JMM Correspondence

The *fms21* (*pilA*)-*fms20* locus encoding one of four distinct pili of *Enterococcus faecium* is harboured on a large transferable plasmid associated with gut colonization and virulence

In the last two decades, enterococci, especially *Enterococcus faecalis* and *Enterococcus faecium*, have emerged as multidrug-resistant opportunistic pathogens causing difficult-to-treat healthcare-associated (HA) infections including endocarditis (Arias & Murray, 2009). Recent surveillances reported a major epidemiological shift of increasing frequency of isolation of *E. faecium* from United States hospitals (Hidron *et al.*, 2008). Furthermore, increased morbidity and mortality due to vancomycin-resistant strains of *E. faecium* have also been reported in patients who have multiple medical problems, such as cancer, transplantation or prior surgery. Using multilocus sequence typing, epidemiological studies have shown that the HA *E. faecium* strains are different sequence types from the commensal strains that colonize the gastrointestinal tract of healthy humans (Top *et al.*, 2008). Further analyses suggested that the recent success of the *E. faecium* hospital-associated genogroup [also referred to as clonal complex 17 (CC17); Top *et al.*, 2008] is due at least in part to the presence of an active form of the *acm* gene encoding collagen adhesin (Nallapareddy *et al.*, 2008) or acquisition of additional genetic loci encoding enterococcal surface protein (*esp*$_{fm}$) (Willems *et al.*, 2001), a hyaluronidase-like glycoside hydrolase (*hyl*$_{Efm}$) (Rice *et al.*, 2003) or *E. faecium* surface proteins (*fms*) (Hendrickx *et al.*, 2007; Sillanpaa *et al.*, 2008, 2009).

Pili and MSCRAMMs (microbial surface components recognizing adhesive matrix molecules) have been implicated in the ability of pathogenic bacteria to adhere to and colonize host tissues, processes important in initiating infections. Recent bioinformatics analyses of the unfinished genome sequence of endocarditis-derived *E. faecium* strain TX0016 (DO) identified 15 of 22 Fms proteins containing MSCRAMM-like characteristics as well as predicted folding into multiple immunoglobulin-like domains (Sillanpaa *et al.*, 2008). Of these, 11 *fms* genes were found clustered in four genomic loci: *fms1* (*ebpA*$_{fm}$)-*fms5* (*ebpB*$_{fm}$)-*fms9* (*ebpC*$_{fm}$), *fms11-fms19-fms16*, *fms14-fms17-fms13* and *fms21 (pilA)-fms20*, and each of these four loci also encodes a class C (subfamily 3) sortase predicted to be involved in pilus biogenesis (Sillanpaa *et al.*, 2008). The proteins encoded by the *ebp*$_{fm}$ genes are orthologues of *E. faecalis* Ebp pili subunit proteins associated with pathogenesis in experimental rat endocarditis and mouse urinary tract infections (Nallapareddy *et al.*, 2006; Singh *et al.*, 2007). The *fms11-fms19-fms16* genes (designated the *orf903-907* gene cluster in Hendrickx *et al.*, 2007) are present on a genomic island flanked by direct repeats, are often found as pseudogenes, and are enriched in the CC17 genogroup (Hendrickx *et al.*, 2007; Sillanpaa *et al.*, 2009). EbpC$_{fm}$ (designated PilB in Hendrickx *et al.*, 2008) and Fms21 (designated PilA in Hendrickx *et al.*, 2008) have been shown experimentally to form two distinct types of *E. faecium* pili (Hendrickx *et al.*, 2008; Sillanpaa *et al.*, 2008). However, the function(s) of these pili are yet to be elucidated.

In our current efforts of sequencing and analysis of the regions surrounding these pilus-encoding loci of TX0016 and TX1330 (a commensal human isolate), we identified a *repA* gene in the vicinity of the *pilA* (*fms21*)-*fms20* locus, suggesting that this locus is likely harboured on a plasmid. The genes encoding PilA (Fms21) and Fms20 of these two strains were separated by two ORFs: SrtC4 and a 252 aa ORF with no cell-wall-anchoring motif exhibiting >25 % similarity to the *E. faecalis* EbpB pilus subunit protein. One class A sortase-encoding gene was also found immediately upstream of *pilA* (*fms21*) in both these strains. To provide experimental proof for the predicted plasmid localization of the *pilA* (*fms21*)-*fms20* locus, we performed *Aspergillus oryzae* S1 nuclease (Sigma) digests of agarose plugs containing genomic DNA of TX0016, TX1330 and TX2158 (a derivative of TX1330RF to which the *hyl*$_{Efm}$-containing plasmid was transferred by conjugation from TX0016; Arias *et al.*, 2009) and four additional endocarditis-derived *E. faecium* strains using a method described earlier (Arias *et al.*, 2009). The S1 nuclease digestion combined with PFGE was previously shown to be useful for detection and estimation of the size of large bacterial plasmids (Barton *et al.*, 1995). The isolates used in this study were selected based on *pilA* (*fms21*), *fms20* and *hyl*$_{Efm}$ (a marker for a transferable plasmid of TX0016; Arias *et al.*, 2009) gene diversity that we previously reported using colony hybridization (Rice *et al.*, 2003; Sillanpaa *et al.*, 2008, 2009); the genotypes of these strains are shown in Table 1.

PFGE followed by Southern hybridizations showed that *pilA* (*fms21*) and *fms20* are indeed located on plasmids of varying sizes, running between 145.5 and 291 kb in PFGE gels, in different *E. faecium* strains (Fig. 1). Southern hybridization profiles for these isolates are in agreement with our previously published colony hybridization data (Sillanpaa *et al.*, 2008, 2009). The extrachromosomal location of *pilA* (*fms21*) and *fms20* genes was further confirmed by I-*Ceu*I (which recognizes a sequence in 23S rRNA genes, thus hybridization with a probe targeting 23S rRNA genes results in six chromosomal bands in each strain) digestion of agarose plugs followed by PFGE and hybridization (data not shown). In a separate experiment, *Eco*RV endonuclease digestion of caesium chloride-extracted, *hyl*$_{Efm}$-containing plasmid (~220 kb) DNA from *E. faecium* strain TC6 (Rice *et al.*, 2009) was performed. Southern hybridizations with

Table 1. Characteristics of *E. faecium* isolates used in this study

Strain	Clinical source; origin; year of isolation/collection	MLST type	hyl_{Efm}*	*pilA* (*fms21*)*	*fms20**
TX0016†	Endocarditis/blood; Houston, TX, USA; 1992	ST18	+	+	+
TX1330RF‡	Rifampicin and fusidic acid resistant version of commensal human isolate TX1330		−	+	+
TX2158‡	Transconjugant of TX1330RF–TX0016		+	+	+
TX0068	Endocarditis/blood; Worcester, MA, USA; 1994	ST18	−	+	+
TX0074	Endocarditis/blood; Valhalla, NY, USA; 1995	ST337	−	+	−
TX0081	Endocarditis/blood; Baltimore, MD, USA; 1996	ST154	+	+	−
TX0082	Endocarditis/blood; Houston, TX, USA; 1999	ST17	−	−	−

*+, Gene present; −, gene absent.
†Also designated DO in previous studies.
‡These isolates were generated in an earlier study (Arias *et al.*, 2009).

pilA (*fms21*) and *fms20* probes of these restriction-digested fragments showed that the hyl_{Efm} and *fms* genes are co-localized in this plasmid as well (data not shown). We have also investigated the location of another CC17 genogroup-enriched (Hendrickx *et al.*, 2007; Sillanpaa *et al.*, 2009) gene cluster, *fms11-fms19-fms16*, by hybridization of Southern blots of S1 nuclease-digested DNA as well as I-*Ceu*I-digested DNA with *fms11*, *fms16* and *fms19* gene probes. The *fms11*, *fms16* and *fms19* genes were found to be on the chromosome of four of four isolates tested and were not present in the transconjugant strain TX2158, indicating that they were not acquired along with transfer of the hyl_{Efm} plasmid.

Of note, we also analysed the *pilA* (*fms21*)-*fms20* surrounding regions of multiple unfinished *E. faecium* whole genome sequences (recently sequenced by the Broad Institute of MIT and Harvard, Cambridge, MA, USA, and available at http://www.ncbi.nlm.nih.gov/sutils/genom_table.cgi) belonging to clinical and non-clinical groups and found *repA* in the

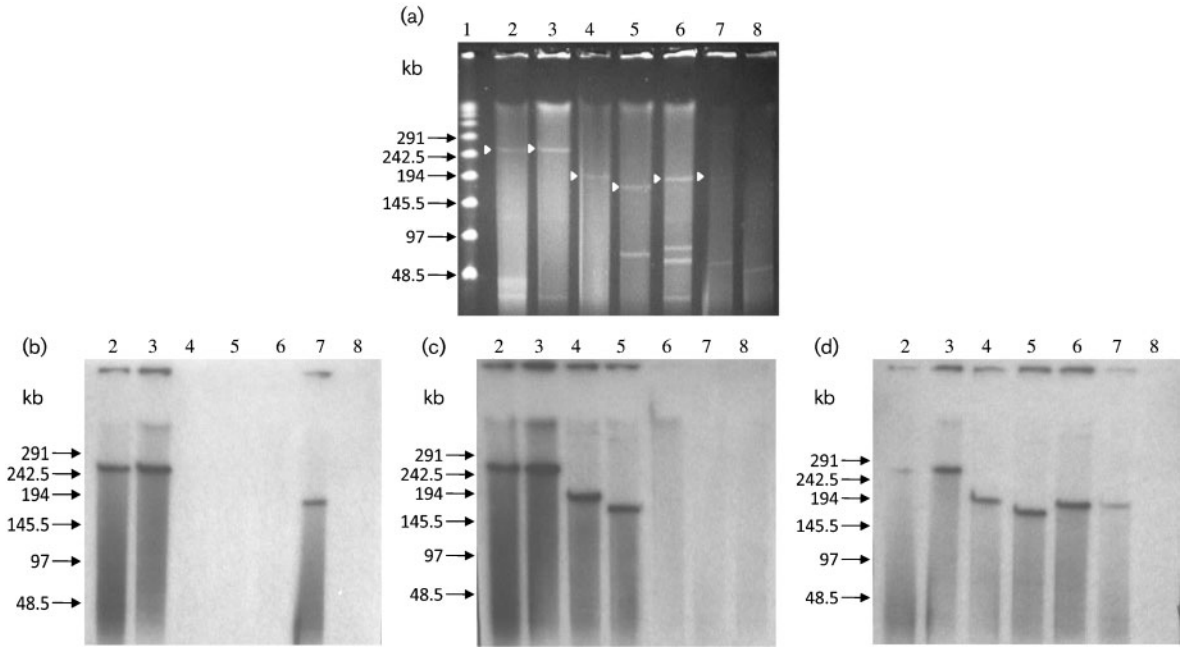

Fig. 1. PFGE of S1 nuclease digested genomic DNA showing large plasmids of *E. faecium* isolates and hybridizations with hyl_{Efm}, *pilA* (*fms21*) and *fms20* probes. (a) PFGE of S1 digestions of total genomic DNA. Plasmid bands are shown as linearized fragments on the gel; the white arrows represent the plasmid bands hybridizing with the different probes tested in the following panels. (b) Hybridizations with the hyl_{Efm} probe. (c) Hybridizations with the *fms20* probe. (d) Hybridizations with the *pilA* (*fms21*) probe. Lanes: 1, lambda ladder; 2, TX0016; 3, TX2158; 4, TX1330RF; 5, TX0068; 6, TX0074; 7, TX0081; 8, TX0082.

neighbourhood [present in the same contig as *pilA* (*fms21*)-*fms20*], further confirming that the *pilA* locus is harboured on a plasmid in multiple genetic backgrounds.

In *E. faecium* strains that are hyl_{Efm}^{+} (TX0016, TX2158 and TX0081), *pilA* (*fms21*) and/or *fms20* probes as well as an hyl_{Efm} probe hybridized to same plasmid band (~250 kb plasmid in TX0016 as well as in the TX2158 transconjugant and to ~170 kb plasmid in TX0081). Interestingly, as shown in Fig. 1, transconjugant TX2158 lost the TX1330RF native plasmid (of size ~190 kb) during acquisition of the hyl_{Efm} plasmid from donor TX0016, possibly due to incompatibility. Sequence analysis of an ~45 kb region spanning *pilA* (*fms21*)-*fms20* and the *repA* region of the native plasmids of TX1330 and TX0016 showed >90 % identity (unpublished results).

Although hyl_{Efm} has been implicated as a possible virulence determinant in multiple worldwide surveys based on its enrichment in the HA genogroup, the exact role of the hyl_{Efm} gene cluster (flanked by insertion elements) in pathogenesis or colonization by *E. faecium* has not been demonstrated. However, it has been recently reported that it is encoded on a large transferable plasmid containing variable antibiotic resistance genes in different isolates (Arias *et al.*, 2009) and the presence of this plasmid enhances the ability of *E. faecium* to colonize the gut of mice (Rice *et al.*, 2009) and increases lethality in a mouse peritonitis model (Arias *et al.*, 2009).

Taken together, we speculate that some ancestral and commensal isolates carried the *pilA* (*fms21*)-*fms20* gene cluster on a large plasmid, explaining the lack of significant differences in its occurrence in CC17 versus non-CC17 isolates (Sillanpaa *et al.*, 2009). We further speculate that the plasmid(s) subsequently acquired multiple antibiotic resistance genes (Arias *et al.*, 2009) and virulence determinants (e.g. hyl_{Efm} gene cluster) via horizontal gene transfer [as also suggested by the striking differences in the size of *pilA* (*fms21*)-*fms20*-containing plasmids found in different clinical strains] and that these traits helped the host strain combat selective pressures in the hospital environment. Our ongoing efforts to completely sequence these large plasmids should unravel important insights into the evolution of these plasmids.

Acknowledgements

This work was supported in part by NIH grant R01 AI067861 from the Division of Microbiology and Infectious Diseases, NIAID, to B. E. M. D. S. K. was supported by a Summer Scholarship Award from the NIH grant T32 AI55449. C. A. A. was supported by NIH Pathway to Independence award 4 R00 AI072961 from the NIAID. D. P. was partially funded by a graduate scholarship from the Instituto Colombiano para el Desarrollo de la Ciencia y Tecnología, 'Francisco José de Caldas', Colciencias.

David S. Kim,[1,2] Kavindra V. Singh,[1,2] Sreedhar R. Nallapareddy,[1,2] Xiang Qin,[3] Diana Panesso,[1,2,4] Cesar A. Arias[1,2,4] and Barbara E. Murray[1,2,5]

[1]Center for the Study of Emerging and Re-Emerging Pathogens, University of Texas Medical School at Houston, Texas, USA

[2]Division of Infectious Diseases, Department of Internal Medicine, University of Texas Medical School at Houston, Texas, USA

[3]Human Genome Sequencing Center, Baylor College of Medicine, Houston, Texas, USA

[4]Molecular Genetics and Antimicrobial Resistance Unit, Universidad El Bosque, Bogota, Colombia

[5]Department of Microbiology and Molecular Genetics, University of Texas Medical School at Houston, Texas, USA

Correspondence: Barbara E. Murray (bem.asst@uth.tmc.edu)

Arias, C. A. & Murray, B. E. (2009). Antibiotic-resistant bugs in the 21st century – a clinical super-challenge. *N Engl J Med* **360**, 439–443.

Arias, C. A., Panesso, D., Singh, K. V., Rice, L. B. & Murray, B. E. (2009). Cotransfer of antibiotic resistance genes and a hyl_{Efm}-containing virulence plasmid in *Enterococcus faecium*. *Antimicrob Agents Chemother* **53**, 4240–4246.

Barton, B. M., Harding, G. P. & Zuccarelli, A. J. (1995). A general method for detecting and sizing large plasmids. *Anal Biochem* **226**, 235–240.

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

Hendrickx, A. P., Bonten, M. J., van Luit-Asbroek, M., Schapendonk, C. M., Kragten, A. H. & Willems, R. J. (2008). Expression of two distinct types of pili by a hospital-acquired *Enterococcus faecium* isolate. *Microbiology* **154**, 3212–3223.

Hidron, A. I., Edwards, J. R., Patel, J., Horan, T. C., Sievert, D. M., Pollock, D. A. & Fridkin, S. K. (2008). NHSN annual update: antimicrobial-resistant pathogens associated with healthcare-associated infections: annual summary of data reported to the National Healthcare Safety Network at the Centers for Disease Control and Prevention, 2006–2007. *Infect Control Hosp Epidemiol* **29**, 996–1011.

Nallapareddy, S. R., Singh, K. V., Sillanpaa, J., Garsin, D. A., Hook, M., Erlandsen, S. L. & Murray, B. E. (2006). Endocarditis and biofilm-associated pili of *Enterococcus faecalis*. *J Clin Invest* **116**, 2799–2807.

Nallapareddy, S. R., Singh, K. V., Okhuysen, P. C. & Murray, B. E. (2008). A functional collagen adhesin gene, *acm*, in clinical isolates of *Enterococcus faecium* correlates with the recent success of this emerging nosocomial pathogen. *Infect Immun* **76**, 4110–4119.

Rice, L. B., Carias, L., Rudin, S., Vael, C., Goossens, H., Konstabel, C., Klare, I., Nallapareddy, S. R., Huang, W. & other authors (2003). A potential virulence gene, hyl_{Efm}, predominates in *Enterococcus faecium* of clinical origin. *J Infect Dis* **187**, 508–512.

Rice, L. B., Lakticova, V., Carias, L. L., Rudin, S., Hutton, R. & Marshall, S. H. (2009). Transferable capacity for gastrointestinal colonization in *Enterococcus faecium* in a mouse model. *J Infect Dis* **199**, 342–349.

Sillanpaa, J., Nallapareddy, S. R., Prakash, V. P., Qin, X., Hook, M., Weinstock, G. M. & Murray, B. E. (2008). Identification and phenotypic characterization of a second collagen adhesin, Scm, and genome-based identification and analysis of 13 other predicted MSCRAMMs, including four distinct pilus loci, in *Enterococcus faecium*. *Microbiology* **154**, 3199–3211.

Sillanpaa, J., Prakash, V. P., Nallapareddy, S. R. & Murray, B. E. (2009). Distribution of genes encoding MSCRAMMs and pili in clinical and natural populations of *Enterococcus faecium*. *J Clin Microbiol* **47**, 896–901.

Singh, K. V., Nallapareddy, S. R. & Murray, B. E. (2007). Importance of the *ebp* (endocarditis- and biofilm-associated pilus) locus in the pathogenesis of *Enterococcus faecalis* ascending urinary tract infection. *J Infect Dis* **195**, 1671–1677.

Top, J., Willems, R. & Bonten, M. (2008). Emergence of CC17 *Enterococcus faecium*: from commensal to hospital-adapted pathogen. *FEMS Immunol Med Microbiol* **52**, 297–308.

Willems, R. J., Homan, W., Top, J., van Santen-Verheuvel, M., Tribe, D., Manzioros, X., Gaillard, C., Vandenbroucke-Grauls, C. M., Mascini, E. M. & other authors (2001). Variant *esp* gene as a marker of a distinct genetic lineage of vancomycin-resistant *Enterococcus faecium* spreading in hospitals. *Lancet* **357**, 853–855.