ISSN: 0975-833X

Available online at http://www.journalcra.com

International Journal of Current Research
Vol. 8, Issue, 11, pp.40933-40939, November, 2016

INTERNATIONAL JOURNAL OF CURRENT RESEARCH

RESEARCH ARTICLE

DRAFT GENOMES OF VANCOMYCIN-RESISTANT ISOLATES OF ANIMAL *ENTEROCOCCUS FAECIUM*

***,[1]Ashwag Shami, [2]Malcolm Horsburgh and [2]Alistair C. Darby**

[1]Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia
[2]The university of Liverpool, Liverpool, UK

ARTICLE INFO

Article History:
Received 15th August, 2016
Received in revised form
05th September, 2016
Accepted 23rd October, 2016
Published online 30th November, 2016

Key words:

Vancomycin-resistant,
Animal *Enterococcus faecium*.

ABSTRACT

Enterococci are Gram-positive bacteria that inhabit the gastrointestinal tract of humans and animals as commensal flora. In recent years two species, *Enterococcus faecalis* and *Enterococcus faecium,* have become an increasing medical concern by virtue of their ability to gain and spread antibiotic resistance. In this study, genomes of vancomycin-resistant isolates of *E. faecium* from pig, chicken and calf were sequenced using 454 platforms. The assembled genomes were annotated and compared with human *E. faecium* isolates to identify their repertoire of genes potentially associated with colonising each host. Phylogenomics of *E. faecium* was used to investigate the relationship between animal and human strains. The genomes of the chicken, pig and calf isolates differed in size (2.5 Mb to 3.3 Mb) with the size difference due to horizontally-acquired elements (mostly phage, transposons and insertion sequences); the chicken isolate genome contained five prophages. The aims were to sequence the genomes of three vancomycin-resistant isolates of *E. faecium* from chicken, calf and pig using next generation pyrosequencing on the Roche 454 titanium platform. These genomes were selected specifically to investigate host adaptation in mammalian hosts.

Citation: Ashwag Shami, Malcolm Horsburgh and Alistair C. Darby, 2016. "Draft genomes of Vancomycin-resistant isolates of animal *Enterococcus faecium*", *International Journal of Current Research*, 8, (11), 40933-40939.

INTRODUCTION

Bacterial diseases represent a major source of morbidity and mortality amongst humans and animals. Pathogenic bacteria comprise a diverse range of species, which have discrete virulence mechanisms. A good knowledge and understanding of these mechanisms is necessary to design successful new therapies against bacterial diseases and manage the emergence of novel isolates. The design of therapies is limited due to the extent of information about the pathogenesis of some diseases being limited or non-existent (Donkor, 2013). Genome sequencing, combined with interpretation using Bioinformatic analyses of genome data, has dramatically extended our understanding of bacterial pathogens, particularly with respect to their ecology, evolution, and pathogenesis (Tang and Holden, 1999, Donkor, 2013). Doolittle (1999) states that the ability to exploit complete genome sequences of microbes offers many opportunities for medicine and delivers an abundance of knowledge for interrogating evolutionary networks. Greater than 1,800 bacterial genomes, including the majority of bacterial pathogens, have now been completely sequenced (Ribeiro, Przybylski *et al*., 2012). The resource of sequenced genomes and the direct access to genome data have advanced studies in biology and has given birth to a new science called genome-based biology (Garcia-Vallve, Romeu *et al*., 2000). The typical bacterial genome consists of a single circular chromosome, however there are exceptions, with several medically significant bacteria having two or more chromosomes, including *Burkholderia*, *Brucella, Vibrio*, and *Leptospira* species; several species have linear chromosomes, for example *Borrelia burgdorferi* (Guzman, Romeu *et al*. 2008). Allen *et al.* (2006) indicated that the majority of bacterial genomes are smaller than 5 Mb in size, although species have been described with genomes up to 30 Mb, for example *Bacillus megaterium* (Allen, Price *et al*., 2006). Guzman *et al.* (2008) establish that the difference in bacterial genome size appears related to lifestyles, whereby obligate pathogen species have smaller genomes than parasitic species, which in turn have smaller genomes than free-living species. The nucleotide composition in bacterial genomes varies across bacteria. The GC (guanosine-cytosine) content may differ locally within a genome, but it is relatively constant within a bacterial genus and species, varying from ~25% GC in *Mycoplasma spp* to ~75% in *Micrococcus* species. The variation in GC content within a single genome was used to determine the acquisition of genomic portions by horizontal

****Corresponding author: Ashwag Shami,***
Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

gene transfer, classically pathogenicity islands, since these frequently have a different GC ratio (Walk, Alm *et al*., 2007).

MATERIALS AND METHODS

Genome sequencing

The genome of three vancomycin-resistant animal *E. faecium* isolates has been sequenced by whole genome shotgun using 454 pyrosequencing. The pyrosequencing were performed by generating standard fragment template 8 Kb DNA libraries, which were multiplex identifier (MID) tagged to allow multiple samples to be run in a single plate region, using the GS-FLX 454 Life Sciences through The Center for Genomic Research (CGR) in Liverpool University.

Plasmid amplification

A colony of each strain was inoculated into 10 ml of THB then incubated overnight with shaking at 37°C. 0.1 ml of the overnight culture was transferred to 50 ml of fresh THB and incubated with shaking at 37°C for 4h. 25 ml of cells were added to 500 ml of THB and incubated with shaking at 37°C for exactly 2 h. Chloramphenicol was added to the culture to a final concentration of 170 µg ml-1 and incubated contrived at 37°C overnight. Plasmids were purified using QIAGEN Plasmid Mini prep as per manufacturer guidelines but with several modifications. 100 ml of the overnight culture was used to extract the plasmid. Lysozyme was added to a final concentration of 100 µg ml-1 and the cells were incubated for 10 min at37°C prior to plasmid purification. For cell lysis, buffer P2 was added and cells were incubated for 5 min at 37°C. Finally, the resulting plasmid DNA extract was electrophoresed on 1% (w/v) agarose gel and electrophoresed for 2 h at 90V.

Genome analysis

Genome annotations were managed using RAST server (http://rast.nmpdr.org) and IMG/ER (Integrated Microbial genomics) (https://img.jgi.doe.gov). Gene structure was assigned by the automated gene-calling algorithm, Prokka (version1.8) (http://www.vicbioinformatics.com/software.prokka.shtml) using default parameters. To validate the prokka gene prediction, the open reading frames (ORF) were compared to published sequences using BLASTn. After the gene-finding progression, different types of investigation were made in order to predict the function of the encoded proteins. BLAST search algorithm was used to examine the homology of the putative ORFs (DNA and protein). Functional classification of ORFs was based on homology search against COGs. Protein function annotation was constructed based on the homology search against NCBI protein database. Mauve progressive alignments to determine conserved sequence segments most likely to be conserved in recombinational events were determined using the Mauve algorithm.

RESULTS AND DISCUSSION

Genome sequencing and assembly

The genome sequences of *E. faecium* strain E429, isolated from chicken, strain E172, isolated from calf and strain E142, isolated from pig, were determined using the GS-FLX sequencing platform (454 Life Sciences. The insert library representing each genome was sequenced extensively to provide reads for each *E. faecium* isolate of 849,986, 366,122 and 335,440, respectively for E429 (chicken), E172 (calf) and E142 (pig) (Table 1). For each respective strain, these reads were assembled into 922, 786 and 136 contigs respectively. The longest scaffold gives the best approximation for the size of the three genomes although the number of scaffolds obtained for strain E429 (chicken), E172 (calf) and E142 (pig) were 19, 18 and 3. The chromosome of the animal strains of *E. faecium* varies in size, therefore, from approximately 3.38 Mb in the chicken strain to 2.94 Mb in the calf strain and 2.52 Mb in the pig strain, with a GC-content of 38.75 %, 38.67 % and 38.13 %, respectively (Table 1). Associated with each genome assembly are 62, 67 and 55 tRNAs respectively for strain E429 (chicken), E172 (calf) and E142 (pig) and markedly different numbers of ribosomal genes with 11 rRNAs (1 x 5S, 3 x 16S and 7x 23S), 14 rRNAs (2 x 5S, 4 x 16S and 8 x 23S) and 3 rRNAs (1 of 5S, 1 of 16S and 1 of 23S), respectively for chicken, pig and calf.

Annotation of the *E. faecium* genome animal strains

The genomes of the chicken, calf and pig strains were annotated using IMG-ER (Integrated Microbial Genomes Expert Review). The initial annotation analysis identified 3,574, 2,892 and 2,641 protein coding genes in chicken, calf and pig, respectively. Approximately 2 %, 2.72 % and 2.15 % of the genes in the animal strain genomes, respectively, determine structural RNAs. The remaining 98% of predicted ORFs in strain E429 (chicken), 97.28% in strain E172 (calf) and 97.85% in strain E142 (pig) were studied using homology analyses with sequence databases, which identified that 74% (2,708), 78% (2,325) and 79% (2,147) of the predicted ORFs, respectively, were likely to be functional proteins. Nearly 10% of the genomes are non-AGCT bases.

Table 1. Structural features and genome composition features associated with the sequenced genomes of *E. faecium* strains E429, E172 and E142

Genomic features	E429 (Chicken)	E172 (Calf)	E142 (Pig)
DNA, total number of bases	3383541	100	2948249
Estimated genome size	3.4 MB	2.9 MB	2.5 MB
Number of scaffolds	19	18	3
Number of contigs scaffolded	179	204	85
Scaffold G + C content	38.1%	38.2%	38.1%
Number of contigs	922	786	136
Total number of assembled bases	3383541	2948249	2525775
DNA coding number of bases	2700854	79.8	2310519
DNA G+C number of bases	1311102	38.7	1140083
Genes total number	3647	100	2973
Protein coding genes	3574	98	2892
Protein coding genes with COGs	2437	66.8	2186

Many bacterial genomes have been described to contain repetitive DNA. These repeat sequences are typically 400 bp in size (Delihas 2011). Analysis of the genomes of the animal *E. faecium* strains using the software package Unipor UGENE determined that there were 1885, 1758 and 1422 short tandemly repeated sequences (STRs) in the chicken, calf and pig strains, respectively. These STRs have a repeat length of 3 bp and tandem size from 9-10 bp. In addition, 750, 550 and 285 short sequence repeats (SSRs) were found in strains E429, E172 and E142, respectively; with a minimum repeat length of 15 bp and a distance between the repeats of 2 bp to 2000kb.

UNIPOR-UGENE displays approximate repeat sequences found in the DNA sequence. The repetitive sequence elements in the animal *E. faecium* genome sequences have a high sequence identity and high copy number. The observed genome inversions could be derived from these repeats. Qin *et al.* (2012) demonstrate that the genome size of *E. faecium* isolated from humans ranges from 2.50 Mb (strain E1039) to 3.14 Mb (strain 1,230,933). The numbers of protein-coding genes range from 2,587 (E1039) to 3,118 (strain TX0133A). By comparing the size of human *E. faecium* sequenced strains with animal *E. faecium* sequenced strains in this study, it is clear that the calf strain has the largest genome among all *E. faecium* strains in the database. The large size of the genome could reflect a capacity of the bacterium to compete and survive in a nutritionally complex niche. The nutritional and physiochemical environment of the gastrointestinal tract might demand increased capability and versatility of this species relative to human isolates. When compared with other *Enterococcus* species including *E. faecalis, E. gallinarum* and *E. casseliflavus*, *E. faecium* isolates were found to have an intermediate genome size. *E. gallinarum* and *E. casseliflavus* have the largest genome size range from 3.4 to 3.6 kb (IMG-Integrated Microbial Genomes, Palmer, Godfrey *et al*. 2012, Qin, Galloway-Pena *et al*., 2012). van Schaik *et al.* (2010) explained that this variation in genome size across *Enterococcus* species was proposed to occur due to expansion within species due to duplication and horizontal gene transfer. The mean genome size of the majority of human infection isolates and epidemic isolates, including the clonal complex 17 (CC17) genogroup, is significantly larger (2.84 to 2.98 Mb) than that of isolates from faeces of non-hospitalised humans (2.71 to 2.84 Mb) or animal isolates and sporadic human infection isolates (2.59 to 2.75 Mb). This difference could represent the effect of cycles of infection and survival in the hospital being correlated with the acquisition of new genes (Lebreton, van Schaik *et al*., 2013).

Mobile genetics elements in animal *E. faecium*

Many plasmids have been described in *Enterococcus* species that confer resistance to antimicrobials and heavy metals. To first investigate the extra-chromosomal plasmid content of the three animal strains of *E. faecium*, plasmid DNA was purified and visualised by gel electrophoresis. Three similarly sized plasmids were observed in the three animal strains, estimated at ~ 4.7 kb in size (Figure 1). The calf strain (E172) potentially contained at least one more plasmid of smaller size (~ 1.5 kb)

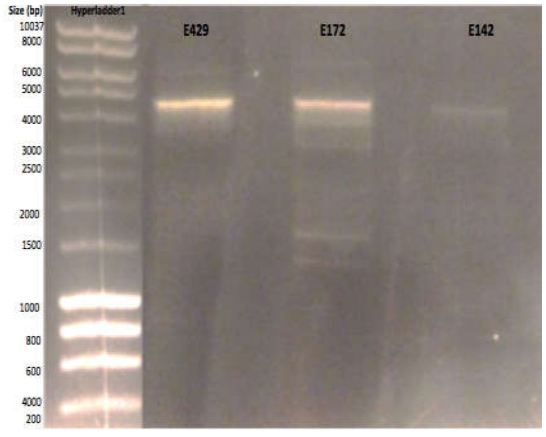

Figure 1. Gel-electrophoresis of plasmid DNA. Lanes from left to right: Hyperladder1; E429 (chicken strain); E172 (calf strain); E142 (pig strain)

To characterise the plasmid complement of the three animal strains *in silico* a comparative analysis was made with the 34 *E. faecium* plasmid sequences that were publicly available. This analysis indicated that the animal strains of *E. faecium* isolated from chicken and pig, each contain DNA corresponding to mega-plasmids present in the closed genomes of *E. faecium* Aus0004, DO and strain Aus00085. Strains E429 (chicken) and E142 (pig) appear to have the same mega plasmid, but located with a different synteny (scaffold 1 and 2, respectively). Strain E172 (calf) only possesses segments of this mega plasmid. The plasmid sequence identified in animal isolates were found to have homology with strain DO plasmids (DO1 (CP003584.1, 36.26 Kb), DO2 (CP003585.1, 66.25 Kb), DO3 (CP003586.1, 251.93 Kb), strain Aus0004 plasmid Aus0004_p1 (CP003352.1, 56.52Kb) and strain Aus0085 plasmids P1 (CP006621.1, 130.72 Kb), P2 (CP006622.1, 67.31 Kb) and P3 (CP006623.1, 31 Kb). The annotation of the DO, Aus0004 and Aus0085 identified plasmids that found in the complete genomes of *E. faecium,* which have homology with animal isolates plasmid, reveal a variety of encoded functions, including toxin–antitoxin, sortase A and an *LPXTG* cell wall anchor protein. In addition, the plasmids contain genes encoding tetracycline resistance and multiple bacteriocin genes. Some of these genes may be found on plasmids but they are not necessarily plasmid genes. *Enterococcus* species harbour plasmids which often mediate resistance to antimicrobials and heavy metals, provide enhance virulence and/or encode DNA repair mechanisms (Arias, Panesso *et al.,* 2009, Garcia-Migura, Hasman *et al.,* 2009). The mega-plasmids identified in chicken (E429) and pig (E142) harbour genes encode potential adhesi with the presence of sortase A and an *LPXTG* cell wall anchor protein. It is known that *LPXTG* surface proteins may play a significant role in the pathogenesis of *E. faecium* in hospital-related infections (Hendrickx, van Wamel *et al.,* 2007, Lam, Seemann *et al.,* 2012).

The mega-plasmid was found in the genome of chicken, calf and pig, which is unique to these strains and it encodes heavy metal resistance genes for resistance to lead, cadmium zinc and mercury. The mega-plasmid (56kb) is apparently integrated into the chromosome of the chicken *E. faecium* strain (E429). Due to the homology between plasmids and the genome an occurrence of a single homologous recombination event can integrate a complete plasmid into the chromosome (Heap, Ehsaan *et al.,* 2012). Homologous recombination following transformation will potentially occur if plasmids are incapable of replication in a specific host. These insertion incidents have been widely detected in *E. faecalis, E. coli*, *B. subtilis*, *S. pneumoniae*, *L. plantarum* and *L. lactis subsp. Lactis* (Casey, Daly *et al.,* 1991). Several of the novel animal genes (22 genes encoding hypothetical proteins) were located on a plasmid. Carbohydrate utilisation operons were identified in chapter 4 as being located on plasmids and these operons were identified with specificities for citrate, and ascorbate, resistance to heavy metal including lead, cadmium, zinc and mercury. These genes form the novel region in the chicken genome map. Analysis of plasmid genome content across all of the *E. faecium* genomes revealed relationships based on shared DNA sequences (Figure 2).Genes carried by plasmids in animal *E. faecium* were found to be common across *E. faecium* strains, including the commensal isolates. The co-occurrence of the plasmid with animal and CC17 strains show strong association since most of the animal strains were located in a clade different from the CC17 strains, which suggested that animal strains contains

plasmid genes specific for animal host. Some of the plasmid genes, for example helix-destabilizing protein, helix-turn-helix domain protein and a sortase (surface protein transpeptidase) were found as core genes in *E. faecium* isolates (Figure 2).

Bacteriophage

Phages have been described that were resident in *E. faecium* strains or that were shown to infect the species (Mazaheri Nezhad Fard, Barton *et al*., 2010, van Schaik, Top *et al*., 2010,

Table 3. Virulence factors in animal *E. faecium*

Virulence gene product	E429 (chicken)	E172 (calf)	E142 (pig)
LPXTG surface protein	10	5	10
Collagen adhesin precursor	1	1	4
Esp	1	1	1
EspA	7	0	0
PilA	2	1	1
PilB	0	0	1
SagA	1	1	1
Hemolysin	3	3	3

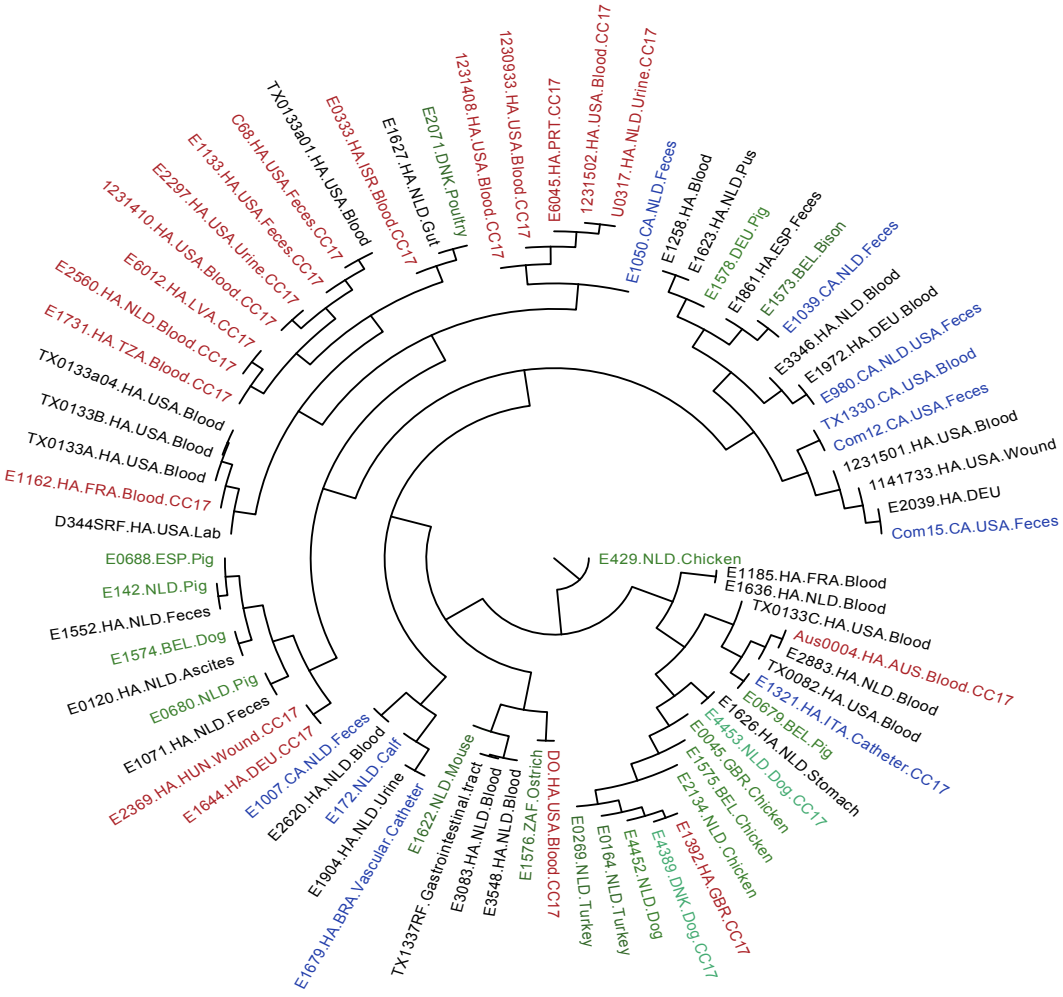

Figure 2. A presence and absence tree of plasmid orthologues in *E. faecium.* The red clade indicates CC17 genotype isolates, blue indicates commensal strains, green indicates animal isolates and black indicates other clinical isolates

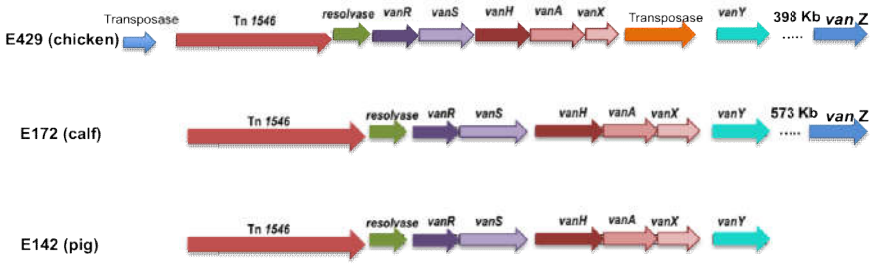

Figure 3. Vancomycin resistance genes in animal *E. faecium*. The arrows show a similar Tn*1546* linked operon that is composed of 6 *van* genes (*van*R, S, H, A, X, and Y)

Yasmin, Kenny *et al*., 2010, Galloway-Pena, Roh *et al*., 2012). Van Schaik *et al.* (2010) indicated that the prophages that have been induced from *E. faecium* are *Siphoviridae* and morphologically identical to prophages induced from *E. faecalis*. The genome sequences of the *E. faecium* strains isolated from chicken (E429), calf (E172) and pig (E142) contain prophages. The genome size differences between the chicken strain and other two animal strains are mostly due to the acquisition of horizontally transferred of genetic material, and a major part of this derives from temperate bacteriophage. Six phage regions were found in chicken strain E429 compared with only one in calf and pig strains (Shami and Horsburg, 2016).

Investigating animal *E. faecium* genomes with regards to virulence, resistance and survival

BLAST analysis of candidate virulence factor genes present in human strains of *E. faecium* confirmed the presence of multiple virulence genes. The enterococcal surface protein (encoded by *esp)*, collagen adhesin precursor (encoded by *acm)*, secreted antigen SagA, pilus (encoded by *pilA* and *pilB*) and hemolysin (Table 3) are variably present among the three sequenced strains revealing that known virulence determinants reside in their genomes. The virulence proteins in animal *E. faeciun* have 93 to 100% similarity with virulence genes in *E. faecium* as a whole, namely collagen adhesin precursor (AAN12397), PilB (ACI49665), PilA (ACI49671) and SagA (AF242196_3). Collagen adhesin precursor gene was found in the chicken strain (position 520769-522466), calf strain (1474869-1477595) and with four copies (1504841-1506127, 1506662-1507090, 1507087-1507290 and 1507290-1507571) in the pig strain. The PilA gene is located in positions 178302-178679 and 180785-182434 in the chicken strain and 2429443-2431419 in the pig strain. Collagen adhesin precursor gene was found in most *E. faecium* isolates including clinical, commensal and animal. However, a novel collagen adhesin precursor homolog was found only in bird isolates and the calf strain (E172). PilB was found in the pig strain only (119904-121781). The SagA gene is located at positions 2654224-2655801, 1761135-1762700 and 1798805-1800376 in chicken, calf and pig, respectively. Hemolysin genes are located at 80020-801583, 971478-972131, and 2349596-2350903 in the chicken strain, 1051269-1051922, 1201273-1202649 and 2072066-2073367 in the calf strain and 1070554-1071207, 1229858-1231234 and 2099474-2100775 in the pig strain. LPXTG family cell-wall anchored proteins were found in the three animal *E. faecium* genomes as multiple copies. At least 5 of these genes are novel since no significant similarity was found in the NCBI database which includes those from Gram-positive species including *Staphylococcus* and *Lactobacillus* species. LPXTG in positions 3145123-3145713 and 3169914-3170672 in chicken strain share high level of similarity (89%) with LPTXG in *Lactobacillus brevis* and (98%) to Cna protein B-type domain protein in *Staphylococcus aureus*, respectively. The gene encoding hyaluronidase was absent from the three animal *E. faecium* isolates, in contrast to its presence in all CC17 genotype isolates, confirming it represents a signature of this CC17 genotype. The gene encoding the enterococcal surface proteins Esp and EspA share low level of similarity (23 to 36%) with Esp (ZP_06678454) and cell wall surface anchor family protein EsbA (ZP_06702708) in *E. faecium* strains E1162 and U0317, respectively. The *Esp* gene is located at positions 1961868-1965038, 104424-107594 and 133123-135006 in the genomes of chicken, calf and pig, respectively. Comparative analyses of antibiotic resistance genes among *E. faecium* isolates were previously reported by Qin *et al.* (2012) and Lebreton *et al.* (2013) and revealed the widespread occurrence of antibiotic genes in *E. faecium* species. A comparative analysis of antibiotic resistance genes in the three sequenced animal *E. faecium* isolates in this study was done by performing BLAST searches against antibiotic resistance sequence databases. Multiple antibiotic resistance genes were identified in the chicken (E429), calf (E172) and pig (E142) strain genomes such as genes encode resistance to antibiotics as follows: *ermA* and *ermB* (erythromycin), *lunB* (lincomycin), *aacA-aphD* (gentamycin), *aad6* (spectinomycin) and *aadE* (streptomycin); *cat* (chloramphenicol), *tetM* and *tetL*(tetracycline), *van A* (vancomycin type A), *van B* (vancomycin type B), *fos* (fosfomycin), *parC* and *g1rA* (fluoroquinolone and ciprofloxacin), *Pbp5-R* (ampicillin), *st* (streptothricin); *azlC*(azaleucine), *ble*(bleomycin), *fmtC* (oxacillin) and *vgb* (streptogramin).

Each of the sequenced animal *E. faecium* strains in this study is vancomycin resistant. To explore the nature of this resistance the *van* operons were identified by homology. In strain E429 (chicken) the *van* operon is about 7.6 kb in size (2874238bp - 2881898pb), with *van*Z located 398 kb distant to the operon (Figure 5.5). The operon is surrounded by mobile elements including transpoase TnA (Tn*1546*), a transcriptional regulator, Tn*916*, DNA topoisomerase and a tetracycline resistance gene is located 1.5kb upstream. Unexpectedly, a second copy of *van*R, *van*S and *van*Y are clustered together in an operon, 2.5 kb in size located about 2 Mb distant (647926-650500). In the calf strain (E172), the Van operon is smaller 5.987 kb (2514921bp - 2520908bp), with *vanZ* is located 573 kb distant. The operon in the pig strain is a similar size as the calf strain operon (located 24141042bp- 2408056bp), however, *vanZ* is absent. Mutations in the *gyrA* or *parC* subunit genes that are responsible for fluoroquinolone and ciprifloxacin resistance were found in the three animal *E. faecium* strains. The described amino acid change E to K occurs in codon 88 of the *gyrA* gene and amino acid change E to K in occurs in codon 86 of *parC*. Fluoroquinolone, streptothricin and azaleucine resistance were found only in the animal strains, which might reflect that these antibiotics are used in animal husbandry. Gentamicin resistance was also found in the three animal isolates. A previous study by Qin *et al* (2012) reported that 15 genes encoding LPXTG family cell wall-anchored proteins with MSCRAMM-like features were present in the complete genome of E. *faecium* (TX16). The LPXTG family cell wall-anchored proteins present in the three animal strains are novel or share homology with other Gram-positive species such as *Staphylococcus* and *Lactobacillus* species. Qin *et al.* (2012) identified that in 21 *E. faecium* draft genomes, all of the MSCRAMM-encoding genes were broadly dispersed, excluding (*esbA*), which was only present in HA-clade isolates. Multiple copies of *esbA*-likegenes were also found with low sequence identity (25-37%) in the three animal *E. faecium* genomes in this study, possibly indicating they are novel MSCRAMMs. Enterococcal surface protein (Esp) and collagen-binding adhesin (Acm) contribute to colonisation and infection, however recent studies have determined that Esp is not fundamental for infection in murine infection models (Heikens, Leendertse *et al.*, 2009). An *esp*-like gene was found in the three animal *E. faecium* genomes but the low percentage identity (24%), possibly indicating it is distinct. Collagen adhesin genes with percentage identity ranging from 61% to 100% were found in the three animal strains. This gene

is present as a pseudo gene in all of the *E. faecium* commensal isolates except 1,141,733 in Qin *et al.* (2012) study and *acm* pseudo genes were also found in clinical *E. faecium* that do not belong to CC17 genotype.

The presence and absence of 19 antibiotic resistance genes across 72 *E. faecium* isolates including clinical, animal and commensal was also searched. These data correspond to previously published frequency data for a smaller set of isolates (Qin, Galloway-Pena *et al.,* 2012, Lebreton, van Schaik *et al.,* 2013). Comparative analysis of antibiotic resistance revealed that commensal, animal and clinical isolates have clear differences in terms of their resistance profile. All of the clinical and animal isolates have multiple resistance determinants, excluding strains 1,231,501 and E1039. The clinical strain (1,231,501) lacks all antibiotic resistances including pbp5-R, may have lost genes through recombination and acquired pbp5-S. Certainly, 1,231,501 was shown to be a hybrid of clinical and commensal genomes by Palmer, *et al.* (2012) and the (hybrid) region including pbp5-S, which could clarify the origin of pbp5-S in this strain. In contrast, Qin *et al.* (2012) stated that all of the commensal-associated isolates (1,141,733, Com12, Com15, E980 and TX1330) lacked genes for antibiotic resistance to chloramphenicol, erythromycin, streptomycin, spectinomycin, gentamycin, vancomycin, ciprofloxacin and ampicillin. Strain E1039, which is a commensal isolate, but genetically closer to the clinical strains, has an ampicillin resistance gene. In 2013, same analysis applied by Lebreton *et al* to two other commensal isolates (E1050 and E1007) showed their resistance to streptomycin and spectinomycin, while E1050 also encoded resistance to fosfomycin. Disease treatment and growth promotion could explain the multiple antimicrobial resistance of most *E. faecium* isolates, including animals strains. The delivery of low levels of antimicrobials has apparently resulted in considerable colonisation of animals with antibiotic resistant bacteria, such as *E. coli* strains and acquisition of resistance in *E. coli* in the intestinal flora of the farmers has been described (Marshall and Levy 2011, Lebreton, van Schaik *et al.,* 2013). Aarestrup (2000) reported that resistance to streptothricin antibiotics has been described in Gram-negative bacteria as a result of using nourseothricin as an antimicrobial feed promoter in industrial animal farms in Germany. In addition, resistance to streptogramins may be related to the use of virginamycin, as a feed promoter combined in agriculture for animal food production. Virginamycin use was prevented in Denmark in 1998 followed by the rest of the EU in 1999. Virginamycin resistance was identified in this study in all three animal *E. faecium* and these strains were isolated from the same geographic region (The Netherlands) and resistance might also have arisen from the historic use of this antibiotic as a feed promoter in Dutch agriculture.

Conclusion

Sequencing, assembly, annotation and homology of three animal strains of *E. faecium* isolated from chicken, calf and pig y has described. A mega plasmid was identified in the genomes of the sequenced chicken, calf and pig *E. faecium* isolates, which is specific to these strains. A second mega plasmid identified in the sequenced chicken and pig genomes was also present in the humans isolate genomes. The study of MGE is challenging since there are many complications with annotating MGE sequences and therefore as a whole they are poorly annotated, particularly as part of bacterial-genome sequencing projects. For example, few phages have previously been well characterised in *E. faecium* and only recently one complete phage genome (IME-EFm1) was reported (Wang *et al*., 2014). The narrow sequence homology among functionally equivalent phage-encoded proteins complicates the study of their function (Ford *et al*., 2003). There is a requirement for developments in bioinformatics of MGEs to identify their unique features.

Acknowledgment

I wish to acknowledge the Royal Saudi Government, Ministry of Higher Education, Princess Nourah bint Abdulrahman University for the support.

REFERENCES

Allen, T. E., N. D. Price, A. R. Joyce and B. O. Palsson 2006. "Long-range periodic patterns in microbial genomes indicate significant multi-scale chromosomal organization." *PLoS Comput Biol*., 2(1): e2.

Arias, C. A., D. Panesso, K. V. Singh, L. B. Rice and B. E. Murray, 2009. "Cotransfer of antibiotic resistance genes and a hylEfm-containing virulence plasmid in Enterococcus faecium." *Antimicrob Agents Chemother*, 53(10): 4240-4246.

Casey, J., C. Daly and G. F. Fitzgerald, 1991. "Chromosomal integration of plasmid DNA by homologous recombination in Enterococcus faecalis and Lactococcus lactis subsp. lactis hosts harboring Tn919." *Appl Environ Microbiol*., 57(9): 2677-2682.

Donkor, E. S. 2013. "Sequencing of bacterial genomes: principles and insights into pathogenesis and development of antibiotics." *Genes (Basel)*, 4(4): 556-572.

Galloway-Pena, J., J. H. Roh, M. Latorre, X. Qin and B. E. Murray, 2012. "Genomic and SNP analyses demonstrate a distant separation of the hospital and community-associated clades of Enterococcus faecium." *PLoS One,* 7(1): e30187.

Garcia-Migura, L., H. Hasman and L. B. Jensen, 2009. "Presence of pRI1: a small cryptic mobilizable plasmid isolated from Enterococcus faecium of human and animal origin." *Curr Microbiol*., 58(2): 95-100.

Garcia-Vallve, S., A. Romeu and J. Palau 2000. "Horizontal gene transfer in bacterial and archaeal complete genomes." *Genome Res*., 10(11): 1719-1725.

Guzman, E., A. Romeu and S. Garcia-Vallve, 2008. "Completely sequenced genomes of pathogenic bacteria: a review." *Enferm Infecc Microbiol Clin*., 26(2): 88-98.

Heap, J. T., M. Ehsaan, C. M. Cooksley, Y. K. Ng, S. T. Cartman, K. Winzer and N. P. Minton, 2012. "Integration of DNA into bacterial chromosomes from plasmids without a counter-selection marker." *Nucleic Acids Res*., 40(8): e59.

Heikens, E., M. Leendertse, L. M. Wijnands, M. van Luit-Asbroek, M. J. Bonten, T. van der Poll and R. J. Willems 2009. "Enterococcal surface protein Esp is not essential for cell adhesion and intestinal colonization of Enterococcus faecium in mice." BMC Microbiol9: 19.

Hendrickx, A. P., W. J. van Wamel, G. Posthuma, M. J. Bonten and R. J. Willems, 2007. "Five genes encoding surface-exposed LPXTG proteins are enriched in hospital-adapted Enterococcus faecium clonal complex 17 isolates." *J Bacteriol*., 189(22): 8321-8332.

Jacob, F. and J. Monod, 1961. "Genetic regulatory mechanisms in the synthesis of proteins." *J Mol Biol*., 3: 318-356.

Lam, M. M., T. Seemann, D. M. Bulach, S. L. Gladman, H. Chen, V. Haring, R. J. Moore, S. Ballard, M. L. Grayson, P. D. Johnson, B. P. Howden and T. P. Stinear 2012. "Comparative analysis of the first complete Enterococcus faecium genome." *J Bacteriol*., 194(9): 2334-2341.

Lebreton, F., W. van Schaik, A. M. McGuire, P. Godfrey, A. Griggs, V. Mazumdar, J. Corander, L. Cheng, S. Saif, S. Young, Q. Zeng, J. Wortman, B. Birren, R. J. Willems, A. M. Earl and M. S. Gilmore, 2013. "Emergence of epidemic multidrug-resistant Enterococcus faecium from animal and commensal strains." *MBio*., 4(4).

Marshall, B. M. and S. B. Levy 2011. "Food animals and antimicrobials: impacts on human health." *Clin Microbiol Rev*., 24(4): 718-733.

Mazaheri Nezhad Fard, R., M. D. Barton and M. W. Heuzenroeder 2010. "Novel Bacteriophages in Enterococcus spp." *Curr Microbiol*., 60(6): 400-406.

Pennisi, E. 1998. "Evolution: Genome Data Shake Tree of Life." *Science*, 280((5364)): 672-674.

Qin, X., J. R. Galloway-Pena, J. Sillanpaa, J. H. Roh, S. R. Nallapareddy, S. Chowdhury, A. Bourgogne, T. Choudhury, D. M. Muzny, C. J. Buhay, Y. Ding, S. Dugan-Rocha, W. Liu, C. Kovar, E. Sodergren, S. Highlander, J. F. Petrosino, K. C. Worley, R. A. Gibbs, G. M. Weinstock and B. E. Murray, 2012. "Complete genome sequence of Enterococcus faecium strain TX16 and comparative genomic analysis of Enterococcus faecium genomes." *BMC Microbiol*., 12: 135.

Ribeiro, F. J., D. Przybylski, S. Yin, T. Sharpe, S. Gnerre, A. Abouelleil, A. M. Berlin, A. Montmayeur, T. P. Shea, B. J. Walker, S. K. Young, C. Russ, C. Nusbaum, I. MacCallum and D. B. Jaffe, 2012. "Finished bacterial genomes from shotgun sequence data." *Genome Res*., 22(11): 2270-2277.

Shami, A. and M. Horsburg, 2016. "Comparative genomics of e. faecium bacteriophages." *American Scientific Research Journal for Engineering, Technology & Sciences*, vol26(1).

Tang, C. and D. Holden 1999. "Pathogen virulence genes--implications for vaccines and drug therapy." *Br Med Bull.,* 55(2): 387-400.

van Schaik, W., J. Top, D. R. Riley, J. Boekhorst, J. E. Vrijenhoek, C. M. Schapendonk, A. P. Hendrickx, I. J. Nijman, M. J. Bonten, H. Tettelin and R. J. Willems, 2010. "Pyrosequencing-based comparative genome analysis of the nosocomial pathogen Enterococcus faecium and identification of a large transferable pathogenicity island." *BMC Genomics*, 11: 239.

Walk, S. T., E. W. Alm, L. M. Calhoun, J. M. Mladonicky and T. S. Whittam 2007. "Genetic diversity and population structure of Escherichia coli isolated from freshwater beaches." *Environ Microbiol*., 9(9): 2274-2288.

Yasmin, A., J. G. Kenny, J. Shankar, A. C. Darby, N. Hall, C. Edwards and M. J. Horsburgh, 2010. "Comparative genomics and transduction potential of Enterococcus faecalis temperate bacteriophages." *J Bacteriol.,* 192(4): 1122-1130.

****