Original Article

Genomic differences among strains of *Corynebacterium cystitidis* isolated from uterus of camels

Haitham Elbir[1], Faisal Almathen[1,2], Faisal Mohammed Almuhasen[3]

[1] Camel Research Center, King Faisal University, Al-Hasa, Saudi Arabia
[2] Department of Veterinary Public Health, College of Veterinary Medicine, King Faisal University, Al-Hasa, Saudi Arabia
[3] Veterinary Teaching Hospital, King Faisal University Al-Hasa, Saudi Arabia

Abstract
Introduction: Members of the *Corynebacterium cystitidis* species are usually isolated from kidney and urine of cow having pyelonephritis. Nevertheless, we have isolated *Corynebacterium cystitidis* for the first time from uterus of camels, extending the type of mammalian host for this species. Furthermore, it remains unknown whether there are significant genetic variations between strains isolated from different host species and anatomic sites. In this perspective, we investigated the genomic diversity of *Corynebacterium cystitidis* species, whose pan genome remain unexplored to date.
Methodology: Thus, we sequenced and compared the genomes of five *Corynebacterium cystitidis* of camel origin and a public genome of cow associated *Corynebacterium cystitidis*.
Results: Results revealed open pan genome of 4,038 gene clusters and horizontal gene transfer played a role in the extensive genetic diversity. Further, we found an obvious distinction between cow and camel associated *C. cystitidis* via phylogenomic analysis and by average nucleotide identity value of 95% between the two distant lineages and > 99% within camel associated *C. cystitidis* strains. Moreover, our data supports the hypothesis that the gene repertoire of cow associated *Corynebacterium cystitidis* developed so as to become more adaptable to the urine milieu. These genetic potentials are specifically evident for genes required for benzoate breakdown, iron transport, citrate and alanine utilization.
Conclusions: Our findings confirm the differentiation of strains into camel lineage and cow lineage. These different niches, comprising the uterus of camel and urinary tract of cow probably played a role in shaping the gene repertoire of strains.

Key words: *Corynebacterium cystitidis*; kidney; uterus.

Introduction

The first described *Corynebacterium cystitidis* infection was reported in 1978 by Yanagawa and Honda, as they found a Gram-positive bacterial rod causing pyelonephritis in cow [1]. This bacterium remained unreported until 2005, when Rosenbaum *et al.*, isolated *C. cystitidis* from bovine kidney with pyelonephritis during a survey in slaughterhouse [2]. Recently, Smith *et al.*, 2020 reported the isolation of *C. cystitidis* from 4 cases of pyelonephritis in cow [3]. Currently, studies on *C. cystitidis* are rare with one study showing bacterium ability to survive in soil for two months [4] and another study describing adhesion to urinary epithelial cells through pili structures existing on the bacteria cell surface [5]. Yet, there is no information about uterine infection caused by *C. cystitidis* in animals nor report of isolation from others site in animals than cow.

Currently, only a single closed genome is available for the type strain NCTC 11863 from an isolate cultured from a cow kidney. The genome is circular with a size of 3.01 Mbp that harbor 2,838 proteins. Nevertheless, it remains unknown whether there are significant genetic variations between strains isolated from different host species and anatomic sites. Furthermore, genetic discrimination of strains based on whole genome sequence has become a valuable tool to scrutiny on the population structure of pathogens [6,7]. Yet, no information about the molecular characteristics of *C. cystitidis* strains.

Uterine contamination by bacteria which can progress to infection such as metritis and endometritis occur mostly after delivery. Natural postpartum events lead to presence of blood inside uterus, therefore the postpartum milieu of the uterine lumen is nutrients rich for growth of bacteria [8]. In contrast, urinary tract (UT)

contain urine that is known as a moderately nutrient-limited growth environment. Metabolome profile of cow urine reveals several components such as Benzoate, Citrate, Choline and Alanine [9]. Thus, the bacteria need to have necessary metabolic pathways to utilize nutrients in the urine as well as efficient iron acquisition tools to cope with limited iron concentration in the urine, as iron is essential for bacterial survival and pathogenicity. The UT of cow and uterus of camel are different niche, therefore we expect to impact gene content of *C. cystitidis* in way that adapt it to the specific niche. Therefore, the aim of this study was to investigate the phylogenetic relationships of the *C. cystitidis* strains collected from the uterus of camels and the strain NCTC 11863 isolated from the kidney of cow. Also, to identify the genomic makeup of strains belonging to either of the two host in order to test the hypothesis of the presence of strain-specific repertoire of genes associated with adaptation to urine milieu. To address these questions, we sequence, describe and compare the genomes of five *C. cystitidis* isolates recovered from the uterus of camels showing uterine infection from different location and time point. In addition, we compare the camel associated *C. cystitidis* strains to already sequenced *C. cystitidis* strain isolated from cow kidney.

Methodology

Samples collection and isolates

All *C. cystitidis* clinical isolates were collected between November 2017 and July-2018 from uterus of five she-camels during routine examination at Veterinary Teaching Hospital with a history of conception failure despite several times mating with fertile camel. Examination of reproductive tract included transrectal palpation and transrectal ultrasound of uterus and ovaries using a 7.5 MHz. linear transducer probe (Aloka, Tokyo, Japan) beside inspection of the vagina and cervix. For uterine swab collection, the perineum area of the camel was cleaned and disinfected with acriflavine 0.1%. Then a double-guarded sterile swab (Kruuse, Langeskov, Denmark) was then passed through the vagina and cervix into the lumen of the uterus, the cotton swab was released from the double guard tube and rub with the endometrium before being retracted into the guard. Swabs were immediately cultured aerobically at 37 °C in 5% sheep blood agar media.

Genome sequencing, assembly and annotation

DNA was extracted from 48 hours cultured bacteria in blood agar plates using the Wizard genomic DNA purification kit (Promega, Madison, USA). DNA libraries for whole genome sequencing were constructed using the TruSeq Nano DNA Library Preparation Kits (Illumina, San Diego, USA) according to the manufacturer's instructions. Then libraries concentration was measured, equimolar volume of each of the libraries was pooled and sequenced by Macrogen Inc (Macrogen, Seoul, South Korea) on an Illumina's NovaSeq sequencer to produce 2x151 bp paired-end reads. Following sequencing, sequence reads were checked by Fastqc 0.11.8 [10] and low-quality bases were trimmed of using fastp software version 0.20.0 [11]. *De novo* assembly of passed filter reads was performed with the ABySS 2.2.1 [12], SPAdes 3.11.1 [13] and CLC genomic workbench 20.0.4 (Qiagen, Copenhagen, Denmark). Scaffolds with a coverage value lower than 10% and length < 500 base pairs (bp) were discarded. The best assembly was selected based on number of scaffolds and N50. Contigs were ordered against the closed reference genome *Corynebacterium cystitidis* strain NCTC11863 using Mauve software [14]. Open reading frames (ORFs) were identified with Prodigal [15] with default parameters. Identified ORFs spanning a sequencing gap region (containing N) were removed. The functional categorization of ORFs was performed using online RPS-BLAST against the cluster of orthologous groups (COG) [16] Pfam databases [17] (E-value $1e^{-02}$). If no hit was found, a search against the non-redundant (NR) database was done using BLASTP with E-value of $1e^{-03}$, coverage 0.7 and an identity percent of 30% on both query and hit. tRNAs were predicted with the Aragorn software [18] and ribosomal RNAs were predicted with RNAmmer [19].

Isolates identification

The identification of isolates to the species level was performed via calculation of average nucleotide identity (ANI) values via OrthoANIu (OrthoANI using USEARCH) [20] and by amplifications of 16S rRNA gene with published universal primers 27f and 1392R [2]. Briefly, two microliters of *Corynebacterium* DNA and 0.3 µL of each primer (10 pmol) (Macrogen, Seoul, South Korea) were added to the PCR mixture, consisting of one unit of Max Taq DNA Polymerase (Vivantis, Shah Alam, Malaysia), 5 µL of 10X ViBuffer (Vivantis, Shah Alam, Malaysia) and 2 µL of dNTPs (10 mM). The volume was completed to 25 µL by adding distilled water. DNA amplification was performed on a Tpersonal Thermocycler (BIOMETRA, Göttingen, Germany) with an initial 15 minutes cycle at 95 °C followed by 35 cycles consisting of 30 seconds at 94 °C, 1 minute at 55 °C and 1 minute extension at 72

°C, followed by a 10 minutes final extension step at 72 °C. To exclude DNA or amplicon contamination, molecular grade water negative control was used during the expermints. The PCR products were sequenced by Macrogen Inc. (Seoul, South Korea) using BigDye (Applied Biosystems, Foster, USA) on ABI3730XL DNA sequencer (Applied Biosystems). The obtained sequences were screened against NCBI non-redundant (nr) database to find the closest type species.

Detection of prophages, CRISPRs and virulence genes

The prediction of prophages in the genomes was performed using PHASTER (Phage Search Tool Enhanced Release) [22]. It uses a scoring technique to classify prophage regions as intact (> 90), questionable (70–90), or incomplete (< 70). The detection of genes that are predicted to be acquired by horizontal gene transfer (HGT) events was done using COLOMBO v4.0 with Sigi-HMM v1.0 and Sigi-CRF [23]. The identification of clustered regularly interspaced short palindromic repeats (CRISPRs) was performed using both CRISPRfinder and CRISPRCasFinder software [24,25], which are used to accurately detect CRISPR-Cas subtype, number of CRISPR and spacers. As for detection of potential virulence genes, we performed homology searching via BLAST tool against virulence factor database (VFDB) [26], which currently contains 36 virulence genes for seven *Corynebacterium* species including *C. diphtheriae*. Notably, homology searching against VFDB could only detect conserved virulence genes but may fail to detect virulence genes that are evolutionary distant from deposited virulence genes in VFDB.

Pan and core genome analysis

The *C. cystitidis* pan and core genome size was calculated in a way related to that previously reported approaches [27]. To calculate the number of orthologous genes within this species, we used the GET_HOMOLOGUES scripts [28] and Ortho Markov Cluster algorithm (OMCL) [29]. Orthologous genes were grouped using sequence identity of > 30%, query coverage of > 70% and E-value of > 1e-05.

lastly, the pan-genome is defined as the entire repertoire of genes present in this species, these genes are grouped into genes conserved in all strains (i.e., core-genome) and set of auxiliary genes (i.e., present in some but not all genomes) and strain specific genes (i.e., present in only one genome). Both auxiliary and strain specific genes represent the variable genomic content. The curve fitting of the core and pan-genome was executed as previously described [27]. Finally, the core and pan-genome were functionally described using COG (Clusters of Orthologous Groups) and KEGG (Kyoto Encyclopedia of Genes and Genomes) annotations [30].

Phylogenomic analysis

For whole-genome phylogenetic analysis of *C. cystitidis* strains, we used two approaches. First, we used genome-wide core- single nucleotide polymorphism (GWC-SNP) based phylogenetic tree approach as follow; the sequence reads were mapped to the *C. cystitidis* reference genome (Strain NCTC11863) using MEM command form Burrows–Wheeler Aligner (BWA) version 0.7.17-r1188 with default parameters [31]. The generated Sequence Alignment Maps (sam files) were sorted and indexed to yield Binary Alignment Maps (bam files) using SAMtools version 1.10 (using htslib 1.10.2) [32]. Duplicates reads were checked using flagstat command in Samtools. SNPs were determined using mpileup and call commands from BCFtools version 1.10.2 (using htslib 1.10.2). The SNPs were filtered according to the following parameters: minimal read mapping quality of 60, quality of variant calls > 40, read depth > 10 and genotyping quality >30. The identified SNPs were annotated using snpEff 4.3t [33]. For each genome, pseudosequences of polymorphic positions were used to infer maximum likelihood tree using MEGA7 [34]. In the second approach we constructed parsimony pangenome tree for the 6 *C. cystitidis* genome derived from presence/absence data in a consensus OMCL pangenome matrix, by using GET_HOMOLOGUES software package.
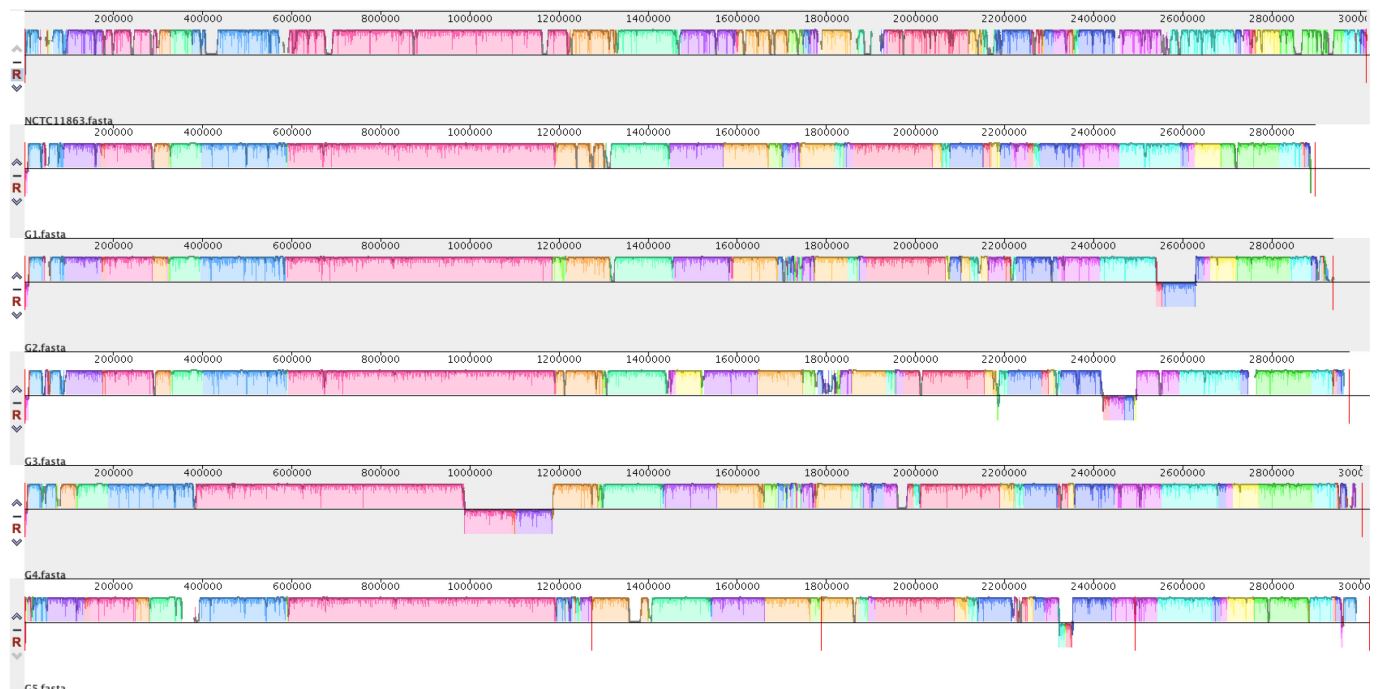
Results

Isolation and genome comparisons of C. cystitidis isolates

Uterine swab samples were taken from 5 she-camels having infertility and uterine infection symptoms (symptomatic n = 4; asymptomatic n = 1) between November 2017 and July 2018. Clinical examination showed accumulation of purulent content into the uterine lumen. Based on clinical signs and ultrasonography, samples were diagnosed as either clinical endometritis or pyometra (Table 1). Culturing uterine swab on blood agar media followed by16S rRNA sequencing revealed that sample2 and 5 were monoculture and named (*C. cystitidis* G2 and *C. cystitidis* G5). As for sample1, 3 and 4 they yielded mixed bacterial culture named (*C. cystitidis* G1, *C. cystitidis* G3 and *C. cystitidis* G4).

Table 1. General genomic features of 5 *C. cystitidis* isolates and cow associated *C. cystitidis* NCTC 11863.

Strains	*C. cystitidis*					
	NCTC11863	G1	G2	G3	G4	G5
Collection date		6/11/17	21/11/17	6/12/17	22/1/18	7/2/18
Host	Cow	Camel	Camel	Camel	Camel	Camel
Bacterial species found with *C. cystitidis*	Unknown	Staphylococcus simulans Corynebacterium efficiens Staphylococcus aureus	Only *C. cystitidis*	*Escherichia* sp	Unclassified bacterial species	Only *C. cystitidis*
Disease	Pyelonephritis	Pyometra	Pyometra	Endometritis	Endometritis	Unknown
Size (Mbp)	3.01	2.89	2.93	2.96	3.00	3.01
GC%	56.9	57	56.97	56.93	56.91	56.92
Genes	2837	2839	2892	2931	2946	2993
Coding density	88.84%	88.52%	89.32%	88.62%	88.98%	89.35%
Type II TA system	RelE/StbE-RelB/StbD YoeB-YefM HipA-HipB PspC-PspB HigB-HigA ParE-ParD Doc-Phd FitB-FitA	RelE/StbE-RelB/StbD YoeB-YefM HipA-HipB PspC-PspB HigB-HigA ParE-ParD	RelE/StbE-RelB/StbD YoeB-YefM HipA-HipB PspC-PspB HigB-HigA	RelE/StbE-RelB/StbD YoeB-YefM HipA-HipB PspC-PspB HigB-HigA	RelE/StbE-RelB/StbD YoeB-YefM HipA-HipB PspC-PspB HigB-HigA	RelE/StbE-RelB/StbD YoeB-YefM HipA-HipB PspC-PspB HigB-HigA
CRISPR-Cas subtype	I-U	I-U	II-C	I-U	II-C	No CRISPR System
No CRISPR	2	1	1	1	1	0
No spacers	9 and 43	27	14	28	12	0
RM	I II IV	I II IV	I II IV	I II IV	I II III IV	I II III IV
GenBank accession	LT906473	JAERQP000000000	JAERQQ000000000	JAERQR000000000	JAERQS000000000	JAERQT000000000

Figure 1. Mauve alignment of *C. cystitidis* NCTC11863 genome with the five strains of *C. cystitidis* genomes. Coloured blocks each represent a locally co-linear block containing sequence conserved between all the genomes being aligned. Blocks beneath the center line indicate sequences that align in the reverse complement (inverse) orientation.

Other identified bacterial species found with sample1,3 and 4 were shown in Table 1.

Hence, the genome of these five newly isolated *C. cystitidis* strains were sequenced using Illumina's NovaSeq platform. *De novo* assembly of the five strains generated a total genome size ranged from 2.89 to 3.01 Mbp with average coverage per scaffold > 100x. As summarized in Table 1, Gene number ranged from 2837 to 2993 and with coding density of 89%. For each of the 5 camel *C. cystitidis* genomes, contigs were ordered by mauve against closed genome of cow associated *C. cystitidis* NCTC11863 and concatenated by N characters yielding single chromosome. Overall, large number of collinear blocks of homologous regions in the genomes were observed (Figure 1). Notable, another publicly available *C. cystitidis* DSM 20524 unfinished genome was excluded since it is same type strain and therefore identical with *C. cystitidis* NCTC 11863. Entire genome sequence comparisons were accomplished in a pairwise manner, through calculation of ANI values for every genome pair. The ANI value calculated from pairwise comparisons within camel associated *C. cystitidis* strains were > 99%. In contrast, the ANI values calculated from comparisons of cow associated *C. cystitidis* strain with camel associated *C. cystitidis* strains were in the ranged between 95.81-95.93% (Figure 2). Four identical copies of 16S rRNA

Figure 2. Genetic similarity of genomes in *C. cystitidis*. BLAST Matrix for the 16Sr RNA and ANI values of six *C. cystitidis* genomes. In the matrix cells, numbers are given as percentage. The red numbers represent the 16Sr RNA value while black numbers represent ANI value. The strain name and their matrix cell containing 16Sr RNA and ANI value have the same colour except for *C. cystitidis* NCTC 11863, where the matrix cells borders are in red dashes.
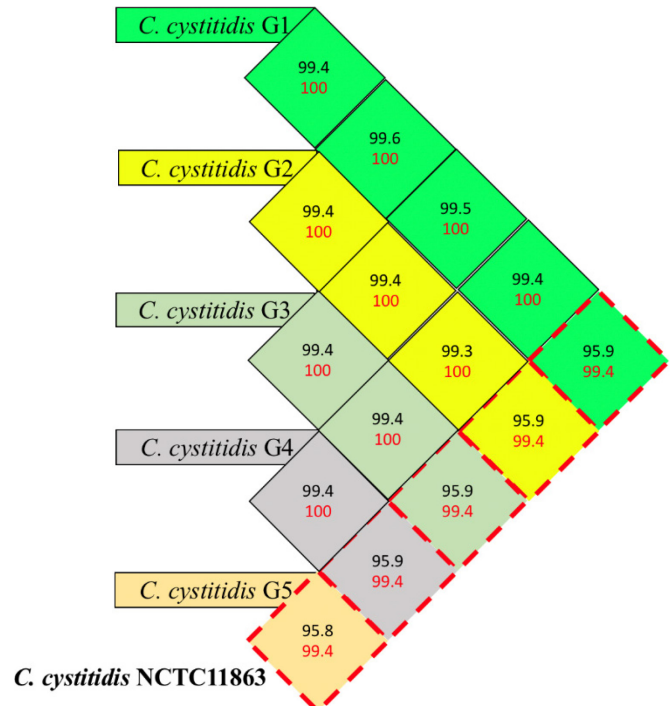

Table 2. Functional classification of pan genome based on COG categories.

COG categories	Core-genes	Strain specific genes						Auxiliary genes						HTG genes					
		G1	G2	G3	G4	G5	NT	G1	G2	G3	G4	G5	NT	G1	G2	G3	G4	G5	NT
Translation, ribosomal structure and biogenesis [J]	**177**	0	1	0	1	1	2	4	4	4	4	4	0	3	1	0	1	1	2
Amino acid transport and metabolism [E]	**157**	8	0	6	1	2	3	17	22	3	21	18	6	1	5	1	2	2	0
Carbohydrate transport and metabolism [G]	**140**	2	0	1	1	0	4	12	14	12	12	14	2	3	7	3	4	2	4
General function prediction only [R]	128	1	5	7	9	0	8	**44**	**47**	**46**	**49**	**43**	**23**	5	4	3	2	2	3
Inorganic ion transport and metabolism [P]	126	1	0	2	6	0	5	20	16	21	16	17	6	0	0	3	0	0	3
Coenzyme transport and metabolism [H]	124	0	1	1	0	0	3	9	8	10	9	9	4	3	2	5	0	2	2
Transcription [K]	118	2	0	0	0	2	14	23	25	23	23	24	4	4	1	1	4	0	1
Energy production and conversion [C]	99	1	2	0	1	2	3	7	11	9	5	5	2	1	1	0	0	1	0
Cell wall/membrane/envelope biogenesis [M]	99	1	2	0	1	4	9	12	11	12	11	11	2	**7**	**7**	**8**	**7**	**4**	**8**
Function unknown [S]	87	3	2	3	3	0	3	13	11	9	13	12	5	0	1	1	1	0	2
Replication, recombination and repair [L]	85	3	2	4	0	5	3	15	18	19	20	25	6	6	7	4	1	1	4
Posttranslational modification, protein turnover, chaperones [O]	84	0	2	0	1	1	1	3	4	3	3	5	2	0	1	0	0	1	1
Nucleotide transport and metabolism [F]	74	1	0	0	2	0	1	2	3	19	3	5	0	0	0	0	0	0	1
Signal transduction mechanisms [T]	65	0	1	0	1	2	8	24	24	24	23	22	2	0	0	1	1	0	1
Lipid transport and metabolism [I]	62	0	1	0	2	0	3	5	10	6	6	5	6	3	2	1	0	0	3
Defense mechanisms [V]	47	2	3	2	0	4	12	**28**	**32**	**32**	**39**	**29**	**10**	**7**	**7**	**7**	**4**	**5**	**5**
Cell cycle control, cell division, chromosome partitioning [D]	29	0	0	0	0	1	1	1	5	3	4	4	0	3	0	1	2	1	0
Secondary metabolites biosynthesis, transport and catabolism [Q]	29	0	1	0	1	0	3	2	7	3	3	2	7	0	0	0	0	0	0
Intracellular trafficking, secretion, and vesicular transport [U]	19	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	1
Cell Motility [N]	9	1	0	0	1	1	0	3	3	2	3	3	0	0	0	0	0	0	0
Mobilome: prophages, transposons [X]	6	**4**	**4**	**9**	**4**	**0**	**6**	**46**	**52**	**37**	**47**	**50**	**27**	**7**	**10**	**6**	**4**	**7**	**10**
Extracellular structures [W]	2	0	0	0	0	0	0	1	1	1	1	0	1	0	0	0	0	0	0
RNA processing and modification [A]	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

High values are highlighted in bold. NT: *C. cystitidis* NCTC 11863.

sequence were found in the genome of *C. cystitidis* strains of camel origin. Pairwise comparison of 16S rRNA sequence between camel associated *C. cystitidis* strains and cow associated *C. cystitidis* were 99.4% while within camel associated strains were 100%. (Figure 2). Moreover, the best second blast match for camel associated *C. cystitidis* strains was with *Corynebacterium mycetoides* DSM 20632 with 16S rRNA sequence similarity value of 95.1%.
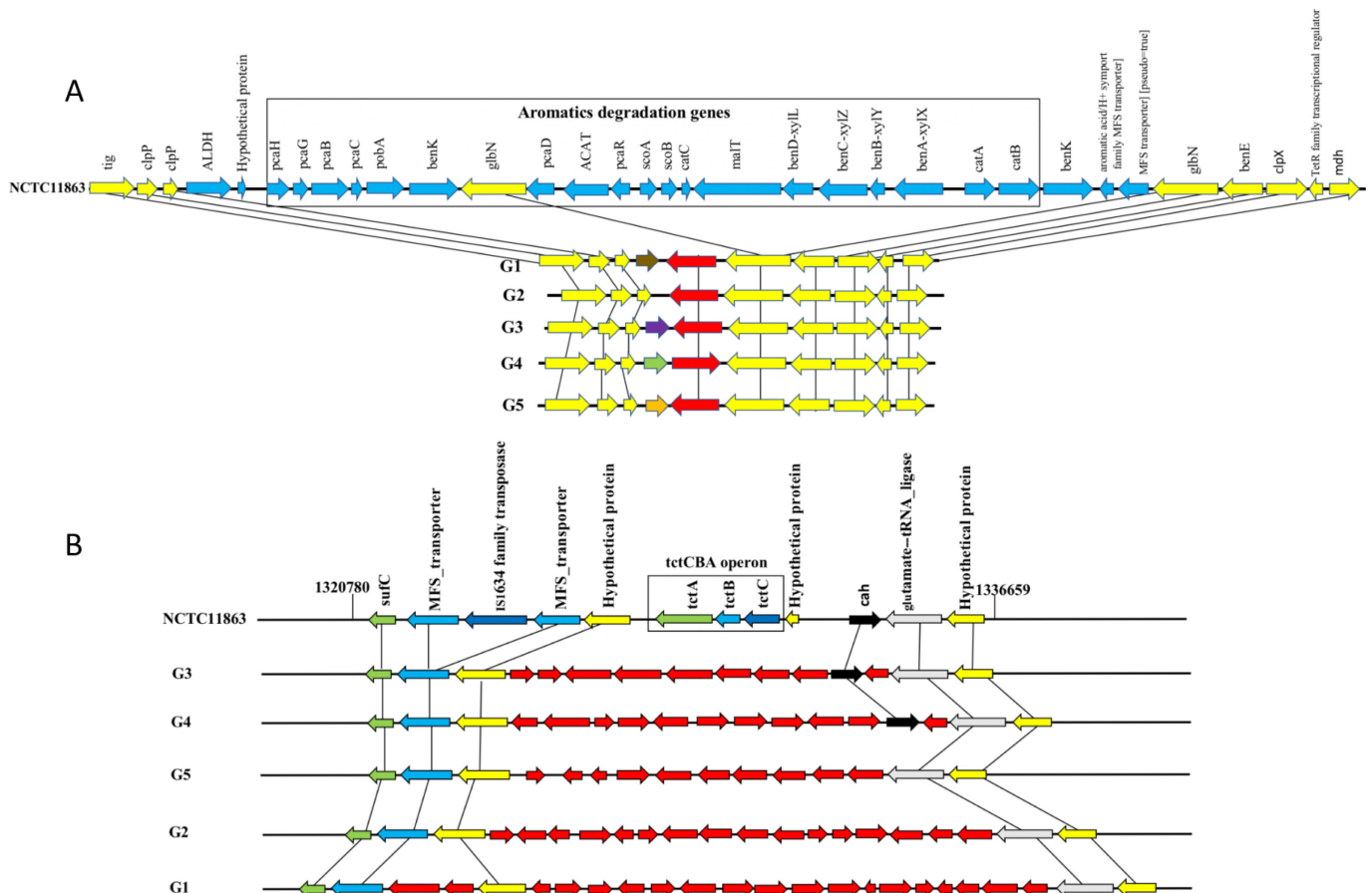
Core genome and pan genome of the C. cystitidis species

We described the core- and pan-genome among 6 *C. cystitidis* strains to explore their genetic diversity. A total of 4,038 pan-genome gene clusters were found. The core gene clusters (2,082, 62%) were abundant in the category Translation, ribosomal structure and biogenesis, (177, 11.1%) followed by the category Amino acid transport and metabolism (157, 9.8%) and the category Carbohydrate transport and metabolism (140, 8.8%) (Table 2). The prevalence of strain-specific

genes in *C. cystitidis* was various, ranging from 73 genes (strain G1) to 518 genes (strain NCTC 11863), suggesting high genomic plasticity. These unique genes encode diverse functions, of these functions which were much represented include the category Amino acid transport and metabolism and category Defense mechanisms (Table 2).

Shared and strain-specific metabolic potentials for survival in urine environment

Genome functional analysis by KEGG revealed shared and specific metabolic pathway. One of the shared traits, urease pathway for decomposition of urea encoded by ureABCEFGD operon. Urea can act as source of organic nitrogen for bacterial growth. The difference between cow and camel associated *C. cystitidis* strains include, for aromatics compound degradation, cow associated *C. cystitidis* strain encode the operon benABCD for degradation of benzoate into catechol and cat genes degrading the resulting catechol (Figure 3A). Thus, cow associated *C. cystitidis* strain

Figure 3. Organization of aromatic degradation genes and tctCBA operon of NCTC 11863 and the absence of these genes in the strains G1, G2, G3, G4 and G5. (A) The locus of the aromatic degradation genes is highlighted by rectangle. Consistent color coding linked by line indicates orthologous genes. (B) The tctCBA operon is highlighted by rectangle. Consistent color coding linked by line indicates orthologous genes. Most red color coding indicates strain specific gene.

seems to have a potential capacity to utilize benzoate, which is present in cow urine, as a carbon source. In contrast, only camel associated *C. cystitidis* strains encode genes for tetrahydrofolate biosynthesis. Major differences in molecule transport are also detected among the two strains: Unlike camel associated *C. cystitidis* strains, cow associated *C. cystitidis* strain carries putative citrate transport systems encoded by the tctCBA operon (Figure 3B). Thus, cow associated *C. cystitidis* seems to have a potential capacity to utilize citrate, which is present in cow urine, as a carbon source. Furthermore, cow associated *C. cystitidis* harbors ribose transport systems encoded by the rbsACBK operon. The amino acid alanine is present in the urine of cow, unlike camel associated *C. cystitidis* strains, cow associated *C. cystitidis* encodes alanine dehydrogenase gene. Thus, cow associated *C. cystitidis* seems to have a potential capacity to synthesize alanine from pyruvate and utilization of alanine as a nitrogen source.

Potential virulence factors

To define the potential virulence capability of *C. cystitidis* strains, we inspected the genomes for the presence of virulence genes using blast search tool against virulence factor database (VFDB) [26]. Virulence factors of pathogenic *Corynebacterium* species in VFDB were classified into 4 main groups, that is, adherence, iron uptake, toxin and regulation. We have identified in all strains two LPXTG motif proteins. Surface proteins that carry an LPXTG motif are covalently link to the cell wall peptidoglycan by membrane-bound sortase enzymes. All *C. cystitidis* strains carried two LPXTG-anchored proteins arranged in a cluster including, SpaH/EbpB family LPXTG-anchored major pilin, sortase A and LPXTG cell wall anchor domain-containing protein. These genes set are highly conserved between camel associated *C. cystitidis* strains (98-99%) but had low homology (<56%) to the cow associated *C. cystitidis* strain. This in turn are probably associated with niche adaptation. Another putative surface protein not mentioned in VFDB and present only in *C. cystitidis* of cow origin include Rib/alpha/Esp surface antigen repeat.

Cow and camel associated *C. cystitidis* strains have several shared sets of iron acquisition systems with some differences. We found seven clusters of iron related genes that are highly conserved between cow and camel associated *C. cystitidis* strains (> 98) but cow associated *C. cystitidis* strain possess extra 13 additional iron related genes involved in iron transport and acquisition. Other commonly detected virulence

factors included metal-dependent transcriptional regulator gene (DtxR). Notably, genes encoding the diphtheria toxin of *C. diphtheriae* and phospholipase D of *C. pseudotuberculosis* were not found among all *C. cystitidis* strain genomes.

The two-component systems (TCSs) aid bacteria to sense, act and adapt to a broad variety of stimuli, such as nutrients state, changes in osmolarity and temperature. Furthermore, some TCSs can regulate gene clusters involved in biofilm formation and virulence in pathogenic bacteria [35]. TCS usually composed of sensor histidine kinase (HK), mostly a membrane linked protein and a cytoplasmic response regulator (RR) [35]. Genomic analysis revealed 4 out of 5 TCSs were shared by all strains that include 1) MtrB-MtrA system, involved in osmotic stress response in *C. glutamicum* [36], 2) SenX3-RegX3 system is required for virulence in *Mycobacterium tuberculosis* [37], 3) MprB-MprA required for persistent infections in *Mycobacterium tuberculosis* and the fourth TCS include sensor kinase-response regulator. Unlike cow associated *C. cystitidis*, the genome of strains *C. cystitidis* of camel origin carry one more TCS system called DesK-DesR. The DesK-DesR system is triggered at low temperature and preserves membrane lipid fluidity upon temperature changes in the bacteria *Bacillus subtilis* [38].

C. cystitidis defense strategies

In order to examine the capability of members of the *C. cystitidis* species to protect themselves against attack by foreign DNA, we examined the existence of CRISPR-Cas systems

using CRISPRfinder and CRISPRCasFinder software. Among the 6 *C. cystitidis* genomes analyzed, we detected CRISPR-Cas system except for strain G5 (Table 1). Based on *cas* gene content and CRISPR length, we identified 3 type I systems and 2 type II systems, the types were further classified into subtypes I-U and II-C When inspecting the CRISPR sequences, the highest number of spacers was found in cow associated *C. cystitidis* and this strain possess two different CRISPR sequences.

Other systems that prevent gaining of foreign DNA are represented by restriction-modification (RM) systems [39]. Exploring the predicted *C. cystitidis* proteome for RM systems shown that type I, II, IV RM systems represent the prevalent gene cluster, being present in all strains, followed by type III RM systems identified on the genomes of C. *cystitidis* G4 and G5 (Table 1). Hence, it seems that the genomes of *C. cystitidis* G4 and G5 were armed with a wider genomic

collection to prevent the attack by foreign DNA sequences.

Toxin-antitoxin (TA) systems are wide spread in bacteria and have been linked to abortive infection (bacteriophage immunity through altruistic suicide of infected cells) [40] and enhancing bacterial survival in environmental stresses such as heat shock, nutrient starvation and antibiotic. Screening *C. cystitidis* of cow origin genome revealed at least 8 type II TA loci, followed by *C. cystitidis* G1 encodes 6 type II TA loci and other strains encode 5 type II TA loci (Table 1). We observed that most of the *C. cystitidis* TA systems are located within core genome, suggesting that it may achieve a function common to all *C. cystitidis* strains. Notably, the dispensable genome of cow associated *C. cystitidis* strain carry two type II TA loci (FitB-FitA and Doc-Phd) that are missing from the camel strains. The TA system, FitB-FitA, is constituted of an antitoxin FitA and a toxin FitB and has been linked to reduce intracellular replication and intracellular trafficking of *Neisseria gonorrhoeae* leading to the formation of persistent populations [41]. As for Doc-Phd, it contributes to persister formation in *Salmonella* Typhimurium [42]. Thus, FitB-FitA and Doc-Phd TA system probably play roles in virulence of cow associated *C. cystitidis*. Another type II TA system, ParE-ParD are involved in plasmid stabilization but the deposited closed genome sequence in the GenBank lack any plasmids associated with *C. cystitidis* genome. Therefore, the function played by this locus in the genome of *C. cystitidis* of cow origin and *C. cystitidis* G1 is unknown. Other loci of TA system are presented in Table 1.

C. cystitidis mobilome

The mobilome is defined as the entire genes that may have been acquired by horizontal genes transfer (HGT). The gaining of foreign genes could transfer new traits, which are essential for pathogenicity and adaptation into various niches [43,44]. The obtained results revealed that 3.7% of the entire gene pools in the case of cow associated *C. cystitidis* were predicted to represent HGT gained genes and up to 2.7%, 2%, 3.3%, 3.2%, 2.7% of that of camel associated *C. cystitidis* G1, *C. cystitidis* G2, *C. cystitidis* G3, *C. cystitidis* G4 and *C. cystitidis* G5 respectively. HGT genes contributed to the auxiliary genes and strain specific genes as follow: we found a percent of 12.6, 7.3, 5.9, 8.3, 7.7 and 5.6 of auxiliary genes were HGT for strain NCTC 11863, G1, G2, G3, G4 and G5 respectively. In addition, 9.5, 15.1, 8, 14, 4.9 and 10.3 and 15.1 to 4.9 of strain specific genes were HGT for strain NCTC 11863, G1, G2, G3,

G4 and G5 respectively. Moreover, a total of 11 potential donor taxa were identified including the genus *Anaplasma, Brevibacillus, Desulfomonile, Grimontia, Hahella, Mycobacterium, Nitrosococcus, Serratia, Syntrophothermus, Talaromyces* and *Thiomicrospira.* Out of the 11 potential donor taxa the genera *Talaromyces* and *Mycobacterium* were the main donor taxa. The COG functional classification of predicted mobilome of *C. cystitidis* revealed genes encoding diverse functions, of these functions which were much represented include the category Cell wall and membrane envelope biogenesis, Defense mechanisms and "Mobilome: prophages, transposons" (Table 2). Closer inspection of cow associated *C. cystitidis* mobilome revealed genes involved in iron acquisition and Type II R-M system. Also, closer inspection of strain G1, G2 and G4 mobilome revealed genes involved in Type IV R-M system. As for the *C. cystitidis* G3 mobilome, genes involved in Type IV R-M system and CRISPR-Cas system were also found.

Finally, we used in silico analysis to detect candidate prophages throughout the genomes of *C. cystitidis* strains. This genomic screen revealed no intact or questionable prophages in any of the six *C. cystitidis* strains, at least one incomplete phage region was found in all strains. As for *C. cystitidis* of cow origin harbor one incomplete prophage region (7.8kb) which match best to protein sequence of PHAGE_Bacill_phiNIT1_NC_021856. As for *C. cystitidis* G1, four different incomplete prophage regions (9kb, 4.7kb, 6.2kb and 7.8 kb) were identified and having match best to protein sequence of PHAGE_Staphy_42E_NC_007052, PHAGE_Shigel_SfIV_NC_022749, PHAGE_Sinorh_PBC5_NC_00332 and PHAGE_Achrom_JWAlpha_NC_023556 respectively. Moreover, cystitidis G2, *C. cystitidis* G3 and *C. cystitidis* G4 harbor similar one incomplete prophage region (~43.4kb) which match best to PHAGE_Coryne_Poushou_NC_042139. In the case of *C. cystitidis* G5, two incomplete prophage regions (45.7kb and 37.9kb) were identified and having match best to protein sequence of PHAGE_Coryne_Poushou_NC_042139 and PHAGE_Gordon_Attis_NC_041883 respectively. Notably strain G5 harbor the largest incomplete prophage regions and it lacks CRISPR-Cas system while strain of cow origin has two CRISPR and harbor the smallest incomplete prophage regions. Thus, CRISPR-Cas system perhaps provide protection against phage infection.

Phylogenomic analyses

Phylogenetic relationship analysis concerning strains of a same species needed to be performed using core or whole genome sequences to reach a higher resolution and a more comprehensive analysis [45]. Therefore, we investigated the genetic diversity within the strains of *C. cystitidis* by constructing the phylogenomic tree based on GWC-SNP. For this purpose, we mapped read sequences of the *C. cystitidis* G1, G2, G3, G4 and G5 against cow associated *C. cystitidis* that yield a total of 95338 SNPs, of these, 83397 SNPs (87.48%) were located in the coding region and 11941 (12.52%) in intergenic region. Further analysis of SNPs reveals that 29.03% were missense mutation, 0.23% were nonsense and 70.74% were silent mutation. We generated high resolution maximum likelihood phylogeny based on 95338 SNPs. As displayed in Figure 4), all of the camel associated *C. cystitidis* strains clustered in a separate branch named lineage1 while the cow associated *C. cystitidis* formed a distant long branch named lineage 2, indicating a clear distinction between camel and cow associated *C. cystitidis* strains. lineage1 is further sub-divided into sub-lineage 1A and 1B (Figure 4). In addition to the aforesaid phylogenetic method, we also inferred parsimony pangenome tree for the 6 *C. cystitidis* genome derived from presence or absence of the 4,038 pan-genes. As shown in Figure 4, the camel associated *C. cystitidis* strains resolved into two branches named lineage 1 and 2, while cow associated *C. cystitidis* placed in a distant branch named lineage 3.
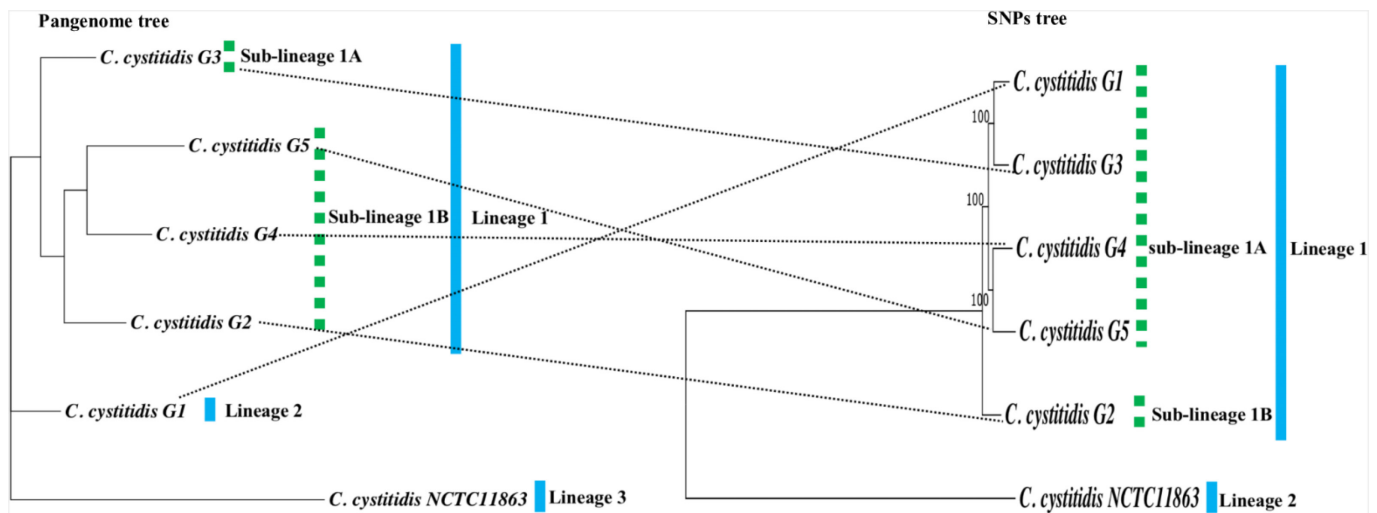
Discussion

Since their isolation, members of the *C. cystitidis* have been associated with the development of pyelonephritis in cow. To our knowledge this is the first report of isolating *C. cystitidis* from the uterus of animals. Whether *C. cystitidis* is a true causative agent of endometritis/pyometra or cause disease synergistically with another uterine microbes remain to be clarified.

ANI is authenticated method for comparing genomes, identification and estimating species genetic relationships [46]. The ANI value within camel associated *C. cystitidis* strains were > 99%, exhibiting high level of relatedness. In contrast, the ANI values calculated from comparisons of cow associated *C. cystitidis* strain with camel associated *C. cystitidis* strains were in the ranged between 95.81-95.93% reflecting genetic diversity between *C. cystitidis* strain isolated from cow and camels. This is similar with a figure of 95.1% for the ANI between two members of *Serratia marcescens*, strains DB11 and SM39 harbored by two unrelated hosts [47]. Thus, both the ANI value which exceeded the recommended 95% threshold for separating two species [46] and the high 16S rRNA value, confirmed that camel associated *Corynebacterium* strains are members of the species *C. cystitidis*.
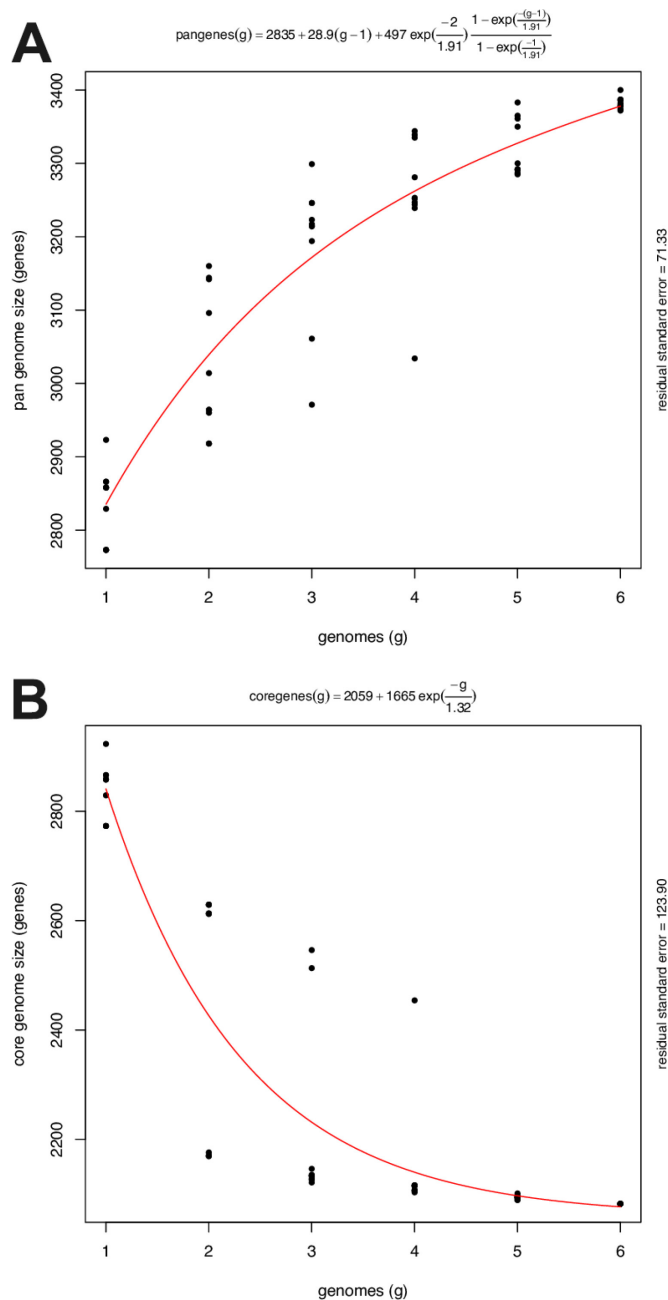
The variable genomic content constituted (48.4%) of the pan-genome, revealing genetic diversity in *C. cystitidis*. The variable genomic content is responsible for species diversity and confer selective benefit to the bacteria, such as niche adaptation, virulence traits and

Figure 4. Comparison of tree constructed from SNPs and parsimony pangenome tree derived from presence/absence data in a consensus OMCL pangenome matrix.

antibiotic resistance. The auxiliary gene clusters were abundant in the category General function prediction only, Mobilome: prophages, transposons and Defense mechanisms. Furthermore, the strain-specific gene existing in only one genome are probably conferring special traits. Additionally, viewing the core genome and pan genome plot of analyzed genomes revealed that a curve has not reached a plateau suggesting that this

Figure 5. Core and pan-genome analysis of 6 *C. cystitidis* strains. (A) Pangenome (B) and Core genome curves exhibiting the ascending trend of the pan gene families and the descending trend of the core-gene families with the rise in the number of genomes.

pangenome is still open. Therefore, more than the six sequenced genomes should be sequenced to estimate the diversity of *C. cystitidis* gene repertoire (Figure 5A and 5B). This is in contrast to the pan-genome of the intracellular endosymbiont *Buchnera aphidicola* has been closed by 6 genomes [48]. This is due the fact that obligate intracellular bacteria have a restricted niche and have low capacity to gain foreign genes. On the contrary, free-living and environment linked bacteria occupy numerous niches, facing many external stresses. Additionally, it exhibits a capacity to gain foreign genes by horizontal transfer [49,50]. To our knowledge, *C. cystitidis* inhibit at least two different niches (urinary tract of cow and camel uterus), thus, the pan-genome probably has capacity to acquire or lose gene to adapt to the different niche.

Cow associated *C. cystitidis* needs to cope with limited nutrient in the urine and antibacterial effect of urea [51]. Metabolome profile of cow urine reveals several components such as urea, benzoate, citrate and alanine [7]. Unlike camel associated *C. cystitidis*, cow associated *C. cystitidis* seems to have a potential capacity to utilize benzoate, citrate [52] and alanine which is present in cow urine, as a carbon source. Knowledge from other bacteria causing UTI such as in *Enterococcus faecalis* revealed upregulation of genes linked to iron transport during urine growth [53]. In *Pseudomonas aeruginosa,* the use of lactate, citrate, and amino acids as carbon sources has been shown during growth in synthetic urine [54]. Together, these genes support the notion that *C. cystitidis* of cow origin is probably more adapted to the nutritional environment in the urinary tract than *C. cystitidis* of camel origin.

As for the virulence factor, we have identified in all strains two LPXTG motif proteins. Several characterized LPXTG motif proteins have role in adhesion and consequently virulence [55]. Another putative surface protein not mentioned in VFDB and present only in *C. cystitidis* of cow origin include Rib/alpha/Esp surface antigen repeat. We hypothesize that *C. cystitidis* use the above-mentioned proteins for adhesion to host cells.

Iron uptake is crucial for bacterial growth inside host organisms because iron is sequestered by the host. Thus, genes associated with iron acquisition are known as virulence factors [26]. The finding of more iron related genes in cow associated *C. cystitidis* strain indicates that effective transport systems could convene a selective growth advantage to cow associated *C. cystitidis* strain by assisting it to efficiently acquire iron from the urine milieu. Other commonly detected virulence factors included metal-dependent

transcriptional regulator gene (DtxR) that is known to controls the transcription of several genes involved in iron homeostasis and toxin gene [56].

The finding of different donor taxa raise question where the HGT events occur. Notably the uterus of camel is normally contaminated by several bacteria species after delivery, which provide opportunities for HGT to occur. Thus altogether, this information suggests that HGT events contributed particularly to iron transport and defense mechanism and generating genetic diversity within the strains. Of note, one of the donor taxa, *Talaromyces* is a fungus, this finding is not surprising as HGT from Bacteria and plants to the Fungus was reported recently (57). Notably, prophages exit as remnant in the genome of *C. cystitidis* strains a finding similar to what have been observed previously in *C. bovis* [58]. In contrast, pathogenic strains of *C. diphtheriae* and *C. ulcerans* harbor the gene encoding diphtheria toxin on a prophage [59,9].

Phylogenomic analysis of *C. cystitidis* strains resolved into two distant branches. Notably, the detailed topologies differed somewhat between GWC-SNP phylogeny and the pan-genome phylogeny. This discordance indicated that variable genomic content played important role in the evolution of *C. cystitidis* species, which was subsequently characterized by determinations of several horizontally transferred genes, mobile genetic elements. Regardless of the differences in tree deep branch topology, both trees shown a clear distinction between *C. cystitidis* strains isolated from camel and the strain of *C. cystitidis* isolated from cow. In addition, our genome phylogeny unveils no clonality of *C. cystitidis* strains, suggesting that *C. cystitidis* strains differences mirror the ecological source of such inspected strains.

Conclusions

In this study, isolation of five *C. cystitidis* strains from the uterus of camels showing signs of uterine disease for the first time, extend the type of mammalian host for this species. Further investigation is required to clarify the role of *C. cystitidis* in causing uterine disease. Analysis of five *C. cystitidis* sequenced genomes and the one publicly available *C. cystitidis* genome, unveiled that more genome sequencing of new *C. cystitidis* strains is required to close the current open pan genome. Nonetheless, phylogenomic trees, which were constructed using two different methods (GWC-SNP and gene presence/absence), showed a clear differentiation between the camel associated *C. cystitidis* strains and cow associated *C. cystitidis* branch. Likewise, ANI support this finding. Anyhow,

camel associated *C. cystitidis* strains showed large diversity. Furthermore, horizontal gene transfer played a role in *C. cystitidis* strains diversity. In fact, genome analysis of each strain showed the existence of SNPs and strain specific genes, representing the genetic variations between these six strains. Interestingly, a portion of strain specific genes of cow associated *C. cystitidis* seems to be involved in benzoate breakdown, iron transport, citrate and alanine utilization, supporting the hypothesis that the gene repertoire of this strain developed to become more adaptable to the urine milieu. Hence, these different niches, comprising the uterus of camel and urinary tract of cow probably played a role in shaping the gene repertoire of strains. Anyhow, our findings confirm the differentiation of strains into camel lineage and cow lineage.

References
1. Yanagawa R, Honda E (1978) *Corynebacterium pilosum* and *Corynebacterium cystitidis*, two new species from cows. Int J Syst Bacteriol 28: 209-216.
2. Rosenbaum A, Guard CL, Njaa BL, McDonagh PL, Schultz CA, Warnick LD, White ME (2005) Slaughterhouse survey of pyelonephritis in dairy cows. Vet Rec157: 652-655.
3. Smith JS, Krull AC, Schleining JA, Derscheid RJ, Kreuder AJ (2020) Clinical presentations and antimicrobial susceptibilities of *Corynebacterium cystitidis* associated with renal disease in four beef cattle. J Vet Intern Med 34: 2169-2174.
4. Hayashi A, Yanagawa R, Kida H (1985) Survival of *Corynebacterium renale*, *Corynebacterium pilosum* and *Corynebacterium cystitidis* in soil. Vet Microbiol 10: 381-386.
5. Sato H, Yanagawa R, Fukuyama H (1982) Adhesion *of Corynebacterium renale*, *Corynebacterium pilosum* and *Corynebacterium cystitidis* to bovine urinary bladder epithelial cells of various ages and levels of differentiation. Infect Immun. 36: 1242-1245.
6. Timms VJ, Nguyen T, Crighton T, Yuen M, Sintchenko V (2018) Genome-wide comparison of *Corynebacterium diphtheriae* isolates from Australia identifies differences in the Pan-genomes between respiratory and cutaneous strains. BMC Genomics 19: 869.
7. Trost E, Blom J, Soares Sde C, Huang IH, Al-Dilaimi A, Schröder J, Jaenicke S, Dorella FA, Rocha FS, Miyoshi A, Azevedo V, Schneider MP, Silva A, Camello TC, Sabbadini PS, Santos CS, Santos LS, Hirata R Jr, Mattos-Guaraldi AL, Efstratiou A, Schmitt MP, Ton-That H, Tauch A (2012) Pangenomic study of *Corynebacterium diphtheriae* that provides insights into the genomic diversity of pathogenic isolates from cases of classical diphtheria, endocarditis, and pneumonia. J Bacteriol 194: 3199-3215.

8. Sheldon IM, Williams EJ, Miller AN, Nash DM, Herath S (2008) Uterine diseases in cattle after parturition. Vet J 176: 115-121.

9. Basoglu A, Baspinar N, Tenori L, Licari C, Gulersoy E (2020) Nuclear magnetic resonance (NMR)-based metabolome profile evaluation in dairy cows with and without displaced abomasum. Vet Q 40: 1-15.

10. Babraham Bioinformatics (2010) FastQC. A quality control tool for high throughput sequence data. Available: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/. Accessed: 27 April 2021.

11. Chen S, Zhou Y, Chen Y, Gu J (2018) fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 134: i884-i890.

12. Simpson JT, Wong K, Jackman SD, Schein JE, Jones SJ, Birol I (2009) ABySS: a parallel assembler for short read sequence data. Genome Res 19: 1117-1123.

13. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA (2012) SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19: 455–477.

14. Darling AC, Mau B, Blattner FR, Perna NT (2004) Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res 14: 1394-1403.

15. Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ (2010) Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics 8: 119.

16. Marchler-Bauer A, Panchenko AR, Shoemaker BA, Thiessen PA, Geer LY, Bryant SH (2002) CDD: a database of conserved domain alignments with links to domain three-dimensional structure. Nucleic Acids Res 30: 281-283.

17. Sonnhammer EL, Eddy SR, Durbin R (1997) Pfam: a comprehensive database of protein domain families based on seed alignments. Proteins 28: 405-420.

18. Laslett D, Canback B (2004) ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. Nucleic Acids Res 32: 11-16.

19. Lagesen K, Hallin P, Rødland EA, Staerfeldt HH, Rognes T, Ussery DW (2007) RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res 35: 3100-3108.

20. Yoon SH, Ha SM, Lim J, Kwon S, Chun J (2017) A large-scale evaluation of algorithms to calculate average nucleotide identity. Antonie Van Leeuwenhoek 110: 1281-1286.

21. Ni K, Wang Y, Li D, Cai Y, Pang H (2015) Characterization, identification and application of lactic acid bacteria isolated from forage paddy rice silage. PLoS One 10: e0121967.

22. Arndt D, Grant JR, Marcu A, Sajed T, Pon A, Liang Y, Wishart DS (2016) PHASTER: a better, faster version of the PHAST phage search tool. Nucleic Acids Res 44: W16-W21.

23. Waack S, Keller O, Asper R, Brodag T, Damm C, Fricke WF, Surovcik K, Meinicke P, Merkl R (2006) Score-based prediction of genomic islands in prokaryotic genomes using hidden Markov models. BMC Bioinformatics 7: 142.

24. Couvin D, Bernheim A, Toffano-Nioche C, Touchon M, Michalik J, Néron B, Rocha EPC, Vergnaud G, Gautheret D, Pourcel C (2018) CRISPRCasFinder, an update of CRISRFinder, includes a portable version, enhanced performance and integrates search for Cas proteins. Nucleic Acids Res 46: W246-W251.

25. Grissa I, Vergnaud G, Pourcel C (2007) CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. Nucleic Acids Res 35: W52–W57.

26. Liu B, Zheng D, Jin Q, Chen L, Yang J (2019) VFDB 2019: a comparative pathogenomic platform with an interactive web interface. Nucleic Acids Res 47: D687-D692.

27. Tettelin H, Masignani V, Cieslewicz MJ, Donati C, Medini D, Ward NL, Angiuoli SV, Crabtree J, Jones AL, Durkin AS, Deboy RT, Davidsen TM, Mora M, Scarselli M, Margarit y Ros I, Peterson JD, Hauser CR, Sundaram JP, Nelson WC, Madupu R, Brinkac LM, Dodson RJ, Rosovitz MJ, Sullivan SA, Daugherty SC, Haft DH, Selengut J, Gwinn ML, Zhou L, Zafar N, Khouri H, Radune D, Dimitrov G, Watkins K, O'Connor KJ, Smith S, Utterback TR, White O, Rubens CE, Grandi G, Madoff LC, Kasper DL, Telford JL, Wessels MR, Rappuoli R, Fraser CM (2005) Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: implications for the microbial "pan-genome". Proc Natl Acad Sci U S A 102: 13950-13955.

28. Contreras-Moreira B, Vinuesa P (2013) GET_HOMOLOGUES, a versatile software package for scalable and robust microbial pangenome analysis. Appl Environ Microbiol 79: 7696–7701.

29. Li L, Stoeckert CJ Jr, Roos DS (2003) Ortho MCL: identification of ortholog groups for eukaryotic genomes. Genome Res 13: 2178–2189.

30. Kanehisa M, Furumichi M, Sato Y, Ishiguro-Watanabe M, Tanabe M (2020) KEGG: integrating viruses and cellular organisms. Nucleic Acids Res 30: gkaa970.

31. Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25: 1754-1760.

32. Li H (2011) A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. Bioinformatics 27: 2987–2993.

33. Cingolani P, Platts A, Wang le L, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden D (2012) A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. Fly (Austin) 6: 80-92.

34. Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol 33: 1870-1874.

35. Tiwari S, Jamal SB, Hassan SS, Carvalho PVSD, Almeida S, Barh D, Ghosh P, Silva A, Castro TLP, Azevedo V (2017) Two-component signal transduction systems of pathogenic bacteria as targets for antimicrobial therapy: an overview. Front Microbiol 8: 1878.

36. Bott M, Brocker M (2012) Two-component signal transduction in *Corynebacterium glutamicum* and other corynebacteria: on the way towards stimuli and targets. Appl Microbiol Biotechnol 94: 1131-1150.

37. Parish T, Smith DA, Roberts G, Betts J, Stoker NG (2003) The senX3-regX3 two-component regulatory system of *Mycobacterium tuberculosis* is required for virulence. Microbiology (Reading) 149: 1423-1435.

38. Albanesi D, Mansilla MC, de Mendoza D (2004) The membrane fluidity sensor DesK of *Bacillus subtilis* controls the signal decay of its cognate response regulator. J Bacteriol 186: 2655-2663.

39. Ershova AS, Rusinov IS, Spirin SA, Karyagina AS, Alexeevski AV (2015) Role of restriction-modification systems in

prokaryotic evolution and ecology. Biochemistry (Mosc) 80: 1373-1386.

40. Dy RL, Przybilski R, Semeijn K, Salmond GP, Fineran PC (2014) A widespread bacteriophage abortive infection system functions through a Type IV toxin-antitoxin mechanism. Nucleic Acids Res 42: 4590-4605.

41. Mattison K, Wilbur JS, So M, Brennan RG (2006) Structure of FitAB from *Neisseria gonorrhoeae* bound to DNA reveals a tetramer of toxin-antitoxin heterodimers containing pin domains and ribbon-helix-helix motifs. J Biol Chem 281: 37942-37951.

42. Helaine S, Cheverton AM, Watson KG, Faure LM, Matthews SA, Holden DW (2014) Internalization of *Salmonella* by macrophages induces formation of nonreplicating persisters. Science 343: 204-208.

43. Ochman H, Lawrence JG, Groisman EA (2000) Lateral gene transfer and the nature of bacterial innovation. Nature 405: 299-304.

44. Gyles C, Boerlin P (2014) Horizontally transferred genetic elements and their role in pathogenesis of bacterial disease. Vet Pathol 51: 328-340.

45. Möller J, Musella L, Melnikov V, Geißdörfer W, Burkovski A, Sangal V (2020) Phylogenomic characterisation of a novel *corynebacterial* species pathogenic to animals. Antonie Van Leeuwenhoek 113:1225-1239.

46. Konstantinidis KT, Ramette A, Tiedje JM (2006) The bacterial species definition in the genomic era. Philos Trans R Soc Lond B Biol Sci 361: 1929-1940.

47. Iguchi A, Nagaya Y, Pradel E, Ooka T, Ogura Y, Katsura K, Kurokawa K, Oshima K, Hattori M, Parkhill J, Sebaihia M, Coulthurst SJ, Gotoh N, Thomson NR, Ewbank JJ, Hayashi T (2014) Genome evolution and plasticity of *Serratia marcescens*, an important multidrug-resistant nosocomial pathogen. Genome Biol Evol 6: 2096-2110.

48. Snipen L, Almøy T, Ussery DW (2009) Microbial comparative pan-genomics using binomial mixture models. BMC Genomics 10: 385.

49. Merhej V, Royer-Carenzi M, Pontarotti P, Raoult D (2009) Massive comparative genomic analysis reveals convergent evolution of specialized bacteria. Biol Direct 4: 13.

50. Mira A., Martín-Cuadrado A.B., D'Auria G., Rodríguez-Valera F (2010) The bacterial pan-genome:a new paradigm in microbiology. Int. Microbiol 13: 45–57.

51. Sobel JD (1985) New aspects of pathogenesis of lower urinary tract infections. Urology 26: 11-16.

52. Brocker M, Schaffer S, Mack C, Bott M (2009) Citrate utilization by *Corynebacterium glutamicum* is controlled by the CitAB two-component system through positive regulation of the citrate transport genes citH and tctCBA. J Bacteriol 191: 3869-3880.

53. Vebø HC, Solheim M, Snipen L, Nes IF, Brede DA (2010) Comparative genomic analysis of pathogenic and probiotic *Enterococcus faecalis* isolates, and their transcriptional responses to growth in human urine. PLoS One 5: e12489.

54. Tielen P, Rosin N, Meyer AK, Dohnt K, Haddad I, Jänsch L, Klein J, Narten M, Pommerenke C, Scheer M, Schobert M, Schomburg D, Thielen B, Jahn D (2013) Regulatory and metabolic networks for the adaptation of *Pseudomonas aeruginosa* biofilms to urinary tract-like conditions. PLoS One 8: e71845.

55. Garandeau C, Réglier-Poupet H, Dubail I, Beretti JL, Berche P, Charbit A (2002) The sortase SrtA of *Listeria monocytogenes* is involved in processing of internalin and in virulence. Infect Immun 70: 1382-1390.

56. Brune I, Werner H, Hüser AT, Kalinowski J, Pühler A, Tauch A (2006) The DtxR protein acting as dual transcriptional regulator directs a global regulatory network involved in iron metabolism of *Corynebacterium glutamicum*. BMC Genomics 7: 21.

57. Li M, Zhao J, Tang N, Sun H, Huang J (2018) Horizontal gene transfer from bacteria and plants to the arbuscular mycorrhizal fungus *Rhizophagus irregularis*. Front Plant Sci 9: 701.

58. Cheleuitte-Nieves C, Gulvik CA, McQuiston JR, Humrighouse BW, Bell ME, Villarma A, Fischetti VA, Westblade LF, Lipman NS (2018) Genotypic differences between strains of the opportunistic pathogen *Corynebacterium bovis* isolated from humans, cows, and rodents. PLoS One 13: e0209231.

59. Sekizuka T, Yamamoto A, Komiya T, Kenri T, Takeuchi F, Shibayama K, Takahashi M, Kuroda M, Iwaki M (2012) *Corynebacterium ulcerans* 0102 carries the gene encoding diphtheria toxin on a prophage different from the *C. diphtheriae* NCTC 13129 prophage. BMC Microbiol 12: 72.

Corresponding author

Haitham Elbir, PhD
Camel Research Center, King Faisal University, 400 Al-Hasa, 31982, Saudi Arabia
Phone:00249967845711
Email: haythamalbur@hotmail.com

Conflict of interests: No conflict of interests is declared.