transformed into TOP10 *Escherichia coli* using heat shock method. Figure 4.4 below shows the TOP10 *E. coli* plated on LB agar plate containing 100 µg/mL ampicillin which act as a negative control. No colonies formed on the plate. Figure 4.5 shows the transformed TOP10 *E. coli* with pTG19-T cloned plasmid from the pellet source while Figure 4.6 shows the transformed TOP10 *E. coli* with pTG19-T cloned plasmid that have not undergone centrifugation. Colonies were observed on plates shown in Figure 4.5 and Figure 4.6.

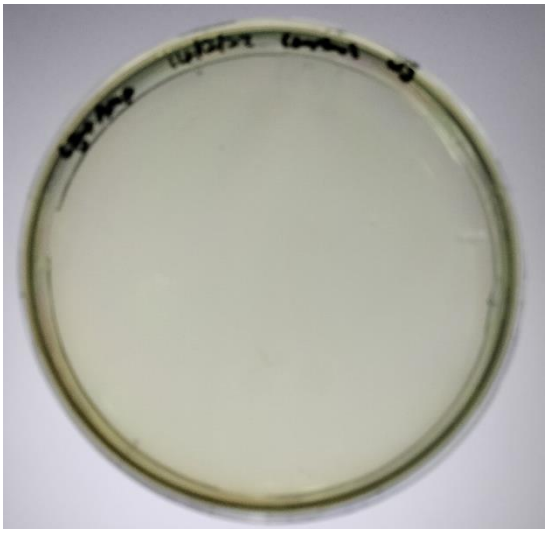

Figure 4.4: TOP10 *E. coli* plated on an ampicillin plate. No colonies were formed as *E. coli* cells were killed by ampicillin due to the absence of ampicillin resistance gene.

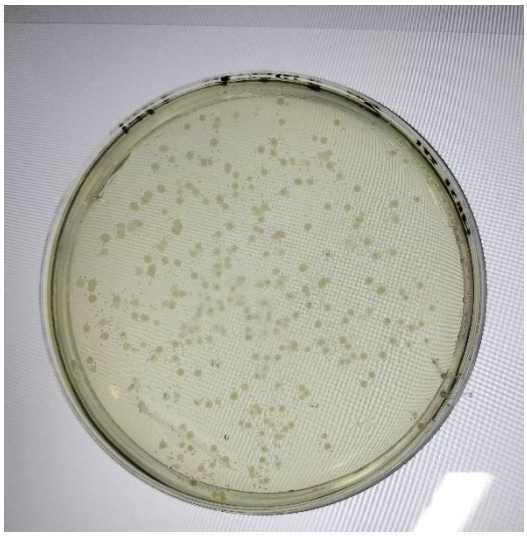

Figure 4.5: Transformed TOP10 *E. coli* concentrated by centrifugation, plated on LB agar plate supplemented with ampicillin. Colonies formed due to presence of ampicillin resistance gene in pTG19-T vector. Satellite colonies formed around the transformed TOP10 *E. coli*.

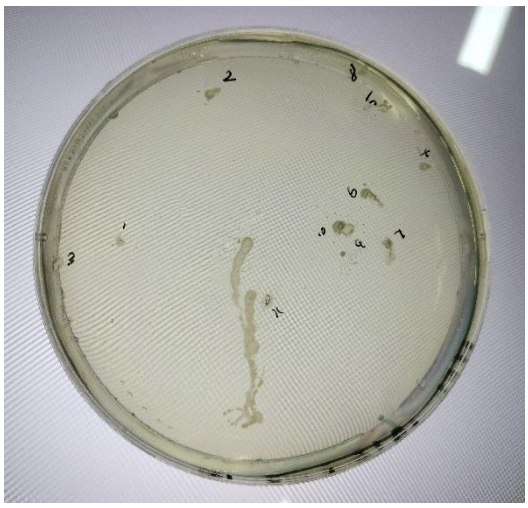

Figure 4.6: Transformed TOP10 *E. coli* plated on LB agar plate supplemented with ampicillin. Colonies formed due to presence of ampicillin resistance gene in pTG19-T vector. Colonies were labelled with numbers “1” to “11”.

4.4 Colony PCR

Colonies from Figure 4.5 and Figure 4.6 were randomly selected and subcultured on LB agar plates supplemented with ampicillin antibiotic 100 µg/mL. Colonies were selected and subcultured carefully without picking up the satellite colonies and maximum 10 colonies for each plates. Figures 4.7, 4.8 and 4.9 below show the colonies that have been subcultured on LB agar plate containing ampicillin. Each plate (plates 1, 2 and 3) has 10 subcultured colonies, labelled with number 1 to 30.

Figure 4.7: Plate 1 with selected colonies number 1 to 10.

Figure 4.8: Plate 2 with selected colonies number 11 to 20.

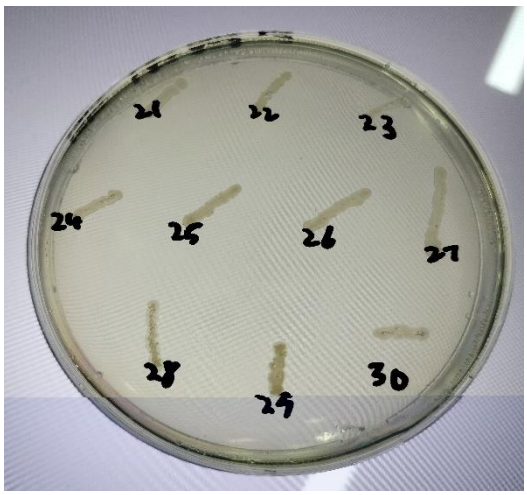

Figure 4.9: Plate 3 with selected colonies number 21 to 30.

Screening of colonies was carried out to confirm the succeed of transformation into TOP10 *E. coli* and carrying the insert of *16S rRNA* gene. Figures 4.10 and 4.11 show the results of colony PCR analysed of using 1% agarose gel electrophoresis.

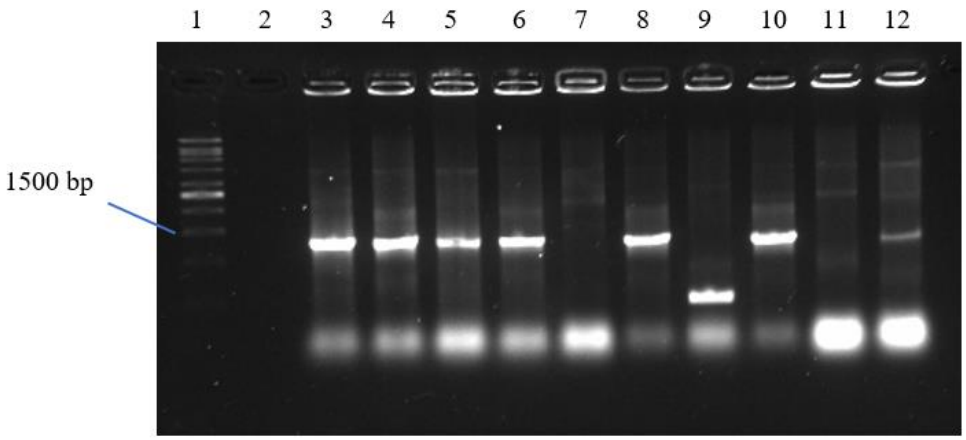

Figure 4.10: Screening of positive clones by colony PCR. Lane 1: 1 kb DNA ladder; lane 2: negative control; lanes 3 to 6, 8, 10 and 12: positive results; lanes 7, 9 and 11: negative results.

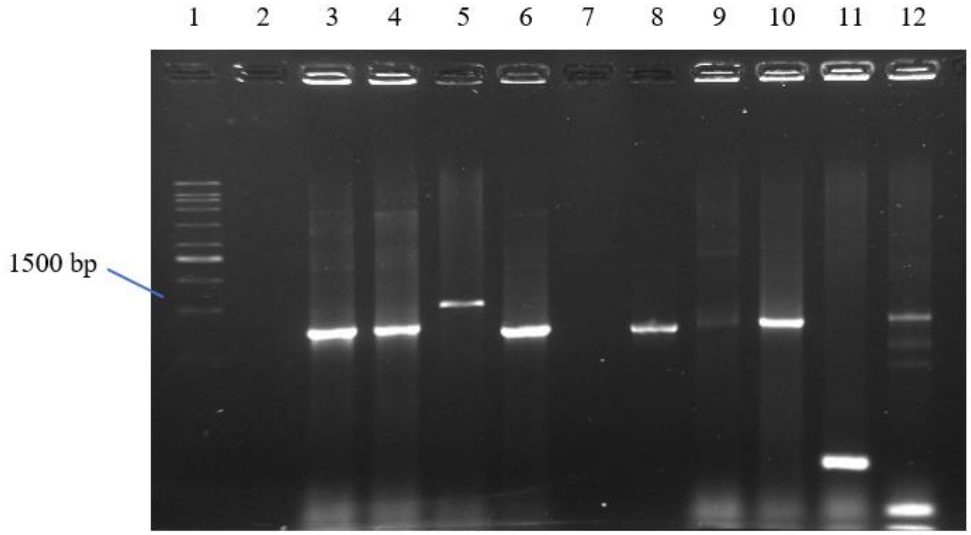

Figure 4.11: Screening of positive clones by colony PCR. Lane 1:1 kb DNA ladder; lane 2: negative control; lanes 3, 4, 6, 8, 10 and 12: positive results; lanes 5, 7, 9 and 11: negative results.

4.5 Plasmid Extraction

Twelve positive clones (colonies 1, 2, 3, 4, 6, 8, 11, 12, 13, 14, 16 and 18) verified by colony PCR were subjected to plasmid extraction using FavorPrep^{TM} Plasmid Extraction Mini Kit (Manufacturer: Favorgen Biotech Corp.). Figures 4.12, 4.13 and 4.14 show the extracted plasmid analysed by 1% agarose gel electrophoresis. Figure 4.13 shows absence of band of sample 13 in lane 6. Therefore, 10 plasmids, samples 1, 2, 3 and 4 shown in Figure 4.12, samples 6, 11, 12 and 14 shown in Figure 4.13 and samples 16 and 18 in Figure 4.14 were selected for DNA sequencing.

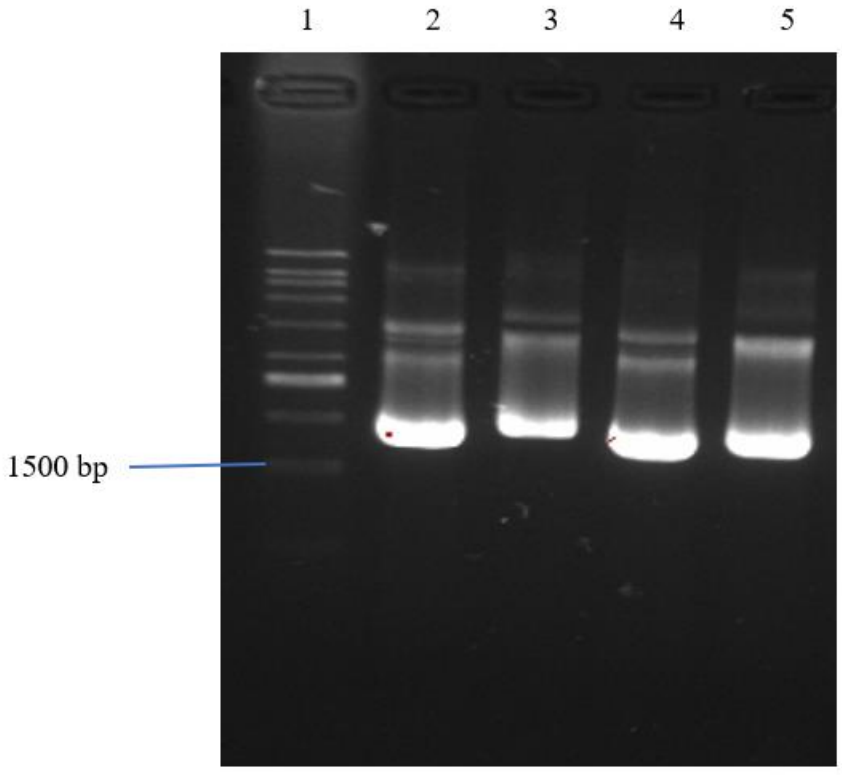

Figure 4.12: 1% agarose gel electrophoresis analysis of extracted plasmids. Lane 1: 1 kb DNA ladder; lane 2: sample 1; lane 3: sample 2; lane 4: sample 3; lane 5: sample 4.

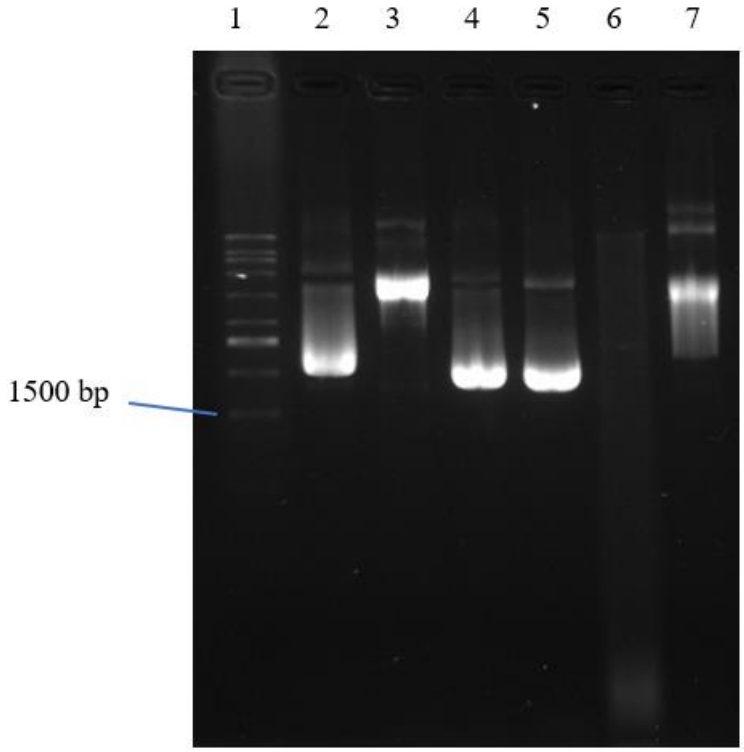

Figure 4.13: 1% agarose gel electrophoresis analysis of extracted plasmids. Lane 1: 1 kb DNA ladder; lane 2: sample 6; lane 3: sample 8; lane 4: sample 11; lane 5: sample 12; lane 6: sample 13; lane 7: sample 14.

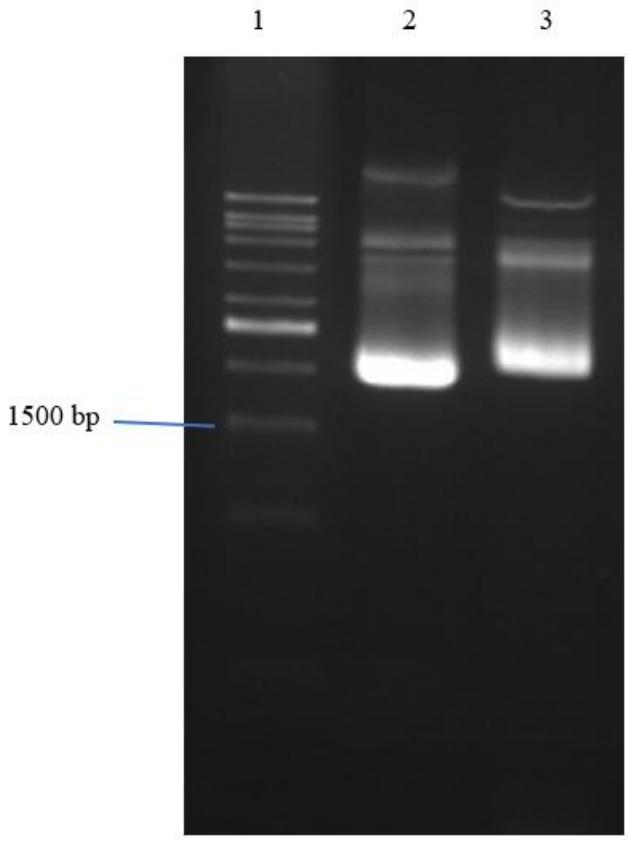

Figure 4.14: 1% agarose gel electrophoresis analysis of extracted plasmids. Lane 1: 1 kb DNA ladder; lane 2: sample 16; lane 3: sample 18.

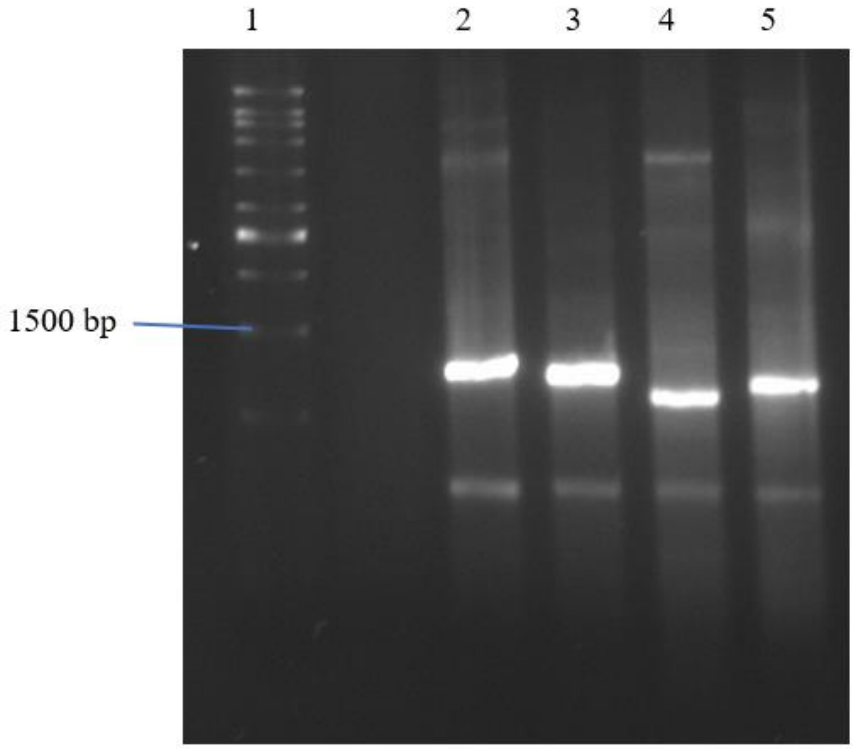

Figure 4.15: 1% agarose gel electrophoresis analysis of *16S rRNA* gene PCR of extracted plasmids. Lane 1: 1 kb DNA ladder; lane 2: sample 1; lane 3: sample 2; lane 4: sample 3; lane 5: sample 4.

Plasmid DNA concentration of each samples were measured using NanoDrop™ 2000 UV-Vis Spectrophotometer (Manufacturer: Thermo Fisher Scientific) to further verify whether meet the basic requirement prior to DNA sequencing. Table 4.1 shows the ten plasmids' DNA concentrations and purities.

Table 4.1: Ten plasmids' DNA concentration and purities.

Extracted plasmids	DNA concentration (ng/µL)	DNA purity ($A_{260/280}$ ratio)	DNA purity ($A_{260/230}$ ratio)
1	302.1	1.89	2.36
2	153.4	1.90	2.31
3	299.6	1.89	2.31
4	197.4	1.90	2.30
6	205.3	1.90	2.33
11	278.9	1.89	2.13
12	337.3	1.91	2.22
14	134.8	1.91	2.31
16	239.0	1.92	2.30
18	142.9	1.96	2.37

4.6 DNA Sequencing

Samples “1, 2, 3, 4, 6, 11, 12, 14, 16 and 18” were shipped out to First BASE Laboratories Sdn Bhd (604944-X) for sequencing. The primer used in DNA sequencing was M13 forward primer. After receiving back the sequencing result, nucleotide BLAST (BLASTn) was used to analyze the sequences of *16S rRNA* gene. For each samples, approximately 800 bases was trimmed and proceeded to BLASTn. Through the program, the bacterial sample ID, sequence ID or accession numbers and identities can be known and recorded in Table 4.2.

Based on the BLASTn result from National Center for Biotechnology Information (NCBI) website, among the 10 identified bacterial samples, 3 samples which are “C2”, “C5” and “C9” have same sequence ID and bacteria identity of *Lactobacillus crispatus* strain TCD56.8 *16S ribosomal RNA* gene, partial sequence.

Table 4.2: BLASTn result of ten clones.

Sample ID	**Sequence ID/ Accession Number**	**Bacteria Identity**	**Identities Percentage, %**
C1	AY959061.1	Uncultured bacterium clone rRNA288 *16S ribosomal RNA* gene, partial sequence	99.88 (829/830)
C2	KU851161.1	*Lactobacillus crispatus* strain TCD56.8 *16S ribosomal RNA* gene, partial sequence	100.00 (637/637)
C3	KU991819.1	*Lactobacillus crispatus* strain 5-1-1 *16S ribosomal RNA* gene, partial sequence	100.00 (927/927)
C4	CP027644.1	*Ligilactobacillus salivarius* strain IBB3154 chromosome, complete genome	99.59 (726/729)
C5	KU851161.1	*Lactobacillus crispatus* strain TCD56.8 *16S ribosomal RNA* gene, partial sequence	100.00 (772/772)
C6	MK713651.1	*Candidatus* Mycoplasma liparidae clone H1 *16S ribosomal RNA* gene, partial sequence	91.85 (811/883)
C7	CP059140.1	*Lactobacillus crispatus* strain B4 chromosome, complete genome	99.77 (851/853)

C8	KU991816.1	*Lactobacillus crispatus* strain N-11 *16S ribosomal RNA* gene, partial sequence	100.00 (772/772)
C9	KU851161.1	*Lactobacillus crispatus* strain TCD56.8 *16S ribosomal RNA* gene, partial sequence	99.76 (842/844)
C10	LT223590.1	*Lactobacillus ingluviei* partial *16S rRNA* gene, strain Marseille-P209	99.88 (825/826)

Subsequently, phylogenetic tree was constructed using MEGA X software, with pre-setted Maximum Likelihood method and Tamua-Nei model. Figure 4.15 shows the phylogenetic tree of the different bacterial samples, which indicating the evolutionary relationship among the bacterial samples. From the Figure 4.15, phylogentic tree shows 8 different bacterial species due to sample ID “C2”, “C5” and “C9” are the same bacteria species and strain.

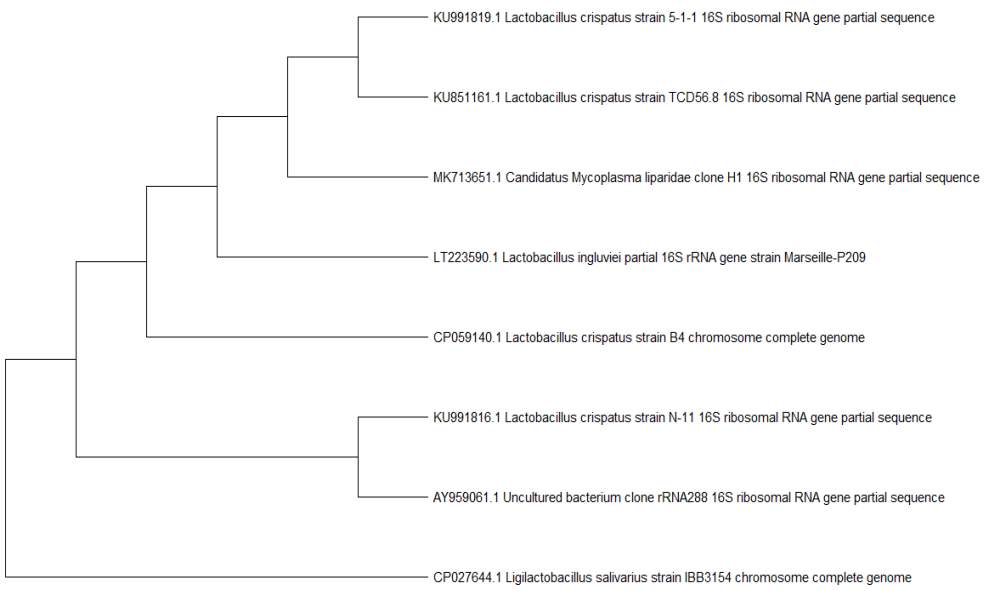

Figure 4.16: Phylogenetic tree showing the evolutionary relationship of 8 different bacterial species.

CHAPTER 5

DISCUSSIONS

5.1 Genomic DNA Extraction

The total genomic DNA in pigeon faecal sample was isolated using GF-1 Soil Sample DNA Extraction Kit (Manufacturer: Vivantis) instead of conventional method. This is because the conventional method is time-consuming and tedious compared to DNA extraction kit (Chauhan, 2018). In addition, the DNA extraction kit used in this study provides numerous benefits like high yield and purity, timesaving, reliable, reproducible, easy to handle, non-toxic or organic-based extraction needed, and the eluted DNA is ready to be used in downstream application (Vivantis technologies Malaysia, n.d.). DNA extraction process can be commonly distinguished into 5 main steps. Firstly, cell lysis to release DNA through chemical method by adding detergent to "punch" a hole on plasma membrane or physical method like pressurization, sonication and shearing cells that based on mechanical force (Gallik, 2013). Second, DNA separation from proteins and cellular debris. Third, precipitation of DNA in which alcohol will be added carefully to precipitate the DNA. Fourth, purification of DNA by adding buffer with slightly alkalotic characteristic. Last, verification and confirmation of DNA concentration, purity and presence (Science Learning Hub - Pokapū Akoranga Pūtaiao, 2009).

In this study, glass beads and Buffer SL1 containing salts were used to break the cells. The salts in Buffer SL1 aid in protecting the phosphate groups of DNA backbone, due to the presence of positively charged sodium ions in Buffer SL1. This process also known as homogenization where the cell components and organelles are released as free suspension (de Araújo, et al., 2015). In addition, absolute ethanol was confirmed to be added into both SPW Wash Buffer and SPW Wash Buffer 2 to dilute the concentrated buffer, the ethanol can aid in precipitation as DNA will be insoluble in alcohol and salt, thus forming white precipitate (Science Learning Hub - Pokapū Akoranga Pūtaiao, 2009).

For the verification of presence of DNA, it can be observed that the DNA bands are presented in lane 3 and lane 5 in Figure 4.1. DNA band in lane 3 is more obvious and brighter than that of lane 5 due to higher DNA concentration of 1st eluted DNA in lane 3 compared to 2nd eluted DNA in lane 5. Moreover, DNA band can be viewed under gel imager with brighter characteristic, due to the presence of SYBR Safe DNA staining dye that fluorescence when exposed to Trans UV light.

For the measured DNA concentration, the 1st DNA elution showed 9.8 ng/µL with A_{260}/A_{280} ratio of 1.76 and A_{260}/A_{230} ratio of 1.60 while 2nd DNA elution showed 2.5 ng/µl with A_{260}/A_{280} ratio of 1.69 and A_{260}/A_{230} ratio of -8.31. The concentration of 1st elution is higher than that of 2nd elution. This is because after eluting the 1st batch of DNA, the remaining DNA at the centre

membrane of column became lesser. When the 2nd batch of Elution Buffer was added, the DNA concentration will be lower or considered as diluted. In addition, the A_{260}/A_{280} ratio of both 1st and 2nd eluted DNA are near to 1.8 which can be accepted as high purity (Thermo Scientific NanoDrop Lite, 2012). For A_{260}/A_{230} ratio, both eluted DNA have lower value than the common range within 2.0 to 2.2. This may be due to presence of contaminant like residual guanidine in the eluted DNA which absorbed and detected at wavelength 230 nm when using the NanoDrop™ 2000 UV-Vis Spectrophotometer (Manufacturer: Thermo Fisher Scientific) (Thermo Scientific NanoDrop Spectrophotometers, n.d.). For negative value -8.31, it may be due to impurities like guanidine salt that were involved when using the kit for extraction.

5.2 *16S rRNA* Gene PCR

16S rRNA gene is ribosomal gene that can be found in 30S subunit of bacteerial ribosome. It functions by providing structural support to ribosomal protein by aiding in the binding of 30S and 50S subunits which lead to ribosome function. It also involves in protein translation process due to the interaction between its' 3' end sequence with other proteins like S1 and S21. *16S rRNA* gene consists of two main domain such as conserved domain and hypervariable region. The conserved domain can be found as same and unchangeable in all types of bacteria while hypervariable region which involve nine regions is different and unique for each type of bacteria, thus making the unknown bacteria are able to be identified and classified by looking at this region (Chauhan, 2020). In

addition, *16S rRNA* gene has total 1500 base pairs (bp) in length, which is sufficient in size for bioinformatics work (Janda and Abbott, 2007). Therefore, *16S rRNA* gene also known as universal gene marker.

In this study, *16S rRNA* gene polymerase chain reaction (PCR) was carried out to screen and select unknown bacterial species from the gene pool which contains both prokaryotes and eukaryotes. Universal primers like 27 Forward and 1492 Reverse were used in PCR process due to high coverage which almost the full length of *16S rRNA* gene (Lane et al., 1985; Frank et al., 2008). The PCR preparation was done in laminar fume hood to ensure sterility. The cycling condition of *16S rRNA* gene PCR was shown above in Figure 3.1. The denaturation phase was at 95 °C purposely to separate doube stranded DNA by breaking the hydrogen bonds between nitrogenous bases (Masny and Plucienniczak, 2003). During annealing phase, 48 °C was applied to let the DNA primers to attach to template. When proceeded to extension phase with 72 °C, the temperature increased to enable synthesize of new complementary DNA strand through the replication by *Taq* polymerase with adding of dNTPs in 5' to 3' direction. 30 cycles were carried out to ensure doubling in DNA copies' numbers (Yourgenome Organization, 2021). The amplified products were then cooled to 10 °C of stable status without any reaction occurred.

For the confirmation of successful of *16S rRNA* gene PCR without contamination, 1% agarose gel electrophoresis was carried out. From the result of agarose gel electrophoresis shown in Figure 4.2 and 4.3, it can be observed

that there is absence of band in negative control lane, indicating no contamination while the DNA bands that present show approximately same size (1500 bp) compared to DNA ladder. This can be concluded that the *16S rRNA* gene PCR was succeed.

For the DNA concentration of 1st eluted DNA and diluted DNA that showed -39.4 ng/µL and -130.2 ng/µL, the presence of negative value in DNA concentration may be due to technical error during the measuring process. For instance, introducing of air bubbles when measuring the samples on NanoDrop™ 2000 UV-Vis Spectrophotometer (Manufacturer: Thermo Fisher Scientific) will affect the result's accuracy (Bradburn, n.d.). Another reason may be due to the contaminated pedestal which was not cleaned properly before introducing blank for measuring process (Yau, 2017).

5.3 Cloning and transformation

Cloning process was carried out by inserting the amplified gene of interest (*16S* rRNA gene) into pTG19-T vector. In this study, pTG19-T vector (Manufacturer: Vivantis) was used due to its' time-saving, rapid, convenient and efficient features. Based on the protocol handbooks for this vector, this vector provides linearized 3'dT overhangs which can prevent vector recircularization and able to recombine with amplified PCR products that contain 3'dA overhangs, thus producing higher yield of recombinant clones. Based on the advantages provided by the manufacturer in the handbooks, this vector is efficient with higher percentage (>80%) of recombinant clones contain the targeted DNA. It has also clone selection like M13 primer sites for PCR screening which can be used in colony PCR. Moreover, *lacZ* gene is available for blue-white selection and restriction enzyme cutting site to release the insert (Vivantis technologies Malaysia., n.d.).

In this study, T4 DNA Ligase and 10 × Buffer Ligase were added with vector and PCR product for cloning. DNA Ligase is enzyme that aid in DNA reparing process. Phosphodiester bond can be formed between 5' end of one side of DNA together with 3' end of another side. In this study, T4 DNA ligase which origined from T4 bacteriophage, was used for the ligation of the amplified *16S rRNA* gene and pTG19-T vector. Adenosine triphosphate (ATP) energy molecule which act as activator was utilized by T4 DNA ligase for its' mechanism of reaction. T4 DNA ligase aid in ligation of sticky ends of amplified *16S rRNA* gene which act as insert, together with that of pTG19-T

vector. 10 × Buffer Ligase was playing role by providing ATP molecule to T4 DNA ligase. The mixture was then incubated in 16 °C for 4 hours as usually the optimum activity of DNA ligase is at 16 °C (Rossi, et al., 1997; Chauhan, 2019). After incubation period, the mixture was stored in 4 °C overnight instead -20 °C, purposely for slower ligation process. This is because ligation still can be carried out in 4 °C, but not available at -20 °C, this lead to higher yield.

The successful of cloning process can be verified by proceeding to transformation process. In this study, heat-shock transformation method was used. Prior to heat-shock transformation, *E. coli* competent cells were prepared using Hanahan's method (Hanahan, 1983). This is because *E. coli* in competent status can take up exoegenous DNA with higher molecular weight easily. During the heat shock transformation, the mixture of plasmid and *E. coli* competent cells was incubated on ice for 30 minutes. The mixture was heat-shocked for 30 seconds at 42°C and immediately incubated on ice for 5 minutes. The reason of incubating the mixture on ice for 30 minutes is to let the DNA binds to *E. coli* cells' surface, due to the characteristics of calcium chloride ($CaCl_2$), that the positively charged calcium ions (Ca^{2+}) that attract the negatively charged DNA. Then, the mixture was heat shocked for 30 seconds in order to let DNA enters cells due to the sudden heating process that lower the membrane permeability and creating smaller pores. The mixture was then incubated on ice for 5 minutes purposely to recover the membrane permeability and pores (Panja, et al., 2008; Chang, et al., 2017). The mixture was then supplemented with 900 µL of LB broth and incubated at 37°C with constant agitation of 220 rpm for 60 minutes for growing purpose and spread onto agar

plate containing sufficient amount and concentration of ampicillin antibiotics for screening and selection purpose.

Antibiotic selection was carried out to determine the successful of transformation. The results of antibiotic selections were shown in Figure 4.4, 4.5, and 4.6. There are no colonies formed in Figure 4.4, as untransformed *E. coli* cells were spread on LB agar plate containing ampicillin, the *E. coli* cells were killed by ampicillin, due to lack of ampicillin resistance gene. This served as negative control to ensure that the untransformed *E. coli* competent cells are not resist to ampicillin. For Figure 4.5 and 4.6, colonies were formed on LB agar plate containing ampicillin, this is because transformed *E. coli* cells that carrying recombinant clones contain ampicillin resistance gene which make them to survive. However, there is difference can be observed when comparing plates in Figure 4.5 and 4.6, where the *E. coli* cells that have been centrifuged had more colonies than that without centrifuged in Figure 4.6. This is because that the centrifuged *E. coli* would remain at pellet form which higher in numbers and concentrations. Therefore, the colonies formed in Figures 4.5 and 4.6 were selected to be proceeded to colony PCR. Besides, it can be observed that satellite colonies were formed at surrounding of the transformed *E. coli* colonies in both Figures 4.5 and 4.6. Satellite colonies were formed as ampicillin near to the area of colony of interest was used up and degraded (Anon., 2013), due to secretion of β-lactamase which suppress ampicillin which belongs to β-lactam antibiotics (Whitehouse, et al., 2018; Oswald, 2019). These colonies were usually bacteria without ampicillin resistant gene.

5.4 Colony PCR

Colony PCR of 30 selected and sub-cultured colonies were done purposely to examine whether the colonies that carrying recombinant clones contain the gene of interest or insert, which is *16S rRNA* gene. From the results part shown in Figures 4.10 and 4.11, in the 1% agarose gel, there are total 13 positive clones were identified, where the DNA bands size were approximately 1500 bp when compared to DNA ladder, thus indicating that the gene of interest (*16S rRNA* gene) was inserted in each positive clones, instead of unspecific region. For those lanes without any DNA bands, it may be due to technical error when picking and introducing colonies into PCR mix, where too little or no colonies were added.

M13 Forward and Reverse primers were used in this colony PCR. These primers were used due to the specific features of pTG19-T vector which contain the M13 Forward primer and Reverse primer binding sites at nucleotide position 575-592 and 745-765 respectively. In addition, the insert site for *16S rRNA* gene is in between both M13 Forward and Reverse primers binding sites, approximately 73 bp away from M13 Forward primer binding site. In colony PCR, the primers used were differed from that of *16S rRNA* gene PCR, in order to prevent targeting the *16S rRNA* gene of *E. coli* competent cells which carrying plasmid. Furthermore, both M13 Forward and Reverse primers are specifically designed to target and amplify specific gene region instead of entire

plasmid. Besides, the cycling condition of colony PCR was differed to *16S rRNA* gene PCR, in which the annealing stage condition is 46 °C and the cycles were 25 cycles. This is due to the used of M13 primers which have different melting temperature compared to 16S universal primers. The annealing temperature is usually approximately 5 °C lower than the melting temperature of primers (Maddocks and Jenkins, 2017). In addition, 25 cycles are enough to produce desired number of copies.

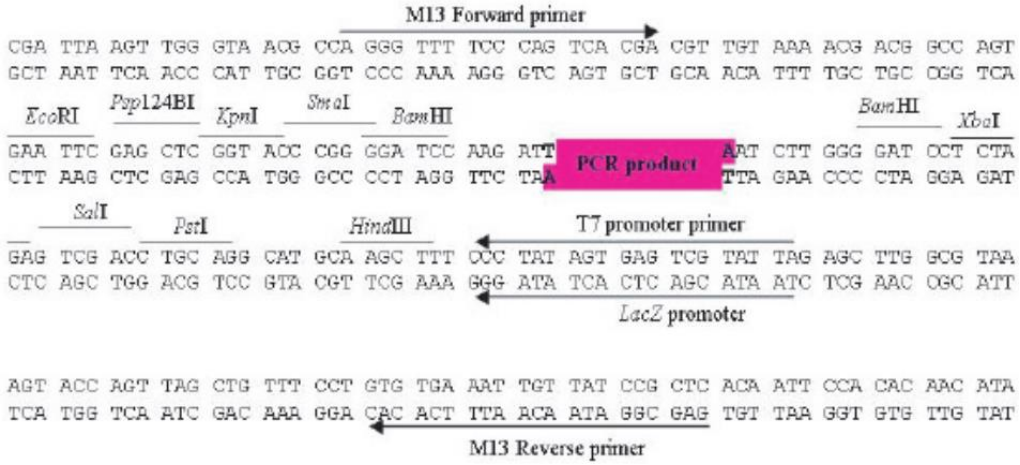

Figure 5.1: Multiple cloning site sequence of pTG19-T vector.

5.5 Plasmid Extraction

Ten colonies with correct size and successful verification from colony PCR were selected for plasmid extraction process. The extraction of plasmids was carried out using FavorPrep^{TM} Plasmid Extraction Mini Kit (Manufacturer: Favorgen Biotech Corp.). From Figures 4.12, 4.13 and 4.14, it can be observed presence of DNA bands in 1% agarose gel, indicating presence of plasmid DNA. However, each plasmid DNA vary in size, possibly due to different conformations, for instance supercoiled, multimer and nicked. The plasmid DNA concentration of each samples were measured using NanoDrop^{TM} 2000 UV-Vis Spectrophotometer (Manufacturer: Thermo Fisher Scientific) to further verify whether meet the basic requirement for DNA sequencing, which are 100 ng/μL. However, sample "8" did not reached the concentration required to be shipped out for DNA sequencing. Therefore, samples "16" and "18" which reached the required concentration were selected to replace the vacancy. The measured concentration of plasmid DNA concentration of each sample are shown below in Table 5.1.

Table 5.1: Measured plasmid DNA concentration of each sample.

Samples	Plasmid DNA Concentration (ng/µL)
1	302.1
2	153.4
3	299.6
4	197.4
6	205.3
8	88.7
11	278.9
12	337.3
13	71.2
14	134.8
16	239.0
18	142.9

5.6 DNA Sequencing and Data Analysis

From the Table 4.1 and constructed phylogenetic tree above in Figure 4.15, it can be observed that most of the bacteria genus are belongs to *Lactobacillus*, under *Lactobacillaceae* family. *Lactobacillus* is usually defined as a genus consists of bacteria with Gram-positive, rod-shaped, catalase-negative, and non-spore-forming characteristics (Hammes and Vogel, 1995). Lactic acid bacteria (LAB) are acidophile and fermentative, with lower GC content than other bacteria. LAB carry out sugar metabolism or carbohydrate fermentation process which produces lactic acid as the major end-product. They are usually found in gastrointestinal (GI) tracts and generally considered as safe. LAB are important in food production and used as probiotics which are beneficial to human health (Salvetti, et al., 2012). Based on the results shown above, there are mostly 8 out of 10 bacteria belong to lactic acid bacteria where 7 are *Lactobacillus crispatus* with another one *Ligilactobacillus salivarius*. These bacteria are found in pigeon faecal sample in this study, indicating the presence of these bacteria in pigeon's GI tract.

Lactobacillus ingluviei strain Marseille-P209 was found in pigeon faecal samples in this study. Based on NCBI, this *L. ingluviei* partial *16S rRNA* gene, strain Marseille-P209 shows total length of 1538 base pairs in linear DNA conformation (Raoult, 2016). *Lactobacillus crispatus* with different strains such as TCD56.8, 5-1-1, B4 and N-11 were found in pigeon faecal samples. *L. crispatus* bacteria are usually found in GI tract of human and chicken, carry out fermentation of substances like galactose, lactose, starch, sucrose and raffinose.

Moreover, *L. crispatus* can be found at vaginal area of human, playing role of maintaining and balancing normal microflora to prevent infections and maintain fertility. *L. crispatus* also involved in food industry to produce probiotics and β-galactosidases (Hahn, 2017).

One of the LAB, *Ligilactobacillus salivarius* has been produced as probiotics and proved to promote human's health by enhancing immune system and antimicrobial activity by regulating the microbiota in intestine (Raftis, et al., 2011; Guerrero Sanchez, et al., 2022). In this study, *Ligilactobacillus salivarius* strain IBB3154 was found in pigeon faecal samples. Based on information from NCBI, this bacterial strain has total size of 1, 921, 419 base pairs in circular DNA conformation and there is record of this isolated bacterial strain from hen faeces (Puzia, et al., 2022).

L. salivarius strain IBB3154 has strong adherence ability and high tolerance to acidic environment, which make it potentially to persist in chicken's GI tract to promote health benefits (Aleksandrzak-Piekarczyk, et al., 2019). Based on a study, *L. salivarius* aid in balancing the intestinal microflora by producing bacteriocins and hydrogen peroxide substances, which are antimicrobial. In the article, the authors further explained that some previous studies have proven the antimicrobial effects of avian *L. salivarius* strain towards pathogenic bacteria like *Salmonella enterica*, *Campylobacter* species and *Staphylococcus aureus* (Dec, et al., 2021). However, based on a study in Poland, several isolated lactobacilli from faecal samples of domesticated

pigeons such as *Lactobacillis ingluviei, L. salivarius* and *L. agilis* can establish drug resistance characteristic and act as a potential reservoir of transferring antibiotic resistance genes horizontally to other microbes in GI tract (Dec, et al., 2020). Therefore, in this study, although these LAB were found in the pigeon's faecal sample and generally classified as "good bacteria", however they may potentially act as reservoir of other microbes for developing antibiotic resistance characteristic. Further research is needed to deeply investigate the role and effect of these bacteria in pigeons' microbiota.

Candidatus Mycoplasma liparidae belongs to mycoplasma species from the *Mycoplasmataceae* family (National Center for Biotechnology Information, n.d.). In this study, *Candidatus* Mycoplasma liparidae clone H1 *16S ribosomal RNA* gene, partial sequence was found in pigeon faecal samples. This bacteria species also known as CML, named purposely due to its' novel lineage between *Mycoplasma* and *Ureaplasma* groups, was previously found in guts of deep sea hadal snailfish and vital for riboflavin biosynthesis. Based on the article, the authors further stated that CML can encode Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) system that can protect snailfishes from virus infection by preserving normal microbiota in gut (Lian, et al., 2020).

The uncultured bacterium clone rRNA288 *16S ribosomal RNA* gene, partial sequence that found in this study. Based on information from NCBI, this bacterial species was firstly isolated from human vaginal epithelium (Hyman, et al., 2005). In this study, this bacterial species known as "uncultured

bacterium clone", due to unavailability of the bacterial species to grow in laboratory. Therefore, the culture independent method used in this study allowed detection and identification of unculturable bacteria. Based on the BLAST result from NCBI, the identifiable bacterial species that has high similarity in percentage to the "uncultured bacterium clone rRNA288 *16S ribosomal RNA* gene, partial sequence" is "*Lactobacillus crispatus* strain 21s.P1 *16S ribosomal RNA* gene, partial sequence" with 99.76%. Culture dependent methods which involving microbiology and culturing bacterial cells on plates are time-consuming, high contamination rate and costly compared to culture independent method like molecular work. In addition, some bacteria need specific requirements to be cultured and growth, this led to declination of efficiency of bacterial identification, even some bacteria are dangerous and infectious to human, to be cultured in lab (Chauhan, 2020).

5.7 Future research work

In short, based on the presence of numerous *Lactobacillus* genus bacteria in this study, it can be hypothesized and assumed that pigeons inhabit at Block C of UTAR Kampar campus are in healthy status where many beneficial bacteria were identified and found in the pigeons' faecal samples. The normal microflora in the pigeons' GI tract aid in preventing the adhesion of pathogenic bacteria, thus preventing pigeons to become carrier of pathogenic bacteria. However, the sample size involved in this study is insufficient. Therefore, larger sample size is needed in future, in order to establish significant

data report. More robust methods like next-generation sequencing could be applied to broaden the sample size. More samples can be collected from different sampling sites of UTAR such as Block G and Block J that having cafeteria. Besides, regarding the bacteria antibiotics resistant issue, antibiotic assay could be used in screening of the culturable antibiotic resistant bacteria in pigeon faecal samples. Moreover, the future study will report on the relationship of interaction between pigeons and health of UTAR community in UTAR Kampar campus. Data will be submitted to UTAR administration to control pigeons' population in campus.

CHAPTER 6

CONCLUSION

The bacterial *16S rRNA* genes were amplified from the genomic DNA extracted from pigeon faecal samples. The amplified gene's sticky ends containing 3'dA overhangs were ligated successfully with linearized 3'dT overhangs of pTG19-T vector with the aid of T4 DNA Ligase. The plasmids were then heat-shock transformed into TOP10 *E. coli*. Positive clones were selected from LB agar plate containing ampicillin. Colony PCR was carried out to screen for the positive clones harboring the plasmid inserted with *16S rRNA* gene. Plasmid extraction was carried out to extract 10 samples for DNA sequencing analysis.

The DNA sequencing results were analysed using BLASTn program to identify the bacterial identities and accession numbers. There are mainly five types of bacteria species were identified such as *Lactobacillus crispatus, Lactobacillus ingluviei, Ligilactobacillus salivarius, Candidatus* Mycoplasma liparidae and uncultured bacterium clone. Mostly belonged to *Lactobacillus* genus. MEGA X software with pre-setted Maximum Likelihood method and Tamua-Nei model was used to construct phylogenetic tree, in order to study the evolutionary relationship between the different bacteria species.

In the nutshell, the bacterial diversity was studied by identifying the bacterial species and their taxonomic relationships. On the other hand, although the identified bacterial species in this study are considered as non-pathogenic, however the sample size of this study was limited and cannot take a part for the whole. More studies are required in future, in order to detect the possible pathogenic bacteria in pigeon faeces that may pose a health risk to UTAR community.

REFERENCES

Aleksandrzak-Piekarczyk, T. et al., 2019. Potential of Lactobacillus plantarum IBB3036 and Lactobacillus salivarius IBB3154 to persistence in chicken after in ovo delivery. *MicrobiologyOpen*, 8(1), p. e00620.

Animals Network Team, 2017. *Pigeon.* [Online] Available at: < https://animals.net/pigeon/> [Accessed 15 October 2022].

Anon., 2013. *Synthetic Biology for Igem - Transformations, Cloning.* [Online] Available at: <https://static.igem.org/mediawiki/2015/4/4e/IGEM4_Transformations.pdf> [Accessed 7 February 2023].

BioMed Central, 2010. *Harmful bacteria carried by pigeons.* [Online] Available at: <https://www.sciencedaily.com/releases/2010/06/100621215944.htm> [Accessed 23 October 2022].

Blanco-Pena, K., Esperon, F., Torres-Mejia, A. M., Torre, A. D. L. Cruz, E. D. L., and Jimenez-Soto, M., 2017. Antimicrobial Resistance Genes in Pigeons from Public Parks in Costa Rica. *Zoonoses Public Health*, 64(7), pp. e23-e30.

Bradburn, S., n.d. *8 Tips For Using The Nanodrop To Measure DNA/RNA.* [Online] Available at: <https://toptipbio.com/tips-using-nanodrop/#:~:text=Since%20the%20Nanodrop%20can%20use%20volumes%20as%20low,the%20pedestal%20arm%20down%20on%20the%20sample%20slowly.> [Accessed 6 February 2023].

Britannica, 2020. *Pigeon.* [Online] Available at: <https://www.britannica.com/animal/pigeon> [Accessed 13 October 2022].

Burt, S. A., Roring, R. E. and Heijne, M., 2018. Chlamydia psittaci and C. avium in feral pigeon (Columba livia domestica) droppings in two cities in the Netherlands. *Veterinary Quarterly,* 38(1), pp. 63-66.

Caprioli, A., Morabito, S., Brugère, H. and Oswald, E., 2005. Enterohaemorrhagic Escherichia coli: emerging issues on virulence and modes of transmission. *Veterinary research,* 36(3), pp. 289-311.

Centers for Disease Control and Prevention, 2019. *Campylobacter (Campylobacteriosis).* [Online] Available at: <https://www.cdc.gov/campylobacter/symptoms.html> [Accessed 23 October 2022].

Centers for Disease Control and Prevention, 2020. *Shigella - Shigellosis.* [Online] Available at: <https://www.cdc.gov/shigella/index.html> [Accessed 14 February 2023].

Centers for Disease Control and Prevention, 2022. *Psittacosis.* [Online] Available at: <https://www.cdc.gov/pneumonia/atypical/psittacosis/index.html> [Accessed 23 October 2022].

Central, B., 2010. *Harmful bacteria carried by pigeons.* [Online] Available at: <https://www.sciencedaily.com/releases/2010/06/100621215944.htm> [Accessed 15 October 2022].

Chang, A., Chau, Y., Landas, J. and Pang, Y., 2017. Preparation of calcium competent Escherichia coli and heat-shock transformation. *JEMI methods,* Volume 1, pp. 22-25.

Chauhan, T., 2018. *Phenol Chloroform DNA Extraction: Basics, Preparation Of Chemicals And Protocol.* [Online] Available at: <https://geneticeducation.co.in/phenol-chloroform-dna-extraction-basics-preparation-of-chemicals-and-protocol/> [Accessed 4 February 2023].

Chauhan, T., 2019. *What Is DNA Ligase? And How T4 DNA Ligase Works?*. [Online] Available at: <https://geneticeducation.co.in/what-is-dna-ligase-and-how-t4-dna-ligase-works/#T4_DNA_ligase> [Accessed 6 February 2023].

Chauhan, T., 2020. *16S RRNA: Gene, Sequencing And Importance.* [Online] Available at: <https://geneticeducation.co.in/16s-rrna-gene-sequencing-and-importance/> [Accessed 5 February 2023].

Conventry Pest Control, n.d. *Pigeon Waste Removal.* [Online] Available at: <https://www.coventrypestcontrol.co.uk/pigeon-dropping-removal/#:~:text=Bird%20faeces%20can%20lead%20to%20a%20series%20of,get%20rid%20of%20your%20bird%20dropping%20problem%20fast.> [Accessed 15 October 2022].

de Araújo, M., Lamberti, G. and Huber, L., 2015. Homogenization of Mammalian Cells. *Cold Spring Harb Protoc.*, Volume 11, pp. 1009-1012.

Dec, M., Stepien-Pysniak, D., Nowaczek, A., Puchalski, A., and Urban-Chmiel, R., 2020. Phenotypic and genotypic antimicrobial resistance profiles of fecal lactobacilli from domesticated pigeons in Poland. *Anaerobe,* Volume 65, p. 102251.

Dec, M., Stepien-Pysniak, D., Puchalski, A., Hauschild, T., Pietras-Ozga, D., Ignaciuk, S., and Urban-Chmiel, R., 2021. Biodiversity of Ligilactobacillus salivarius Strains from Poultry and Domestic Pigeons. *Animals (Basel),* 11(4), p. 972.

Food and Drug Administration, 2019. *Salmonella (Salmonellosis).* [Online] Available at: <https://www.fda.gov/food/foodborne-pathogens/salmonella-salmonellosis> [Accessed 23 October 2022].

Frank, J. A., Reich, C. I., Sharma, S., and Weisbaum, J. S., 2008. Critical Evaluation of Two Primers Commonly Used for Amplification of Bacterial 16S rRNA Genes. *Applied and Environmental Microbiology,* 74(8), pp. 2461-2470.

Gallik, S., 2013. *Experiment 1 - CellBiologyOLM.* [Online] Available at: <http://histologyolm.stevegallik.org/cellbiologyolm_fractionation.html> [Accessed 4 February 2023].

Guerrero Sanchez, M., Passot, S., Campoy, S., Olivares, M., and Fonseca, F., 2022. Ligilactobacillus salivarius functionalities, applications, and manufacturing challenges. *Applied Microbiology and Biotechnology,* Volume 106, pp. 57-80.

Gyimesi, Z. S., 2015. Columbiformes. In: R. E. Miller & M. E. Fowler, eds. *Fowler's Zoo and Wild Animal Medicine, Volume 8.* s.l.:Saunders, pp. 164-171.

Hahn, A., 2017. *Lacctobacillus crispatus.* [Online] Available at: <https://microbewiki.kenyon.edu/index.php/Lactobacillus_crispatus> [Accessed 12 February 2023].

Hammes, W. and Vogel, R., 1995. The genus Lactobacillus. In: B. Wood & W. Holzapfel, eds. *The Genera of Lactic Acid Bacteria. The Lactic Acid Bacteria, vol 2.*. Boston, MA: Springer, pp. 19-54.

Hanahan, D., 1983. Studies on Transformation of Escherichia coli with Plasmids. *Journal of Molecular Biology,* Volume 166, pp. 557-580.

Health and Safety Executive, n.d. *Construction micro-organisms: Psittacosis and other diseases from work involving bird droppings.* [Online] Available at: <https://www.hse.gov.uk/construction/healthrisks/hazardous-substances/harmful-micro-organisms/other-diseases.htm> [Accessed 15 October 2022].

Hyman, R. W., Fukushima, M., Diamond, L., Kumm, J., Giudice, L. C., and Davis, R. W., 2005. Microbes on the human vaginal epithelium. *Proceedings of the National Academy of Sciences of the United States of America,* 102(22), pp. 7952-7957.

James, G., 2010. Chapter 28 Universal Bacterial Identification by PCR and DNA Sequencing of 16S rRNA Gene. In: M. Schuller, et al. eds. *PCR for Clinical Microbiology.* Westmead: Springer, pp. 209-214.

Janda, J. and Abbott, S., 2007. 16S rRNA Gene Sequencing for Bacterial Identification in the Diagnostic Laboratory: Pluses, Perils, and Pitfalls. *Journal of Clinical Microbiology,* 45(9), pp. 2761-2764.

Johnston, R. F. and Janiga, M., 1995. *Feral Pigeons.* New York: Oxford University Press.

Karim, S. J. I., Islam, M., Sikder, T., Rubaya, R., Halder, J., and Alam, J., 2020. Multidrug-resistant Escherichia coli and Salmonella spp. isolated from pigeons. *Veterinary World,* 13(0), pp. 2156-2165.

Kimpe, A., Decostere, A., Martel, A., Haesebrouck, F., and Devriese, L. A., 2002. Prevalence of antimicrobial resistance among pigeon. *Avian Pathology,* 31(4), pp. 393-397.

Koh, X. P. et al., 2012. Characterization of Shigella sonnei in Malaysia, an increasingly prevalent etiologic agent of local shigellosis cases. *BMC Infectious DIseases,* 12(122).

Kumar, A., Tiwary, B. K., Kachhap, S., Nanda, A. K., Chakraborty, R., 2015. An Escherichia coli Strain, PGB01, Isolated from Feral Pigeon Faeces, Thermally Fit to Survive in Pigeon, Shows High Level Resistance to Trimethoprim. *PLoS ONE,* 10(3), p. e0119329.

Lane, D. J., Pace, B., Olsen, G. J., Stahl, D. A., Sogin, M. L., and Pace N. R., 1985. Rapid determination of 16S ribosomal RNA sequences for phylogenetic analyses. *Proc Natl Acad Sci U S A.,* 82(20), pp. 6955-6959.

Lian, C.-A., Yan, G.-Y., Huang, J.-M., Danchin, A., Wang, Y., and He, L.-S., 2020. Genomic Characterization of a Novel Gut Symbiont From the Hadal Snailfish. *Frontiers in Microbiology,* Volume 10, p. 2978.

Luca, G., Stefan, S. and David, H. L., 2004. Pet animals as reservoirs of antimicrobial-resistant bacteria: Review. *Journal of Antimicrobial Chemotherapy,* 54(2), pp. 321-332.

Maddocks, S. and Jenkins, R., 2017. Designing and Ordering Your Polymerase Chain Reaction Primers. In: *Understanding PCR.* s.l.:Academic Press, pp. 11-30.

Magino, S., Haag-Wackernagel, D., Geigenfeind, I., Helmecke, S., Dovc, A., Prukner-Radovcic, E., Residbegovic, E., Ilieski, V., Laroucau, K., Donati, M., Martinov, S., and Kaleta E. F., 2009. Chlamydial infections in feral pigeons in Europe: Review of data and focus on public health implications. *Veterinary Microbiology,* 135(1-2), pp. 54-67.

Masny, A. and Pluciennicz ak, A., 2003. Ligation mediated PCR performed at low denaturation temperatures - PCR melting profiles. *Nucleic Acids Research,* 31(18), p. e114.

MicropeMaster, n.d. *Bacterial Transformation - Gene Transfer, Heat Shock, Vs Transduction.* [Online] Available at: <https://www.microscopemaster.com/bacterial-transformation.html> [Accessed 6 February 2023].

National Center for Biotechnology Information, 2022. *Ligilactobacillus salivarius strain IBB3154 chromosome, complete genome.* [Online] Available at: <https://www.ncbi.nlm.nih.gov/nuccore/NZ_CP027644.1> [Accessed 11 February 2023].

National Center for Biotechnology Information, n.d. *Candidatus Mycoplasma liparidae.* [Online] Available at: <https://www.ncbi.nlm.nih.gov/data-hub/taxonomy/2698225/> [Accessed 12 February 2023].

Oswald, N., 2019. *What's The Problem With Ampicillin Selection?*. [Online] Available at: <https://bitesizebio.com/10188/whats-the-problem-with-ampicillin-selection/> [Accessed 7 February 2023].

Panja, S., Aich, P., Jana, B. and Basu, T., 2008. How does plasmid DNA penetrate cell membranes in artificial transformation process of Escherichia coli?. *Molecular Membrane Biology*, 25(5), pp. 411-422.

Pathogen Regulation Directorate, Public Health Agency of Canada, 2010. *Pathogen Safety Data Sheets: Infectious Substances – Shigella spp.*. [Online] Available at: <https://www.canada.ca/en/public-health/services/laboratory-biosafety-biosecurity/pathogen-safety-data-sheets-risk-assessment/shigella.html> [Accessed 14 February 2023].

Paton, A. and Paton, J., 1999. Direct Detection of Shiga Toxigenic Escherichia coli Strains Belonging to Serogroups O111, O157, and O113 by Multiplex PCR. *Journal of Clinical Microbiology*, 37(10), pp. 3362-3365.

Pigeon Control Resource Centre, 2009. *PIGEONS - EVERYTHING THERE IS TO KNOW ABOUT THE PIGEON*. [Online] Available at: <https://www.pigeoncontrolresourcecentre.org/html/about-pigeons.html> [Accessed 13 October 2022].

Science Learning Hub - Pokapū Akoranga Pūtaiao 2009. *DNA extraction.* [Online] Available at: <https://www.sciencelearn.org.nz/resources/2036-dna-extraction> [Accessed 4 February 2023].

Radimersky, T., Frolkova, P., Janoszowska, D., Dolejska, M., Svec, P., Roubalova, E., Cikova, P., Cizek, A., and Literak, I., 2010. Antibiotic resistance in faecal bacteria (Escherichia coli, Enterococcus spp.) in feral pigeons. *Journal of Applied Microbiology*, 109(5), pp. 1687-1695.

Raftis, E. J., Salvetti, E., Torriani, S., Felis, G. E., and O'Toole, P. W., 2011. Genomic diversity of Lactobacillus salivarius. *Applied and Environmental Microbiology*, 77(3), pp. 954-965.

Raoult, D., 2016. *Lactobacillus ingluviei partial 16S rRNA gene, strain Marseille-P209.* [Online] Available at: <https://www.ncbi.nlm.nih.gov/nuccore/LT223590.1?report=genbank&to=1538> [Accessed 12 February 2023].

Rossi, R., Montecucco, A., Ciarrocchi, G. and Biamonti, G., 1997. Functional characterization of the T4 DNA ligase: a new insight into the mechanism of action. *Nucleic Acids Research*, 25(11), pp. 2106-2113.

Salvetti, E., Torriani, S. and Felis, G., 2012. The Genus Lactobacillus: A Taxonomic Update. *Probiotics and Antimicrobial Proteins*, Volume 4, pp. 217-226.

Santos, H. M., Tsai, C.-Y., Catulin, G. E. M., Trangia, K. C. G., Tayo, L. L., Liu, H.-J., and Chuang, K. P., 2020. Common bacterial, viral, and parasitic diseases in pigeons (Columba livia): A review of diagnostic and treatment strategies. *Veterinary Microbiology*, Volume 247, p. 108779.

Sato, G., Oka, C., Asagi, M. and Ishiguro, N., 1978. Detection of conjugative R plasmids conferring chloramphenicol resistance in Escherichia coli isolated from domestic and feral pigeons and crows.. *Zentralblatt fur Bakteriologie, Parasitenkunde, Infektionskrankheiten und Hygiene. Erste Abteilung Originale. Reihe A: Medizinische Mikrobiologie und Parasitologie*, 241(4), pp. 407-417.

Silva, V. L., Nicoli, J. R., Nascimento, T. C. and Diniz, C. G., 2009. Diarrheagenic Escherichia coli strains recovered from urban pigeons (Columba livia) in Brazil and their antimicrobial susceptibility patterns.. *Curr Microbiol.*, 59(3), pp. 302-308.

Singh, H., 2021. *BACTERIAL DIVERSITY IN PIGEON FAECAL SAMPLES*, s.l.: UTAR Institutional Repository.

Thermo Scientific NanoDrop Lite, 2012. *Interpretation of Nucleic Acid 260/280 Ratios.* [Online] Available at: <https://tools.thermofisher.com/content/sfs/brochures/T123-NanoDrop-Lite-Interpretation-of-Nucleic-Acid-260-280-Ratios.pdf> [Accessed 4 February 2023].

Thermo Scientific NanoDrop Spectrophotometers, n.d. *260/280 and 260/230 Ratios.* [Online] Available at: <https://dna.uga.edu/wp-content/uploads/sites/51/2019/02/Note-on-the-260_280-and-260_230-Ratios.pdf> [Accessed 4 February 2023].

Vasconcelos, R. H., Teixeira, R., Silva, I., Lopes, E., and Maciel W., 2018. Feral pigeons (Columba livia) as potential reservoirs of Salmonella sp. and Escherichia coli. *Arq. Inst. Biol. (Sao. Paulo)*, 85(0).

Vasconcellos, H. V. G. d.., Silva, K. F. B., Montenegro, H., Miguel, C. B., Tizioto, P., Agostinho, F., Araujo, M. C., Ribas, R. M., Silva, M. V. d., Castro Soares, S. d., Junior, V. R., Fonseca Batistao, D. W. d., Oliveira, C. J. F., Rodrigues, W. F., 2022. Staphylococcus aureus and Enterococcus faecium isolated from pigeon droppings (Columba livia) in the external environment close to hospitals. *Revista da Sociedade Brasileira de Medicina Tropical,* Volume 55, pp. e0353-2021.

Vasquez, B., Esperon, F., Neves, E., Lopez, J., Ballesteros, C., and Munoz, M. J., 2010. Screening for several potential pathogens in feral pigeons (Columba livia) in Madrid. *Acta Veterinaria Sccandinavica,* 52(1), p. 45.

Vivantis technologies Malaysia., n.d. *pTG19-T PCR Cloning Vector.* [Online] Available at: <https://www.vivantechnologies.com/index.php?option=com_content&view=article&id=1160&Itemid=110> [Accessed 6 February 2023].

Vivantis technologies Malaysia, n.d. *GF-1 Soil Sample DNA Extraction Kit.* [Online] Available at: <https://www.vivantechnologies.com/index.php?option=com_content&view=

article&id=865:gf-1-soil-sample-dna-extraction-kit&Itemid=44> [Accessed 4 February 2023].

Whitehouse, C. A., Zhao, S. and Tate, H., 2018. Antimicrobial Resistance in Campylobacter Species: Mechanisms and Genomic Epidemiology. In: S. Sariaslani & G. M. Gadd, eds. *Advances in Applied Microbiology.* s.l.:Academic Press, pp. 1-47.

Yau, Y., 2017. *Negative Results on Nanodrop..* [Online] Available at: <https://www.researchgate.net/post/Negative_Results_on_Nanodrop/59480465615e27abd6659c27/citation/download.> [Accessed 6 February 2023].

Yourgenome Organization, 2021. *What is PCR (polymerase chain reaction)?.* [Online] Available at: <https://www.yourgenome.org/facts/what-is-pcr-polymerase-chain-reaction/#:~:text=Annealing%20stage&text=This%20enables%20the%20primers%20to,to%2030%20bases%20in%20length.> [Accessed 9 February 2023].

APPENDICES

C1

Uncultured bacterium clone rRNA288 16S ribosomal RNA gene, partial sequence

Sequence ID: AY959061.1 Length: 1514 Number of Matches: 1

Range 1: 675 to 1504 GenBank Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
1528 bits(827)	0.0	829/830(99%)	0/830(0%)	Plus/Minus

```
Query  1     CGGTTACCTTGTTACGACTTCACCCCAGTCATCTGCCCTGCCTTAGACGGCTCCTTCCCG  60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1504  CGGTTACCTTGTTACGACTTCACCCCAGTCATCTGCCCTGCCTTAGACGGCTCCTTCCCG  1445

Query  61    AAGGTTAGGCCACCGGCTTTGGGCATTGCAGACTCCCATGGTGTGACGGGCGGTGTGTAC  120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1444  AAGGTTAGGCCACCGGCTTTGGGCATTGCAGACTCCCATGGTGTGACGGGCGGTGTGTAC  1385

Query  121   AAGGCCCGGGAACGTATTCACCGCGGCGTGCTGATCCGCGATTACTAGCGATTCCAGCTT  180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1384  AAGGCCCGGGAACGTATTCACCGCGGCGTGCTGATCCGCGATTACTAGCGATTCCAGCTT  1325

Query  181   CGTGCAGTCGAGTTGCAGACTGCAGTCCGAACTGAGAACAGCTTTCAGAGATTCGCTTGC  240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1324  CGTGCAGTCGAGTTGCAGACTGCAGTCCGAACTGAGAACAGCTTTCAGAGATTCGCTTGC  1265

Query  241   CTTCGCAGGCTCGCTTCTCGTTGTACTGCCCATTGTAGCACGTGTGTAGCCCAGGTCATA  300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1264  CTTCGCAGGCTCGCTTCTCGTTGTACTGCCCATTGTAGCACGTGTGTAGCCCAGGTCATA  1205

Query  301   AGGGGCATGATGACTTGACGTCATCCCCACCTTCCTCCGGCTTGTCACCGGCAGTCTCAT  360
             ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Sbjct  1204  AGGGGCATGATGACTTGACGTCATCCCCACCTTCCTCCGGTTTGTCACCGGCAGTCTCAT  1145

Query  361   TAGAGTGCCCAACTTAATGCTGGCAACTAATAACAAGGGTTGCGCTCGTTGCGGGACTTA  420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1144  TAGAGTGCCCAACTTAATGCTGGCAACTAATAACAAGGGTTGCGCTCGTTGCGGGACTTA  1085

Query  421   ACCCAACATCTCACGACACGAGCTGACGACAGCCATGCACCACCTGTCTTAGCGTCCCCG  480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1084  ACCCAACATCTCACGACACGAGCTGACGACAGCCATGCACCACCTGTCTTAGCGTCCCCG  1025

Query  481   AAGGGAACTTTGTATCTCTACAAATGGCACTAGATGTCAAGACCTGGTAAGGTTCTTCGC  540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1024  AAGGGAACTTTGTATCTCTACAAATGGCACTAGATGTCAAGACCTGGTAAGGTTCTTCGC  965

Query  541   GTTGCTTCGAATTAAACCACATGCTCCACCGCTTGTGCGGGCCCCCGTCAATTCCTTTGA  600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  964   GTTGCTTCGAATTAAACCACATGCTCCACCGCTTGTGCGGGCCCCCGTCAATTCCTTTGA  905

Query  601   GTTTCAACCTTGCGGTCGTACTCCCCAGGCGGAGTGCTTAATGCGTTAGCTGCAGCACTG  660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  904   GTTTCAACCTTGCGGTCGTACTCCCCAGGCGGAGTGCTTAATGCGTTAGCTGCAGCACTG  845

Query  661   AGAGGCGGAAACCTCCCAACACTTAGCACTCATCGTTTACGGCATGGACTACCAGGGTAT  720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  844   AGAGGCGGAAACCTCCCAACACTTAGCACTCATCGTTTACGGCATGGACTACCAGGGTAT  785

Query  721   CTAATCCTGTTCGCTACCCATGCTTTCGAGCCTCAGCGTCAGTTGCAGACCAGAGAGCCG  780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  784   CTAATCCTGTTCGCTACCCATGCTTTCGAGCCTCAGCGTCAGTTGCAGACCAGAGAGCCG  725

Query  781   CCTTCGCCACTGGTGTTCTTCCATATATCTACGCATTCCACCGCTACACA  830
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  724   CCTTCGCCACTGGTGTTCTTCCATATATCTACGCATTCCACCGCTACACA  675
```

C2

Lactobacillus crispatus strain TCD56.8 16S ribosomal RNA gene, partial sequence

Sequence ID: KU851161.1 Length: 955 Number of Matches: 1

Range 1: 1 to 637 GenBank Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
1177 bits(637)	0.0	637/637(100%)	0/637(0%)	Plus/Plus

```
Query  1    AGAGTTTGATCATGGCTCAGGACGAACGCTGGCGGCGTGCCTAATACATGCAAGTCGAGC  60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1    AGAGTTTGATCATGGCTCAGGACGAACGCTGGCGGCGTGCCTAATACATGCAAGTCGAGC  60

Query  61   GAGCGGAACTAACAGATTTACTTCGGTAATGACGTTAGGAAAGCGAGCGGCGGATGGGTG  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  61   GAGCGGAACTAACAGATTTACTTCGGTAATGACGTTAGGAAAGCGAGCGGCGGATGGGTG  120

Query  121  AGTAACACGTGGGGAACCTGCCCCATAGTCTGGGATACCACTTGGAAACAGGTGCTAATA  180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  121  AGTAACACGTGGGGAACCTGCCCCATAGTCTGGGATACCACTTGGAAACAGGTGCTAATA  180

Query  181  CCGGATAAGAAAGCAGATCGCATGATCAGCTTTTAAAAGGCGGCGTAAGCTGTCGCTATG  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  181  CCGGATAAGAAAGCAGATCGCATGATCAGCTTTTAAAAGGCGGCGTAAGCTGTCGCTATG  240

Query  241  GGATGGCCCCGCGGTGCATTAGCTAGTTGGTAAGGTAAAGGCTTACCAAGGCGATGATGC  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  241  GGATGGCCCCGCGGTGCATTAGCTAGTTGGTAAGGTAAAGGCTTACCAAGGCGATGATGC  300

Query  301  ATAGCCGAGTTGAGAGACTGATCGGCCACATTGGGACTGAGACACGGCCCAAACTCCTAC  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  301  ATAGCCGAGTTGAGAGACTGATCGGCCACATTGGGACTGAGACACGGCCCAAACTCCTAC  360

Query  361  GGGAGGCAGCAGTAGGGAATCTTCCACAATGGACGCAAGTCTGATGGAGCAACGCCGCGT  420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  361  GGGAGGCAGCAGTAGGGAATCTTCCACAATGGACGCAAGTCTGATGGAGCAACGCCGCGT  420

Query  421  GAGTGAAGAAGGTTTTCGGATCGTAAAGCTCTGTTGTTGGTGAAGAAGGATAGAGGTAGT  480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  421  GAGTGAAGAAGGTTTTCGGATCGTAAAGCTCTGTTGTTGGTGAAGAAGGATAGAGGTAGT  480

Query  481  AACTGGCCTTTATTTGACGGTAATCAACCAGAAAGTCACGGCTAACTACGTGCCAGCAGC  540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  481  AACTGGCCTTTATTTGACGGTAATCAACCAGAAAGTCACGGCTAACTACGTGCCAGCAGC  540

Query  541  CGCGGTAATACGTAGGTGGCAAGCGTTGTCCGGATTTATTGGGCGTAAAGCGAGCGCAGG  600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  541  CGCGGTAATACGTAGGTGGCAAGCGTTGTCCGGATTTATTGGGCGTAAAGCGAGCGCAGG  600

Query  601  CGGAAGAATAAGTCTGATGTGAAAGCCCTCGGCTTAA  637
            ||||||||||||||||||||||||||||||||||||
Sbjct  601  CGGAAGAATAAGTCTGATGTGAAAGCCCTCGGCTTAA  637
```

C3

Lactobacillus crispatus strain 5-1-1 16S ribosomal RNA gene, partial sequence

Sequence ID: KU991819.1 Length: 1528 Number of Matches: 1

Range 1: 600 to 1526 GenBank Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
1712 bits(927)	0.0	927/927(100%)	0/927(0%)	Plus/Minus

```
Query  1     CGGTTACCTTGTTACGACTTCACCCCAGTCATCTGCCCTGCCTTAGACGGCTCCTTCCCG  60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1526  CGGTTACCTTGTTACGACTTCACCCCAGTCATCTGCCCTGCCTTAGACGGCTCCTTCCCG  1467

Query  61    AAGGTTAGGCCACCGGCTTTGGGCATTGCAGACTCCCATGGTGTGACGGGCGGTGTGTAC  120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1466  AAGGTTAGGCCACCGGCTTTGGGCATTGCAGACTCCCATGGTGTGACGGGCGGTGTGTAC  1407

Query  121   AAGGCCCGGGAACGTATTCACCGCGGCGTGCTGATCCGCGATTACTAGCGATTCCAGCCT  180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1406  AAGGCCCGGGAACGTATTCACCGCGGCGTGCTGATCCGCGATTACTAGCGATTCCAGCCT  1347

Query  181   CGTGCAGTCGAGTTGCAGACTGCAGTCCGAACTGAGAACAGCTTTCAGAGATTCGCTTGC  240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1346  CGTGCAGTCGAGTTGCAGACTGCAGTCCGAACTGAGAACAGCTTTCAGAGATTCGCTTGC  1287

Query  241   CTTCGCAGGCTCGCTTCTCGTTGTACTGCCCATTGTAGCACGTGTGTAGCCCAGGTCATA  300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1286  CTTCGCAGGCTCGCTTCTCGTTGTACTGCCCATTGTAGCACGTGTGTAGCCCAGGTCATA  1227

Query  301   AGGGGCATGATGACTTGACGTCATCCCCACCTTCCTCCGGTTTGTCACCGGCAGTCTCAT  360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1226  AGGGGCATGATGACTTGACGTCATCCCCACCTTCCTCCGGTTTGTCACCGGCAGTCTCAT  1167

Query  361   TAGAGTGCCCAACTTAATGCTGGCAACTAATAACAAGGGTTGCGCTCGTTGCGGGACTTA  420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1166  TAGAGTGCCCAACTTAATGCTGGCAACTAATAACAAGGGTTGCGCTCGTTGCGGGACTTA  1107

Query  421   ACCCAACATCTCACGACACGAGCTGACGACAGCCATGCACCACCTGTCTTAGCGTCCCCG  480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1106  ACCCAACATCTCACGACACGAGCTGACGACAGCCATGCACCACCTGTCTTAGCGTCCCCG  1047

Query  481   AAGGGAACTTTGTATCTCTACAAATGGCACTAGATGTCAAGACCTGGTAAGGTTCTTCGC  540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1046  AAGGGAACTTTGTATCTCTACAAATGGCACTAGATGTCAAGACCTGGTAAGGTTCTTCGC  987

Query  541   GTTGCTTCGAATTAAACCACATGCTCCACCGCTTGTGCGGGCCCCCGTCAATTCCTTTGA  600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  986   GTTGCTTCGAATTAAACCACATGCTCCACCGCTTGTGCGGGCCCCCGTCAATTCCTTTGA  927

Query  601   GTTTCAACCTTGCGGTCGTACTCCCCAGGCGGAGTGCTTAATGCGTTAGCTGCAGCACTG  660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  926   GTTTCAACCTTGCGGTCGTACTCCCCAGGCGGAGTGCTTAATGCGTTAGCTGCAGCACTG  867

Query  661   AGAGGCGGAAACCTCCCAACACTTAGCACTCATCGTTTACGGCATGGACTACCAGGGTAT  720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  866   AGAGGCGGAAACCTCCCAACACTTAGCACTCATCGTTTACGGCATGGACTACCAGGGTAT  807

Query  721   CTAATCCTGTTCGCTACCCATGCTTTCGAGCCTCAGCGTCAGTTGCAGACCAGAGAGCCG  780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  806   CTAATCCTGTTCGCTACCCATGCTTTCGAGCCTCAGCGTCAGTTGCAGACCAGAGAGCCG  747

Query  781   CCTTCGCCACTGGTGTTCTTCCATATATCTACGCATTCCACCGCTACACATGGAGTTCCA  840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  746   CCTTCGCCACTGGTGTTCTTCCATATATCTACGCATTCCACCGCTACACATGGAGTTCCA  687

Query  841   CTCTCCTCTTCTGCACTCAAGAAAAACAGTTTCCGATGCAGTTCCTCGGTTAAGCCGAGG  900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  686   CTCTCCTCTTCTGCACTCAAGAAAAACAGTTTCCGATGCAGTTCCTCGGTTAAGCCGAGG  627

Query  901   GCTTTCACATCAGACTTATTCTTCCGC  927
             |||||||||||||||||||||||||||
Sbjct  626   GCTTTCACATCAGACTTATTCTTCCGC  600
```

C4

Ligilactobacillus salivarius strain IBB3154 chromosome, complete genome

Sequence ID: CP027644.1 Length: 1921419 Number of Matches: 7

Range 1: 123746 to 124474 GenBank Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
1330 bits(720)	0.0	726/729(99%)	0/729(0%)	Plus/Plus

```
Query  1       AGAGTTTGATCCTGGCTCAGGACGAACGCTGGCGGCGTGCCTAATACATGCAAGTCGAAC  60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  123746  AGAGTTTGATCCTGGCTCAGGACGAACGCTGGCGGCGTGCCTAATACATGCAAGTCGAAC  123805

Query  61      GAAACTTTCTTACACCGAATGCTTGCATTCACCGTAAGAAGTTGAGTGGCGGACGGGTGA  120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  123806  GAAACTTTCTTACACCGAATGCTTGCATTCACCGTAAGAAGTTGAGTGGCGGACGGGTGA  123865

Query  121     GTAACACGTGGGTAACCTACCTAAAAGAAGGGGATAACACTTGGAAACAGGTGCTAATAC  180
               |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Sbjct  123866  GTAACACGTGGGTAACCTGCCTAAAAGAAGGGGATAACACTTGGAAACAGGTGCTAATAC  123925

Query  181     CGTATATCTCTAAGGATCGCATGATCCTTAGATGAAAGATGGTTCTGCTATCGCTTTTAG  240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  123926  CGTATATCTCTAAGGATCGCATGATCCTTAGATGAAAGATGGTTCTGCTATCGCTTTTAG  123985

Query  241     ATGGACCCGCGGCGTATTAACTAGTTGGTGGGGTAACGGCCTACCAAGGTGATGATACGT  300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  123986  ATGGACCCGCGGCGTATTAACTAGTTGGTGGGGTAACGGCCTACCAAGGTGATGATACGT  124045

Query  301     AGCCGAACTGAGAGGTTGATCGGCCACATTGGGACTGAGACACGGCCCAAACTCCTACGG  360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  124046  AGCCGAACTGAGAGGTTGATCGGCCACATTGGGACTGAGACACGGCCCAAACTCCTACGG  124105

Query  361     GAGGCAGCAGTAGGGAATCTTCCACAATGGACGCAAGTCTGATGGAGCAACGCCGCGTGA  420
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  124106  GAGGCAGCAGTAGGGAATCTTCCACAATGGACGCAAGTCTGATGGAGCAACGCCGCGTGA  124165

Query  421     GTGAAGAAGTTCTTCGGATCGTAAAACTCTGTTGTTAGAGAAGAACACGAGTGAGAGTAA  480
               ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  124166  GTGAAGAAGGTCTTCGGATCGTAAAACTCTGTTGTTAGAGAAGAACACGAGTGAGAGTAA  124225

Query  481     CTGTTCATTCGATGACGGTATCTAACCAGCAAGTCACGGCTAACTACGTGCCAGCAGCCG  540
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  124226  CTGTTCATTCGATGACGGTATCTAACCAGCAAGTCACGGCTAACTACGTGCCAGCAGCCG  124285

Query  541     CGGTAATACGTAGGTGGCAAGCGTTGTCCGGATTTATTGGGCGTAAAGGGAACGCAGGCG  600
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  124286  CGGTAATACGTAGGTGGCAAGCGTTGTCCGGATTTATTGGGCGTAAAGGGAACGCAGGCG  124345

Query  601     GTCTTTTAAGTCTGATGTGAAAGCCTTCGGCTTAACCGGAGTAGTGCATTGGAAACTGAA  660
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Sbjct  124346  GTCTTTTAAGTCTGATGTGAAAGCCTTCGGCTTAACCGGAGTAGTGCATTGGAAACTGGA  124405

Query  661     AGACTTGAGTGCAGAAGAGGAGAGTGGAACTCCATGTGTAGCGGTGAAATGCGTAGATAT  720
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  124406  AGACTTGAGTGCAGAAGAGGAGAGTGGAACTCCATGTGTAGCGGTGAAATGCGTAGATAT  124465

Query  721     ATGGAAGAA  729
               |||||||||
Sbjct  124466  ATGGAAGAA  124474
```

C5

Lactobacillus crispatus strain TCD56.8 16S ribosomal RNA gene, partial sequence

Sequence ID: KU851161.1 Length: 955 Number of Matches: 1

Range 1: 1 to 772 GenBank Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
1426 bits(772)	0.0	772/772(100%)	0/772(0%)	Plus/Plus

```
Query  1    AGAGTTTGATCATGGCTCAGGACGAACGCTGGCGGCGTGCCTAATACATGCAAGTCGAGC  60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1    AGAGTTTGATCATGGCTCAGGACGAACGCTGGCGGCGTGCCTAATACATGCAAGTCGAGC  60

Query  61   GAGCGGAACTAACAGATTTACTTCGGTAATGACGTTAGGAAAGCGAGCGGCGGATGGGTG  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  61   GAGCGGAACTAACAGATTTACTTCGGTAATGACGTTAGGAAAGCGAGCGGCGGATGGGTG  120

Query  121  AGTAACACGTGGGGAACCTGCCCCATAGTCTGGGATACCACTTGGAAACAGGTGCTAATA  180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  121  AGTAACACGTGGGGAACCTGCCCCATAGTCTGGGATACCACTTGGAAACAGGTGCTAATA  180

Query  181  CCGGATAAGAAAGCAGATCGCATGATCAGCTTTTAAAAGGCGGCGTAAGCTGTCGCTATG  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  181  CCGGATAAGAAAGCAGATCGCATGATCAGCTTTTAAAAGGCGGCGTAAGCTGTCGCTATG  240

Query  241  GGATGGCCCCGCGGTGCATTAGCTAGTTGGTAAGGTAAAGGCTTACCAAGGCGATGATGC  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  241  GGATGGCCCCGCGGTGCATTAGCTAGTTGGTAAGGTAAAGGCTTACCAAGGCGATGATGC  300

Query  301  ATAGCCGAGTTGAGAGACTGATCGGCCACATTGGGACTGAGACACGGCCCAAACTCCTAC  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  301  ATAGCCGAGTTGAGAGACTGATCGGCCACATTGGGACTGAGACACGGCCCAAACTCCTAC  360

Query  361  GGGAGGCAGCAGTAGGGAATCTTCCACAATGGACGCAAGTCTGATGGAGCAACGCCGCGT  420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  361  GGGAGGCAGCAGTAGGGAATCTTCCACAATGGACGCAAGTCTGATGGAGCAACGCCGCGT  420

Query  421  GAGTGAAGAAGGTTTTCGGATCGTAAAGCTCTGTTGTTGGTGAAGAAGGATAGAGGTAGT  480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  421  GAGTGAAGAAGGTTTTCGGATCGTAAAGCTCTGTTGTTGGTGAAGAAGGATAGAGGTAGT  480

Query  481  AACTGGCCTTTATTTGACGGTAATCAACCAGAAAGTCACGGCTAACTACGTGCCAGCAGC  540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  481  AACTGGCCTTTATTTGACGGTAATCAACCAGAAAGTCACGGCTAACTACGTGCCAGCAGC  540

Query  541  CGCGGTAATACGTAGGTGGCAAGCGTTGTCCGGATTTATTGGGCGTAAAGCGAGCGCAGG  600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  541  CGCGGTAATACGTAGGTGGCAAGCGTTGTCCGGATTTATTGGGCGTAAAGCGAGCGCAGG  600

Query  601  CGGAAGAATAAGTCTGATGTGAAAGCCCTCGGCTTAACCGAGGAACTGCATCGGAAACTG  660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  601  CGGAAGAATAAGTCTGATGTGAAAGCCCTCGGCTTAACCGAGGAACTGCATCGGAAACTG  660

Query  661  TTTTTCTTGAGTGCAGAAGAGGAGAGTGGAACTCCATGTGTAGCGGTGGAATGCGTAGAT  720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  661  TTTTTCTTGAGTGCAGAAGAGGAGAGTGGAACTCCATGTGTAGCGGTGGAATGCGTAGAT  720

Query  721  ATATGGAAGAACACCAGTGGCGAAGGCGGCTCTCTGGTCTGCAACTGACGCT  772
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  721  ATATGGAAGAACACCAGTGGCGAAGGCGGCTCTCTGGTCTGCAACTGACGCT  772
```

C6

Candidatus Mycoplasma liparidae clone H1 16S ribosomal RNA gene, partial sequence

Sequence ID: MK713651.1 Length: 1473 Number of Matches: 1

Range 1: 591 to 1471 GenBank Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
1229 bits(665)	0.0	811/883(92%)	3/883(0%)	Plus/Minus

```
Query  1     CGGTTACCTTGTTACGACTTAACTCCAGTTACCAATCCTACCCTAGACGTATGCCTCCGA  60
             ||| |||||||||||||||||||||| ||| | ||||||||||||||||||||||||||||
Sbjct  1471  CGGCTACCTTGTTACGACTTAACTCTAGTCATCAATCCTACCCTAGACGTATGCCTCCGA  1412

Query  61    GGTTAGCGATACGGTTTTAGATATTATCGGTTTCCATAGTTCGACGGGCGGTGTGTACAA  120
              |||||||||||||||| |||||||| |   | |||||||| ||||||||||||||||||||
Sbjct  1411  AGTTAGCGATACGGTTTCAGATATTACCAACTCCCATAGTTTGACGGGCGGTGTGTACAA  1352

Query  121   GACCTGGGAACGTATTCACCGCGACATGGCTGATTCGCGATTACTAGTGATTCCAACTTC  180
             ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Sbjct  1351  GACCTGGGAACGTATTCACCGCGACATGTCTGATTCGCGATTACTAGTGATTCCAACTTC  1292

Query  181   AAGAGGACGAATTGCAGCCCTCTATCCGAACTGAGATCGGCTTTTTGTGATTTGCATCTT  240
             | |||| ||| |||||||||||| ||||||||||||||||||||  ||||||||| |||
Sbjct  1291  ATGAGGGCGAGTTGCAGCCCTCAATCCGAACTGAGATCGGCTTTAAGTGATTTGCGTCTC  1232

Query  241   ATTACTAAGTAGCGACACTTTGTACCGACCATTGTAGCACGTTTGCGGCCCTAGATATAA  300
              |||| ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Sbjct  1231  CTTACGAAGTAGCGACACTTTGTACCGACCATTGTAGCACGTTTGCAGCCCTAGATATAA  1172

Query  301   GGGGCATGATGATTTGACGTCGTCCCCACCTTCCTCTGCCTTGCAGCAGCAGTCTTGTTA  360
             |||||||||||||||||||| |||||||||||||||| |||||  ||||||||| ||||
Sbjct  1171  GGGGCATGATGATTTGACGTCATCCCCACCTTCCTCTGTCTTGCGACAGCAGTCTCGTTA  1112

Query  361   GATAAGGTAACTAACAATAGGGGTTGCGCTCGTTGCGGGACTTAACCCAACATCTCACGA  420
             ||||| ||||||||| | | ||||||||||||||||||||||||||||||||||||||||||
Sbjct  1111  GATAAAGTAACTAACGACAAGGGTTGCGCTCGTTGCGGGACTTAACCCAACATCTCACGA  1052

Query  421   CACGAGCTGACGACAACCATGCACCACCTGTCACTCGGTTAACCTCCACTATATTTCTAT  480
             |||||||||||||||||||||||||||||||| | | |||||||| |  |||| |||||
Sbjct  1051  CACGAGCTGACGACAACCATGCACCACCTGTCACCCTGATAACCTCTATCATATCTCTAT  992

Query  481   AGCTTTGCCAAGGATGTCAAATCTAGGTAAGGTTTTACGTGTATTGTCAAATTAAGCAAC  540
               ||||||   |||||||||| ||||||||||||||||||||||||||||||||||||||
Sbjct  991   GACTTTGCAGGGGATGTCAAACCTAGGTAAGGTTTTACGTGTATTGTCAAATTAAGCAAC  932

Query  541   ATGCTCCACCACTTGTGCAGGTCCCCGTCAATTCCGTTTGAGTTTCATTCTTGCGAATGT  600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  931   ATGCTCCACCACTTGTGCAGGTCCCCGTCAATTCCGTTTGAGTTTCATTCTTGCGAATGT  872

Query  601   ACTACCCAGGCAGGTTATTTAATGCGTTAGCTGCAACACCAACACTTACGTGCTGACATT  660
             |||||||||| || |||||||||||||||||||||||||| |||| | ||||| || |||
Sbjct  871   ACTACCCAGGTAGATTATTTAATGCGTTAGCTGCAACACCGACACAT-CGTGCCGATATT  813

Query  661   TAATAACCATCGTTTACGGTGTGGACTACTAGGGTATCTAATCCTATTTGCTCCCCACAC  720
             |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  812   TAATAATCATCGTTTACGGTGTGGACTACTAGGGTATCTAATCCTATTTGCTCCCCACAC  753

Query  721   TTTCGAGCCTAAGCGTCAGTAATAGTCCAAGTAATCGCCTTCGCCTCTAGTGTTCTTCCA  780
             |||||||||||||||||| |||||||| |||| |||||||||||| ||||||||||||||
Sbjct  752   TTTCGAGCCTAAGCGTCAATAATAGTCTAAGTTATCGCCTTCGCCACTAGTGTTCTTCCA  693

Query  781   TATATCTACGCATTCCACTGCTCCACATGGAGTTCCATTACTCTCTACTACACTCTAGA-  839
             |||||||||||||||||| |||||||||||||||||||| |||||||||||| | ||||||
Sbjct  692   TATATCTACGCATTCCACCGCTCCACATGGAGTTCCATAACTCTCTACTATATTCTAGAC  633

Query  840   TTAGCAGTTTCCAATGCATACAATAGTTAAGCTACTGCATTTA  882
             |||  |||||||||||||  || | ||| ||| |||||| ||||
Sbjct  632   TTAT-AGTTTCCAATGCAATCAGTTGTTGAGCAACTGCCTTTA  591
```

C7

Lactobacillus crispatus strain B4 chromosome, complete genome

Sequence ID: CP059140.1 Length: 2039590 Number of Matches: 5

Range 1: 82112 to 82964 GenBank Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
1565 bits(847)	0.0	851/853(99%)	0/853(0%)	Plus/Minus

```
Query  1      CGGTTACCTTGTTACGACTTCACCCCAGTCATCTGCCCTGCCTTAGACGGCTCCTTCCCG  60
              ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  82964  CGGCTACCTTGTTACGACTTCACCCCAGTCATCTGCCCTGCCTTAGACGGCTCCTTCCCG  82905

Query  61     AAGGTTAGGCCACCGGCTTTGGGCATTGCAGACTCCCATGGTGTGACGGGCGGTGTGTAC  120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  82904  AAGGTTAGGCCACCGGCTTTGGGCATTGCAGACTCCCATGGTGTGACGGGCGGTGTGTAC  82845

Query  121    AAGGCCCGGGAACGTATTCGCCGCGGCGTGCTGATCCGCGATTACTAGCGATTCCAGCTT  180
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Sbjct  82844  AAGGCCCGGGAACGTATTCACCGCGGCGTGCTGATCCGCGATTACTAGCGATTCCAGCTT  82785

Query  181    CGTGCAGTCGAGTTGCAGACTGCAGTCCGAACTGAGAACAGCTTTCAGAGATTCGCTTGC  240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  82784  CGTGCAGTCGAGTTGCAGACTGCAGTCCGAACTGAGAACAGCTTTCAGAGATTCGCTTGC  82725

Query  241    CTTCGCAGGCTCGCTTCTCGTTGTACTGCCCATTGTAGCACGTGTGTAGCCCAGGTCATA  300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  82724  CTTCGCAGGCTCGCTTCTCGTTGTACTGCCCATTGTAGCACGTGTGTAGCCCAGGTCATA  82665

Query  301    AGGGGCATGATGACTTGACGTCATCCCCACCTTCCTCCGGTTTGTCACCGGCAGTCTCAT  360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  82664  AGGGGCATGATGACTTGACGTCATCCCCACCTTCCTCCGGTTTGTCACCGGCAGTCTCAT  82605

Query  361    TAGAGTGCCCAACTTAATGCTGGCAACTAATAACAAGGGTTGCGCTCGTTGCGGGACTTA  420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  82604  TAGAGTGCCCAACTTAATGCTGGCAACTAATAACAAGGGTTGCGCTCGTTGCGGGACTTA  82545

Query  421    ACCCAACATCTCACGACACGAGCTGACGACAGCCATGCACCACCTGTCTTAGCGTCCCCG  480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  82544  ACCCAACATCTCACGACACGAGCTGACGACAGCCATGCACCACCTGTCTTAGCGTCCCCG  82485

Query  481    AAGGGAACTTTGTATCTCTACAAATGGCACTAGATGTCAAGACCTGGTAAGGTTCTTCGC  540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  82484  AAGGGAACTTTGTATCTCTACAAATGGCACTAGATGTCAAGACCTGGTAAGGTTCTTCGC  82425

Query  541    GTTGCGTCGAATTAAACCACATGCTCCACCGCTTGTGCGGGCCCCCGTCAATTCCTTTGA  600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  82424  GTTGCGTCGAATTAAACCACATGCTCCACCGCTTGTGCGGGCCCCCGTCAATTCCTTTGA  82365

Query  601    GTTTCAACCTTGCGGTCGTACTCCCCAGGCGGAGTGCTTAATGCGTTAGCTGCAGCACTG  660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  82364  GTTTCAACCTTGCGGTCGTACTCCCCAGGCGGAGTGCTTAATGCGTTAGCTGCAGCACTG  82305

Query  661    AGAGGCGGAAACCTCCCAACACTTAGCACTCATCGTTTACGGCATGGACTACCAGGGTAT  720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  82304  AGAGGCGGAAACCTCCCAACACTTAGCACTCATCGTTTACGGCATGGACTACCAGGGTAT  82245

Query  721    CTAATCCTGTTCGCTACCCATGCTTTCGAGCCTCAGCGTCAGTTGCAGACCAGAGAGCCG  780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  82244  CTAATCCTGTTCGCTACCCATGCTTTCGAGCCTCAGCGTCAGTTGCAGACCAGAGAGCCG  82185

Query  781    CCTTCGCCACTGGTGTTCTTCCATATATCTACGCATTCCACCGCTACACATGGAGTTCCA  840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  82184  CCTTCGCCACTGGTGTTCTTCCATATATCTACGCATTCCACCGCTACACATGGAGTTCCA  82125

Query  841    CTCTCCTCTTCTG  853
              |||||||||||||
Sbjct  82124  CTCTCCTCTTCTG  82112
```

C8

Lactobacillus crispatus strain N-11 16S ribosomal RNA gene, partial sequence

Sequence ID: KU991816.1 Length: 1527 Number of Matches: 1

Range 1: 754 to 1525 GenBank Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
1426 bits(772)	0.0	772/772(100%)	0/772(0%)	Plus/Minus

```
Query  1     CGGTTACCTTGTTACGACTTCACCCCAGTCATCTGCCCTGCCTTAGACGGCTCCTTCCCG  60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1525  CGGTTACCTTGTTACGACTTCACCCCAGTCATCTGCCCTGCCTTAGACGGCTCCTTCCCG  1466

Query  61    AAGGTTAGGCCACCGGCTTTGGGCATTGCAGACTCCCATGGTGTGACGGGCGGTGTGTAC  120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1465  AAGGTTAGGCCACCGGCTTTGGGCATTGCAGACTCCCATGGTGTGACGGGCGGTGTGTAC  1406

Query  121   AAGGCCCGGGAACGTATTCACCGCGGCGTGCTGATCCGCGATTACTAGCGATTCCAGCTT  180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1405  AAGGCCCGGGAACGTATTCACCGCGGCGTGCTGATCCGCGATTACTAGCGATTCCAGCTT  1346

Query  181   CGTGCAGTCGAGTTGCAGACTGCAGTCCGAACTGAGAACAGCTTTCAGAGATTCGCTTGC  240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1345  CGTGCAGTCGAGTTGCAGACTGCAGTCCGAACTGAGAACAGCTTTCAGAGATTCGCTTGC  1286

Query  241   CTTCGCAGGCTCGCTTCTCGTTGTACTGCCCATTGTAGCACGTGTGTAGCCCAGGTCATA  300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1285  CTTCGCAGGCTCGCTTCTCGTTGTACTGCCCATTGTAGCACGTGTGTAGCCCAGGTCATA  1226

Query  301   AGGGGCATGATGACTTGACGTCATCCCCACCTTCCTCCGGTTTGTCACCGGCAGTCTCAT  360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1225  AGGGGCATGATGACTTGACGTCATCCCCACCTTCCTCCGGTTTGTCACCGGCAGTCTCAT  1166

Query  361   TAGAGTGCCCAACTTAATGCTGGCAACTAATAACAAGGGTTGCGCTCGTTGCGGGACTTA  420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1165  TAGAGTGCCCAACTTAATGCTGGCAACTAATAACAAGGGTTGCGCTCGTTGCGGGACTTA  1106

Query  421   ACCCAACATCTCACGACACGAGCTGACGACAGCCATGCACCACCTGTCTTAGCGTCCCCG  480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1105  ACCCAACATCTCACGACACGAGCTGACGACAGCCATGCACCACCTGTCTTAGCGTCCCCG  1046

Query  481   AAGGGAACTTTGTATCTCTACAAATGGCACTAGATGTCAAGACCTGGTAAGGTTCTTCGC  540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1045  AAGGGAACTTTGTATCTCTACAAATGGCACTAGATGTCAAGACCTGGTAAGGTTCTTCGC  986

Query  541   GTTGCTTCGAATTAAACCACATGCTCCACCGCTTGTGCGGGCCCCCGTCAATTCCTTTGA  600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  985   GTTGCTTCGAATTAAACCACATGCTCCACCGCTTGTGCGGGCCCCCGTCAATTCCTTTGA  926

Query  601   GTTTCAACCTTGCGGTCGTACTCCCCAGGCGGAGTGCTTAATGCGTTAGCTGCAGCACTG  660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  925   GTTTCAACCTTGCGGTCGTACTCCCCAGGCGGAGTGCTTAATGCGTTAGCTGCAGCACTG  866

Query  661   AGAGGCGGAAACCTCCCAACACTTAGCACTCATCGTTTACGGCATGGACTACCAGGGTAT  720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  865   AGAGGCGGAAACCTCCCAACACTTAGCACTCATCGTTTACGGCATGGACTACCAGGGTAT  806

Query  721   CTAATCCTGTTCGCTACCCATGCTTTCGAGCCTCAGCGTCAGTTGCAGACCA  772
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  805   CTAATCCTGTTCGCTACCCATGCTTTCGAGCCTCAGCGTCAGTTGCAGACCA  754
```

C9

Lactobacillus crispatus strain TCD56.8 16S ribosomal RNA gene, partial sequence

Sequence ID: KU851161.1 Length: 955 Number of Matches: 1

Range 1: 1 to 844 GenBank Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
1548 bits(838)	0.0	842/844(99%)	0/844(0%)	Plus/Plus

```
Query  1    AGAGTTTGATCATGGCTCAGGACGAACGCTGGCGGCGTGCCTAATACATGCAAGTCGAGC  60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1    AGAGTTTGATCATGGCTCAGGACGAACGCTGGCGGCGTGCCTAATACATGCAAGTCGAGC  60

Query  61   GAGCGGAACTAACAGATTTACTTCGGTAATGACGTTAGGAAAGCGAGCGGCGGATGGGTG  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  61   GAGCGGAACTAACAGATTTACTTCGGTAATGACGTTAGGAAAGCGAGCGGCGGATGGGTG  120

Query  121  AGTAACACGTGGGGAACCCGCCCCATAGTCTGGGATACCACTTGGAAACAGGTGCTAATA  180
            |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Sbjct  121  AGTAACACGTGGGGAACCTGCCCCATAGTCTGGGATACCACTTGGAAACAGGTGCTAATA  180

Query  181  CCGGATAAGAAAGCAGACCGCATGATCAGCTTTTAAAAGGCGGCGTAAGCTGTCGCTATG  240
            |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Sbjct  181  CCGGATAAGAAAGCAGATCGCATGATCAGCTTTTAAAAGGCGGCGTAAGCTGTCGCTATG  240

Query  241  GGATGGCCCCGCGGTGCATTAGCTAGTTGGTAAGGTAAAGGCTTACCAAGGCGATGATGC  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  241  GGATGGCCCCGCGGTGCATTAGCTAGTTGGTAAGGTAAAGGCTTACCAAGGCGATGATGC  300

Query  301  ATAGCCGAGTTGAGAGACTGATCGGCCACATTGGGACTGAGACACGGCCCAAACTCCTAC  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  301  ATAGCCGAGTTGAGAGACTGATCGGCCACATTGGGACTGAGACACGGCCCAAACTCCTAC  360

Query  361  GGGAGGCAGCAGTAGGGAATCTTCCACAATGGACGCAAGTCTGATGGAGCAACGCCGCGT  420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  361  GGGAGGCAGCAGTAGGGAATCTTCCACAATGGACGCAAGTCTGATGGAGCAACGCCGCGT  420

Query  421  GAGTGAAGAAGGTTTTCGGATCGTAAAGCTCTGTTGTTGGTGAAGAAGGATAGAGGTAGT  480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  421  GAGTGAAGAAGGTTTTCGGATCGTAAAGCTCTGTTGTTGGTGAAGAAGGATAGAGGTAGT  480

Query  481  AACTGGCCTTTATTTGACGGTAATCAACCAGAAAGTCACGGCTAACTACGTGCCAGCAGC  540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  481  AACTGGCCTTTATTTGACGGTAATCAACCAGAAAGTCACGGCTAACTACGTGCCAGCAGC  540

Query  541  CGCGGTAATACGTAGGTGGCAAGCGTTGTCCGGATTTATTGGGCGTAAAGCGAGCGCAGG  600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  541  CGCGGTAATACGTAGGTGGCAAGCGTTGTCCGGATTTATTGGGCGTAAAGCGAGCGCAGG  600

Query  601  CGGAAGAATAAGTCTGATGTGAAAGCCCTCGGCTTAACCGAGGAACTGCATCGGAAACTG  660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  601  CGGAAGAATAAGTCTGATGTGAAAGCCCTCGGCTTAACCGAGGAACTGCATCGGAAACTG  660

Query  661  TTTTTCTTGAGTGCAGAAGAGGAGAGTGGAACTCCATGTGTAGCGGTGGAATGCGTAGAT  720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  661  TTTTTCTTGAGTGCAGAAGAGGAGAGTGGAACTCCATGTGTAGCGGTGGAATGCGTAGAT  720

Query  721  ATATGGAAGAACACCAGTGGCGAAGGCGGCTCTCTGGTCTGCAACTGACGCTGAGGCTCG  780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  721  ATATGGAAGAACACCAGTGGCGAAGGCGGCTCTCTGGTCTGCAACTGACGCTGAGGCTCG  780

Query  781  AAAGCATGGGTAGCGAACAGGATTAGATACCCTGGTAGTCCATGCCGTAAACGATGAGTG  840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  781  AAAGCATGGGTAGCGAACAGGATTAGATACCCTGGTAGTCCATGCCGTAAACGATGAGTG  840

Query  841  CTAA  844
            ||||
Sbjct  841  CTAA  844
```

C10

Lactobacillus ingluviei partial 16S rRNA gene, strain Marseille-P209

Sequence ID: LT223590.1 Length: 1538 Number of Matches: 1

Range 1: 712 to 1537 GenBank Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
1520 bits(823)	0.0	825/826(99%)	0/826(0%)	Plus/Minus

```
Query  1     CGGTTACCTTGTTACGACTTCACCCCAGTCATCTGCCCTGCCTTAGGCGGCTGGCTCCAA  60
             ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1537  CGGCTACCTTGTTACGACTTCACCCCAGTCATCTGCCCTGCCTTAGGCGGCTGGCTCCAA  1478

Query  61    AAGGTTACCCCACCGACTTTGGGCATTGCAAACTCCCATGGTGTGACGGGCGGTGTGTAC  120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1477  AAGGTTACCCCACCGACTTTGGGCATTGCAAACTCCCATGGTGTGACGGGCGGTGTGTAC  1418

Query  121   AAGGCCCGGGAACGTATTCACCGCGGCATGCTGATCCGCGATTACTAGCGATTCCGACTT  180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1417  AAGGCCCGGGAACGTATTCACCGCGGCATGCTGATCCGCGATTACTAGCGATTCCGACTT  1358

Query  181   CGTGCAGTCGAGTTGCAGACTGCAGTCCGAACTGAGAACGGTTTTAAGAGATTAGCTTGC  240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1357  CGTGCAGTCGAGTTGCAGACTGCAGTCCGAACTGAGAACGGTTTTAAGAGATTAGCTTGC  1298

Query  241   CCTCGCGAGTTCGCGACTCGTTGTACCGTCCATTGTAGCACGTGTGTAGCCCAGGTCATA  300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1297  CCTCGCGAGTTCGCGACTCGTTGTACCGTCCATTGTAGCACGTGTGTAGCCCAGGTCATA  1238

Query  301   AGGGGCATGATGATCTGACGTCGTCCCCACCTTCCTCCGGTTTGTCACCGGCAGTCTCAC  360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1237  AGGGGCATGATGATCTGACGTCGTCCCCACCTTCCTCCGGTTTGTCACCGGCAGTCTCAC  1178

Query  361   TAGAGTGCCCAACTGAATGCTGGCAACTAGTAACAAGGGTTGCGCTCGTTGCGGGACTTA  420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1177  TAGAGTGCCCAACTGAATGCTGGCAACTAGTAACAAGGGTTGCGCTCGTTGCGGGACTTA  1118

Query  421   ACCCAACATCTCACGACACGAGCTGACGACGACCATGCACCACCTGTCATTGCGTCCCCG  480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1117  ACCCAACATCTCACGACACGAGCTGACGACGACCATGCACCACCTGTCATTGCGTCCCCG  1058

Query  481   AAGGGAACGCCCTATCTCTAGGGTTGGCGCAAGATGTCAAGACCTGGTAAGGTTCTTCGC  540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1057  AAGGGAACGCCCTATCTCTAGGGTTGGCGCAAGATGTCAAGACCTGGTAAGGTTCTTCGC  998

Query  541   GTAGCTTCGAATTAAACCACATGCTCCACCGCTTGTGCGGGCCCCCGTCAATTCCTTTGA  600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  997   GTAGCTTCGAATTAAACCACATGCTCCACCGCTTGTGCGGGCCCCCGTCAATTCCTTTGA  938

Query  601   GTTTCAACCTTGCGGTCGTACTCCCCAGGCGGAGTGCTTAATGCGTTAGCTCCGGCACTG  660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  937   GTTTCAACCTTGCGGTCGTACTCCCCAGGCGGAGTGCTTAATGCGTTAGCTCCGGCACTG  878

Query  661   AAGGGCGGAAACCCTCCAACACCTAGCACTCATCGTTTACGGCATGGACTACCAGGGTAT  720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  877   AAGGGCGGAAACCCTCCAACACCTAGCACTCATCGTTTACGGCATGGACTACCAGGGTAT  818

Query  721   CTAATCCTGTTCGCTACCCATGCTTTCGAGCCTCAGCGTCAGTTACAGACCAGGCAGCCG  780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  817   CTAATCCTGTTCGCTACCCATGCTTTCGAGCCTCAGCGTCAGTTACAGACCAGGCAGCCG  758

Query  781   CCTTCGCCACTGGTGTTCTTCCATATATCTACGCATTCCACCGCTA  826
             ||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  757   CCTTCGCCACTGGTGTTCTTCCATATATCTACGCATTCCACCGCTA  712
```