Ministry of Higher Education
And Scientific Research
University of Babylon
College of Medicine

Phylogeny Diversity and Polyketide Synthetase Pathogenicity Island Among *Escherichia coli* Isolated from Vaginitis

A Thesis

Submitted to the Council of College of Medicine-University of Babylon in Partial Fulfillment of the Requirements for the Doctor of Philosophy in Science / Medical Microbiology

By

Noor Ismeal Nasser Hussein
B.Sc. Microbiology / College of Science / 2005/University of Baghdad
M.Sc. Medical Microbiology / College of Medicine / 2010/University of Al-Qadisiyah

Supervised by

Professor
Dr.Mohammed Sabri Abdul-razaq

Professor
Dr. Nadia Mudher Al-Hali

2023 A.D. **1444 A.H.**

بسم الله الرحمن الرحيم

﴿قالوا سبحانك لا علم لنا إلا ما علمتنا﴾

صدق الله العلي العظيم
سورة البقرة – الآية(32-33)

Supervision Certificate

We certify that this thesis was prepared under our supervision at the Department of Medical Microbiology, College of Medicine, University of Babylon, in a partial fulfillment for the requirements of the degree of Doctorate of philosophy

in Medical Microbiology.

Prof. Dr. Mohammed Sabri Abdul-razaq
of Babylon University
College of Medicine
2023

Prof. Dr. Nadia Mudher Al-Hafi
of Babylon University
College of Medicine
2023

Recommendation of Head of Department of Microbiology

According to the available recommendations, I forward this thesis for debate by the examining committee.

Prof. Dr. Hayam Khalis Al-Masoudi

Head of Department of Medical Microbiology

College of Medicine

University of Babylon

Decision of Examination Committee

We, the examination committee, certify that we have read the thesis entitled **(Phylogeny Diversity and Polyketide Synthetase Pathogenicity Island Among *Escherichia coli* Isolated from Vaginitis)** and have examined the student **(Noor Ismeal Nasser AL- Rubaie)** in its contents, and that in our opinion it is

accepted as a thesis for Degree of Doctorate of Philosophy in Medical Microbiology with excellent estimation.

Signature
Prof. Dr. Adnan H. AL-Hamadani
College of Medicine,University of Al-Qadisiya
(Chairman)

Signature
Prof. Dr. Lamees Abdul-Razzaq
College of Medicine, University of Babylon
(Member)

Signature
Prof. Dr. Suhaila Fadhil Al-Shaikh
College of Medicine, University of Babylon
(Member)

Signature
Prof. Dr. Ausama Abed Alkadhum Mahdi
College of Sciences for woman , University of Babylon (Member)

Signature
Assist. Prof. Dr. Fatima Malik Abood
Faculty of Dentistry, University of Babylon
(Member)

Signature
Prof. Dr. Mohammed Sabri Abdul-razaq
College of Medicine, University of Babylon
(Member, Supervisor)

Signature
Prof. Dr. Nadia Mudher Al-Hali
College of Medicine, University of Babylon
(Member, Supervisor)

Approved for the college committee of graduate studies

Signature

Prof. Dr. Mohend Abbass Nori Al-Shalah

Dean of College of Medicin

University of Babylon

Dedication

To my family

for their deep love and Endless Support

Noor\2023

Acknowledgements

First and foremost, I would like to thank Allah the Most Great and most Merciful. I am thankful to my supervisors, **Prof. Dr. Mohammed Sabri** and **Prof. Dr. Nadia Mudar Al-Hali** for their guidance,

encouragement and patience. I will always be grateful for their support and mentorship.

Faithful thanks to the dean of the College of Medicine **Prof.Dr.*Mohend Abbass Nori Al-Shalah*** and the Head of Department of Medical Microbiology **Prof.Dr**. **Hayam Khalis Al-Masoudi** for their help and great cooperation.

Honest thanks to the staff of the Department of Medical Microbiology at the College of Medicine, University of Babylon for their efforts to communicate the information's during our study.

Finally, my sincere thanks go to all the people and patients who helped me in one way or another in the completion of my work.

Summary:

Out of 150 high vaginal swap samples, 27(18%) *E. coli* isolates were detected by 16s rDNA. The phylogenetic analysis reveal that 20(74%) of the tested isolates have *chuA* gene, 20(74%) of isolates have *yjaA* gene and 19 (70.3%) of the isolates have TspE4C2 DNA fragment so The phylogenetic distribution of the isolates showed that most of the isolates belonged to phylogenetic groups $B2_2$ 21 (77.7 %), while the rest belonged to phylogenetic groups B_1 and A 3 (11.1%) for each. Screening for *E. coli* ST 131 indicated that 11 (40.7%) isolates belonged to sequence type 131. Phylogenetic analysis showed that most belonged to phylogroup B_2 (90.9%), and one isolate belonged to phylogroup B1 (9.09%). Nine (33.3%) of the 27 *E. coli* isolates carried the pks island, with the majority belonging to phylogroup B_2 8 (88.8%) and one isolate belonging to phylogroup B_1 1 (11.1%). The alignment results of the *clbB* gene 579 bp samples revealed the presence of one nucleic acid variant represented by one nucleic acid substitution (62 C>T) compared with their reference nucleic acid sequences.

A comprehensive phylogenetic tree was generated based on the nucleic acid variations observed in the 579 bp region of the *clbB* amplicons. A slight tilt was observed between samples S2, S3, S6, and S7 and the related strains. This slightly different positioning was attributed to the detected variation (62C>T) in the investigated samples. The alignment results of the *clbN* gene 733 bp samples revealed the presence of one nucleic acid variant represented by one nucleic acid substitution (243C>T) compared with their reference nucleic acid sequences. A comprehensive phylogenetic tree was generated based on the nucleic acid variations observed in the 733 bp region of the *clbN* amplicons. A slight tilt was observed between samples S1 and S6 and their related strains. This slightly different

positioning was attributed to the detected variation (243C>T) in the investigated samples.

The antibiograms indicated that most *E. coli* isolates were resistant to ciprofloxacin, norfloxacin, and levofloxacin. However, low resistance rates were observed at higher concentrations of gatifloxacin and moxifloxacin. Among the B_2 phylogroup isolates, resistance was observed for norfloxacin, ciprofloxacin, levofloxacin, Gatifloxacin (75µg/ml), gatifloxacin (30 µg/ml), and moxifloxacin (30 µg/ml). Most of the B_1 and A phylogenetic group isolates were resistant to almost all fluoroquinolones used in the study. The *E. coli* ST131 isolates displayed a high pattern of resistance to norfloxacin, ciprofloxacin, and levofloxacin, while resistance to gatifloxacin and moxifloxacin was dose-dependent

List of Contents

List of Table

List of figure

List of Abbreviations

Abbreviation	Meaning
CTX-M-class	extended-spectrum beta-lactamase
ExPEC	Extra intestinal pathogenic *Escherichia coli*
GABA	Gamma-Aminobutyric Acid
IFN-γ	Interferon-gamma
IL-1β	interleukin-1-beta
IL-6	Interleukin-6
IL-8	Interleukin-8
MDR	Multiple drug resistance
ml	Milliliter
NRPS-PKS	hybrid non-ribosomal peptide synthetase- polyketide synthetase
PAIs	Pathogenicity islands
PCR	Polymerase chain reaction
PKs	Polyketide synthetase pathogenicity island
ST131	Sequence type 131
TBE	Trise base boric acid
TNF-α	Tumor necrosis factor- alpha
UTI	Urinary tract infection
µl	Microliter
EDTA	Ethylene diamine tetra acetic acid

1.1. Introduction:

Genital tract infections represent a series of public health challenges for females in both developed and developing countries. Different microbial infections of the vaginal tract can cause serious medical complications, such as preterm labor, premature rupture, amniotic fluid infection, fetal membranes, and low birth weight of the neonate, which leads to high perinatal morbidity and mortality (Kaambo *et al.*,2017; Yasin *et al*.,2021).

Vaginitis is a general term used to describe inflammation of the vagina, and it is essential to distinguish it from bacterial vaginosis, the vaginal microflora plays a crucial role in maintaining a healthy environment in the vagina. Normally, the vagina is colonized by a diverse population of bacteria, with *Lactobacillus* species being the predominant ones. *Lactobacilli* help to maintain the vaginal pH in the acidic range (around 3.5-4.5), which creates an inhospitable environment for many harmful bacteria. However, in cases of vaginitis, the balance of the vaginal microflora is disrupted, leading to a decrease in *Lactobacillus* species, inconstant levels of vaginal inflammation, and inadequate epithelial maturation and an increase in other bacteria, particularly aerobic and enteric bacteria (Donders et al.,2002; Raheem *et al*.,2022).

When the number of lactobacilli decreases in the vaginal microflora, it disrupts the normal balance of bacteria and creates an environment that facilitates the growth of potentially pathogenic or harmful bacteria, such as *Enterococcus faecalis*, *Escherichia coli*, *Staphylococcus aureus*, and *Staphylococcus epidermidis*, in addition to *Streptococcus agalactiae*., *Pseudomonas* spp., *Acinetobacter* spp., and *Klebsiella pneumonia* at a lower rate (Fan *et al*.,2013; Raheem *et al.*,2022).

Patients with vaginitis commonly experience various symptoms, including itching, purulent (abnormal, pus-like) discharge, inflammation of the vaginal epithelium, and redness of the vaginal wall. These symptoms are indicative

of an imbalanced vaginal microflora and an overgrowth of potentially harmful bacteria. (Hassan *et al*.,2020).

However, the most common cause of vaginitis among gram-negative bacteria is *E. coli*, which can cause infection owing to its ability to produce different virulence factors, In such cases, *E. coli* may ascend from the rectal or perineal area to the vagina, leading to infection. These bacteria may also be commonly found in healthy vaginas owing to contamination from the anal region. These gram-negative bacteria can produce multiple factors and enzymes that contribute in their infection. These include chelating agents, hemolysin, Antibiotic resistance, Adhesins and biofilm formation (Yasin *et al*.,2021; Ma *et al*.,2022).

Escherichia coli (E. coli) strains can be classified into different phylogenetic groups based on their source and site of infection. These groups are labeled as A, B_1, B_2, C, D, E, F, and cryptic clade known as clade I, they represent distinct genetic lineages within the *E. coli* species. Each phylogenetic group may have specific characteristics and associations with certain types of infections. also, it can be further categorized into pathotypes, which are groups of strains that cause specific types of infections or have distinct virulence factors.

Specific strains of *E. coli* have been identified and characterized using serological and molecular techniques. Serotyping involves the identification of specific surface antigens that can help differentiate between different strains. Molecular techniques, such as polymerase chain reaction (PCR), allow for the detection of specific genes or genetic elements associated with particular pathotypes (Terlizzi *et al*.,2017., Erjavec,2019., Sarowska *et al*.,2019).

E. coli strains can also contain pathogenicity islands, which are genomic regions that contain clusters of genes encoding virulence factors. These pathogenicity islands are acquired through horizontal gene transfer and play a significant role in the ability of bacteria to cause disease. These islands are dispersed

along the bacterial genome and are distinct from the core genome, which contains genes necessary for basic cellular functions (Terlizzi *et al*.,2017., Erjavec,2019., Sarowska *et al*.,2019).

Pathogenicity islands can contain various virulence factors, such as adhesins, toxins, and invasins, which enhance the bacterial ability to colonize host tissues, evade the immune system, and cause damage. These virulence factors contribute to the increased pathogenicity of the bacteria, allowing them to cause various infections in different host environments. Furthermore, *E. coli* strains may have fitness islands, which are genomic regions containing genes that provide the bacteria with a selective advantage, allowing them to survive and thrive in specific environmental conditions (Terlizzi *et al*.,2017., Erjavec,2019., Sarowska *et al*.,2019).

On the other hand Colibactin is a cyclomodulin, a toxin produced by certain strains of *E. coli* and other members of the Enterobacteriaceae family. The pks genomic island synthesizes it and is mainly observed in *E. coli* strains belonging to the B_2 phylogroup. The pks island contains a cluster of 19 genes responsible for producing colibactin and other functions. Colibactin acts as a genotoxin that interferes with the eukaryotic cell cycle of the host. It induces DNA double-strand breaks in host cells, which can lead to genomic instability and potentially contribute to the long-term development of cancer (Sarowska *et al*.,2019; Wami *et al*.,2021).

The accumulation of DNA damage triggered by colibactin can have various consequences on host cells, such as cell cycle arrest, in which affected cells temporarily halt their cell cycle progression, this response allows time for DNA repair mechanisms to attempt to repair the damage before the cell divides, also it cause cellular senescence is a state of irreversible cell cycle arrest in which cells are still alive but are no longer capable of dividing. Additionally, colibactin can cause chromosomal aberrations. Although colibactin is associated with potential

pathogenic effects, pK-positive bacteria have also been found to possess some beneficial effects. These effects include (Anti-inflammatory properties, Antibiotic effects, and Analgesic effects) (Sarowska *et al*.,2019; Wami *et al*.,2021).

Aim of study:

Molecular detection of colibactin and PKs islands in clinical isolates of *E. coli*.

The objective of the study:

1- Isolation and identification of *Escherichia coli* isolates using traditional techniques and 16 rRNA.

2- Identification of phylogenetic groups of *E. coli* isolates using PCR.

3- Screening for *E. coli ST131* among the strains isolated from vaginal infections.

4- Molecular detection of PKs islands Colibactin genes (clb A, clb Q, clb B, clb N). among the clinical isolates.

5- Sequencing of Colibactin genes (*clb B and clb N*).

6- Show the effect of some antibiotics (fluoroquinolones) on bacterial isolates.

1.2 Literature Review

1.2.1 General Characteristics of *Escherichia coli*:

Escherichia coli bacteria was discovered by the German-Austrian Pediatrician Dr. Theodor Escherich (1885), who observed a slender short rod its dimensions of (1–5 μm) in length and (0.4 μm) in width, which he named Bacterium coli commune. Subsequently, in 1919, this bacterium was renamed by Castellani and Chalmers and became known as *Escherichia coli* (Erjavec,2019). This bacterium is one of the members of the family Enterobacteriaceae and has many traits, including non-sporulation, non-motile or motile by peritrichous flagella, chemoorganotrophic, facultatively anaerobic, producing acid from the fermentation of glucose, catalase positive, oxidase negative, and mesophilic (Murray *et al*.,2016).

The bacterium *E. coli* is normally found in the microflora of the human and animal digestive tracts and has a symbiotic relationship with its hosts. It has several benefits in that it is involved in the digestion and synthesis of certain vitamins. Despite this, it can cause many diseases inside and outside the gastrointestinal tract, for example, but not limited to gastrointestinal tract infections, urinary tract infections, hemolytic-uremic syndrome, hospital-acquired pneumonia, sepsis, surgical site infection and meningitis (Erjavec,2019).

E. coli strains can be classified serologically based on the presence of different surface antigens. These antigens are known as O antigens, K antigens, and H antigens, and they are used to differentiate and classify *E. coli* strains into specific serotypes. O antigens are part of the lipopolysaccharide (LPS) molecules present on the outer surface of the bacterial cell. They are heat-stable and are essential for the recognition of *E. coli* strains. There are numerous types of O antigens, and currently, over 173 different types have been identified. K antigens are capsular antigens that

are involved in the formation of the bacterial capsule. The capsule helps the bacterium evade the host's immune response. Like O antigens, K antigens are diverse, and there are around 80 different types. H antigens are flagellar antigens, located in the bacterial flagella, which are responsible for bacterial motility. These antigens allow *E. coli* strains to move in liquid environments. Similar to O and K antigens, H antigens exhibit a high level of diversity, and approximately 56 types have been identified. The presence of a large number of different antigens and their potential combinations leads to an enormous diversity of *E. coli* serotypes. This complexity can make serological classification challenging and contribute to the vast number of possible serotypes, estimated to be between 50,000 and 100,000. It is also important to note that the serotype of an *E. coli* strain does not necessarily correlate directly with its genetic relatedness. closely related strains may have different serotypes, while genetically distant strains may have the same serotype. This discrepancy is due to the independent evolution of surface antigens and the core genetic content of *E. coli* strains (Erjavec,2019; Liu *et al*.,2020).

1.2.2 phylogenetic groups of *E. coli*

E. coli exhibits a diverse genetic substructure, and it has evolved in a clonal manner over time. Genetic studies have revealed that *E. coli* can be categorized into several phylogenetic groups based on their genetic characteristics. One commonly used method for determining the phylogenetic groups of *E. coli* is the triplex PCR (polymerase chain reaction) method. This genotyping technique relies on the amplification of three specific gene fragments; *chuA*, *yjaA*, and *TSPE4.C2*.

The *chuA* gene fragment (279 bp) is a gene associated with heme transport in *E. coli*. This gene is commonly used as a marker for phylogenetic group B_2 and D, *yjaA* gene fragment (211 bp) is involved in an unknown function in *E. coli*. It serves as a marker for phylogenetic group A, *TSPE4.C2* fragment (152 bp) is a non-

coding area of the *E. coli* genome, and its presence or absence is used as a marker for phylogenetic group B_1. by using the combination of these three genetic markers, can determine which phylogenetic group an *E. coli* strain belongs to. The four major phylogenetic groups identified in *E. coli* are A, B_1, B_2, and D. Each of these groups represents a distinct lineage within the *E. coli* species, with unique genetic characteristics. Phylogenetic groups analysis used for understanding the genetic diversity and evolutionary relationships among different *E. coli* strains. It helps to identify the genetic substructure of *E. coli* populations and study the distribution and prevalence of specific phylogenetic groups in various environments and clinical settings (Clermont *et al*., 2000; Fratamico *et al*.,2016).

with the development of multi-locus sequence typing (MLST) and whole genome sequencing, it's given the opportunity to analyze more extensive genetic data, leading to the discovery of additional phylogenetic groups in *E. coli*. The total number of recognized phylogenetic groups increased to eight, which now include A, B_1, B_2, C, D, E, F, and a cryptic clade known as clade I. To distinguish the newly identified phylogenetic groups, additional genetic markers were employed. For example, the F phylogenetic group is distinguished using a Quadruplex PCR method that targets an additional gene called *arpA*. Single PCR reactions are used to differentiate between phylogroups E and I or phylogroups A and C. Moreover, two phylogroups, D and E, are also identified using specific genetic markers (Clermont *et al*.,2013; Logue et al.,2017).

The commensal *E. coli* which are non-pathogenic strains that colonize the gastrointestinal tract mucosa and do not cause any significant disease, they are most commonly found in phylogenetic groups A and B_1. The Pathogenic *E. coli* that responsible for intestinal infections, such as diarrhea and gastroenteritis, can belong to phylogenetic groups A, B_1, or D. The *E. coli* strains responsible for infections

outside the intestinal tract, such as urinary tract infections and bloodstream infections, are primarily found in phylogenetic groups B_2 and D. The group E is related to group D and includes some pathogenic strains, including the well-known E. coli O157: H7, which is a major cause of foodborne illnesses. The group F is related to the main phylogenetic group B_2, which is associated with extraintestinal infections. Additionally, there is a cryptic clade known as clade I, These strains are indistinguishable based on their observable traits but can be classified into this cryptic clade based on their genetic similarity (Logue et al.,2017).

1.2.3 Extra intestinal pathogenic *Escherichia coli* (ExPEC).

The extraintestinal pathogenic *Escherichia coli* (ExPEC) strains possess unique characteristics that make them capable of causing a wide range of infections outside the intestinal tract.This strains are genetically diverse and can belong to multiple phylogenetic groups, most notably groups B_2 and D. It's have a flexible genome that allows them to acquire and exchange genetic material through horizontal gene transfer. This capability enables them to gain new virulence traits and adapt to changing environments. and hosts. Also, its carry a diverse array of virulence factors, including adhesins, toxins, and iron-acquisition systems, that enable them to colonize and invade host tissues, evade the immune system, and cause disease. ExPEC strains are proficient in acquiring mobile genetic elements, such as plasmids and pathogenicity islands, which carry genes encoding virulence factors. These elements can facilitate the spread of virulence traits within bacterial populations.

The ExPEC strains typically reside in the intestinal microbiota of humans as commensal organisms. They can emerge from this reservoir and ascend to cause infections in various body sites, such as the urinary tract, bloodstream, and

respiratory tract. It was responsible for both healthcare-associated infections (e.g., catheter-associated urinary tract infections, surgical site infections) and infections in the community (e.g., urinary tract infections in otherwise healthy individuals). Despite their genetic diversity, only a limited number of ExPEC lineages are responsible for the majority of infections. This highlights the importance of specific clones with enhanced virulence potential. ExPEC strains are a significant public health concern due to their ability to cause a broad spectrum of infections and their potential for antimicrobial resistance (Sarowska *et al*.,2019; Sora *et al*.,2021).

The ExPEC group encompasses several pathogenic strains of *E. coli* that are associated with a variety of extraintestinal infections. Some of the most notable members of the ExPEC group include; Uropathogenic *E. coli* strains which are responsible for the majority of urinary tract infections (UTIs) in humans. Its can colonize and invade the urinary tract, leading to conditions such as cystitis (bladder infection) and pyelonephritis (kidney infection). Neonatal Meningitis *E. coli* strains which are a leading cause of meningitis in newborns and infants. These bacteria can cross the blood-brain barrier and cause inflammation of the membranes surrounding the brain and spinal cord. Sepsis-associated *E. coli* strains can also cause bloodstream infections, leading to sepsis which is a life-threatening condition characterized by a systemic inflammatory response to infection that can result in organ failure. Each of these subgroups within the ExPEC group possesses specific sets of virulence factors that enable them to colonize and cause disease in distinct host niches. For example, Uropathogenic *E. coli* strains have specialized adhesins that allow them to attach to and invade the cells of the urinary tract. Neonatal Meningitis *E. coli* strains have factors that facilitate their crossing of the blood-brain barrier and surviving within the central nervous system. ExPEC strains are

particularly concerning due to their ability to cause severe and recurrent infections. (Sarowska *et al*.,2019; Manges *et al*.,2019; Sora *et al*.,2021).

This strains associated with the acquisition of new and troubling antibiotic resistance genes. The clinical and economic impact of these infections and their optimal management, especially in the face of increasing antibiotic resistance, are challenging and underappreciated. Resistance mechanisms to all major antibiotic classes exist among this strains. These include the production of ESBLs, and the resistance to fluoroquinolones and aminoglycosides occurs via plasmid- or chromosomally encoded resistance gene or via mechanisms that reduce antimicrobial uptake into the cell (Manges *et al*.,2019**).**

Multilocus sequence typing (MLST) has been instrumental in understanding the genetic diversity and distribution of major ExPEC sequence types (STs). Some of the most prevalent and clinically significant ExPEC STs identified are sequence type 131 is particularly notorious for its global dissemination and association with antimicrobial resistance. It is one of the most common and clinically relevant STs and is responsible for a significant proportion of urinary tract infections and other extraintestinal infections, (ST69) Another widespread ExPEC ST associated with diverse infections, including urinary tract infections and bloodstream infections, (ST10) This sequence type has also been frequently reported in various studies and is associated with different types of extraintestinal infections, ST405, ST38, ST95, ST648, ST73, ST410, ST393, ST354, ST12, ST127, ST167, ST58, ST617, ST88, ST23, ST117, and ST1193: These are additional sequence types that have been identified and play roles in causing a wide range of infections. Among these STs, ST131 stands out as the most prevalent and concerning. It has been reported in over 90% of the studies, underscoring its global importance and clinical significance. ST131 is notorious for its ability to acquire antimicrobial resistance

genes, leading to multidrug-resistant infections, which further complicates treatment (Hojabri *et al*.,2019).

1.2.4 Pathogenesis of extra intestinal pathogenic *Escherichia coli*

ExPEC is the most common gram-negative bacterial pathogen, and it is present in various tissues and sites throughout the human body. This strains have the ability to colonize and invade different tissues, leading to a wide range of extraintestinal infections. These infections include neonatal meningitis, sepsis, urinary tract infection (UTI), prostatitis, aerobic vaginitis, and peritonitis. The urinary tract is typically a sterile space, with the exception of the distal urethra. UTIs occur when ExPEC strains colonize and ascend the urinary tract, leading to infections of the bladder (cystitis) or the kidneys (pyelonephritis). The UTI infection, caused predominantly by ExPEC, is one of the most prevalent disease. That affects a high percentage of women between the ages of 16 and 35. Also, the incidence of ExPEC-related UTIs varies with age, and older individuals, especially men with enlarged prostate glands, are at higher risk of UTIs due to urinary flow obstruction. The ability of ExPEC to cause such a diverse range of infections underscores its adaptability and virulence as a pathogen (Poolman *et al*.,2016; Tan and Chlebicki .,2016).

Urinary tract infections (UTIs) can affect the pediatric population, and there are some notable differences in the prevalence and clinical presentation compared to adults. In the pediatric population, UTIs are more common in boys, accounting for approximately 60% of cases. However, this tendency is reversed after the age of 2-3 months, and UTIs become more prevalent in girls. In young infants, especially in the first few months of life, UTIs may not present with typical clinical symptoms that are observed in older children or adults. This lack of characteristic symptoms can make the diagnosis challenging and may result in delayed identification and

treatment. The most common pathogens associated with UTIs in both adults and the pediatric population are Gram-negative Enterobacteriaceae, with uropathogenic *E. coli* being the predominant causative agent. Uropathogenic *E. coli* strains possess virulence factors that enable them to adhere to and colonize the urinary tract, leading to infection (Chakraborty *et al*.,2017).

E. coli strains are responsible for a substantial majority of uncomplicated UTIs, accounting for approximately 75 to 95 percent of cases. Uncomplicated UTIs typically involve infections of the lower urinary tract (bladder) and are more common in otherwise healthy individuals, mainly women. This strains are also a significant cause of complicated UTIs, contributing to about 40 to 50 percent of cases. Complicated UTIs are more challenging to treat and often involve infections in individuals with structural abnormalities or underlying medical conditions, such as kidney stones, urinary tract obstructions, or weakened immune systems. One of the challenges in managing UTIs caused by *E. coli* is the rising prevalence of antimicrobial resistance. Some *E. coli* strains have acquired resistance to commonly used antibiotics, making treatment more complicated and emphasizing the importance of appropriate antibiotic stewardship. UTIs can cause significant discomfort and have a considerable impact on the quality of life for affected individuals. Moreover, if left untreated or inadequately treated, UTIs can lead to severe complications, including kidney infections and sepsis (Flores-Mireles *et al.*,2015).

Uropathogenic *Escherichia coli* possess a wide range of virulence factors that contribute to their ability to cause urinary tract infections and extraintestinal infections. This strains have specific adhesins and fimbriae that enable them to adhere to the epithelial cells lining the intestines, kidneys, and lower urinary tract. This adhesion is a critical step in the initiation of infection, allowing the bacteria to

establish themselves on the urinary tract mucosa. Also, its can trigger the production of cytokines by T cells, which are immune system components. These cytokines play a crucial role in inflammation and immune response during the infection. This inflammatory response can contribute to the clinical symptoms of UTIs, such as pain and swelling. Additionally, these bacteria have the ability to invade and multiply within the host's epithelial cells. This intracellular survival is a unique feature that allows the bacteria to evade the immune system and antibiotics, making treatment more challenging. These virulence factors collectively enhance their ability to evade the host's defenses, adhere to and colonize the urinary tract tissues, and cause infection. The combination of these factors allows this strains to establish a persistent infection and cause recurrent UTIs in some individuals (Baldy-Chudzik *et al*.,2015; Terlizzi *et al*.,2017).

The majority of neonatal invasive infections, including meningitis and bacteremia/sepsis, are caused by *E. coli* K1 strains, The K1 antigen, part of the bacterial capsule, is closely associated with the severity of these diseases. Healthcare workers can serve as a potential source of infection with these strains, leading to neonatal infections (the membranes surrounding the brain and spinal cord) in neonates, causing neonatal meningitis which are a serious condition that poses significant risks to newborns. The infection can lead to high rates of morbidity (approximately 30%) and mortality (approximately 10%) (Wijetunge *et al*.,2015).

Neonatal meningitis *E. coli* strains possess several crucial virulence components that contribute to their ability to cause invasive infections, particularly meningitis, The K1 capsular antigen is a critical virulence factor that shields neonatal meningitis *E. coli* from phagocytosis by the host's immune cells. The capsule helps the bacteria to evade the host's defense mechanisms and facilitates bacterial dissemination, allowing them to spread within the host's tissues. The presence of

Invasion genes (*ibeA, ibeB*, and *ibeC*) involved in the host tissue invasion. These genes play a crucial role in the ability of the bacteria to cross the blood-brain barrier and infect the meninges, leading to neonatal meningitis. The Iss protein is another important virulence factor found in this strains. It helps protect the bacteria from phagocytosis by immune cells and enhances their survival in the host's bloodstream and tissues. Also, Colony-stimulating factor V expressed by this bacteria likely contributes to the bacteria's ability to cause meningitis by promoting their survival and persistence in the central nervous system.The combination of these virulence factors allows Neonatal meningitis *E. coli* strains to evade the host's immune system, invade host tissues, and establish infection in the central nervous system, resulting in neonatal meningitis. These factors contribute to the severity and persistence of Neonatal Meningitis *E. coli* infections and the associated high rates of morbidity and mortality. (Lemaître *et al*.,2014; Guzman-Hernandez *et al*.,2016; Sarowska *et al*.,2019).

1.2.5 *E. coli* ST131.

The emergence and global dissemination of *Escherichia coli* sequence type 131 have been a major concern in the field of infectious diseases. In 2008, *E. coli* ST131 isolates were first identified in Asia, Europe, and North America. Subsequently, these strains rapidly spread worldwide, becoming a significant pathogen responsible for various diseases in humans. It is associated with urinary tract infections, bloodstream infections, respiratory infections, and other extraintestinal infections, and its resistance to multiple antibiotics have led to increased morbidity, mortality, and healthcare costs (Mathers *et al*.,2015).

E. coli ST131 has emerged as a dominant and truly pathogenic extraintestinal pathogenic *E. coli* clone that characterized by possessing a large number of virulence-associated genes, such as *sat, iutA*, *malX*, *usp*, and *ompT*. These genes contribute to enhance its adaptability, competitiveness, and colonization capabilities. This allows it to efficiently establish infections in different host environments. This strians is capable of causing infections in both community and healthcare settings. This dual ability to cause infections in various environments poses a significant public health concern, as it can lead to outbreaks and the spread of multidrug-resistant strains in both settings (Matsumura *et al*.,2016).

Fluoroquinolones, such as ciprofloxacin and levofloxacin, are commonly prescribed antibiotics used to treat various infections, including those caused by *E. coli* ST131. However, these strains have developed resistance to fluoroquinolones, which has become a major challenge in the management of infections caused by this pathogen. Its resistance to fluoroquinolones due to specific mutations in the chromosomal *gyrA* and *parC* genes. These genes encode enzymes known as DNA gyrase and DNA topoisomerase IV, respectively, which are the targets of fluoroquinolones. Coupled with its recent spread, ST131 itself has been responsible for much of the observed rise in antibiotic resistance (Banerjee *et al*.,2014; Stoesser *et al*.,2016; Biggel *et al*.,2022).

The accessory contains genes that are variably present among ST131 isolates, providing them with additional traits, such as antibiotic resistance and virulence factors, the phylogenetic clustering of ST131 is largely driven by accessory genes content. this sequence type appears to have undergone compensatory mutations that allow it to acquire multidrug resistance and virulence plasmids without incurring a significant fitness cost. This means that these additional genetic elements do not impose a significant burden on the bacterium, allowing it to

thrive and compete effectively in various environments. Furthermore ST131 has been found to have a broad host range, affecting not only humans but also capable of colonizing and causing infections in various animal species, expanding their ecological niche beyond humans (Matsumura *et al*.,2016; Manges *et al*.,2019; Pajand *et al*.,2021).

Investigations of capsular variation in *E. coli* ST131 have revealed signs of several recombination events over time. Capsular variation refers to changes in the genes responsible for the bacterial capsule, which can influence the bacterium's ability to evade the host's immune system. While in vitro tests may not directly link capsular variation to virulence, it is possible that changes in the capsule can impact bacterial fitness and adaptation to different hosts and environments. Therefore, capsular variation could still play a role in the widespread distribution of *E. coli* ST131. A comparative genomic analysis has been conducted to compare *E. coli* ST131 with other extraintestinal pathogenic *E. coli* (ExPEC) sequence types. These analyses revealed a substantial association between *E. coli* ST131 and the presence of the serine protease autotransporter of the Enterobacteriaceae virulence gene (*espC*). This gene is a virulence factor that can enhance the pathogenicity of *E. coli* by promoting bacterial adherence, invasion of host cells, and evasion of the host's immune response. The presence of the *espC* gene and potentially other virulence factors in *E. coli* ST131 may contribute to its ability to cause a large fraction of total infections in both community and healthcare settings. The possession of such virulence genes may provide ST131 with a competitive advantage over other ExPEC strains (Alqasim *et al*.,2014; Freire *et al*.,2020).

1.2.5.1 Classification of *E. coli* sequence type 131.

E. coli ST131 is mostly associated with phylogenetic group B_2. This group is known to include both strains that cause extraintestinal infections and those frequently isolated from the feces of asymptomatic individuals. The Phylogenetic group B_2 can be further subdivided into several subgroups based on genetic differences among the strains. *E. coli* ST131 belongs to subgroup I within group B_2. Subgroup I has been proposed as the basal subgroup of B_2 strains, suggesting that it represents an ancestral lineage from which other B_2 strains have diverged over time. (Johnson *et al*.,2014; Denamur *et al*.,2021).

E. coli ST131 isolates are commonly associated with certain serotypes of O antigens, particularly O25b and O16. However, it's worth noting that some of these isolates may not be typeable for the O antigen, meaning that they lack a clear serotype designation (Mathers *et al*.,2015; Hajihasani *et al*.,2022). The fimH gene, which encodes the type 1 fimbrial adhesin, is known to be highly allelic diverse in *E. coli* ST131 isolates. Among the different fimH alleles observed, fimH30 is the most prevalent allele. The predominance of the fimH30 allele has led to the classification of a specific lineage within ST131 known as the H30 lineage. The H30 lineage of *E. coli* ST131 can be further divided into two significant sublineages (H30-R) This sublineage is characterized by specific genetic markers and represents a major and widespread sublineage within the H30 lineage. It has been associated with high levels of antimicrobial resistance, contributing to its clinical significance and challenges in treatment, (H30-Rx) The H30-Rx sublineage is a related sublineage with additional genetic variations. It is also associated with multidrug resistance and has been linked to outbreaks and the spread of resistant strains (Mathers *et al*.,2015).

Population genetics studies have revealed that the phylogeny of *E. coli* ST131 can be divided into three primary clades: A, B, and C. The clades A and B

are smaller subsets of ST131 and represent a minority of the overall phylogenetic diversity within this sequence type. The clade C is the largest clade within the ST131 phylogeny. It is highly prevalent and includes sub-lineages C1 (H30-R) and C2 (H30-Rx). (Matsumura *et al*.,2016; Li *et al*.,2021). Clade C, is uniformly resistant to fluoroquinolone antibiotics. This resistance is conferred by specific amino acid substitutions in the quinolone-resistance-determining region (QRDR) of the gyrA and parC genes, which are the targets of fluoroquinolones. The H30-Rx sub-lineage is strongly associated with the presence of CTX-M β-lactamases. CTX-M β-lactamases are enzymes that confer resistance to extended-spectrum cephalosporins, a class of β-lactam antibiotics, and contribute to the multidrug resistance observed in this sub-lineage (Mathers *et al*.,2015; Ben Zakour *et al*.,2016; Denamur *et al*.,2021).

1.2.6 The pathogenicity islands(PAIs).

Bacterial evolution mostly takes place in two directions: vertically and horizontally. Although horizontal transfer has a considerable impact on some regions of the genome, vertical transfer is slow and inconsistent. horizontal gene transfer has a big impact on how bacteria evolve, especially when it comes to gaining pathogenic properties. Transduction, transformation, and conjugation are the three mechanisms of horizontal gene transfer. Plasmids and more significant parts of the genome, such as genomic islands, can be conjugated from one bacterium to another (Messerer *et al*.,2017). PAI apparently have been acquired during the speciation of pathogens from their nonpathogenic or environmental ancestors. The acquisition of PAI not only is an ancient evolutionary event that led to the appearance of bacterial pathogens on a timescale of millions of years but also may represent a mechanism that contributes to the appearance of new pathogens within a human life span (Schmidt and Hensel,2004; Hallstrom and McCormick,2015).

PAIs are specific regions on the bacterial chromosome where virulence genes accumulate. PAIs and their associated virulence genes spread among bacterial populations by horizontal transfer. Pathogenic bacteria, like extra-intestinal pathogenic *Escherichia coli*, host a variety of PAIs, some of which are bigger than 100 kb in size. Pathogenicity islands are a subclass of genomic islands that have a considerable impact on the development of these bacteria (Sarowska *et al*.,2019; Nougayrède *et al*.,2021).

With DNA often ranging from 10-200 kb, these islands exhibit distinctive traits, structural properties, and substantial gene clusters. Pathogenic strains have pathogenicity islands, which may be found on chromosomes or as a part of virulence plasmids, but nonpathogenic strains of the same species do not. PAI often differ from the core genome in their base composition and also show a different codon usage. The base composition is expressed as percentage of guanine and cytosine (G+C) bases, and the average G+C content of bacterial DNA can range from 25 to 75%. Most pathogenic bacterial species have G+C contents between 40 and 60%. The reasons for that variation are not known, but the conservation of a genus- or species-specific base composition is a remarkable feature of bacteria(Messerer *et al*.,2017; Desvaux *et al*.,2020).

It is considered that the horizontally acquired PAI still has the base composition of the donor species. On the other hand, it is also observed that the base composition of horizontally acquired DNA will gravitate to the base composition of the recipient's genome during evolution. Thus it is difficult to explain why “ancient” PAI still show a different base composition. Further factors such as DNA topology or specific codon usage of the virulence genes in PAI may also account for the maintenance of the divergent base composition (Messerer *et al*.,2017; Desvaux *et al*.,2020).

PAI are frequently located adjacent to tRNA genes. This observation gave rise to the hypothesis that tRNA genes serve as anchor points for insertion of foreign DNA that has been acquired by horizontal gene transfer. The frequent insertion at tRNA loci may be explained by the observation that genes encoding tRNAs are highly conserved between various bacterial species. After acquisition by horizontal gene transfer, a DNA fragment that contains a tRNA gene can insert into the recipient's genome by recombination between the tRNA genes. The second observation is that certain bacteriophages use tRNA genes as specific insertion points in the host genome. tRNA genes may represent specific anchor points for the integration of foreign DNA. (Messerer *et al*.,2017; Desvaux *et al*.,2020).

It is frequently associated with mobile genetic elements. They are often flanked by direct repeats which are defined as DNA sequences of 16 to 20 bp (up to 130 bp) with a perfect or nearly perfect sequence repetition. It may be served as recognition sites for the integration of bacteriophages, and their integration resulted in the duplication of the direct repeats. Furthermore, its act as recognition sequences for enzymes involved in excision of mobile genetic elements, thus contributing to the instability of a PAI flanked by it (Hallstrom and McCormick,2015).

Also, its often carry cryptic or even functional mobility genes such as integrases or transposases. Integrases, which may have been derived from lysogenic bacteriophages, mediate the integration of the phage genome into the genome of the host bacteria, such genes are still functional in certain PAI, and the encoded proteins can mediate the excision of the PAI and its loss (Hallstrom and McCormick,2015; Naderi *et al*.,2016).

Other PAI contain genes that are similar to integrase and resolvase genes of transposons, these mobile genetic elements can change their location within the chromosome, but transposons can also jump from a chromosomal location into a

plasmid and vice versa. Insertion sequence elements are frequently observed in PAI. Insertion of IS elements can result in the inactivation of genes, but the combination of two or more IS elements can also result in the mobilization of larger portions of DNA (Messerer *et al*.,2017; Desvaux *et al*.,2020).

The pathogenicity island carry clusters of virulence genes whose products contribute to the pathogenicity of the bacterium such as fimbriae, yersiniabactin siderophore system, Hemolysin, P fimbriae, aerobactin, and P-related fimbriae *ets*. In the case of *E. coli*, such islands have allowed the bacteria to adapt to specific environments and to cause disease (Naderi *et al*.,2016; Desvaux *et al*.,2020).

1.2.6.1 Polyketide islands in *E .coli* (pKs islands).

The polyketide pathogenicity island was first identified in *E. coli* strain IHE3034, which was isolated from neonatal bacterial meningitis. These *E. coli* strains, mostly from the B_2-phylogroup, harbor pathogenic islands known as Polyketide islands (pks). The sequence of this gene cluster is highly conserved among bacterial species. Epidemiological studies have shown that the species *Klebsiella pneumoniae, Citrobacter koseri, and Enterobacter aerogenes* all contain pks. (Auvray *et al*.,2021; Rehm *et al*.,2022).

The polyketide pathogenicity island exhibits a typical characteristic of horizontally acquired genomic components, suggesting that it was acquired from another source and inserted into the bacterial chromosome. This island has a different GC content compared to the chromosomal backbone of the bacteria. This difference in GC content is a common feature of horizontally acquired elements. This island is physically connected to a gene encoding a phage-type integrase. Integrases are enzymes that facilitate the insertion of mobile genetic elements, such as phages or genomic islands, into the host chromosome. In this case, the integrase

likely played a role in integrating of the pks gene island into the bacterial chromosome. It was inserted into the bacterial chromosome at the tRNA (asnW) locus. Additionally, the pks gene island flanked by two short direct repeats. These repeats are similar to those produced during the insertion of mobile genetic elements by integrases, further supporting the idea that the pks gene island was inserted by an integrase-mediated mechanism, the pks gene island does not contain genes involved in mobilization or transfer, meaning it cannot move independently between bacterial cells. Instead, it relies on transfer vehicles (e.g., plasmids, phages, or other mobile genetic elements) to be transferred between bacteria. The widespread distribution, prevalence, and evolutionary persistence of *pks* genes in this bacterial family suggest that this biosynthetic gene cluster may promote host fitness. (Dougherty *et al*.,2021; Rehm *et al*.,2022).

The pks contains four polycistronic transcripts, out of a total of seven, that are all oriented in a single direction. However, there is one exception, which is a single polycistron that encodes two important elements: the pantetheinyl transferase (encoded by the gene *clbA*, and the LuxR-type transcriptional activator encoded by the gene *clbR*. These two genes play a crucial role in promoting the expression of downstream pks genes (Auvray *et al*.,*20*21; Tang *et al*.,2022).

The pantetheinyl transferase enzyme is involved in the activation of the polyketide synthase complex, which is responsible for the biosynthesis of colibactin. The pantetheinyl transferase transfers a pantetheinyl moiety to specific domains within the PKS proteins, which is essential for their proper function in colibactin biosynthesis(Dougherty *et al*.,2021; Rehm *et al*.,2022).

The LuxR-type transcriptional activator (*clbR*) is a regulatory protein that controls the expression of other genes within the pks island. It acts as a transcription

factor, binding to specific DNA sequences in the promoter regions of downstream pks genes and enhancing their transcription. This activation of gene expression by ClbR is critical for the proper assembly and function of the colibactin biosynthetic machinery (Wami *et al*.,2021).

The expression of these regulatory elements in the pks island is important for governing colibactin production. The levels of expression of *clbA* and *clbR* are directly influenced by environmental metabolites. This means that the presence or absence of specific metabolites in the bacterial environment can impact the transcription of *clbA* and *clbR*, which in turn affects the expression of downstream pks genes involved in colibactin biosynthesis (Auvray *et al*.,*20*21; Tang *et al*.,2022). Moreover, some studies showed the presence of the Variable Number Tandem Repeat (VNTR) region located within the clb gene clusters, specifically at a region upstream of the *clb R* gene, which is comprise a varying number of a repeating octanucleotide sequence (5'-ACAGATAC-3') depending on the isolates. This region may play a role in the diagnosis and identification of different strains of *E. coli* relatedness according to the VNTR copy number (Putze *et al*.,2009; Wami *et al*.,2021).

The acquisition of pks islands in *E. coli* has been studied using whole-genome sequencing and phylogenetic analyses. These techniques have revealed two distinct pathways that differ based on the presence or absence of a phage-type integrase (integrase-mediated pks Insertion Pathway). This pathway occurs in B_2 phylogroup strains of *E. coli* and involves the integration of the pks island into one of three specific asn-tRNA genes(*asnU, asnV, or asnW*). Integrase enzymes facilitate the integration process. Integrase is a type of enzyme that mediates the incorporation of foreign DNA into the host genome, typically by catalyzing recombination events (Messerer *et al*.,2017; Auvray *et al*.,2021).

(Non-Integrase-Mediated pks Insertion Pathway) This pathway differs from the first one in that it does not involve the activity of a phage-type integrase. This pathway occurs in *E. coli* strains belonging to phylogroups B_1 and A. In contrast to the integrase-mediated pathway in phylogroup B_2 strains, chromosomal insertion of the pks island in these phylogroups occurs at a non-tRNA locus. It is believed that integrating the pks island in strains of these phylogroups involves the participation of insertion elements, particularly (IS66); insertion elements are mobile genetic elements that can move within the genome, causing genetic rearrangements. In this case, intact or truncated copies of IS66 were present, flanking the pks island, indicating their potential role in the integration process. These two distinct pathways suggest that the pks island's acquisition may have occurred through different genetic events and evolutionary processes. The integrase-mediated pathway appears to be more common in B_2 phylogroup strains, whereas other mechanisms might be involved in pks acquisition in non-B_2 strains. (Messerer *et al*.,2017; Auvray *et al*.,2021).

1.2.6.2 Colibactin (genotoxin).

The interaction between the complex cellular machinery of the human host and microorganisms, particularly the compounds produced by the human microbiota, plays a crucial role in maintaining health and influencing illness. Among the various microbiome compounds, colibactin stands out as one of the most intriguing and perplexing. Ii is a genotoxin that is encoded by commensal and extraintestinal microorganisms, including certain strains of *E. coli*. It has been linked to the development of colorectal cancer, a significant health concern worldwide. For more than 15 years since the colibactin cluster was first identified, a considerable amount of research has focused on understanding its biosynthesis, the precise chemical structure of the genotoxin, and its mechanism of carcinogenesis.

Researchers have been working to unravel the complex pathway involved in the production of colibactin and the steps it takes to synthesize this cryptic-encoded molecule (Bian *et al*.,2015;Zha *et al*.,2016; Zha *et al*.,2017).

Colibactin have a various effect on both eukaryotic and prokaryotic cells. Colibactin causes DNA crosslinks, this damage can lead to genetic mutations and alterations in the DNA structure. In prokaryotic cells, such as bacteria, the DNA damage caused by colibactin activates the SOS response pathway. This response is a cellular stress response that triggers the expression of various DNA repair and mutagenic genes to deal with the genotoxic stress. However, the activation of the SOS response in bacterial lysogens (bacteria harboring integrated bacteriophage DNA) can lead to the induction of prophages, that lead to the production of new phage particles, potentially lysing the bacterial host cell and releasing more phages. Also, Colibactin has been shown to modulate the expression of virulence genes in pathogenic bacterial species. Virulence genes are involved in the expression of factors that contribute to the pathogenicity and virulence of bacteria, enabling them to cause harm to the host. By affecting both eukaryotic and prokaryotic cells and modulating gene expression in pathogenic bacteria, colibactin can have significant consequences on human health (Balskus *et al*.,2015; Mousa *et al*., 2022).

Colibactin is a genotoxin produced as a secondary metabolite by polyketide island genes that encode hybrid non-ribosomal peptide synthetase, polyketide synthetase, and accessory proteins. This NRPS-PKS protein assembly produces linear biosynthetic intermediates termed precolibactin, which are characterized by an N-myristoyl-D-Aspergen prodrug motif (Dougherty *et al*.,2021; Auvray *et al*.,2021). Precolibactins are translocated into the periplasm via the multidrug and toxin extrusion transporter *ClbM* and then converted to genotoxic colibactin through the removal of prodrug motifs via the membrane-bound peptidase *ClbP*. The

removal of prodrug motifs causes spontaneous four-fold cyclization of linear precolibactins to yield the bioactive colibactin structure (Dougherty *et al*.,2021; Auvray *et al*.,2021).

Healy *et al*., (2019) proposed that some precolibactins, specifically precolibactin-712 and precolibactin-815, may not be solely derived from the clb assembly line offloading, as previously thought. Instead, the researchers suggest that these compounds may be partly derived from the degradation of downstream products.

Colibactin is composed of two nearly symmetrical subunits, each containing an electrophilic cyclopropane warhead with a high binding affinity for adenine-rich motifs within the DNA minor groove to generate interstrand cross-links, which have the potential to trigger genomic instability, as cells with deficiencies in DNA repair pathways or those utilizing error-prone repair pathways have a high risk of mutational acquisition following exposure to genotoxic stressors in their environment(García *et al*.,2016; Tang *et al*.,2022).

Acute exposure of human colonic organoids to *E. coli* harboring pks induces functional mutations related to p53 gene, providing direct evidence of oncogenic mutations following infection with this bacterium. Also, *E. coli* strains that are positive for pks may be relevant to cancer outside the intestine. These bacteria continually colonize extra-intestinal regions, and a specific category of individuals with head and neck, urinary tract, neuroendocrine tumors, and ovarian cancer have been found to have colibactin mutational signatures (Massip *et al*.,2019; Dougherty *et al*.,2021; Tang *et al*.,2022).

These findings suggest that colibactin may play a role in a variety of cancers beyond colorectal cancer, However, the presence of pks positive bacteria alone is

likely insufficient for cancer development, and the colibactin mutational signature is found in a large number of healthy individuals as well as cancer patients, further implying that the carcinogenic capacity of these microbes is derived from the combinatory production of colibactin and other aspects of ecology(Dougherty *et al*.,2021; Tang *et al*.,2022)

In order to exert a mutagenic influence of colibactin on the host epithelium, experimental evidence suggests that *pks*$^{+}$ *E. coli* likely requires two physiologic factors, include transcriptional activation of all *clb* genes (besides *clbS*), and cell-to-cell contact, a close association between colonizing *pks*$^{+}$ strains and host epithelial cells. Bacterial metabolism and biosynthesis of specific compounds is regulated through transcriptional mechanisms influenced by metabolic conditions in the environment (Chagneau *et al*., 2019; Oliero *et al*., 2021; Rehm *et al*.,2022).

1.2.6.3 Colibactin and DNA damage.

The pks island encodes a group of enzymes that work together to synthesize and release colibactin. Once produced, colibactin can exert its genotoxic effects on DNA, The exact mechanism of colibactin-induced DNA damage is not yet fully understood; however, it is believed to involve the formation of DNA adducts, which are covalent complexes between colibactin and DNA bases (Vizcaino.,2015; Bossuet-Greif *et al*.,2018; Li *et al*.,2019).

Colibactin has a high binding affinity for adenine-rich motifs (AAAATT) within the DNA minor groove, which triggers the formation of interstrand crosslinks where the two strands of DNA are covalently bonded, disrupting the typical structure and function of the DNA molecule. This can interfere with important cellular processes, such as replication and transcription (Tang *et al*.,2022).

Colibactin have the potential capacity to trigger genomic instability, as cells with deficiencies in DNA repair pathways or those utilizing error-prone repair pathways have a high risk of mutational acquisition following exposure to its genotoxic stressors in their environment (Vizcaino.,2015; Bossuet-Greif *et al*.,2018; Li *et al*.,2019).

DNA adducts formed by colibactin are abnormal chemical structures formed when colibactin binds to DNA, altering its typical structure and function that indirectly lead to DNA double-strand breaks. This suggests that DNA breaks may be the product of endogenous changes occurring after the formation of alkylated lesions generated by colibactin-mediated cross-linking. These breaks are severe forms of DNA damage that can lead to mutations that disrupt critical cellular processes, including cell cycle regulation and DNA repair mechanisms, eventually contributing to the accumulation of genetic changes that can drive genomic instability and potential long-term health consequences, and potentially contribute to the development of cancer (Gagnaire *et al.*,2017; Xue et al.,2020).

Colibactin-producing *E. coli*-induced inhibition of the MLH1 protein, a crucial component of the DNA mismatch repair pathway, can lead to genomic instability. The mismatch repair pathway is responsible for correcting errors that occur during DNA replication, and its impairment can result in the accumulation of mutations and chromosomal aberrations(Vizcaino.,2015; Bossuet-Greif *et al*.,2018; Li *et al*.,2019).

However, colibactin induces cell cycle arrest. Although cell cycle arrest is generally considered a protective mechanism to prevent the proliferation of damaged cells, it is also associated with cellular senescence. This is a state of permanent cell cycle arrest that can be triggered by various stressors, including DNA damage (Cougnoux *et al.*,2014; Eklöf *et al.*,2017).

The induction of microRNA-20a-5p by pks *E. coli* and its effect on post-translational modification of the p53 protein is another intriguing aspect. The p53 protein is a well-known tumor suppressor that plays a central role in regulating cell cycle arrest, DNA repair, and apoptosis (programmed cell death) in response to cellular stress, including DNA damage. However, the modification of p53 by microRNA-20a-5p in the context of colibactin-producing *E. coli* infection leads to cellular senescence, which contradicts the promotion of cancer. One suggested explanation for these conflicting observations is that cellular senescence can have both tumor-suppressive and tumor-promoting effects; it can act as a tumor-suppressive mechanism by preventing the proliferation of damaged cells, and it can also have paradoxical effects on cancer development. By promote tumor growth and aggressiveness through the senescence-associated secretory phenotype, which can create a pro-tumorigenic environment (Secher *et al*.,2013; Gagnière *et al*.,2017; Chagneau *et al*.,2021).

1.2.6.4. The assembly of colibactin

Colibactin is a member of a broad family of compounds with numerous activities and is produced via a hybrid non-ribosomal peptide synthetase-polyketide synthase assembly line. colibactine biosynthesis assembly line encoded within the pks island, representing 19 genes (*clbA*–*clbS*). This machinery comprises three polyketide mega synthases (PKS: ClbC, ClbI, and ClbO); three non-ribosomal peptide mega synthases (NRPS: ClbH, ClbJ, and ClbN); two hybrid NRPS/PKS megasynthases (ClbB and ClbK); and nine auxiliary, tailoring, and editing enzymes. In which the synthesized compound moves from one enzymatic module to the next. Colibactin is a component of a category of hybrid polyketide-non-ribosomal peptides that undergoes a prodrug activation process requiring the addition of a

structural motif at the N-terminus, which is then deleted in the last stage of manufacture(Healy *et al*.,2016; Trautman *et al*.,2017).

The NRPS and PKS enzymes are activated at the start of the assembly process by ClbA, a phosphopantetheinyl (PPant) transferase. Interestingly, ClbA participates in the synthesis of siderophores, enterobactin, and yersiniabactin. The possibility of trans-complementation between PPants in *E. coli* is thus supported by evidence, which suggests a link between the production of colibactin and siderophores (Brachmann *et al*.,2015).

Then, a building block was added to the enzyme modules in accordance with the domain architecture: for PKS, acetyl, malonyl, or methyl malonyl-CoA monomers; amino acid monomers; and amino acids that are either proteinogenic or non-proteinogenic for NRPS. During progression along the colibactin-assembly line, these building pieces are sequentially integrated into the synthesized molecule and changed by editing enzymes. It was discovered that ClbN utilizes asparagine to create N-myristoyl-d-Asparagine, a prodrug motif that is recognized by ClbB, which in turn adds L-amino acid to generate (malonyl-CoA) that integrates into the intermediate (Li *et al*.,2016)

The synthesis of precolibactin continues with the activity of two polyketide synthases (ClbC and ClbI), two non-ribosomal peptide synthases (ClbH and ClbJ), and hybrid PKs-NRPS (Clbk), which use malonyl-CoA as a substrate and amino acids, including Glysin, Cystin, and the L-Methionine-derived cyclopropane-containing amino acid (Zha *et al*.,2016; Zha *et al*.,2017). Cyclopropane (C3H3) production, which causes DNA alkylation by colibactin, involves the collaboration of ClbH and ClbI (Bian *et al*.,2013).

The aminomalonyl unit is recognized by ClbG, which then passes it to ClbK so that it can be incorporated into colibactin (Li *et al*.,2015) Interestingly, the ClbG enzyme can transmit the aminomalonyl unit to several pks assembly line enzymes (ClbC, ClbK, ClbO, and ClbI), raising the possibility that the aminomalonyl unit may be incorporated into colibactin more than once (Healy *et al*.,2016; Bossuet-Greif *et al*.,2018) Thiazole rings, which are heterocyclic structures with both sulfur and nitrogen (C3H3NS) produced by ClbK, are another structure thought to be significant for the biological function of colibactin(Garcie *et al*.,2016).

The final enzymatic module of the colibactin-assembly pathway may be the pKs enzyme, ClbO. However, colibactin production can result in several derailment. It is unclear how substances from the colibactin-assembly line are off-loaded. However, ClbQ, a type II thioesterase thought to be involved in this process, may be crucial for regulating the flux of colibactin synthesis (Zhou *et al*.,2021) The prodrug is taken up by ClbM(a multidrug and toxic compound extrusion transporter) and released into the periplasmic region once precolibactin production is complete (Jiang *et al*.,2019).

Precolibactin is matured by the ClbP peptidase once it enters the periplasmic region (Cougnoux *et al*.,2016) by removing the N-myristoyl-d-Asn side chain (Xue *et al*.,2019; Li *et al*., 2019)Nevertheless, it is still unknown how colibactin is exported from bacteria. to outside. Bacteria activate prodrugs in the periplasm to protect their DNA from the harmful effects of colibactin. To avoid the toxic effects of colibactin in the cytoplasm, bacteria produce ClbS, which can sequester colibactin and protect DNA (Molan *et al*.,2019) According to functional investigations, ClbS exhibits cyclopropane hydrolase activity, which converts genotoxic colibactin into a harmless substance products or precolibactin intermediates (Xue *et al*.,2018;Iftekhar *et al*.,2021).

Chromosome abnormalities and the generation of DNA double-strand breaks, both in vitro and in vivo, are hallmarks of colibactin. Colibactin may be either very unstable or weakly diffusible, as the genotoxic action of pks *E. coli* requires live bacteria and direct contact with epithelial cells (Zhou *et al*.,2021). Therefore, it has been speculated that colibactin-producing bacteria may have a significant impact on host health.

1.2.6.5. Other bioactivities of colibactin and *clb* metabolites.

Colibactin and related moieties conferring *E. coli* with a competitive advantage and increasing their colonization capacity in the intestine. Additionally, the pks island, which is responsible for colibactin production can exert various effects beyond genotoxicity due to the production of different colibactin and precolibactin molecules. These compounds exhibit a wide range of bioactivities, including cytotoxic, antibacterial, pro-inflammatory, and analgesic effects. PKs posative *E. coli* strains can display cytotoxic effects on various cell types. For instance, precolibactin-886 has been found to exhibit moderate cytotoxicity to HeLa cells at high concentrations (Faïs *et al*.,2018).

Colibactin and related moieties display antibiotic activity against pathogenic bacteria such as *Staphylococcus aureus*. The small molecule N-myristoyl-D-Aspergen, which is released by *ClbP* from precolibactin, has been shown to inhibit the growth of *Bacillus subtilis*. This antibacterial activity can provide *E. coli* with a niche advantage leading to enhanced colonization capacity. Also, precolibactin-derived compounds may have analgesic effects such as C12-Asn-GABA, which depend on *clbA*, *clbB*, and *clbN*. These molecules have the ability to cross the epithelial barrier in the intestine and interact with sensory

neurons. By blocking the GABAB receptor on these neurons, they can decrease visceral pain in the host (Pleguezuelos-Manzano *et al*.,2020).

Additionally, these compound have Inflammation-promoting and anti-inflammation bioactivities. Different pks-positive *E. coli* strains may exhibit varying effects on inflammation, ranging from promoting inflammation to having anti-inflammatory properties. These diverse bioactivities could be attributed to the variant colibactin moieties produced by various colibactin synthesis gene clusters. Colibactin-producing *E. coli* strains can create a pro-inflammatory microenvironment, leading to the production of pro-inflammatory cytokines and other immune responses. *E. coli* strain CFT073: This specific pks-positive *E. coli* strain has been studied extensively for its interactions with the host immune system. It has been found that CFT073 induces the release of IL-1β from human neutrophils, leading to an inflammatory response. such as increasing reactive oxygen species (ROS) production and phagocytosis. These enhanced neutrophil responses may contribute to the host's defense against *E. coli* infection. Conversely, proinflammatory cytokines such as TNF-α, IL-1β, IL-6, IL-8, and IFN-γ can trigger increased growth of *E. coli* strain CFT073. This response facilitates the colonization of the bacterium in the inflamed intestine, which could be relevant during periods of intestinal inflammation (Koh *et al*,2015;.Koh *et al*,2017).

The pks island has been associated with not only pro-inflammatory effects but also anti-inflammatory activities. This is due to the involvement of ClbA, a protein encoded by the *clbA* gene in the pks island, in the synthesis of other anti-inflammatory products. One of the key functions of ClbA is its involvement in siderophore biosynthesis. Siderophores are small molecules that act as iron chelators, which means they have a high affinity for binding iron. In iron-limited environments, such as the gut, *E. coli* can use siderophores to scavenge and import

iron, an essential nutrient for its survival and growth. By acquiring iron through siderophores, E. coli can enhance its ability to persist and thrive in iron-restricted conditions. Furthermore, the presence of self-resistance gene clbS inactivates colibactin and protects the bacterial DNA from nucleolytic degradation, potentially acting as a mechanism to limit the genotoxic effects of colibactin and dampen inflammation.The complex interplay between colibactin production, siderophore biosynthesis, and other activities of the pks island leads to diverse effects on inflammation. This intricate balance of pro-inflammatory and anti-inflammatory properties likely contributes to the overall outcome of the host-microbe interaction and can impact the progression of disease or the establishment of commensalism (Koh *et al*,2015; Koh *et al*,2017; Molan *et al*,2019).

1.2.7 Quinolones.

Quinolones are a class of synthetic antibiotics that were initially developed in the 1960s. The prototypical quinolone antibiotic is nalidixic acid, which was discovered accidentally when chloroquine was being synthesized. Its act by inhibiting bacterial topoisomerase type II enzymes, which are essential for bacterial replication. Initially, their use was limited due to the narrow spectrum of activity, low serum concentrations achieved, high inhibitory concentration required, and several adverse effects (Bidell *et al*.,2016). The basic structure of almost all quinolones is a (1,4-dihydro-4-oxo-pyridine) molecule, Fluoroquinolones are a subclass of quinolones that contain an additional fluorine atom in their chemical structure. The addition of fluorine enhances their potency and broadens their spectrum of activity against bacteria. They are effective against a wide range of bacteria, including Gram-positive, Gram-negative, and atypical bacteria. Examples of fluoroquinolones include ciprofloxacin, levofloxacin, moxifloxacin, and norfloxacin (Pham *et al*.,2019).

1.2.7.1. Fluroquinolone.

Fluoroquinolones are a class of synthetic antibiotics that are widely used to treat various bacterial infections. They are called "fluoroquinolones" because of the presence of a fluorine atom in their chemical structure, which enhances their activity against bacteria. These antibiotics work by inhibiting the action of bacterial enzymes called DNA gyrase and topoisomerase IV, which are essential for DNA replication and repair in bacteria. fluoroquinolones are commonly classified into four generations based on their antimicrobial spectrum and pharmacological properties. Each generation has seen improvements in terms of potency and spectrum of activity (Sousa *et al*.,2014; Correia *et al;*.,2017; Brar *et al*.,2020).

All quinolones have a carboxylic substituent at position 3, which, together with the carbonyl group at position 4, appears to be fundamental for their activity. The first generation of fluroquenolone, represented by (ciprofloxacin and norfloxacin, displays two major changes with respect to nalidixic acid: the first was the presence of a positively charged group at position seven (piperazine group in ciprofloxacin) and the second was the presence of fluoride at position six,This generation of quinolones was active against all aerobe gram-negative bacteria and showed moderate activity against some gram-positive bacteria(Correia *et al;*.,2017; Brar *et al*.,2020).

Levofloxacin is a second-generation fluoroquinolone that exhibits improved activity against Gram-positive bacteria compared to the first-generation. It retains good activity against Gram-negative bacteria as well. Due to its broad-spectrum coverage, levofloxacin is commonly used to treat respiratory tract infections, skin and soft tissue infections, urinary tract infections, and some sexually transmitted infections. and the third generation showed activity against anaerobes. The representative of this generation is moxifloxacinis which commonly used to

treat respiratory tract infections, sinusitis, pneumonia, skin infections, and intra-abdominal infections. (Sousa *et al*.,2014; Bidell *et al.*,2016).

1.2.7.1.1 Fluoroquinolone mechanism of action.

DNA gyrase is an essential bacterial enzyme responsible for catalyzing negative supercoiling of DNA. It helps in the maintenance of the DNA's helical structure by introducing negative supercoils ahead of the replication fork during DNA replication and transcription. The negative supercoiling is crucial for various cellular processes, including DNA compaction, gene regulation, and DNA repair. By inhibiting DNA gyrase, fluoroquinolones disrupt the normal supercoiling process, leading to the accumulation of positive supercoils, DNA damage, and eventually inhibiting bacterial DNA replication and transcription (Correia *et al*.,2017).

Topoisomerase IV is another type II topoisomerase that plays a critical role in DNA replication and separation of daughter DNA molecules after replication is completed. It is responsible for decatenation, which is the process of untangling and unlinking the interlinked DNA strands that arise during replication. Topoisomerase IV resolves the intertwined DNA strands and separates them into two separate daughter DNA molecules, allowing the bacteria to complete the cell division process and form two independent daughter cells. By inhibiting topoisomerase IV, fluoroquinolones interfere with this crucial step, preventing the successful separation of daughter DNA molecules, leading to DNA damage, cell cycle arrest, and eventually bacterial cell death(Gubaev *et al*.,2016).

Both enzymes (DNA gyrase & Topoisomerase IV) have a similar structure, consisting of two-A subunits and two B subunits. The A subunit of DNA gyrase is encoded by *gyrA* and the B subunit by *gyrB*. The A subunit of DNA gyrase is responsible for the catalytic activity of breaking and rejoining the DNA strands, while the B subunit possesses ATPase activity. The ATPase activity of the B subunit hydrolyzes ATP to provide the energy required for the breaking and resealing reactions catalyzed by the A subunit. Topoisomerase IV contains an A subunit encoded by the *parC* gene and a B subunit encoded by the *parE* gene. The A subunit carries out the catalytic reaction of breaking and rejoining the DNA strands, while the B subunit provides ATPase activity to supply the necessary energy for these reactions. (Correia *et al*., 2017; Kumar *et al*., 2017).

Fluoroquinolones have a high affinity for the active site of the A subunit of DNA gyrase and topoisomerase IV. They specifically interact with a conserved region within the enzyme's active site, which is essential for its catalytic activity, and thus fluoroquinolones prevent the enzyme from catalyzing the breakage and rejoining of DNA strands. This inhibits the enzyme's ability to introduce negative supercoiling into DNA during replication and transcription, disrupting crucial cellular processes. The binding of fluoroquinolones induces a conformational change in these enzymes molecule. This change alters the interaction between DNA gyrase and DNA, causing the DNA itself to become another binding site for fluoroquinolones. This dual binding further stabilizes the fluoroquinolone-DNA gyrase complex. The stable complex of fluoroquinolones with DNA gyrase interferes with the normal progression of DNA replication and transcription. Bacterial DNA cannot be efficiently replicated or transcribed, leading to impaired cell division and eventual bacterial cell death (Brar *et al*.,2020; Pham *et al.*,2019).

1.2.7.1.2 Fluoroquinolone mechanisms of resistance.

Fluoroquinolone resistance in bacteria, including *E. coli*, has been on the rise over the past few decades. Various mechanisms have emerged to confer resistance to these antibiotics, making their treatment more challenging.

Chromosomal mutations in the genes encoding the A subunit of DNA gyrase (*gyrA*) and the A subunit of topoisomerase IV (*parC*) are a significant mechanism of fluoroquinolone resistance. The regions in these genes where mutations occur are referred to as quinolone resistance-determining regions (QRDRs). These mutations alter the structure and function of the enzymes, reducing the binding affinity of fluoroquinolones to their target sites. As a result, the fluoroquinolones become less effective in inhibiting DNA gyrase and topoisomerase IV, leading to bacterial resistance (Röderova *et al*.,2017; Todorović *et al*.,2018; Vinué *et al*.,2018).

The amino acid codons Ser-83 and Asp-87 in the gyrA gene are frequently affected by mutations in fluoroquinolone-resistant *E. coli* strains. Substitutions in these codons can lead to changes in the structure of the DNA gyrase enzyme, reducing the binding affinity of fluoroquinolones to the enzyme's active site, Similarly, mutations at the amino acid codons Ser-80 and Glu-84 in the parC gene are commonly associated with fluoroquinolone resistance in E. coli. These mutations lead to altered topoisomerase IV activity, However, mutations in the (*gyrB and parE*) genes play a minor role in the acquisition of resistance to quinolones in *E. coli* (Röderova *et al*.,2017; Todorović *et al*.,2018; Vinué *et al*.,2018).

Another mechanism of resistance to fluoroquinolones, is related to the reduced intracellular accumulation of these antibiotics. This reduced accumulation can result from alterations in the bacterial cell's outer membrane permeability or increased efflux of the drug from the bacterial cell. Some fluoroquinolones, particularly the hydrophilic ones like ciprofloxacin, rely on the presence of porins to

pass through the outer membrane of Gram-negative bacteria, including E. coli. Porins are protein channels that facilitate the entry of hydrophilic molecules into the bacterial cell. Alterations or reductions in the expression of porins can hinder the entry of certain fluoroquinolones, leading to decreased intracellular accumulation and reduced efficacy against the bacteria.

Bacteria possess efflux pumps, which are specialized transport systems that actively pump out various substances, including antibiotics, from inside the cell to the external environment. Overexpression of efflux pumps can lead to the rapid removal of fluoroquinolones from the bacterial cell before they can exert their bactericidal effect. This efflux process results in reduced intracellular drug concentrations and contributes to antibiotic resistance (Correia *et al.*,2017). The expression of the AcrAB-TolC efflux pump, encoded by the acrAB operon, is a significant mechanism of fluoroquinolone resistance in *Escherichia coli* and other bacteria. This efflux pump is an integral part of the bacterial defense system, as it actively pumps out a wide range of compounds, including fluoroquinolones, from inside the bacterial cell to the external environment (Hernando-Amado *et al*.,2016; Reygaert *et al*.,2018).

Plasmid-mediated mechanisms of fluoroquinolone resistance play a significant role in the dissemination of resistance genes among bacteria. Three important plasmid-mediated mechanisms of fluoroquinolone resistance have been described. Some bacteria can express aminoglycoside-modifying enzymes that can also modify fluoroquinolones. These enzymes acetylate an amino group located in the piperazine ring of the fluroquinolone structure(Röderova *et al*.,2017; Aworh *et al*.,2023).

This modification affects all fluroquinolones, except those that have a blocked amino group, such as levofloxacin. By modifying fluoroquinolones, the

aminoglycoside-modifying enzymes reduce the antibiotics' binding affinity to their target enzymes, DNA gyrase and topoisomerase IV, leading to resistance. The *qnr* gene family encodes a peptide known as the Qnr protein,which can protect DNA gyrase and topoisomerase IV complexes from binding with fluoroquinolones, reducing the inhibitory effect of these antibiotics on the enzymes. This protection allows the enzymes to continue their normal functions. Efflux pumps, such as OqxAB and QepA, are plasmid-encoded transporters that actively pump fluoroquinolones out of the bacterial cell. These pumps recognize and expel fluoroquinolones, reducing the intracellular drug concentration and leading to decreased antibiotic efficacy (Vila *et al*.,2016; Rodríguez-Martínez *et al*.,2016; Correia *et al*.,2017; Decano *et al*.,2021). Escherichia coli sequence type 131 (ST131) has gained attention as a globally disseminated and clinically important clone due to its high rates of resistance to fluoroquinolone antibiotics. making it a major public health concern. (Price *et al*.,2013; Ludden *et al*.,2020).

2 Materials and methods:

2.1 Materials

2.1.1 Laboratory Equipment and Instrument

The main equipment's and instruments used throughout this study are listed in Table (2-1).

Table (2-1) Laboratory Equipment and Instrument

No	Instrument	Company	Country
1	Autoclave	Stermite Olympus A&B	Japan
2	Eppendrof centrifuge	Fisons	England
3	Micropipettes 10-100 µl ,100-1000 µl, 0.5 – 10 µl	Eppendorf	Germany
4	Digital camera	redmi	china
5	Distillator	GFL	Germany
6	DNA extraction tubes	Eppendorf	Germany
7	Freezer	Aucma	Japan
8	Gel electrophoresis	Clever	China
9	Hood	Labogene	Danemark
10	Incubator	Memmert	Germany
11	Light microscope	Stermite Olympus A&D	Japan
12	Benson burner		India
13	Platinum wire loop	Himedia	India
14	Oven	Memmert	Germany
15	Sterile swab for streaking	Lab. Service	S.P.A.

16	PCR thermocycling	Invitrogen	USA
17	PCR tubes	Eppendorf	Germany
18	Refrigerator	Concord	Italy
19	UV-trans illuminator	Clever	USA
20	Vortex	Germmy	Twain
21	Water bath	Memmert	Germany
22	Sensitive electron balance	Stermite Olympus A&D	Japan
23	Millipore filter (0.45mm)	Satorins membrane filter Gm ,BH ,W.	Germany

2.1.2 Chemical and Biological Materials:

The main chemical materials that used in this study are listed in Table (2-2).

Table (2-2) Chemical Materials

No	Chemicals	Company/Country
1	Absolute ethyl alcohol	Fluka/ Germany
2	Agar-agar	Oxoid/UK
3	DNA ladder marker	Bioneer /Korea
4	Ethidium bromid , Agarose , Master mix	Bioneer /Korea
5	Glycerol	Fluka /England
6	Gram Stain Kit	Crescent/KSA
7	NH_4Cl, $MgSO_4$, NaCl, glucose	Merk Darmstadt/ Germany

8	DdH2o (50ml)	Bioneer /Korea
9	Nuclease,protienase free water (100) ml	Himedia\ india
10	Tris-Borate-EDTA (TBE10X) buffer	

2.1.3 Culture Media:

The media have been used for culture in this study are listed in table (2-3).

Table (2-3):- Culture media and their manufacturer.

No.	Culture Media	Manufacture / Origin
1	Brain Heart Infusion Agar	BBL /France
2	Brain Heart Infusion Broth	BBL /France
3	MacConky Agar	Himedia /India
4	Agarose	Himedia /India
5	Mueller-Hinton Agar	Mast /UK
6	Nutrient Agar	Himedia /India
7	Nutrient Broth	Himedia /India
8	Blood agar base	Himedia /India
9	Eosin methylene blue	Himedia /India

2.1.4 Commercial Kits:

The commercial kits used in the present study in the table (2-4).

Table (2-4)Commercial kits used in the present study

No.	*Type of Kit*	*Company/Country*
1	DNA extraction Kit	Geneaid/UK
2	DNA ladder 2000bp	Bioneer /Korea
	DNA ladder 10000bp	Solgent\Korea
3	AccuPower PCR PreMix	Bioneer /Korea
	Solg^{tm}2x Taqplus PCR Smart mix 1	Solgent\Korea
4	Primer of 16S rDNA, *chuA*, *yjaA*, *TspE4C2*, mdh36, gyrB47, *clbA*, *clbQ*, *clbB*, *clbN*	Bioneer /Korea

Table (2-5): DNA extraction kit (Geneaid/UK).

Materials
GT buffer 30 ml
GB buffer 40 ml
W1 buffer 45 ml
Wash buffer 25 ml +100 ml Ethanol
Elution buffer 30 ml
Proteinase K.
bsolute ethanol
RNase A

Table (2-6): Master Mix Used in PCR.

Materials	**Company/Country**
Top DNA polymerase (1U)	Bioneer /Korea
DNTP(each250µM)	
Reaction buffer with 1.5Mm-$MgCl_2$(1x)	
Solg^{tm}2x Taqplus PCR Smart mix 1	Solgent\Korea

2.2 Methods

2.2.1 Preparation solutions

2.2.1.1 1x TBE from 10XTBE:

The solution prepared by diluting a 10X TBE (Tris-Borate-EDTA) solution with water (D.W.) to create a 1X TBE solution. To calculate the concentration of the final solution, Initial volume of 10X TBE added = 10 ml to final volume after dilution = 10 ml (10X TBE) + 90 ml (D.W.) = 100 ml, TBE is a commonly used buffer in molecular biology and electrophoresis, and it is usually prepared at a higher concentration (e.g., 10X) and then diluted to the desired working concentration (e.g., 1X).

2.2.3 Preparation of cultural media:

2.2.3.1. MacConkey agar

This medium was used to isolate and differentiate between bacteria that fermented and non-fermented lactose, prepared by Measure 51.5 grams of the medium's components, dissolve the measured medium in 1 liter of distilled water, heat the mixture to boiling with frequent agitation to ensure complete dissolution of

the medium components, sterilize the medium by autoclaving, which uses high pressure and steam to kill any microorganisms present in the medium, After autoclaving, cool the medium to around 55 °C, Incubate the cooled medium at 37 °C for 24 hours to check for any signs of contamination. If contamination is present, it will become apparent during the incubation period (MacFaddin, 2000).

2.2.3.2. Brain heart infusion broth

Brain Heart Infusion (BHI) broth is a liquid medium used for the cultivation and propagation of a wide range of microorganisms, especially used for culturing fastidious microorganisms that have specific nutritional requirements, propagating bacteria for diagnostic testing and research purposes, Preparing bacterial inocula for susceptibility testing (antibiotic testing) and growing bacteria for serological and biochemical identification. which was prepared by dissolving (37 g) of the medium in (1 L.) of distilled water (MacFaddin 2000).

2.2.3.3. Brain heart infusion agar

This medium was used for activation, prepared by weight (52 g) from media, dissolved in (1 L.) of distilled water, heated to boiling with frequent agitation to completely dissolve the medium, sterilized by autoclaving, cooled to 55 °C, and poured into Petri dishes, which were then incubated at 37 °C for 24 h to remove any contaminating medium (MacFaddin, 2000).

2.2.3.4. Nutrient agar

The medium was prepared according to the method suggested by the manufacturing instructions. The medium (28 g) was dissolved in (1 L.) of distilled water, heated to boiling with frequent agitation to completely dissolve the medium,

sterilized by autoclaving, cooled to 55 °C, poured into Petri dishes, and used for cultivation of bacterial isolates when necessary (McFaddin, 2000).

2.2.3.5. Eosin methylene blue (EMB) medium:

Eosin methylene blue (EMB) agar medium was prepared according to the method recommended by the manufacturer's instructions. When lactose-fermenting bacteria grow on EMB agar, they produce acid as a byproduct of lactose fermentation. The acid causes the dyes (eosin Y and methylene blue) to form a complex, leading to the appearance of colonies with a purple or dark color. *E. coli* that vigorously ferment lactose produce large amounts of acid and form colonies with a characteristic greenish metallic sheen under reflected light (Murray et al., 2003).

2.2.3.6. Blood agar medium:

Blood agar medium was prepared by dissolving 40 g of blood agar base in 1000 ml distal water. The mixture was autoclaved at 121 °C for 15 min and then cooled to 50 °C. Fresh human blood (5 %) was added. This medium was used to cultivate bacterial isolates and to determine the ability of bacteria to induce hemolysis in blood cells (McFaddin, 2000).

2.2.3.7. Maintenance medium

Maintenance media for bacterial isolates: The bacterial isolates were preserved on brain heart infusion agar slants at 4 °C. The isolates were maintained monthly during the study by culturing in new culture media. For long-term preservation, brain heart infusion broth supplemented with 20% glycerol was used, and the isolates were preserved frozen (-20 ºC) for the long term (several months) (Collee *et al*., 1996).

2.2.3.8 Agarose Gel:

The preparation of an agarose gel according to the method described by Sambrook and Rusell (2001), Add 1.8 grams of agarose to 200 ml of 1x TBE buffer, Heat the mixture to boiling until all agarose particles are dissolved, Allow the solution to cool down to 50-60 ºC, Mix the cooled agarose solution with 2 µl of ethidium bromide. Once the agarose gel solution is mixed with ethidium bromide, it is ready to be poured into the gel casting tray with an appropriate comb (for creating wells) and allowed to solidify.

2.2.4. Study Design

A Cross-sectional study was designed to include 150 patient's women with vaginitis, who attended Babylon Teaching Hospital for Maternity and Children and private clinic during a period of six months (from February 2022 to July 2022).

2.2.5 Ethical approval: -

This study was approved by the Committee of publication ethics at College of Medicine, University of Babylon, Iraq. Verbal consent is taken from each patients before sampling.

2.2.6 patients and specimens collection.

A total of 150 high vaginal and cervical swabs were collected from female's patients with excessive vaginal discharge, irritation, pain and itching, in proper ways to avoid contamination, the swabs were inserted into the posterior fornix and upper part of the vagina and rotated there before the withdrawal. A Cusco's speculum was used to provide a clear view of the cervix, and the swabs were rubbed in and around the cervix and withdrawn without contaminating the vaginal wall, the swabs were then placed in tubes containing transport medium to maintain sample moisture until they were taken to the laboratory. The swabs were inoculated on the culture media

and incubated aerobically at 37°C for 24h. Pregnant, women subjected to chemotherapy and women with chronic diseases have been excluded from this study.

2.2.7. Isolation and identification of *E. coli*

2.2.7.1. Culturing of samples.

The specimens are swabbed onto MacConkey agar plates, which is a selective and differential medium used to isolate and differentiate gram-negative bacteria, particularly those that can ferment lactose. Then agar plates are incubated at 37°C for 24 hours. After the incubation period, the plates are examined. Suspected colonies with specific characteristics, such as rosy or pink coloration, are selected. These colors indicate lactose fermentation, a key characteristic used to distinguish certain groups of bacteria. A single colony from each selected primary culture on the MacConkey agar is then transferred to EMB agar. Which is a selective and differential medium used for the isolation and differentiation of gram-negative bacteria based on their ability to ferment lactose. The identification process starts with observing the morphological properties of the bacterial colonies on EMB agar. This includes noting characteristics like colony size, shape, color, translucency, edge, and elevation of texture. These properties can provide initial clues about the bacteria's identity (Forbes *et al*., 2007).

2.2.7.2. Gram's stain

A single colony from nutrient agar was taken by a loop and mixed with a drop of distilled water, spread on a clean slide, waited to dry, fixed by heat, stained with gram stains, and then examined the slide under the microscope (oil immersion) to observed the bacterial cell according to (Jawetz *et al*., 2015) **.**

2.2.7.3. Hemolysin production:

Hemolysin production was carried out by inoculating blood agar medium with bacterial isolates at 37 °C for 24 h. The appearance of a clear zone around the colonies is referred to as complete hemolysis (β-hemolysis). The appearance of a greenish zone around the colonies referred to partial hemolysis (α-hemolysis), whereas no change in the zone referred to non-hemolysis (γ-hemolysis) (Forbes *et al*., 2007).

2.2.8 Antibiotic sensitivity test

A- Bacterial isolates

a total of 27 isolates belong to the species of *E. coli* were tested for fluoroquinolone susceptibility test.

B- Preparation media for Antibiotic sensitivity test.

To prepare Mueller-Hinton agar for performing antibiotic sensitivity tests, Weigh the quantities of the required media based on the desired volume of the agar medium and the manufacturer's instructions (38 g\L). Typically, the ingredients are weighed and dissolved in distilled water in a suitable flask then heat the mixture to dissolve the ingredients completely. Once dissolved, bring the medium to a boil to ensure that all components are well-mixed. Sterilize the medium by autoclaving at 121°C and 15 psi for about 15-20 minutes. Allow the medium to cool down to approximately 45-50°C. At this temperature, the medium is still in a liquid state and is safe to handle. Before the medium solidifies, pour it into sterile, flat Petri dishes to create agar plates. Allow the agar plates to cool and solidify completely at room temperature. Once solidified, the agar plates are ready for use. If not used immediately, the prepared Mueller-Hinton agar plates should be stored in a refrigerator (4°C) until needed.

C- Inoculum preparation

Preparing the bacterial inoculum for antibiotic sensitivity testing involves picking up four to five bacterial isolates selected for testing. Each bacterial isolate is streaked onto separate BHI agar plates using a sterile loop. The inoculated BHI agar plates are incubated at the appropriate temperature (usually 35-37°C) for 16 to 18 hours. After incubation, the bacterial growth from each isolate's colony is harvested. This can be achieved by gently scraping the surface of the agar using a sterile loop or swab and suspending the harvested cells in a sterile saline solution. The bacterial density of each harvested isolate is adjusted to match a specific turbidity standard, typically equivalent to a 0.5 McFarland standard (Miniatis *et al*. 1982; Mahon &Lehman.,2022).

D- Disk diffusion method.

To performing the disk diffusion method, using a sterile swab or inoculation loop, streak the bacterial suspension evenly across the entire surface of a Mueller-Hinton agar plate. Allow the plate to dry for a few minutes to let the bacteria adhere to the agar surface, by sterile forceps, place the antibiotic disks containing specific antibiotics onto the agar surface. Gently press down on the disks to ensure they are in contact with the agar, Allow the plates to stand for a few minutes at room temperature to let the antibiotics diffuse into the agar. Afterward, the plates are incubated at 37°C for a defined period (usually 16-18 hours), After the incubation period, observe the plates for zones of inhibition around the antibiotic disks. Measure and record the zone diameters using a ruler or a zone reader. Zones are measured from the edge of the disk to the point where bacterial growth is inhibited. Compare the zone diameters obtained with interpretive criteria provided by the Clinical and Laboratory Standards Institute (CLSI) Results are categorized as "Sensitive," "Intermediate," or "Resistant" based on the size of the inhibition zone and the specific antibiotic. The antibiotic diffuses from the disk into the surrounding agar,

creating a concentration gradient. If the bacteria are susceptible to the antibiotic, they will not grow near the disk, creating a clear zone of inhibition around it. The larger the zone of inhibition, the more effective the antibiotic is against the bacteria (Miniatis *et al*. 1982; Mahon &Lehman.,2022).

Table (2-7): Antibiotic disc potency was used in this study.

Name of antibiotic	Symbol	Disc potency	Company-country
Ciprofloxacin	**Cip**	**5 µg\ml**	**Liofilchem-italy**
levofloxacin	**Levo**	**5 µg\ml**	**Liofilchem-italy**
Norofloxacin	**Noro**	**10 µg\ml**	**Liofilchem-italy**

E- Prepare of modified Mueller-Hinton agar with Gatifloxacin.

Follow the manufacturer's instructions to prepare the Mueller-Hinton agar base. Typically, mix the agar powder with distilled water to create a homogeneous mixture. Autoclave the Mueller-Hinton agar medium to sterilize it. Allow the sterilized Mueller-Hinton agar to cool to approximately 45-50°C. At this temperature, add the gatifloxacin antibiotic to the medium. The concentration of gatifloxacin will depend on the intended use and the sensitivity of the bacteria being tested. Two concentrations have been used (75-30 µg/mL.). Stir or swirl the medium gently to ensure even distribution of gatifloxacin throughout the agar. Pour the gatifloxacin-supplemented Mueller-Hinton agar into sterile Petri dishes. Use appropriate aseptic techniques to avoid contamination. The prepared agar can be used for antibiotic susceptibility testing to determine the susceptibility or resistance of bacterial isolates to gatifloxacin.

F- **Prepare modified Mueller-Hinton agar with Moxifloxacin.**

Follow the manufacturer's instructions to prepare the Mueller-Hinton agar base. Typically, mix the agar powder with distilled water to create a homogeneous mixture. Autoclave the Mueller-Hinton agar medium to sterilize it. Allow the sterilized Mueller-Hinton agar to cool to approximately 45-50°C. At this temperature, add the Moxifloxacin antibiotic to the medium. The concentration of Moxifloxacin will depend on the intended use and the sensitivity of the bacteria being tested. Two concentrations have been used (75-30 µg/mL.). Stir or swirl the medium gently to ensure even distribution of Moxifloxacin throughout the agar. Pour the Moxifloxacin -supplemented Mueller-Hinton agar into sterile Petri dishes. Use appropriate aseptic techniques to avoid contamination. The prepared agar can be used for antibiotic susceptibility testing to determine the susceptibility or resistance of bacterial isolates to Moxifloxacin.

2.2.9 The identification of the tested organism (*E. coli*) conformed by using 16s rDNA.

The 16S rDNA gene is a widely used target for bacterial identification and phylogenetic analysis. After the amplification, the PCR product can be analyzed using gel electrophoresis or other methods to verify the successful amplification of the target 16S rDNA gene fragment (Magray *et al*.,2011).

A- Primers:

- Forward primer: 5’GATTAGATACCCTGGTAG3’
- Reverse primer: 5’AGTCACTTAACCATACAACCC3’

B- PCR Reaction Mixture (50 µl):

- 7 µl of genomic DNA: This is the template DNA used for amplification.
- 25 µl of smart mix 1: Provides essential ions and pH control for the PCR reaction.
- 2.5 U of Taq DNA polymerase: The enzyme responsible for DNA synthesis during PCR.
- 1.5 mM MgCl2: Magnesium chloride, an essential cofactor for Taq DNA polymerase activity.

C- PCR Amplification Conditions:

1. Initial denaturation: 94ºC for 3 minutes.
    - This step ensures complete denaturation of the double-stranded DNA to create single-stranded templates for the subsequent PCR cycles.
2. Denaturation: 94ºC for 30 seconds (for each cycle).
    - High temperature denaturation separates the DNA strands, making the template available for primer binding.
3. Annealing: 55ºC for 30 seconds (for each cycle).
    - At the lower annealing temperature, the primers bind (anneal) to the complementary regions on the template DNA.
4. Extension: 72ºC for 10 minutes (for each cycle).
    - At the extension temperature, Taq DNA polymerase synthesizes a new DNA strand using the annealed primers as starting points.
5. Number of cycles: 30 cycles.
    - The denaturation, annealing, and extension steps are repeated 30 times to amplify the target DNA fragment.
6. Final extension: 72ºC for 10 minutes.
    - This step allows any remaining incomplete DNA strands to be fully extended.

2.2.7. Molecular analysis steps

2.2.7.1 Extraction of genomic DNA from bacterial culture

It's important to strictly follow the instructions provided by the DNA purification kit manufacturer (Geneaid Genomic DNA Purification Kit\bioneer, Korea), to ensure the best results and maintain the integrity of the genomic DNA. Additionally, proper aseptic techniques should be employed throughout the process to prevent contamination of the DNA sample.

A- Bacterial Culture: inoculating the *E. coli* isolates into 10 ml of Brain Heart Infusion (BHI) broth. The BHI broth provides a rich medium to support bacterial growth.

B- Incubation: Incubate the bacterial culture at 37°C overnight in a shaking incubator. The incubation period allows the bacteria to grow and reach an adequate cell density for DNA extraction.

C- Harvesting the Bacterial Cells: After overnight incubation, the bacterial culture is ready for DNA extraction. The culture is harvested by transferring the entire 10 ml of the overnight culture into a centrifuge tube.

D- Centrifugation: Centrifuge the culture at a suitable speed and time to pellet the bacterial cells. This step separates the bacterial cells from the liquid supernatant.

E- DNA Extraction steps.

1- By using a sterile pipette, transfer the appropriate volume of the bacterial suspension (containing up to 1 x 10^9 cells) to a 1.5 ml microcentrifuge tube. Be sure to handle the suspension aseptically to avoid contamination. Place the

microcentrifuge tube containing the bacterial suspension into the centrifuge rotor. Centrifuge the tube at a speed of 14,000-16,000 x g (gravitational force) for 1 minute. The high centrifugal force causes the bacterial cells to pellet at the bottom of the tube. After centrifugation, carefully remove the microcentrifuge tube from the centrifuge. Using a pipette or by gently pouring, carefully discard (remove and dispose of) the supernatant from the microcentrifuge tube. Take care not to disturb the bacterial pellet at the bottom of the tube during this step.

2- The bacterial cell pellet is obtained after centrifugation and add 180 µl of GT buffer to it. The GT buffer contains guanidine thiocyanate, which is a chaotropic agent that helps in cell lysis and the denaturation of cellular proteins. Resuspend the bacterial cell pellet in GT buffer thoroughly by either using a vortex mixer or pipetting up and down. Ensure that the cell pellet is well-mixed and fully resuspended in the GT buffer ,prepare a Proteinase K solution by adding appropriate volume of sterile ddH_2O. The Proteinase K enzyme helps in digesting the cellular proteins, allowing for the release of genomic DNA. Add the prepared Proteinase K solution to the re-suspended bacterial cell suspension (in GT buffer). Ensure that the Proteinase K is well-mixed with the bacterial cell suspension. Then incubate the tube at 60ºC for at least 10 minutes. This incubation period at a moderate temperature allows the Proteinase K to digest the cellular proteins, facilitating cell lysis and releasing the genomic DNA. During the incubation, invert the tube every 3 minutes. This gentle mixing ensures even distribution of the Proteinase K and helps in maximizing cell lysis and DNA release. After the incubation period, the genomic DNA is released into the solution.

3- Only 200 µl of GB Buffer was added then mixed by vortex for 10 seconds and incubated at 70ºC for at least 10 min. to ensure the sample lysate is clear during the incubation, then inverted the tube every 3 min.

4- absolute ethanol 200 µl is directly added to the lysed bacterial sample. Ethanol is used to precipitate the genomic DNA, allowing it to be purified from the rest of the sample. Mix the sample immediately by shaking the tube vigorously. This helps in ensuring thorough mixing of the ethanol with the sample, allowing DNA precipitation. If a precipitate appears break it up as much as possible by using a pipette. This step helps ensure that the genomic DNA is fully dissolved in the ethanol, maximizing the DNA yield. Place the GD Column in a 2 ml Collection tube. Then, transfer the entire mixture (including any insoluble precipitate) to the GD Column. The GD Column is a chromatography column designed to capture and purify genomic DNA. Centrifuge the GD Column at 14,000-16,000 RPM for 2 minutes. The centrifugation step separates the DNA from the rest of the sample components, allowing the genomic DNA to bind to the GD Column. After centrifugation, discard the 2 ml Collection tube containing the flow-through. The flow-through is the liquid that has passed through the GD Column and contains components other than genomic DNA. Transfer the GD Column to a new 2 ml Collection tube. This tube will collect the purified genomic DNA during the next steps.

5- W1 Buffer 400 µl is added directly to the GD Column. W1 Buffer is a wash buffer used to remove impurities and contaminants that may have bound to the column during the previous steps. Centrifuge the GD Column at 14,000-16,000 RPM for 30 seconds. After centrifugation, discard the 2 ml Collection tube containing the flow-through. Transfer the GD Column back to a clean 2 ml Collection tube. This tube will collect the purified genomic DNA during

the elution step. The washing step with W1 Buffer helps ensure the removal of residual contaminants, salts, and other impurities that may interfere with downstream applications. After the washing step, the purified genomic DNA remains bound to the GD Column.

6- In this step 600 µl of Wash Buffer (with ethanol) was added to the GD Column, and then centrifuged at 14-16,000 RPM for 30 seconds, discarded the flow-through and placed in the GD Column was back in the 2 ml Collection tube, and centrifuged again for 3 min. at 14-16,000 RPM to dry the column matrix. The drying step is essential because ethanol or other solvents can interfere with downstream applications, such as PCR or DNA sequencing. By completely removing the ethanol, you ensure the purity and quality of the genomic DNA for further experiments.

7- Pre-heat the Elution Buffer to 70ºC. It is essential to use a pre-heated elution buffer to ensure efficient elution of the genomic DNA. Transfer the dried GD Column to a clean 1.5 ml microcentrifuge tube. The GD Column contains the purified genomic DNA bound to the column matrix. Add 100 µl of the pre-heated Elution Buffer directly to the center of the column matrix. The elution buffer will release the genomic DNA from the column matrix. Allow the Elution Buffer to stand in the column for at least 3 minutes. During this time, the elution buffer is absorbed by the column matrix, facilitating the release of the genomic DNA. Centrifuge the GD Column at 14,000-16,000 x g for 30 seconds. The centrifugation step will elute the purified genomic DNA from the column matrix into the microcentrifuge tube. The eluted DNA will be present in the liquid phase at the bottom of the tube. The DNA elution step is crucial for obtaining a highly concentrated and purified genomic DNA sample.

2.2.7.2 The mixture of PCR reaction:

Amplification of DNA was carried out in final volume of 20 µl and 50 µl containing the following as mentioned in Table (2-8):

Table (2-8) Contents of the Reaction Mixture

No.	Contents of reaction mixture	Volume
1.	AccuPower PCR PreMix	Dried pellet
2.	Forwarded primer	2 µl
3.	Reverse primer	2 µl
4.	DNA template	5µl
5.	Nuclease free water	11µl
Total volume		20µl
1.	Solg™ 2x Taq PLUS pcr Smart mix 1	25 µl
2.	Forwarded primer	2.5 µl
3.	Reverse primer	2.5 µl
4.	DNA template	7 µl
5.	Nuclease free water	13 µl
Total volume		50 µl

2.2.7.3 Primer Sequences:

Molecular assay in this study includes 10 genes, each one has specific nucleotide and product size. The primer sequences and PCR conditions that used are listed in Table (2-10),(2-11),(2-12).

Table (2-9) The primer sequences and PCR conditions of E. coli Phylogenetic groups

Target gene	P	Primer sequence (5′-3′)	bp	PCR condition	Reference
chuA	F	ATGGTACCGGACGAACCAAC	288	94 ºC 4min 94 ºC 5 Sec 59 ºC 20Sec 30x 72 ºC 5min 72 ºC 1 min	Clermont et al.,2013
	R	TGCCGCCACTACCAAAGACA			
yjaA	F	CAAACGTGAAGTGTCAGGAG	211		
	R	AATGCGTTCCTCAACCTGTG			
TspE4C2	F	CACTATTCGTAAGGTCATCC	152		
	R	AGTTTATCGCTGCGGGTCGC			

Table (2-10) The primer sequences and PCR conditions of PKs for E. coli .

Target gene	P	Primer sequence (5′-3′)	bp	PCR condition	Reference
clbA	F	CTAGATTAT CCGTGGCGA TTC	1000	94 ºC 1min 94 ºC 30Sec 58 ºC 30Sec 30x 72 ºC 1min 72 ºC 5 min	**(Johnson et al., 2000; Kawanishi et al.,2020)**
	R	CAGATACAC AGATACCAT TCA			
clbQ	F	TTATCCTGTTAGCTT TCG TTC	850		
	R	CTTGTATAG TTACACAAC TAT TTC			

Table (2-11) The primer sequences and PCR conditions of colibactin of E. coli

Target gene	P	Primer sequence (5′-3′)	bp	PCR condition
clbB	F	GATTTGGAT ACTGGCGAT AAC CG	579	95 ºC 5min 95 ºC 30Sec 55 ºC 30Sec 30x 72 ºC 1min 72 ºC 5 min
	R	CCATTTCCC GTTTGAGCA CAC		
clbN	F	GTTTTGCTC GCCAGATAGTCA TTC	733	95 ºC 5min 95 ºC 30Sec 55ºC 30Sec 30x 72 ºC 1min 72 ºC 5 min
	R	CAGATACAC AGATACCAT TCA		

Table (2-12): Combination of *chuA*, *yjaA*, and *TspE4.C2* Genes for Determination of Phylogenetic Subgroups of *Escherichia coli* (Barzan *et al*.,2017).

Phylogenetic subgroups	Genes combination profile		
	Chu A	Yja A	TspE4
A0	-	-	-
A1	-	+	-
B1	-	-	+
B22	+	+	-
B23	+	+	+
D1	+	-	-
D2	+	-	+
untenable	-	+	+

2.2.7.4 Detection of Amplified Products by Agarose Gel Electrophoresis:

analysis of the PCR products by agarose gel electrophoresis. Which is a widely used technique in molecular biology to separate and visualize DNA fragments based on their size.

A- Preparation of Agarose Gel:

- Dissolve 1.8 g of agarose powder in 200 ml of 1x TBE buffer (10 ml 10x TBE + 90 ml sterile distilled water).
- Heat the mixture on a hot plate until it reaches boiling to ensure the agarose is completely dissolved.
- Allow the agarose solution to cool down to approximately 45°C.
- Add 2 µl of Ethidium bromide stain to the agarose solution. Ethidium bromide is a DNA intercalating dye that binds to DNA and allows visualization of DNA bands under UV light.

B- Creating Wells in the Gel:

- Fix the comb at one end of the gel tray to create wells for loading DNA samples.
- Gently pour the agarose solution into the tray, making sure no air bubbles are trapped.
- Allow the agarose to solidify at room temperature.
- Carefully remove the comb from the solidified gel to leave wells for sample loading.

C- Electrophoresis Setup:

- Place the solidified agarose gel in an electrophoresis chamber.
- Fill the chamber with TBE buffer to cover the surface of the gel.

D- Loading DNA Samples:

- Carefully transfer 5 µl of each DNA sample into the designated wells in the agarose gel. In the first well, load the DNA. The DNA ladder contains DNA fragments of known sizes used as molecular weight markers for size estimation.

E- Electrophoresis Run:

- Connect the electrophoresis chamber to a power supply.
- Run the electrophoresis at 65 volts for five minutes to allow the samples to settle in the wells.
- Continue the electrophoresis at 70 volts for approximately 30 minutes or until the DNA fragments have migrated to the desired positions on the gel.

F- DNA Visualization:

- After electrophoresis, remove the gel from the chamber.
- Place the gel on a UV transilluminator, which emits UV light to visualize the DNA bands.
- Observe the DNA bands under UV light, and if available, photograph the gel using a digital camera or gel documentation system.

2.2.8 Sequencing method

2.2.8.1 Optimization of PCR conditions for *clb B*, and *clb N* colibactin gene.

This step preformed to optimize the annealing temperature and cycle number, to find the conditions that result in the most specific and efficient amplification of the target DNA (*clb B*, and *clb N*). The selected conditions should produce a strong and clear band of the expected size on the gel.

A- Annealing Temperature Optimization:

- A gradient PCR was performed, where multiple reactions were set up, each with a different annealing temperature (55, 58, 60, 63 & 66). The range typically spans a few degrees below and above the estimated Tm (melting temperature) of the primers.
- PCR reactions were run for each annealing temperature, and the products were analyzed by gel electrophoresis as shown in Figure (2-1)&(2.2).
- The annealing temperature that resulted in the strongest, specific PCR product band without non-specific amplification (primer dimers or smearing) was selected as the optimal annealing temperature.

B- Amplification Cycle Optimization:

- Several PCR reactions were set up with varying numbers of amplification cycles. Common cycle numbers tested might include 25, 27, 30, 35, and 40 cycles.
- The PCR reactions were run for the different cycle numbers, and the products were analyzed as before.
- The cycle number that provided sufficient amplification of the target DNA without reaching the plateau phase (where amplification levels off and nonspecific products may increase) was chosen as the optimal cycle number. the best thermo-cycling conditions for (*clb B*, and *clb N*) were listed in table (2-13) these conditions produce the most specific and sufficient PCR product.

Table (2-14): Optimized thermo-cycling condition for *clb B*, and *clb N* gene.

Target gene	P	Primer sequence (5′-3′)	bp	PCR condition	Reference
clbB	F	GATTTGGAT ACTGGCGAT AAC CG	579	95 ºC 5min 95 ºC 30Sec 63ºC 30Sec 27x 72 ºC 1min 72 ºC 5 min	(Johnson et al., 200; Kawanishi et al.,2020)
	R	CCATTTCCC GTTTGAGCA CAC			
clbN	F	GTTTTGCTC GCCAGATAGTCA TTC	733	95 ºC 5min 95 ºC 30Sec 60 ºC 30Sec 30x 72 ºC 1min 72 ºC 5 min	
	R	CAGATACAC AGATACCAT TCA			

2.2.8.2 The mixture of PCR reaction for sequencing:

Amplification of DNA was carried out in final volume of 25 µl containing the following as mentioned in Table (2-13):

Table (2-15) Contents of the Reaction Mixture

No.	Contents of reaction mixture	concentration	Volume
1.	Master mix	2.5x	10µl
2.	forward primer	10 PM	1 µl
3.	revers primer	10PM	1 µl
4.	DNA template	10-20 ng\µl	2 µl
	$Mgcl_2$	25mm	0.5µl
5.	Nuclease free water		10.5µl
Total volume			25µl

2.2.8.3 Nucleic acids sequencing of PCR amplicons.

A- PCR Amplification: Initially, PCR was performed to amplify the specific DNA region of (*clb B*, and *clb N*) using the extracted DNA samples as templates. This PCR step aimed to create multiple copies of the target region for subsequent sequencing.

B- Sequencing from Both Terminations: After PCR amplification, the PCR amplicons were sent to (Macrogen Inc. in Seoul, South Korea) The sequencing company used Sanger sequencing, a widely used method for DNA sequencing.

D- Chromatographs and ABI Sequence Files: During the sequencing process, chromatographs were generated for each sequenced fragment. Chromatographs are graphical representations of the DNA sequence data, showing peaks corresponding to the different nucleotides.The chromatographs were obtained from ABI (Applied Biosystems) sequence files.

E- Quality Control: To ensure the accuracy and reliability of the sequencing results, performed quality control checks on the chromatographs. They likely used specific criteria to select only the high-quality chromatographs for further analysis.

F- Annotation and Variant Analysis: The high-quality chromatographs were analyzed to identify the nucleic acid sequences of the PCR amplicons. compared the observed nucleic acid sequences from the chromatographs with the expected sequences to confirm the accuracy of the sequencing.

G- Virtual Positioning: The "virtual positions" of the retrieved PCR fragments refer to their relative positions within the larger DNA sequence (e.g., a gene or a genome). By aligning the retrieved sequences with the known reference sequences, could determine where the PCR fragments belonged in the context of the entire DNA sequence.

2.2.8.4 Interpretation of sequencing data

With the help of BioEdit Sequence Alignment Editor Software Version 7.1 (DNASTAR, Madison, WI, USA), the sequencing results of the PCR products of the targeted samples were edited, aligned, and analyzed. Variations observed in each sequenced sample were mapped to their corresponding positions within the referring genome as well as their PCR amplicons. The observed nucleic acids were numbered both in the PCR amplicons and their corresponding positions within the referring genome. SnapGene Viewer ver. 4.0.4 (https://www.snapgene.com) was used to annotate each variant within the bacterial sequences.

2.2.8.5 Comprehensive phylogenetic tree construction

In this study, a specific comprehensive tree was created using the neighbor-joining procedure provided by (Sarhan *et al*., 2019). Using the NCBI-BLASTn server, the detected variations were compared to their nearby homologous reference sequences (Zhang et al. 2000). Finally, using the iTOL suit, a full inclusive tree, including the observed variant, was constructed and shown as a circular cladogram (Letunic and Bork, 2019). The sequences of each included species in the comprehensive tree were color-coded.

الخلاصة:

تم جمع ١٥٠ مسحة مهبلية من النساء المصابات بالتهابات مهبلية، تم فحص العينات لتحديد وعزل الاشريشية القولونية باستخدام تقنية (PCR technology that targets 16s rDNA) ، تم تحديد المجموعات التطورية باستخدام تقنية PCRالثلاثية بناء على وجود او عدم وجود ثلاث جينات chuA, yjaA, and TspE4.C2)) ، كذلك تم تحديد الجينات الدالة على وجود جزر الامراضية pks والمنتجة للكوليبكتين (clb A, clbQ, clbB and clbN)،تم اختيار سبعة عزلات من مجموع تسعة إيجابية لل pks لتحديد المتغيرات الجينية في تسلسل قطعتين محددتين من PCR جزئيا جينات (clbB and clbN) ، تم تحديد التسلسل الجيني ١٣١ في عينات الاشؤيشيا القولونية المعزولة باستخدام تقنية PCR التي تستهدف التغاير في جينات كل من (mdh and gyrB) ، كما تم اختبار الحساسية للمضادات الحيوية الفلوركوينولون لتحديد أنماط المقاومة. كانت نسبة الاشريشية القولونية المعزولة ٢٧(١٨%)من العينات المهبلية من بين ١٥٠ عينة ،وقد أوضحت النتائج ان معضم العزلات تنتمي الى مجموعة النشوء والتطور B2 ،٢١(٧٧,٧%)، وان باقي العزلات تنتمي الى مجموعتي النشوء والتطور A و b_1 .٣(١١,١%)، تسعة (٣٣,٣٪) من ٢٧ عزلة *E. coli*حملت جزيرة pks ، معظمها تنتمي إلى مجموعة النشوء والتطور B2(٨) وعزلة واحدة تنتمي الى مجموعة النشوء والتطور B_1.أظهرت نتائج المحاذاة لجين clbB 579 bp وجود متغير واحد للحمض النووي يمثله استبدال واحد للحمض النووي (٦٢ (C> T مقارنة بتسلسل الحمض النووي المرجعي. لم ينتج عن هذا المتغير بدائل الأحماض الأمينية في البروتين المشفر . تم إنشاء شجرة نسالة شاملة بناءً على اختلافات الحمض النووي التي لوحظت في منطقة ٥٧٩ نقطة أساس من أمبليكون .clbBلوحظ ميل طفيف بين العينات S2 و S3 و S6 و S7 والسلالات ذات الصلة. يُعزى هذا الاختلاف الطفيف في الموضع إلى الاختلاف المكتشف (٦٢ درجة مئوية> (Tفي العينات التي تم فحصها. أظهرت نتائج المحاذاة لعينة الجين clbN 733 bp وجود متغير واحد للحمض النووي يمثله استبدال واحد للحمض النووي (٢٤٣ (C> Tمقارنة بتسلسل الحمض النووي المرجعي. تم إنشاء شجرة نسالة شاملة بناءً على اختلافات الحمض النووي التي لوحظت في منطقة ٧٣٣ نقطة أساس من أمبليكون .clbNلوحظ ميل طفيف بين العينات S1 و S6 والسلالات المرتبطة بها. يُعزى هذا الاختلاف الطفيف في الموضع إلى التباين المكتشف (٢٤٣ درجة مئوية> (Tفي كلتا العينات التي تم فحصها. اظهر فحص بكتريا *E. coli st 131* أن ١١ عزلة (٤٠,٧٪) تنتمي لنوع التسلسل ١٣١. أظهر التحليل الوراثي أن الغالبية تنتمي إلى phylogroup B2 (90.9٪ (، وأن عزلة واحدة تنتمي إلى phylogroup B1 (9.09٪ .(أظهرت المضادات الحيوية أن معظم عزلات الإشريكية القولونية كانت مقاومة للفلوروكينولونات

والسيبروفلوكساسين والنورفلوكساسين والليفوفلوكساسين. في المقابل ، أظهروا حساسية للمضادات الحيوية gatifloxacin و moxifloxacin ، حيث كانت بعض العزلات المرتبطة بالإشريكية القولونية ST131 مقاومة لاثنين أو أكثر من مضادات الفلوروكينولون.

وزارة التعليم العالي والبحث العلمي

جامعة بابل كلية الطب

التنوع الشجيري التطوري لجزر الامراضية لمجموعات مركبات البولكتايد بين عزلات *E.coli* المعزولة من التهابات المهبل

رسالة

مقدمة الى مجلس كليه الطب / جامعة بابل

وهي جزء من متطلبات نيل درجة الدكتوراه فلسفة في

العلوم/الأحياء المجهرية الطبية

من قبل

نور إسماعيل ناصر حسين

بكالوريوس في الاحياء المجهرية /كلية العلوم/جامعة بغداد/2005

ماجستير في الاحياء المجهرية الطبية/كلية الطب/جامعة القادسية/٢٠١٠

اشراف

أ. د. محمد صبري عبد الرزاق

اشراف

أ. د. نادية مضر الحلي

1444ه

2023 م