Characterization of siderophores from *Escherichia coli* strains through genome mining tools: an antiSMASH study

Levent Cavas¹,²* and Ibrahim Kirkiz¹

**Abstract**

Although urinary tract infections (UTIs) affect many people, they are usually a disease observed in women. UTIs happen when exogenous and endogenous bacteria enter the urinary tract and colonize there. Cystitis and pyelonephritis occur when bacteria infect the bladder and the kidneys, respectively. UTIs become much serious if the bacteria causing the infection are antibiotic resistant. Since the pathogenic microorganisms have been adopted to current antibiotics via genetic variations, UTIs have become an even more severe health problem. Therefore, there is a great need for the discovery of novel antibiotics. Genome mining of nonpathogenic and pathogenic *Escherichia coli* strains for investigating secondary metabolites were conducted by the antiSMASH analysis. When the resulting secondary metabolites were examined, it was found that some of the siderophores are effective in UTIs. In conclusion, since the siderophore production in *E. coli* is directly related to UTIs, these molecules can be a good target for development of future pharmaceutical approaches and compounds. Siderophores can also be used in industrial studies due to their higher chelating affinity for iron.

**Key points**

- Genome mining on nonpathogenic and pathogenic *E. coli* was studied.
- Comprehensive and comparative analysis of siderophores were investigated.
- The results may open a new gate on the development of new drugs on pathogenic *E. coli*-based diseases.

**Keywords:** Bioinformatics, *Escherichia coli*, Genome mining, Siderophores, Urinary tract infections

**Introduction**

Urinary tract infections (UTIs) are the most common among infections transmitted by pathogenic *Escherichia coli* strains. UTIs have become a global health problem due to the difficulty of detecting antibiotic-resistant pathogens (Ndzime et al. 2021). Almost half of the women and 12% of men have this infection once in their lifetime (Tabatabaie et al. 2022). Due to the hormonal and anatomical changes that occur during pregnancy, women are more likely to get UTIs (Kalinderi et al. 2018). Although there has been no evidence of human-to-human transmission of this infection, consumption of water and food contaminated with *E. coli* can cause outbreaks in communities (Manges et al. 2001). Because UTIs caused by *E. coli* do not show obvious symptoms, they are often seen as a harmless or short-term affliction. If it is not taken seriously, they can cause cystitis, pyelonephritis, and bacteriuria (Foxman 2002). The *E. coli* strains are presented

*Correspondence: levent.cavas@deu.edu.tr

¹ The Graduate School of Natural and Applied Sciences, Department of Biotechnology, Dokuz Eylul University, Kaynaklar Campus, 35390 Izmir, Türkiye

Full list of author information is available at the end of the article
by pathogenic and nonpathogenic characteristics. Nonpathogenic and commensal *E. coli* strains are usually in a symbiotic relationship with its host and have important roles in the human gastrointestinal tract. Therefore, the nonpathogenic strains rarely cause diseases (Bien et al. 2012). However, the pathogenic strains can also colonize the intestines. These strains differ from each other according to whether they contain the gene region responsible for virulence factors. *E. coli* Nissle 1917 (EcN) is a nonpathogenic and commensal *E. coli* strain that benefits humans without any harm. EcN is used for treating diseases such as gastrointestinal and UTIs by competing with pathogens (Schultz 2008; Ou et al. 2016). On the other hand, Beatson et al. (2015) reported that *E. coli* VR50 causes urinary tract infections via gene modifications. UPEC strains contain many genes that produce virulence factors to increase pathogenicity. The more virulence factors an organism expresses, the more serious infections it can cause. The organisms use virulence factors to attach, invade and attack the host. The virulence factors include adhesins, toxins, siderophores, protective polysaccharide coatings, invasins to aid colonization and increase the severity of the infection (Yun et al. 2014; Beaton et al. 2015).

Iron is essential for living organisms as it is a cofactor for many cellular processes. Electron transport system, oxidative phosphorylation, citric acid cycle, superoxide metabolism and DNA/RNA synthesis are among these processes. It is also involved in the production of toxins, antibiotics and siderophores (Fardeau et al. 2011; Negash et al. 2019). The organisms produce siderophores to acquire essential iron from the environment in response to iron deficiency. Siderophores are low molecular weight organic chelators that have a very specific affinity for Fe (III). The iron affinity of the siderophores is so high that they remove iron from molecules, which iron binds to such as ferritin, transferrin and lactoferrin (Ratledge and Dover 2000; Li et al. 2016). Because of this feature majority of the siderophores have gained importance due to their virulence effects on pathogens. In fact, pathogens that produce an excessive number of siderophores have a higher virulence (Yun et al. 2014; Beaton et al. 2015). Siderophores have been identified as virulence factors by experimental procedures that vary and require a single enzymatic reaction. All NIS enzymes carry a N-terminal iron uptake chelate (*IucA/IucC*) domain and have a C-terminal domain related to iron transport or metabolism (Oves-Costales et al. 2009).

Nowadays, natural compounds form the basis of new therapeutic drugs. Microorganisms producing secondary metabolites contain biosynthetic gene clusters (BGCs) in which more than one gene is located close to each other. Therefore, BGCs and the natural compounds they produce have great therapeutic potential. Bacterial, fungal and plant secondary metabolites are pharmacologically effective compounds that are used for developing new drugs (Prihoda et al. 2021). Genome-based drug discovery approach is used to reveal gene clusters that synthesize bioactive compounds and to propose novel therapeutic drugs from these bioactive compounds. At this point, tools such as antiSMASH that identifies BGCs have gained great importance (Mushtaq et al. 2018). When the whole genome sequence of a microorganism is enlightened, it may be a pioneer for secondary metabolite studies (Kim et al. 2017). For this reason, the importance of genome mining studies in secondary metabolite production has been recognized (Albarano et al. 2020). The determination of secondary metabolites is carried out by experimental procedures that vary and require time. However, bioinformatics tools such as antiSMASH provide results quickly. AntiSMASH is a rapid and reliable source for finding gene clusters responsible for the biosynthesis of secondary metabolites (Villebro et al. 2019; Medema et al. 2011). AntiSMASH analysis also give detailed information of the secondary metabolites predicted. Thus, the types of secondary metabolites that cause pathogenicity are revealed (Zotchev et al. 2012). This could lead to novel methods of treating people with diseases such as UTIs. In diseases that are difficult to diagnose such as urinary tract infections, changing the treatment method according to the type of the metabolite will make it easier to get rid of the pathogens that cause the disease.
No scientific publication has so far been reported on the identification of siderophores in UPEC by using antiSMASH technology. The PUBMED Search results are given in Table 1.

When the whole genome of a microorganism is revealed, the regions related to secondary metabolite production can be predicted. Genome mining has an important place in predicting these metabolites (Villebro et al. 2019; Albarano et al. 2020). As the importance of secondary metabolites becomes clear, genome mining tools begin to be developed. With the increase in genomic data, genome mining tools used for secondary metabolite production has become indispensable (Kim et al. 2017).

The discovery of biosynthetic gene clusters related to siderophore synthesis in pathogenic and nonpathogenic E. coli strains by using genome mining technologies was aimed at investigating in the present study.

**Materials and methods**

**Prediction of the siderophore gene clusters**

antiSMASH webserver was used to investigate siderophore biosynthetic genes (Version 6.0.1). antiSMASH has been created by Blin et al. (2021) for the discovery of secondary metabolites in complete genome or metabolic gene clusters. The default parameters were used for the antiSMASH analysis with relaxed detection strictness. AntiSMASH known clusters and sub clusters were also investigated in the present study.

**BLASTp analysis**

Also, BLASTp webserver was used to find sequences with similar gene regions. BLASTp analysis compares nucleotide or protein sequences to sequence databases and calculates the statistical significance of matches. BLAST can be used to infer functional and evolutionary relationships between sequences and help identify members of gene families (Johnson et al. 2008). BLASTp analysis was performed to search in the non-redundant protein sequences from the NCBI database. The criteria used to determine the sequences according to the BLASTp results were the e-value is \( \leq 0.01 \) and percent identity is \( \geq 98\% \).

**Nucleotide sequence accession numbers**

The complete genome sequence of E. coli Nissle 1917, E. coli K-12 strain MG1655 and E. coli VR50 were retrieved from the GenBank (www.ncbi.nlm.nih.gov) database under accession numbers CP007799.1, U00096.3 and CP011134.1, respectively.

**Results**

AntiSMASH analysis of pathogenic E. coli VR50 genome

E. coli VR50 has four genomic regions for the biosynthesis of secondary metabolites according to the antiSMASH analysis (Fig. 1). When these regions are studied one-by-one, Region 1.1 was found to be responsible for the biosynthesis of NRPS. NRPS synthesize nonribosomal peptides, which are secondary metabolites produced by bacteria. The NRPS in this region is responsible for

---

**Table 1** The searched key words of the study (04.04.2022)

| Searched Key Words                                                                 | Items in PUBMED |
|-----------------------------------------------------------------------------------|-----------------|
| (Urinary tract infection)                                                         | 76,193          |
| ((Urinary tract infection) AND (siderophile))                                     | 200             |
| (((Urinary tract infection) AND (siderophile)) AND (Escherichia coli))           | 121             |
| (((Urinary tract infection) AND (siderophile)) AND (Escherichia coli)) AND (antiSMASH) | 0               |

---

**Fig. 1** Identified secondary metabolite regions in E. coli VR50 (NCBI Accession number: CP011134.1) and similar gene clusters
enterobactin biosynthesis. Additionally, there are gene clusters related to enterobactin biosynthesis and transport in this region. In Region 1.2, gene clusters responsible for the biosynthesis of thiopeptide were found. Thiopeptides are natural products with antibiotic effects to eliminate competitive microorganisms in the same medium (Chan and Burrows 2021). Region 1.3 contains the NRPS/PKS domain. These domains show NRPS and PKS related functions. Polyketide synthases (PKSs) are proteins or protein complexes that produce a large variety of secondary metabolites found in bacteria, fungi, plants (Weng and Noel 2012). In this region, there are gene clusters responsible for yersiniabactin biosynthesis. Region 1.4 is responsible for the biosynthesis of aerobactin. Additionally, Fig. 1 shows gene clusters containing genes like *E. coli* VR50 genomic regions and their metabolites are shown with similarity rates.

**antiSMASH analysis of nonpathogenic *E. coli* strains**

antiSMASH analysis was performed to compare the secondary metabolites of pathogenic *E. coli* VR50 and nonpathogenic *E. coli* Nissle 1917 (O6:K5:H1) and *E. coli* K-12 strain MG1655. Complete genome sequences of *E. coli* Nissle 1917 and *E. coli* K-12 were retrieved from the NCBI database.

As Fig. 2 shows, *E. coli* Nissle 1917 has 6 genomic regions responsible for different secondary metabolites. Region 1.1 synthesizes NRPSs and these NRPSs are responsible for enterobactin biosynthesis and transport. Region 1.2 contains genes for thiopeptide synthesis. Region 1.3 is aryl polycenes biosynthetic gene cluster. Aryl polycenes are polyunsaturated carboxylic acids that can be found in both nonpathogenic and pathogenic strains (Johnston et al. 2021). Region 1.4 is responsible for NRPS and PKS synthesis. There are genes related to aerobactin biosynthesis in Region 1.5. The latter contains the same lysine 6-monoxygenase and siderophore gene clusters (*IucA/IucC*). These gene clusters were also observed in the *E. coli* VR50 genome. Region 1.6 is involved in the synthesis of NRPS-like molecules.

In Fig. 3A, enterobactin-related genes are located at 660,180–671,984 nt. Region 1.1 contains TonB-dependent receptor family, enterobactin esterase, enterobactin synthase and six enterobactin transport genes.

As for the Region 1.5 aerobactin biosynthetic region, there are three aerobactin biosynthetic genes, TonB-dependent receptor, and lysine/ornithine N-monoxygenase gene (Fig. 3B).

There are 2 genomic regions in *E. coli* K-12 strain MG1655. *E. coli* K-12 strain MG1655 does not have a genomic region in its genome that is responsible for NRPS-independent siderophores (Fig. 4). Although when the two genomic regions were examined, genes responsible for enterobactin transport were found in Region 1.1 (Fig. 5A. These regions are blue colored genes between 610,254 and 624,510 nt.). TonB-dependent siderophore receptor family, enterobactin/ferric enterobactin esterase, ferric enterobactin transport proteins (FepE) and ATP-binding protein are the genes responsible for enterobactin transport. TonB-dependent siderophore receptor family and ATP-binding protein are also found in pathogenic *E. coli* VR50 genome.

As can be seen from Fig. 5B, this region has two thiopeptide biosynthetic genes and many putative genes.

**Comprehensive antiSMASH analysis of *E. coli* strains**

To see the comparability of the study, 7 pathogenic and 4 nonpathogenic *E. coli* strains were analyzed with the antiSMASH (Table 2). According to the results, it was observed that similar siderophores were produced in different strains. As can be seen in Table 2, nonpathogenic *E. coli* strains synthesize fewer types of siderophores, and
among these siderophores, yersiniabactin is less common. On the other hand, all examined UPEC strains contain yersiniabactin. Enterobactin siderophore was found in all pathogenic and nonpathogenic strains.

**Siderophore genomic region in E. coli VR50**
Region 1.1 of the *E. coli* VR50 genome contains a core NRPS biosynthetic gene and multiple other genes. There is an enterobactin synthase gene and eight enterobactin transport-related genes (Fig. 6A). These are blue colored genes between 608,930 and 623,796 nt.

As can be seen from Fig. 6B, Region 1.2 has two thiopeptide biosynthetic genes and many enzymes related genes like dimethyl sulfoxide reductase and transferases. NRPS and PKS in Region 1.3 are yersiniabactin biosynthetic genes. As we mentioned before, yersiniabactin is the most synthesized siderophore by pathogenic strains. Other than NRPS and PKS regions, there are three yersiniabactin biosynthetic protein genes between 2,149,729 and 2,153,210 nt and one TonB-dependent receptor gene at 2,153,341–2,155,362 nt (Fig. 6C).

As can be seen from Fig. 6D, there are 12 genes in the Region 1.4 siderophore biosynthetic gene cluster in *E. coli* VR50 due to antiSMASH analysis. This siderophore biosynthetic gene cluster is located at 3,374,667–3,389,081 nt. In this region there are three siderophore biosynthesis protein genes, one *IutA* and one *IucD* gene. These 12 genes contain hypothetical proteins, TonB-dependent siderophore receptor family, lysine/ornithine N-monooxygenase, *IucA/IucC*, putative siderophore biosynthesis protein and transposases (Table 3).
Hypothetical proteins are predicted to be expressed in an organism, but the corresponding translation product has not been characterized yet (Ijaq et al. 2019). IutA is defined as the ferric aerobactin receptor. IutA has been associated with the pathogenicity of UPEC strains (Landgraf et al. 2012). IucD catalyzes L-lysine 6-monooxygenase reaction. This enzyme is the initial step of the aerobactin biosynthesis pathway. Aerobactin is a bacterial iron chelating agent found in E. coli (Thariath et al. 1993). IucA/IucC is an iron uptake chelate domain that is involved in the biosynthesis of the siderophores (Oves-Costales et al. 2009). Transposases are enzymes that move non-expressed transposon regions to different parts of the genome by cut-and-paste mechanism (Lewis et al. 2004).

**KnownClusterBlast and SubClusterBlast analysis**

Figure 7 shows similarity percentages of biosynthetic gene clusters with respect to E. coli VR50 siderophore region. These similarities as percentages in colicin V (E. coli chi7122, BGC0001555), aerobactin (Pantoea ananatis, BGC0001499), aerobactin (Xenorhabdus szentirmaii DSM 16,338, BGC0001498) and aerobactin (Grimontia hollisae, BGC0000939) were 6%, 100%, 66% and 22%, respectively.
respectively. No match was found in the SubClusterBlast analysis.

**BLASTp analysis and results**

In addition to antiSMASH, BLASTp analysis was performed to identify organisms with similar sequences. Two of the core siderophore biosynthetic genes in Region 1.4 of the *E. coli* VR50 were analyzed with the BLAST webserver. The first gene is between 3,379,667 and 3,381,409 nt. (Total 1,743 nt.) and the second gene is between 3,382,357–3,384,081 nt. (Total 1725 nt.). To decide the DNA sequences after the BLASTp analysis, e-value and percent identities were selected as ≤0.01 and ≥98%, respectively (Tables 4, 5).

To compare with the data in Tables 4, 5 the effect of microorganisms on UTIs was investigated.

---

**Fig. 6** Region 1.1 responsible for the biosynthesis of enterobactin and transport (A); Region 1.2 thiopeptide biosynthetic gene cluster (B); Region 1.3 responsible for NRPS and PKS genes related to yersiniabactin biosynthesis (C); Region 1.4 siderophore biosynthetic gene cluster (D) in *E. coli* VR50.
Enterobacteriaceae are pathogens responsible for pneumonia, UTIs and sepsis (Zilberberg et al. 2017). Carbapenem-resistant Enterobacteriaceae (CRE) especially cause major health problems. It is very difficult to treat, as CREs are not affected by the carbapenem antibiotic thanks to the carbapenemase enzyme they produce (Eshetie et al. 2015). There are few reports of Shigella flexneri causing UTIs and UTIs caused by Shigella sonnei are very unusual. It is not known exactly how Shigella species infect the urinary tract (Papasian et al. 1995). However, Shigella spp. and E. coli are similar in terms of phenotype and genotype. Therefore, it is thought that the virulence factors of E. coli may also be present in Shigella species (Tufon et al. 2020).

| Gene | Domain annotations | Location |
|------|-------------------|----------|
| Hypothetical protein | Unknown | 3,375,431–3,375,559 (Total: 129 nt) |
| *IutA* | TonB-dependent siderophore receptor family | 3,376,110–3,378,254 (Total: 2,145 nt) |
| L-lysine 6-monooxygenase *IucD* | Lysine/ornithine N-monooxygenase | 3,378,393–3,379,670 (Total: 1,278 nt) |
| Siderophore biosynthesis protein | Siderophore: *IucA/IucC* | 3,379,667–3,381,409 (Total: 1,743 nt) |
| Siderophore biosynthesis protein | Putative siderophore biosynthesis protein | 3,381,409–3,382,356 (Total: 948 nt) |
| Siderophore biosynthesis protein | Siderophore: *IucA/IucC* | 3,382,357–3,384,081 (Total: 1,725 nt) |
| Putative membrane transport protein | Unknown | 3,384,217–3,385,410 (Total: 1,194 nt) |
| IS2 transposase B | Transposase | 3,385,523–3,386,428 (Total: 906 nt) |
| IS2 transposase A | Unknown | 3,386,421–3,386,750 (Total: 330 nt) |
| Hypothetical protein | Unknown | 3,386,809–3,386,994 (Total: 186 nt) |
| IS629 transposase B | Transposase | 3,387,603–3,388,493 (Total: 891 nt) |
| IS629 transposase A | Transposase IS3/IS911 family protein | 3,388,490–3,388,765 (Total: 276 nt) |
### Table 4  The microorganisms showing similar sequences to the first siderophore biosynthetic gene in *E. coli* VR50

| Sequence                                      | Organism                  | Accession number | E-value | Percent identity |
|------------------------------------------------|---------------------------|------------------|---------|------------------|
| NIS family aerobactin synthetase *iucC*        | Enterobacteriaceae        | WP_001296374.1   | 0.0     | 100.00%          |
| NIS family aerobactin synthetase *iucC*        | Enterobacteriaceae        | WP_001015713.1   | 0.0     | 99.31%           |
| NIS family aerobactin synthetase *iucC*        | *Escherichia coli*         | WP_0063268075.1  | 0.0     | 99.83%           |
| Aerobactin synthase *iucC*                     | *Escherichia coli*         | N976833.1        | 0.0     | 99.83%           |
| TPA: aerobactin synthase *iucC*                | *Escherichia coli*         | HAG9679402.1     | 0.0     | 99.83%           |
| *iucA/iucC* family siderophore biosynthesis protein | *Shigella flexneri*       | EFX8038752.1     | 0.0     | 99.14%           |
| Aerobactin synthase *iucC*                     | *Shigella flexneri*        | EFZ8886218.1     | 0.0     | 99.14%           |

### Table 5  The microorganisms showing similar sequences to the second siderophore biosynthetic gene in *E. coli* VR50

| Sequence                                      | Organism                  | Accession number | E-value | Percent identity |
|------------------------------------------------|---------------------------|------------------|---------|------------------|
| NIS family aerobactin synthetase *iucA*        | Enterobacteriaceae        | WP_001015715.1   | 0.0     | 100.00%          |
| NIS family aerobactin synthetase *iucA*        | Enterobacteriaceae        | WP_001015713.1   | 0.0     | 99.31%           |
| *iucA/iucC* family siderophore biosynthesis protein | *Escherichia coli*        | WP_063268075.1   | 0.0     | 99.83%           |
| TPA: aerobactin synthase *iucA*                | *Escherichia coli*         | N976833.1        | 0.0     | 99.83%           |
| Aerobactin synthase *iucA*                     | *Escherichia coli*         | HAG9679402.1     | 0.0     | 99.83%           |
| Putative siderophore synthetase component      | *Shigella flexneri*        | EFX8038752.1     | 0.0     | 99.14%           |
| aerobactin synthase *iucA*                     | *Shigella flexneri*        | EFX8038752.1     | 0.0     | 99.14%           |
| *iucA/iucC* family siderophore biosynthesis protein | *Shigella sonnei*         | EFX8038752.1     | 0.0     | 99.14%           |

### Discussion

The relation between the siderophore synthesized by the bacteria and its pathogenicity has been shown in previous studies (Feldmann et al. 2007; Henderson et al. 2009; Holden and Bachman 2015). It was found that *E. coli* VR50, whose genome was analyzed, contains gene regions responsible for enterobactin and yersiniabactin biosynthesis. Based on this, it can be concluded that this strain causes UTIs by the effect of these siderophores. Also, nonpathogenic *E. coli* Nissle 1917 contains enterobactin and aerobactin biosynthetic gene regions. It is very interesting that yersiniabactin was existed in pathogenic *E. coli* VR50 but was not found in nonpathogenic *E. coli* Nissle 1917. Moreover, nonpathogenic *E. coli* K-12 strain MG1655 does not contain any gene cluster related to yersiniabactin and aerobactin.

Some *E. coli* strains infect the urinary tract and colonize there, causing UTIs (Kot 2019). One of these strains, *E. coli* VR50’s genome was analyzed to understand how it causes infection. When the genome of *E. coli* VR50 was examined, it was observed that it synthesizes some secondary metabolites that could ensure its survival in the urinary tract. The presence of siderophores among these metabolites attracted our attention. Although siderophore biosynthetic and siderophore-related genes are found in both pathogenic and nonpathogenic strains, they affect the pathogenicity of a strain. In the absence of siderophores, UPEC strains cause low-level local symptoms (Holden and Bachman 2015).

By inhibiting the iron uptake pathways, the pathogenicity of the microorganism can be reduced. Some of the studies focused on the reduction or inhibiting these metabolic pathways are shortly reviewed here. Chelating agents with higher iron affinity than siderophores was targeted to reduce the iron uptake of the microorganism by Qiu et al. (2011). The authors found that iron (III)-selective 3-hydroxypyridin-4-one chelator with very high affinity for iron shows antimicrobial effect and this compound is proposed to treat open wounds. This indicates that the bacterial cell wall structure and physical properties of the chelating agent are also important in inhibiting bacterial growth. Coulanges et al. (1997) conducted some studies on the strain *Listeria monocytogenes* which is known not to produce siderophores but uses external siderophores for iron uptake. In the study, it was found that Pt (II) was an inhibitor for the ferric reductase enzyme found in *L. monocytogenes*. Iron uptake was completely inhibited after treating the microorganism with Pt (II). Thus, ferric reductase activity was found to be fundamental for bacterial iron uptake. Schalk (2018) reported that antibiotics can be attached to siderophores and transport new antibiotics into bacteria using the iron intake mechanism. In the research of Schalk (2018), siderophores are used as a “Trojan Horse” transporting
antibiotics into bacteria without being detected. Since iron is found in hemoglobin, it is responsible for oxygen transport in humans. However, excess iron increases the risk of cancer. Cancer cells need more nutrients than normal cells because they multiply continuously and rapidly. Since iron is necessary for growth, tumor cells require more iron. Considering this situation, siderophores, which are small iron chelators, are used for treating cancer (Saha et al. 2016). It is very interesting to note that siderophore production in E. coli strains isolated from wild- and factory-raised turkeys are different, reported by Craft et al. (2022). Also, it was stated in the study that E. coli strains isolated from factory-raised turkeys produce more siderophores and other virulence factors compared to wild types. Siderophores are not specific for E. coli strains, these molecules are also reported from various species for many different actions. Some of the recent papers are reviewed here. In an interesting study carried out by Shah et al. (2022), the production of siderophores was investigated from five endophytic bacterial strains of the Pakistani wheat varieties. Endophytic bacteria are recently considered as a support for the plant productivity and defense system. Shah et al. (2022) found out that some of the strains have an ability to synthesize siderophores which could be associated with growth-promoting effects for the commercial plants. Podgórskak-Kryszczyk et al. (2022) also studied biological control of pathogenic Fusarium spp. by using antagonistic yeasts. The production of siderophores is also underlined by the authors as one of the action mechanisms against these species. Roskova et al. (2022) studied the use of siderophores as a bioremediation tool. Although the main role of siderophores is to bind iron, they can also chelate other heavy metals. This feature is of great importance in the solubilization and transport of heavy metals in the soil. Bioremediation methods are being developed by using microorganisms and plants that synthesize siderophores. Although the pathogenicity of some microorganisms has been known for a long time, it has been found later that they synthesize siderophores (Courcol et al. 1997). Comprehensive studies on E. coli VR50 should be conducted to give a more precise information about the pathogenicity. Therefore, new studies should be carried out on siderophores and their pathogenic properties.

The studies were found to show that the results obtained in antiSMASH analyzes were compatible with laboratory results. Lv et al. (2014) studied with non-pathogenic K12 strain MG1655 and uropathogenic UTI89 to reveal which kinds of siderophores they synthesize. Stable isotope dilution LC–MS/MS method is used to quantify the siderophores. According to their result, both strains produce enterobactin and uropathogenic UTI89 also produces yersiniabactin and salmochelin. Porcheron et al. (2014) studied the roles of iron regulators RhyB and Fur in the UPEC strain CFT073. Different amounts of siderophores were synthesized with deletion of RhyB and Fur genes. As a result of the study, it was found that aerobactin, enterobactin, and salmochelin siderophores were synthesized in a decreasing manner in the unmutated CFT073 strain. Garcia et al. (2011) conducted a study about the effect of outer membrane iron receptors in urinary tract infections. When the mutated strains were examined, they found that some siderophores were more effective in urinary tract colonization. When wild-type strains were examined enterobactin, salmochelin, aerobactin were found in E. coli CFT053 and enterobactin, salmochelin, yersiniabactin were found in E. coli 536. Cui et al. (2022) developed an indirect competitive ELISA (ic-ELISA) for enterobactin quantification. The ic-ELISA can detect enterobactin in different microorganisms. E. coli AN102, E. coli ATCC 25922, Salmonella enteritidis CVCC 1806, Campylobacter jejuni NCTC 11168, Bacillus cereus ATCC 14579, and Staphylococcus aureus ATCC 29213 strains were grown in iron-restricted medium. Results show that, E. coli AN102, E. coli ATCC 25922, and S. enteritidis CVCC 1806 can produce enterobactin, while other strains cannot. The effect of siderophores produced by E. coli strain 83972 on urinary tract infections was studied by Watts et al. (2012). Siderophores synthesized by E. coli strain 83972 were determined by HPLC/LC–MS analysis. As a result of the analysis, it was found that E. coli 83972 produces the enterobactin, salmochelin, aerobactin, and yersiniabactin siderophores.

A literature search reveal that some of the strains that were analyzed in Table 2 also contains salmoche

lin siderophore, but in antiSMASH analysis salmochelin gene clusters were absent. Salmochelin is an enterobactin that is C-glycosylated. Salmochelin is produced by Salmonella species and some UPEC strains (Hantke et al. 2003). IroN, IroD and IroB genes in Salmonella species are similar to the enterobactin receptor FepA, enterobactin esterase Fes, and glycosyltransferases, respectively (Müller et al. 2009). Bister et al. (2004) found that IroB gene is responsible for the glycosylation of enterobactin.

In summary, secondary metabolites produced by E. coli VR50, which causes UTIs, were examined. Siderophores have been found to be effective in the pathogenicity of the bacteria. However, its main role in infection varies according to the type and amount of siderophore.

In conclusion, genome mining tools which have been developed under in silico methodologies provide big contributions to understanding the etiology of the diseases. Moreover, identification and suppression of specific genes in pathogenic microorganisms that are responsible for secondary metabolites could be a new gate
in the therapy of diseases that are based on pathogenic microorganisms. By using recombinant DNA technology, siderophore-based molecules can be expressed in recombinant bacteria for agriculture and other industrial processes.

Acknowledgements
Mr. Ibrahim Kirkiz is a Master of Science Student in Dokuz Eylül University, Department of Biotechnology (Izmir, Türkiye).

Author contributions
Conceptualization: LC; Investigation: LC and IK; writing—original draft: IK; writing—review and editing: LC; supervision: LC. All authors read and approved the final manuscript.

Funding
Not applicable.

Availability of data and materials
The data presented in this study is not request from the corresponding author.

Declarations
Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
1 The Graduate School of Natural and Applied Sciences, Department of Biotechnology, Dokuz Eylül University, Kaynaklar Campus, 35390 Izmir, Türkiye.
2 Dokuz Eylül University, Faculty of Science, Department of Chemistry, 35390, Kaynaklar Campus, Izmir, Türkiye.

Received: 15 April 2022   Accepted: 8 June 2022
Published online: 15 June 2022

References
Albarano L, Esposito R, Ruocco N, Costantini M (2020) Genome mining as new challenge in natural products discovery. Mar Drugs 17:471–481. https://doi.org/10.3390/md171047199
Avasthi TS, Kumar N, Badamr R, Hussain A, Nandanwar N, Jadhav S, Ahmed N (2011) Genome of multidrug-resistant uropathogenic Escherichia coli strain NA114 from India. J Bacteriol 193:4272–4273. https://doi.org/10.1128/JB.05413-11
Barry SM, Challis GL (2009) Recent advances in siderophore biosynthesis. Curr Opin Chem Biol 13:205–215. https://doi.org/10.1016/j.cbpa.2009.03.008
Beatson SA, Ben Zakour NL, Totsika M, Forde BM, Watts RF, Mabbert AN, Szubert JM, Sarkar S, Phan MD, Peters KM, Petty NK, Alkhian NF, Sullivan MJ, Gawthorne JA, Stanton-Cook M, Nhu NTK, Chong TM, Yin WF, Chan KG, Hancock V, Ussery DW, Ulett GC, Schembri MA (2015) Molecular analysis of asymptomatic bacteruria Escherichia coli strain VR50 reveals adaptation to the urinary tract by gene acquisition. Infect Immun 83:1749–1764. https://doi.org/10.1128/IAI.00810-14
Bien J, Sokolova O, Bozko P (2012) Role of uropathogenic Escherichia coli virulence factors in development of urinary tract infection and kidney damage. Int J Nephrol 2012:1–15. https://doi.org/10.1155/2012/681473
Bister B, Bischoff D, Nicholson GJ, Valdebenito M, Schneider K, Winkelmann G, Hantke K, Sussmuth RD (2004) The structure of salmochelins: C-glucosylated enterobactins of Salmonella enterica. Bio metabol 17:471–481. https://doi.org/10.1023/B:BIOM.0000029432.69418.6a
Blin K, Shaw S, Kloosterman AM, Charlton-Powers Z, van Weezel GP, Medema MH, Weber T (2021) antiSMASH 6.0: improving cluster detection and comparison capabilities. Nucleic Acids Res 49:29–35. https://doi.org/10.1093/nar/gkab335
Chan DC, Burrows LL (2021) Thiopeptides: antibiotics with unique chemical structures and diverse biological activities. J Antibiot 74:161–175. https://doi.org/10.1038/s41429-020-00387-x
Coulanges V, Andre P, Ziegler O, Buchheit L, Vidon DJ (1997) Utilization of iron-catecholamine complexes involving ferric reductase activity in Listeria monocytogenes. Infect Immun 65:2778–2785. https://doi.org/10.1128/iai.65.7.2778-2785.1997
Courec RL, Trivier D, Bissinger MC, Martin GR, Brown MR (1997) Siderophore production by Staphylococcus aureus and identification of iron-regulated proteins. Infect Immun 65:1944–1948. https://doi.org/10.1128/iai.65.5.1944-1948.1997
Craft J, Edington H, Christman ND, Chasten J, Erickson D, Wilson E (2022) Increased microbial diversity and decreased prevalence of common pathogens in the gut microbiomes of wild turkeys compared to domestic turkeys. Appl Environ Microbiol. https://doi.org/10.1128/AEM.01423-21
Cui Y, Wang H, Guo F, Cao X, Wang X, Zeng X, Cui G, Lin J, Xu F (2022) Monoclonal antibody-based indirect competitive ELISA for quantitative detection of Enterobacteriaceae siderophore enterobactin. Food Chem 391:133241. https://doi.org/10.1016/j.foodchem.2022.133241
Dobrindt U, Chowdary MG, Krumholz H, Hacker J (2010) Genome dynamics and its impact on evolution of Escherichia coli. Med Microbiol Immunol 199:145–154. https://doi.org/10.1007/s00430-010-0161-2
Esthete S, Unakal C, Gelaw A, Ayeleyn B, Endris M, Mogens F (2015) Multidrug resistant and carbapenemase producing Enterobacteriaceae among patients with urinary tract infection at referral Hospital, Northwest Ethiopia. Antimicrob Resist Infect Control 4:1–8. https://doi.org/10.1186/s13756-015-0054-7
Fardeau S, Mullie C, Dassonville-Klimpt A, Audic N, Sonnet P (2011). Bacterial iron uptake: a promising solution against multidrug resistant bacteria. Science against microbial pathogens: communicating current research and technological advances, pp 695–705
Feldmann F, Sorsa LJ, Dassonville-Klimpt A, Audic N, Sonnet P (2011) The salmochelin siderophore receptor IroN contributes to invasion of urothelial cells by extraintestinal pathogenic Escherichia coli in vitro. Infect Immun 75:3183–3187. https://doi.org/10.1128/IAI.00656-06
Foxman B (2002) Epidemiology of urinary tract infections: incidence, morbidity, and economic costs. Am J Med 113:5–13. https://doi.org/10.1016/S0002-9343(02)01054-9
Garcia EC, Brumbaugh AR, Mobyte HL (2011) Redundancy and specificity of Escherichia coli iron acquisition systems during urinary tract infection. Infect Immun 79:1225–1235. https://doi.org/10.1128/IAI.01222-10
Goswami K, Chen C, Xiao L, Eaton KA, Dudley EG (2015) Cocculture of Escherichia coli O157:H7 with a nonpathogenic E. coli strain increases toxin production and virulence in a germfree mouse model. Infect Immun 83:4185–4193. https://doi.org/10.1128/IAI.00663-15
Hamasha K, Mohaidat QJ, Putnam RA, Woodman RC, Palchaudhuri S, Rehse SJ (2013) Sensitive and specific discrimination of pathogenic and nonpathogenic Escherichia coli using Raman spectroscopy—a comparison of two multivariate analysis techniques. Biomed Opt Express 4:481–489. https://doi.org/10.1364/BOE.4.000481
Hanke K, Nicholson G, Raboch W, Winkelmann G (2003) Salmochelins, siderophores of Salmonella enterica and uropathogenic Escherichia coli strains, are recognized by the outer membrane receptor IroN. PNAS 100:3677–3682. https://doi.org/10.1073/pnas.0737682100
Henderson JP, Crowley JR, Pinkner JS, Walker JN, Tsukayama P, Stamm WE, Hooton TM, Hultgren SJ (2009) Quantitative metabolomics reveals an epigenetic blueprint for iron acquisition in uropathogenic Escherichia coli. PLoS Pathog 5:e1000305. https://doi.org/10.1371/journal.ppat.1000305
Holden VM, Bachman MA (2015) Diverging roles of bacterial siderophores during infection. Metallophans 7:986–995. https://doi.org/10.1007/s41907-013-0333k
Ilijja J, Malik G, Kumar A, Das PS, Meena N, Bethi N, Sundararajan VS, Suravajhala P (2019) A model to predict the function of hypothetical proteins through a nine-point classification scoring schema. BMC Bioinform 20:1–8. https://doi.org/10.1186/s12859-018-2554-y
*Escherichia coli* strains isolated from clinical specimens in Iran. Rev Res Med Microbiol 33:20–30. https://doi.org/10.1097/MRM.0000000000000252

Thariath A, Socha D, Valvano MA, Viswanatha T (1993) Construction and biochemical characterization of recombinant cytoplasmic forms of the luC protein (lysine: N6-hydroxylase) encoded by the pCoV-K30 aerobactin gene cluster. J Bacteriol 175:589–596. https://doi.org/10.1128/JB.175.3.589-596.1993

Tufon KA, Fokam DPY, Kouanou YS, Meriki HD (2020) Case report on a swift shift in uropathogens from *Shigella flexneri* to *Escherichia coli*: a thin line between bacterial persistence and reinfection. Ann Clin Microbiol Antimicrob 19:1–6. https://doi.org/10.1186/s12941-020-00374-y

Tufon KA, Fokam DPY, Kouanou YS, Meriki HD (2020) Case report on a swift shift in uropathogens from *Shigella flexneri* to *Escherichia coli*: a thin line between bacterial persistence and reinfection. Ann Clin Microbiol Antimicrob 19:1–6. https://doi.org/10.1186/s12941-020-00374-y

Villebro R, Shaw S, Blin K, Weber T (2019) Sequence-based classification of type II polyketide synthase biosynthetic gene clusters for antiSMASH. J Ind Microbiol Biotechnol 46:469–475. https://doi.org/10.1007/s10295-018-02131-9

Watts RE, Totsika M, Challinor VL, Mabbett AN, Ulett GC, De Voss JJ, Schembri MA (2012) Contribution of siderophore systems to growth and urinary tract colonization of asymptomatic bacteriuria. *Escherichia coli*. Infect Immun 80:333–344. https://doi.org/10.1128/IAI.05594-11

Weng JK, Noel JP (2012) Structure–function analyses of plant type III polyketide synthases. Methods Enzymol 515:317–335. https://doi.org/10.1016/B978-0-12-394290-6.00014-8

Yun KW, Kim HY, Park HK, Fan W, Shorr AF (2017) Carbapenem resistance, inappropriate empiric treatment and outcomes among patients hospitalized with Enterobacteriaceae urinary tract infection, pneumonia and sepsis. BMC Infect Dis 17:1–13. https://doi.org/10.1186/s12879-017-2383-z

Zilberberg MD, Nathanson BH, Sulham K, Fan W, Shorr AF (2017) Carbapenem resistance, inappropriate empiric treatment and outcomes among patients hospitalized with Enterobacteriaceae urinary tract infection, pneumonia and sepsis. BMC Infect Dis 17:1–13. https://doi.org/10.1186/s12879-017-2383-z

Zotchev SB, Sekurova ON, Katz L (2012) Genome-based bioprospecting of microbes for new therapeutics. Curr Opin Biotechnol 23:941–947. https://doi.org/10.1016/j.copbio.2012.04.002

**Publisher’s Note**

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at [springeropen.com](https://springeropen.com)