Functional genomics of *Enterococcus faecium*: antibiotic resistance and niche adaptation

Xinglin Zhang

PhD Thesis, University Utrecht, the Netherlands
ISBN: 9789461084071

Cover design: Xiao Dong and Xinglin Zhang
DNA structure © Yunxiang987 | Stock Free Images & Dreamstime Stock Photos

Print: Gildeprint drukkerijen, Enschede, The Netherlands

The studies in this thesis were financially supported by the European Union Sixth Framework Programme "Approaches to Control multi-resistant Enterococci (ACE): Studies on molecular ecology, horizontal gene transfer, fitness, and prevention" under contract LSHE-CT-2007-037410, and by the European Union Seventh Framework Programme (FP7-HEALTH-2011-single-stage) "Evolution and Transfer of Antibiotic Resistance" (EvoTAR) under grant agreement number 282004.

Printing of this thesis was financially supported by the Netherlands Society of Medical Microbiology (NVMM) and the Royal Netherlands Society for Microbiology (KNVM), the Infection & Immunity Center Utrecht, Gilead Sciences and BGI.

Functional genomics of *Enterococcus faecium*: antibiotic resistance and niche adaptation

Functionele genomica van *Enterococcus faecium*:

antibioticum resistentie en niche-adaptatie

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit van Utrecht op gezag van de rector magnificus, prof.dr. G.J. van der Zwaan, ingevolge het besluit van het college voor promoties in het openbaar te verdedigen op vrijdag 12 april 2013 des middags te 4.15 uur

door

Xinglin Zhang

geboren op 26 maart 1982 te Linyi, China

Promotor:	Prof. Dr. M.J.M. Bonten
Co-promotors:	Dr. Ir. W. van Schaik Dr. R.J.L. Willems

献给我的家人

Commissie:	Prof. Dr. J.P.M. van Putten
	Prof. Dr. C.M.J. Pieterse
	Prof. Dr. J. Kok
	Prof. Dr. M. Kleerebezem
	Prof. Dr. J. Hübner

Paranimfen:	Elena Bülow
	Fernanda Paganelli

CONTENTS

Chapter 1

General introduction

Enterococci are Gram-positive facultative anaerobic cocci that belong to the lactic acid bacteria in the phylum Firmicutes. They are generally considered to be inhabitants of the gastrointestinal tracts of humans and other warm-blooded animals as members of the normal intestinal microbiota, but are also found in reptiles, insects, plants, soil, water and food [1, 2]. The first description of *Enterococcus* was provided in 1899 by Thiercelin [3], who described a Gram-positive bacterium of intestinal origin, with a spherical shape and occurring as diplococci. This bacterium was called "*Enterococcus*" from the Greek word '*enteron*' or intestine. Enterococci which were isolated from an intestinal origin had similar characteristics to streptococci and were initially classified as a group of the *Streptococcus* genus by Sherman in 1937 [4]. However, since the 1980s, with the development of molecular and immunological techniques, enterococci were shown to be different from streptococci in many aspects. Accordingly, *Enterococcus* was classified as a separate genus in 1984 [5]. Recent studies based on 16S rRNA gene sequence and whole genome sequence analysis further demonstrated that enterococci clearly differ from streptococci (Figure 1). Currently, the *Enterococcus* genus consists of over 40 species (www.bacterio.cict.fr/e/enterococcus.html), including *E. faecalis* and *E. faecium* which have emerged as important nosocomial pathogens around the world [6, 7].

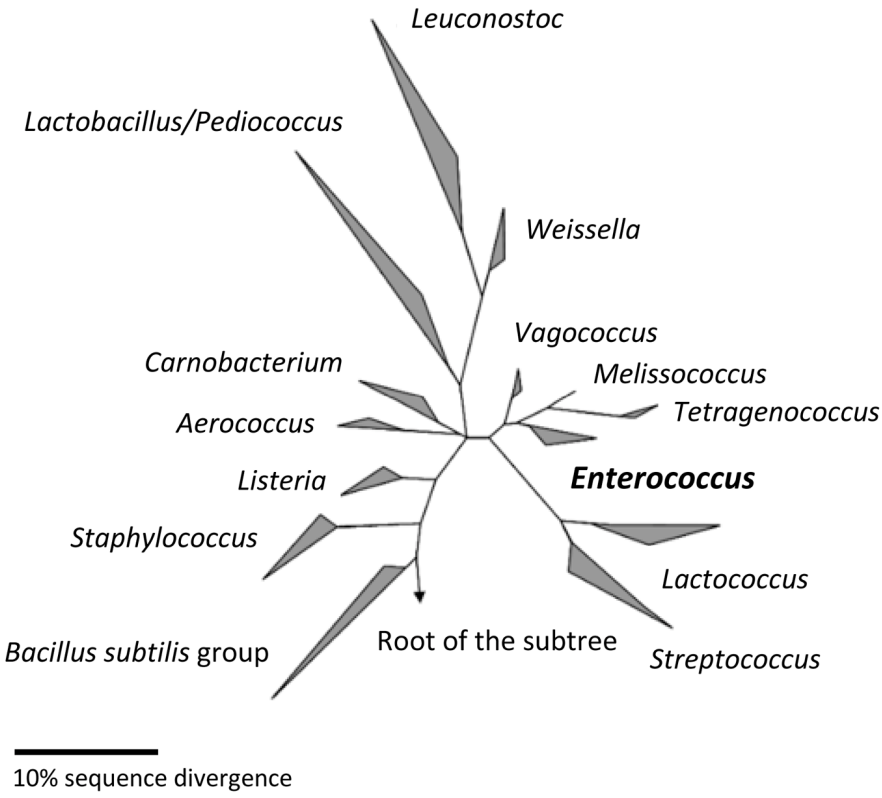


Figure 1. Phylogenetic tree with the position of the genus *Enterococcus* in relation to *Streptococcus* and related Gram-positive genera. This figure is based on comparative sequence analysis of 16S rRNA and is adapted from reference [8].

Enterococci are facultative anaerobes that produce energy primarily through the homolactic fermentation pathway [9]. *E. faecalis* is also capable of aerobic respiration when grown in the presence of heme due to the expression of a cytochrome *bd*-type respiratory oxidase which has not been found in other *Enterococcus* species [10, 11]. Though they are non-spore-forming bacteria, enterococci can survive or grow at a wide ranging of conditions including 10 to 45°C, in 28% NaCl broth and at pH 4.8 to pH 9.6. In addition, enterococci are resistant to bile salts, azide, detergents, sodium hypochlorite, heavy metals, ethanol and prolonged exposure to desiccation [4, 12, 13]. Enterococci have flexible metabolic capabilities that enable them to utilize a broad spectrum of energy sources including diverse carbohydrates, malate, citrate, the diamino acids arginine and agmatine, and many α-keto acids [12]. These characteristics contribute to enterococcal fitness and competitiveness in diverse environments, especially in the intestinal tract of humans and animals [12].

The emergence of enterococci as nosocomial pathogens

The first description of an enterococcal infection dates back to 1899 and is part of a case report of a patient suffering from acute endocarditis [14, 15]. Generally enterococci display low levels of virulence and up to relatively recent times rarely caused human infections. Consequently, enterococci were for many years considered to be harmless gut commensals [14]. Because they frequently produce bacteriocins, *Enterococcus* species even have been used in the last decades by the food industry as probiotics or as starter cultures [16]. However, since the first identification of vancomycin-resistant enterococci (VRE) in 1986 in the United Kingdom and France [17], the number of enterococcal infections has gradually risen throughout the world. Currently, enterococci are the third leading agent of healthcare-associated infections in the USA [6, 7] and can cause bacteraemia, endocarditis, urinary tract and surgical site infections. The majority of enterococcal infections are caused by *E. faecalis* and *E. faecium*. Up to the late 1980s, *E. faecalis* was responsible for the large majority of enterococcal infections, but starting from the 1990s nosocomial infections with *E. faecium* became more frequent [7, 18, 19]. Currently *E. faecium* is almost as common as *E. faecalis* as a causative agent of infections that are acquired during hospital stay [7, 18, 19]. The rapid accumulation of antibiotic resistance genes, including those for clinically important antibiotics such as ampicillin and vancomycin, in clinical *E. faecium* isolates has been proposed as the main reason for this change in epidemiology [7]. For example, in the USA, the percentage of vancomycin-resistant *E. faecium* isolates rose from 0 % before the mid-1980s to more than 80 % by 2007 [20]; while only ~5 % of *E. faecalis* isolates are vancomycin resistant [6]. In addition, clinical *E. faecium* isolates have also acquired dozens of other genes that appear to contribute to fitness in the hospital environment [21]. Consequently *E. faecium* is currently recognized as a nosocomial pathogen of major importance [14].

Molecular epidemiology of *E. faecium*

A large number of molecular epidemiological analyses on *E. faecium* have been performed using Multi-Locus Sequence Typing (MLST), which is based on the detection of allelic differences in seven housekeeping genes [7, 22]. These studies, using *E. faecium* isolates originating from healthy, non-hospitalized humans, hospitalized patients, and non-human sources have revealed the existence of a distinct genetic subpopulation, which was designated clonal complex 17 (CC17) and which is responsible for the majority of nosocomial infections and hospital outbreaks [22-24]. These analyses on the *E. faecium* MLST dataset were initially performed by using the eBURST algorithm [25] . However, eBURST may not be the best tool for the study of *E. faecium* population structure [7, 23, 26] and a recent study [27] showed that the use of the algorithm Bayesian Analysis of Population Structure (BAPS) [28-30] is a more appropriate method for the analysis of the genetic population structure of *E. faecium*. Unlike eBURST [25], BAPS does not cluster isolates by using a phylogenetic model but rather uses a statistical genetic model that takes into account both clonal ancestry and recombination patterns as identified from DNA sequence data. This analysis grouped *E. faecium* isolates into 13 BAPS (sub)groups, with the large majority (80%) of nosocomial isolates clustering in two subgroups (2-1 and 3-3), and revealing the existence of three evolutionarily separate hospital lineages originating from MLST sequence types 17, 18, and 78 [27].

A comparative genomic analysis [21], using mixed whole-genome microarray hybridizations, showed that hospital-acquired *E. faecium* isolates appear to be genetically distinct from indigenous intestinal isolates and exclusively possess more than 100 genes including genes involved in carbohydrate metabolism, insertion sequence elements and putative virulence elements including a pathogenicity island containing the enterococcal surface protein gene *esp* [31, 32]. Recent comparative genomic studies of *E. faecium*, based on 8 to 22 whole genome sequences, identified a deep phylogenetic split between two *E. faecium* clades that were designated clade A and clade B [33] or commensal (CA) and hospital (HA) clades [34, 35]. This confirmed that most clinical, outbreak, and hospital-associated strains (including sequence types 16, 17, 18, and 78), are evolutionally more closely related to each other by phylogenetic and gene content similarity analyses compared to strains isolated from animals or healthy humans.

Antibiotic resistance in enterococci

Enterococci, particularly vancomycin-resistant *Enterococcus* (VRE), are currently considered to be among the most important antibiotic-resistant nosocomial pathogens [36-38].

Enterococci are intrinsically resistant to a broad range of antibiotics and have also gained resistance to many clinically important antibiotics, through the acquisition of new genetic determinants or through sporadic mutations in genes of the *Enterococcus* core genome [14, 39]. This is limiting the choice of therapeutic agents to treat enterococcal infections. Moreover, their resistance to commonly used antibiotics allows enterococci to adapt to and be selectively enriched in the hospital environment and in hospitalized patients, in which they are exposed to multiple antibiotics, including recently introduced drugs (e.g. linezolid and daptomycin), which may promote the evolution of novel antibiotic resistance mechanisms (Table 1).

Table 1. Mechanisms of antibiotic resistance to *E. faecium* and *E. faecalis*[a]

Antibiotic	Mechanism of resistance	Associated enzyme	Phenotype	Type and/or genetic element	Host[b]
Aminoglycosides					
	Low cell wall permeability	-	Low-level aminoglycoside resistance, preserved synergy with cell-wall-active antibiotics	Intrinsic	*Efs*
	Ribosome mutation	-	High-level aminoglycoside resistance	Sporadic	*Efs Efm*
	Aminoglycoside-modifying enzyme (AME)	Aac(6′)-Ii	Low-level tobramycin and kanamycin resistance	Intrinsic	*Efm*
	AME	Aph(3′)-IIIa	Low-level kanamycin resistance	pJH1	*Efm*
	AME	Ant(4′′)-Ia	Low-level resistance to kanamycin, tobramycin, amikacin and neomycin	pIP810	*Efm*
	AME	Aph(2′′)-Ia-Aac(6′)Ie	High-level gentamicin resistance	Tn*5281*	*Efs Efm*
	AME	Aph(2′′)-Ib	High-level gentamicin resistance	Unknown	*Efm*
	AME	Aph(2′′)-Ic	High-level gentamicin resistance	pYN134	*Efs Efm*
	AME	Aph(2′′)-Id	High-level gentamicin resistance	Unknown	*Efm*
	AME	Ant(6′)-Ia	High-level streptomycin resistance	Tn*1546*, Inc.18, Tn*5382*	*Efs Efm*
	AME	Ant(3′′)-Ia	High-level streptomycin resistance	pR538–1	*Efm*
	Ribosome-modifying methyltransferase	EfmM	Tobramycin and kanamycin resistance	Intrinsic	*Efm*
β-lactams and cephalosporins					
	PBP4/5 production	-	Low-level penicillin resistance; moderate to high-level cephalosporin resistance	Intrinsic	*Efs Efm*
	PBP4/5 point mutation	-	High-level ampicillin and imipenem resistance	Sporadic	*Efs Efm*
	Altered cell wall	L,D-transpeptidase	β-lactam resistance	Intrinsic	*Efm*
Glycopeptides					
	Synthesis of alternative cell wall	VanA-E, VanG, VanL-N	Resistance to vancomycin +/− teicoplanin depending on the genotype	Transposons, plasmids	*Efs Efm*
Lincosamides					

Antibiotic	Mechanism of resistance	Associated enzyme	Phenotype	Type and/or genetic element	Host[b]
	ABC-efflux pump	Lsa	Resistance to clindamycin, streptogrammin A and B	Intrinsic	*Efs*
	ABC-efflux pump	MsrC	Low-level resistance to streptogramin B compounds	Intrinsic	*Efm*
	ABC-efflux pump	VgaD	Streptogramin A resistance	Putative transposon	*Efm*
	Acetyltransferase	VatD (SatA)	Streptogramin A resistance	Putative transposon	*Efm*
	Acetyltransferase	VatH	Streptogramin A resistance	Putative transposon	*Efm*
	Acetyltransferase	*VgbA*	Streptogramin B resistance	Unknown	*Efm*
	Acetyltransferase	VatE (SatG)	Streptogramin A resistance	Unknown	*Efm*
	Altered ribosome	ErmA	MLS_A phenotype	Tn*554*	*Efs Efm*
	Altered ribosome	ErmB	MLS_B phenotype	Tn*917*, Tn*1545*	*Efs Efm*
Linezolid					
	rRNA point mutations	G2576T, G2505A, L4(F101L)	Linezolid resistance	Sporadic	*Efs Efm*
	Methylated rRNA	Cfr	Linezolid, lincosamides, streptogramin A resistance	pEF-01	*Efs Efm*
Daptomycin					
	Altered membrane-bound protein	Cardiolipin sythetase	Contributes to daptomycin resistance through an unknown mechanism	Sporadic	*Efs Efm*
	Altered membrane-bound protein	GdpD	Daptomycin resistance, effect is amplified in combination *liaF* mutation	Sporadic	*Efs Efm*
	Altered membrane-bound protein	LiaF	Daptomycin resistance when combined with *gdpD* mutation	Sporadic	*Efs Efm*
(Fluoro)quinolones					
	Altered DNA gyrase	GyrA	(Fluoro)quinolone resistance	Sporadic	*Efs Efm*
	Altered topoisomerase IV	ParC	(Fluoro)quinolone resistance	Sporadic	*Efs Efm*
	Binding to gyrase and/or topoisomerase IV	Qnr-like protein	(Fluoro)quinolone resistance	Intrinsic	*Efs*

[a] This table is adapted from reference [39]. The mechanisms of quinolone resistance are from references [40-44].

[b] *Efs*: *E. faecalis*; *Efm*: *E. faecium*.

Ampicillin resistance

The intrinsic resistance to β-lactam antibiotics of enterococci was already reported 60 years ago, soon after the introduction of penicillin in the early 1940s, when enterococci were found to be considerably more resistant to β-lactams than streptococci [45]. Nowadays, more than 90% of *E. faecium* strains (and only about 5% of *E. faecalis* isolates) recovered from healthcare associated infections in the U.S. are resistant to ampicillin [6]. β-lactam antibiotics are structural analogs of pentapeptide precursors and target the penicillin binding proteins (PBPs), which are enzymes responsible for the cross-linking of peptidoglycan. Inactivation of PBPs results in impaired cell wall synthesis and ultimately cell death [46]. Mutations in the high-molecular weight class B penicillin binding protein PBP5 have been

considered to be the main mechanism for the resistance to β-lactams in *E. faecium*. Upregulated expression of *pbp5* and/or mutations in the 3′ end of the gene lead to a further reduced susceptibility to ampicillin [47-49]. Mainardi *et al.* [50, 51] identified another mechanism leading to ampicillin resistance as a result of laboratory evolution of an *E. faecium* strain by selection on agar media containing increasing concentrations of ampicillin. In this resistant mutant, the D,D-transpeptidase activity of the PBPs could be bypassed by a β-lactam resistant L,D-transpeptidase (Ldt_{fm}) that catalyses the formation of 3→3 cross-links between peptidoglycan side chains instead of the classical 4→3 cross-links. It is as yet unclear whether this resistance mechanism also contributes to ampicillin resistance in natural *E. faecium* isolates.

Vancomycin resistance

For many years, the glycopeptide antibiotic vancomycin has been considered a "last resort" antibiotic that was set aside for the treatment of serious and multi-drug-resistant infections caused by Gram-positive bacteria. However the effectiveness of vancomycin has diminished since the emergence of VRE as a nosocomial pathogen in the 1980s [17]. Like β-lactam antibiotics, vancomycin also acts on peptidoglycan cross-linking, but does so via a different mechanisms. Vancomycin binds to the D-Ala-D-Ala terminus of the pentapeptide precursor, thereby inhibiting cell wall synthesis [52]. Nine genotypes (gene clusters) of vancomycin resistance has been described in enterococci: *vanA, vanB, vanC, vanD, vanE, vanG, vanL, vanM* and *vanN* [53-56]. Of these genotypes, the *vanA* and *vanB* types are by far the most frequently reported [57-59]. *vanA* and *vanB* gene clusters mediate vancomycin resistance (MIC 4 to 1000 μg/ml) in enterococci due to the modification of the pentapeptide precursors, in which the terminal D-Ala can be replaced with D-Lac, resulting in an up to 1,000-fold lower binding affinity to vancomycin [60]. The *vanA* gene cluster is usually acquired through the Tn*1546* transposon or a related transposon of the Tn*3* family, which are commonly found on plasmids but can also be integrated in the chromosome [61, 62]. The *vanB* gene cluster is usually present on the chromosome but may also be carried on plasmids [63]. The transfer of *vanB* gene cluster is associated with the Tn*1547* [64, 65] and Tn*5382* [66] transposons.

Intestinal colonization

As a successful colonizer in the gut of both healthy humans and hospitalized patients, *E. faecium* must have evolved mechanisms to adapt, survive and colonize in the mammalian gut. For instance, *E. faecium* must be tolerant to the low pH and the antimicrobial effects of bile to which it is exposed in the upper digestive tract, and must be able to utilize the

carbohydrates that are available in the gut in competition with the host and other gut microorganisms.

The antibiotic-mediated perturbation of the intestinal microbiota and its adverse effect on human health are increasingly recognized [67-69]. Antibiotic treatment alters intestinal microbial diversity by eliminating susceptible organisms, which provides an opportunity for multi-resistant enterococci to colonize these niches [70, 71]. Ubeda *et al.* [71] demonstrated that during antibiotic treatment the intestinal microbiota of patients displayed a striking shift, in which the genus *Enterococcus* shifted from a minor component of intestinal microbiota into a dominantly occurring (>97%) group of bacteria. These patients subsequently developed VRE bacteremia caused by the strains that originated from the patient's gut [71]. The antibiotic-mediated perturbation of the intestinal microbiota and its link to *E. faecium* infections could be replicated in a mouse model [71]. The VRE that dominantly colonize the intestinal tract can also serve as reservoir for fecal contamination of the urinary tract and the skin, ultimately leading to catheter-related infections [14].

Using a mouse model of VRE colonization, Rice *et al.* showed that some broad-spectrum antimicrobial agents (cefotetan and ceftriaxone) promote high-level enterococcal colonization while others (aztreonam, cefazolin, cefepime, and piperacillin-tazobactam) do not [72]. Donskey *et al.* showed both in human and mouse model that antimicrobial agents with activity against the anaerobic gut-flora promote high-level enterococcal fecal colonization [73, 74]. Antibiotic therapy also indirectly enables colonization by multi-resistant enterococci, because antibiotic-mediated depletion of the Gram-negative microbiota affects the host intestinal immune defenses, particularly through the reduced expression of RegIIIγ, a secreted C-type lectin that is bactericidal for Gram-positive bacteria and which is produced upon immune stimulation by the microbiota through bacterial products such as LPS and flagellin [75, 76].

Several attempts have been made in the past to identify the genetic determinants of nosocomial *E. faecium* strains that promote intestinal colonization during antibiotic therapy. One obvious candidate is the enterococcal surface protein Esp [77, 78] (described below), since the gene encoding this protein is significantly enriched among nosocomial *E. faecium* strains [31, 32]. However, Heikens *et al.* found no difference in the gut colonization capacity between the clinical *E. faecium* strain E1162 and its isogenic *esp* mutant in a mouse colonization model [79]. A study by Rice *et al.* provided experimental evidence that a large conjugative plasmid carrying the hyl_{Efm} gene might be important for gastrointestinal colonization in a mouse model [80]. The study demonstrated that the transconjugants which obtained the large plasmid (pLRM23) also gained an enhanced ability to colonize the gut. However it remains unresolved which genetic determinants are responsible for providing this colonization benefit, since pLRM23 is a megaplasmid (approximately 220 kb in size) and

two additional plasmids (approximately 40 kb and 15 kb in size) also co-transferred to the transconjugants along with pLRM23 [80]. Therefore, the genetic determinants of nosocomial *E. faecium* strains that contribute intestinal during antibiotic therapy remain to be determined.

Virulence

E. faecium is essentially avirulent in healthy individuals and infections with *E. faecium* are typically seen in critically ill and immunocompromised patients. Only four determinants of *E. faecium* have been experimentally verified to contribute to virulence in animal models. These are the large enterococcal surface protein Esp [77, 78] and the pilus subunit Ebp_{fm} [81], which both contribute to biofilm formation and urinary tract infection; the MSCRAMM (Microbial Surface Components Recognizing Adhesive Matrix Molecules) Acm, which is an adhesin of collagen and contributes to endocarditis [82, 83]; and the Gls-like proteins Gls20 and Gls33 that play a role in bile salt resistance and peritonitis [84].

Outline of this thesis

Since their identification in the late 19th century, enterococci have been studied for more than a century, during which, particularly in the past two decades, the commensal bacterium *E. faecium* has emerged globally as an important nosocomial pathogen. Nowadays, microbiological research is rapidly entering the post-genomic era. With the development of nucleotide sequencing technologies, novel bacterial genome sequences are being generated almost routinely, which provides a vast reservoir of genomic information that is revolutionizing our understanding of bacterial life.

However genome sequences often raise more questions than they answer. Functional genomics strives to use the data produced by genome sequencing to experimentally investigate, identify, and verify gene functions and interactions. In *E. faecium*, such studies have long been hampered by the lack of appropriate genetic tools. The aim of the research described in this thesis was to develop and use novel functional genomic approaches to identify and characterize genetic determinants that contribute to *E. faecium* niche adaptation, with a specific focus on antibiotic resistance and intestinal colonization. In **chapter 2** a set of molecular tools were developed for functional genomic studies of *E. faecium*, which enabled us to perform both high-throughput genome-wide analysis and specific gene-targeted investigations in this organism. This led to the identification of *E. faecium* genes that contribute to ampicillin resistance. In **Chapter 3**, we describe a genetic

determinant that was identified by comparative genome analysis as being specifically enriched in clinical *E. faecium* isolates and which contributes to intestinal colonization in clinical *E. faecium* strains. In **Chapter 4** the genetic determinants of bile salt resistance and the transcriptional response to bile salt stress is studied in *E. faecium*. **Chapter 5** describes a genetic element which is carried on transferable megaplasmids and which allows *E. faecium* to utilize the non-digestible sugar raffinose. This genetic element was more prevalent in isolates from community sources than those from clinical infections. **Chapter 6** describes the identification and transcriptional analysis of genetic determinants involved in maltodextrin utilization of *E. faecium*. Finally, the findings and future perspectives of the research described in this thesis are discussed in **Chapter 7.**

References

1. Devriese LA, Collins MD, Wirth R: The genus *Enterococcus.* In The Prokaryotes: A Handbook on the Biology of Bacteria: Ecophysiology, Isolation, Identification, Applications, pp. 1465–1481. Edited by A. Balows, H. G. Trüper, M. Dworkin, W. Harder & K. H. Schleifer. New York: Springer. 1992.
2. Byappanahalli MN, Nevers MB, Korajkic A, Staley ZR, Harwood VJ: Enterococci in the environment. *Microbiol Mol Biol Rev* 2012, 76(4):685-706.
3. Thiercelin ME: Sur un diplocoque saprophyte de l'intestin susceptible de devenir pathogène. *C R Séances Soc Biol* 1899, 50:269-271.
4. Sherman JM: The Streptococci. *Bacteriol Rev* 1937, 1(1):3-97.
5. Schleifer KH, Kilpper-Bälz R: Transfer of *Streptococcus faecalis* and *Streptococcus faecium* to the Genus *Enterococcus* nom. rev. as *Enterococcus faecalis* comb. nov. and *Enterococcus faecium* comb. nov. *Int J Syst Bacteriol* 1984, 34:31-34.
6. Hidron AI, Edwards JR, Patel J, Horan TC, Sievert DM, Pollock DA, Fridkin SK: 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* 2008, 29(11):996-1011.
7. Willems RJ, van Schaik W: Transition of *Enterococcus faecium* from commensal organism to nosocomial pathogen. *Future Microbiol* 2009, 4(9):1125-1135.
8. Klein G: Taxonomy, ecology and antibiotic resistance of enterococci from food and the gastro-intestinal tract. *Int J Food Microbiol* 2003, 88(2-3):123-131.
9. Devriese, L. A., M. D. Collins, and R. Wirth. 1992. The genus *Enterococcus*, p. 1465–1481. *In* The prokaryotes (2nd ed.). A. Balows, H. G. Truper, M. Dworkin, W. Harder, and K.-H. Schleifer (ed.). Springer-Verlag, New York, N.Y.
10. Ritchey TW, Seely HW, Jr.: Distribution of cytochrome-like respiration in streptococci. *J Gen Microbiol* 1976, 93(2):195-203.
11. Winstedt L, Frankenberg L, Hederstedt L, von Wachenfeldt C: *Enterococcus faecalis* V583 contains a cytochrome *bd*-type respiratory oxidase. *J Bacteriol* 2000, 182(13):3863-3866.
12. Huycke M: 2002. Physiology of enterococci, p 133-176. *In*: The Enterococci: Pathogenesis, Molecular Biology, and Antibiotic Resistance. Gilmore MS, Clewell, DB, Courvalin P, Dunny GM, Murray BE, Rice LB (ed), ASM Press, Washington, DC.
13. Facklam RR, Collins MD: Identification of *Enterococcus species* isolated from human infections by a conventional test scheme. *J Clin Microbiol* 1989, 27(4):731-734.
14. Arias CA, Murray BE: The rise of the *Enterococcus*: beyond vancomycin resistance. *Nat Rev Microbiol* 2012, 10(4):266-278.
15. Maccallum WG, Hastings TW: A Case Of Acute Endocarditis Caused By *Micrococcus Zymogenes* (Nov. Spec.), With A Description Of The Microorganism. *J Exp Med* 1899, 4(5-6):521-534.
16. Foulquie Moreno MR, Sarantinopoulos P, Tsakalidou E, De Vuyst L: The role and application of enterococci in food and health. *Int J Food Microbiol* 2006, 106(1):1-24.
17. Leclercq R, Derlot E, Duval J, Courvalin P: Plasmid-mediated resistance to vancomycin and teicoplanin in *Enterococcus faecium*. *N Engl J Med* 1988, 319(3):157-161.
18. Ammerlaan HS, Troelstra A, Kruitwagen CL, Kluytmans JA, Bonten MJ: Quantifying changes in incidences of nosocomial bacteraemia caused by antibiotic-susceptible and antibiotic-resistant pathogens. *J Antimicrob Chemother* 2009, 63(5):1064-1070.
19. Top J, Willems R, Blok H, de Regt M, Jalink K, Troelstra A, Goorhuis B, Bonten M: Ecological replacement of *Enterococcus faecalis* by multiresistant clonal complex 17 *Enterococcus faecium*. *Clin Microbiol Infect* 2007, 13(3):316-319.

20. Arias CA, Murray BE: Emergence and management of drug-resistant enterococcal infections. *Expert Rev Anti Infect Ther* 2008, 6(5):637-655.
21. Leavis HL, Willems RJ, van Wamel WJ, Schuren FH, Caspers MP, Bonten MJ: Insertion sequence-driven diversification creates a globally dispersed emerging multiresistant subspecies of *E. faecium*. *PLoS Pathog* 2007, 3(1):e7.
22. Homan WL, Tribe D, Poznanski S, Li M, Hogg G, Spalburg E, Van Embden JD, Willems RJ: Multilocus sequence typing scheme for *Enterococcus faecium*. *J Clin Microbiol* 2002, 40(6):1963-1971.
23. Willems RJ, Hanage WP, Bessen DE, Feil EJ: Population biology of Gram-positive pathogens: high-risk clones for dissemination of antibiotic resistance. *FEMS Microbiol Rev* 2011, 35(5):872-900.
24. Willems RJ, Top J, van Santen M, Robinson DA, Coque TM, Baquero F, Grundmann H, Bonten MJ: Global spread of vancomycin-resistant *Enterococcus faecium* from distinct nosocomial genetic complex. *Emerg Infect Dis* 2005, 11(6):821-828.
25. Feil EJ, Li BC, Aanensen DM, Hanage WP, Spratt BG: eBURST: inferring patterns of evolutionary descent among clusters of related bacterial genotypes from multilocus sequence typing data. *J Bacteriol* 2004, 186(5):1518-1530.
26. Turner KM, Hanage WP, Fraser C, Connor TR, Spratt BG: Assessing the reliability of eBURST using simulated populations with known ancestry. *BMC Microbiol* 2007, 7:30.
27. Willems RJ, Top J, van Schaik W, Leavis H, Bonten M, Siren J, Hanage WP, Corander J: Restricted gene flow among hospital subpopulations of *Enterococcus faecium*. *MBio* 2012, 3(4):e00151-00112.
28. Corander J, Marttinen P, Siren J, Tang J: Enhanced Bayesian modelling in BAPS software for learning genetic structures of populations. *BMC Bioinformatics* 2008, 9:539.
29. Corander J, Tang J: Bayesian analysis of population structure based on linked molecular information. *Math Biosci* 2007, 205(1):19-31.
30. Tang J, Hanage WP, Fraser C, Corander J: Identifying currents in the gene pool for bacterial populations using an integrative approach. *PLoS Comput Biol* 2009, 5(8):e1000455.
31. Leavis H, Top J, Shankar N, Borgen K, Bonten M, van Embden J, Willems RJ: A novel putative enterococcal pathogenicity island linked to the *esp* virulence gene of *Enterococcus faecium* and associated with epidemicity. *J Bacteriol* 2004, 186(3):672-682.
32. Willems RJ, Homan W, Top J, van Santen-Verheuvel M, Tribe D, Manzioros X, Gaillard C, Vandenbroucke-Grauls CM, Mascini EM, van Kregten E *et al*: Variant esp gene as a marker of a distinct genetic lineage of vancomycin-resistant *Enterococcus faecium* spreading in hospitals. *Lancet* 2001, 357(9259):853-855.
33. Palmer KL, Godfrey P, Griggs A, Kos VN, Zucker J, Desjardins C, Cerqueira G, Gevers D, Walker S, Wortman J *et al*: Comparative genomics of enterococci: variation in *Enterococcus faecalis*, clade structure in *E. faecium*, and defining characteristics of *E. gallinarum* and *E. casseliflavus*. *MBio* 2012, 3(1):e00318-00311.
34. Galloway-Pena J, Roh JH, Latorre M, Qin X, Murray BE: Genomic and SNP analyses demonstrate a distant separation of the hospital and community-associated clades of *Enterococcus faecium*. *PLoS One* 2012, 7(1):e30187.
35. Qin X, Galloway-Pena JR, Sillanpaa J, Roh JH, Nallapareddy SR, Chowdhury S, Bourgogne A, Choudhury T, Muzny DM, Buhay CJ *et al*: Complete genome sequence of *Enterococcus faecium* strain TX16 and comparative genomic analysis of *Enterococcus faecium* genomes. *BMC Microbiol* 2012, 12:135.
36. Rice LB: Progress and challenges in implementing the research on ESKAPE pathogens. *Infect Control Hosp Epidemiol* 2010, 31 Suppl 1:S7-10.
37. Cattoir V, Leclercq R: Twenty-five years of shared life with vancomycin-resistant enterococci: is it time to divorce? *J Antimicrob Chemother* 2012.
38. Arias CA, Murray BE: Antibiotic-resistant bugs in the 21st century--a clinical super-challenge. *N Engl J Med* 2009, 360(5):439-443.

39. Hollenbeck BL, Rice LB: Intrinsic and acquired resistance mechanisms in *enterococcus*. *Virulence* 2012, 3(5):421-433.
40. Kanematsu E, Deguchi T, Yasuda M, Kawamura T, Nishino Y, Kawada Y: Alterations in the GyrA subunit of DNA gyrase and the ParC subunit of DNA topoisomerase IV associated with quinolone resistance in *Enterococcus faecalis*. *Antimicrob Agents Chemother* 1998, 42(2):433-435.
41. Onodera Y, Okuda J, Tanaka M, Sato K: Inhibitory activities of quinolones against DNA gyrase and topoisomerase IV of *Enterococcus faecalis*. *Antimicrob Agents Chemother* 2002, 46(6):1800-1804.
42. Leavis HL, Willems RJ, Top J, Bonten MJ: High-level ciprofloxacin resistance from point mutations in *gyrA* and *parC* confined to global hospital-adapted clonal lineage CC17 of *Enterococcus faecium*. *J Clin Microbiol* 2006, 44(3):1059-1064.
43. Arsene S, Leclercq R: Role of a *qnr*-like gene in the intrinsic resistance of *Enterococcus faecalis* to fluoroquinolones. *Antimicrob Agents Chemother* 2007, 51(9):3254-3258.
44. Hegde SS, Vetting MW, Mitchenall LA, Maxwell A, Blanchard JS: Structural and biochemical analysis of the pentapeptide repeat protein EfsQnr, a potent DNA gyrase inhibitor. *Antimicrob Agents Chemother* 2011, 55(1):110-117.
45. Murray BE: The life and times of the *Enterococcus*. *Clin Microbiol Rev* 1990, 3(1):46-65.
46. Zapun A, Contreras-Martel C, Vernet T: Penicillin-binding proteins and beta-lactam resistance. *FEMS Microbiol Rev* 2008, 32(2):361-385.
47. Fontana R, Aldegheri M, Ligozzi M, Lopez H, Sucari A, Satta G: Overproduction of a low-affinity penicillin-binding protein and high-level ampicillin resistance in *Enterococcus faecium*. *Antimicrob Agents Chemother* 1994, 38(9):1980-1983.
48. Fontana R, Grossato A, Rossi L, Cheng YR, Satta G: Transition from resistance to hypersusceptibility to beta-lactam antibiotics associated with loss of a low-affinity penicillin-binding protein in a *Streptococcus faecium* mutant highly resistant to penicillin. *Antimicrob Agents Chemother* 1985, 28(5):678-683.
49. Williamson R, le Bouguenec C, Gutmann L, Horaud T: One or two low affinity penicillin-binding proteins may be responsible for the range of susceptibility of *Enterococcus faecium* to benzylpenicillin. *J Gen Microbiol* 1985, 131(8):1933-1940.
50. Mainardi JL, Morel V, Fourgeaud M, Cremniter J, Blanot D, Legrand R, Frehel C, Arthur M, Van Heijenoort J, Gutmann L: Balance between two transpeptidation mechanisms determines the expression of beta-lactam resistance in *Enterococcus faecium*. *J Biol Chem* 2002, 277(39):35801-35807.
51. Mainardi JL, Legrand R, Arthur M, Schoot B, van Heijenoort J, Gutmann L: Novel mechanism of beta-lactam resistance due to bypass of DD-transpeptidation in *Enterococcus faecium*. *J Biol Chem* 2000, 275(22):16490-16496.
52. Reynolds PE: Structure, biochemistry and mechanism of action of glycopeptide antibiotics. *Eur J Clin Microbiol Infect Dis* 1989, 8(11):943-950.
53. Courvalin P: Vancomycin resistance in gram-positive cocci. *Clin Infect Dis* 2006, 42 Suppl 1:S25-34.
54. Werner G, Coque TM, Hammerum AM, Hope R, Hryniewicz W, Johnson A, Klare I, Kristinsson KG, Leclercq R, Lester CH *et al*: Emergence and spread of vancomycin resistance among enterococci in Europe. *Euro Surveill* 2008, 13(47).
55. Xu X, Lin D, Yan G, Ye X, Wu S, Guo Y, Zhu D, Hu F, Zhang Y, Wang F *et al*: *vanM*, a new glycopeptide resistance gene cluster found in *Enterococcus faecium*. *Antimicrob Agents Chemother* 2010, 54(11):4643-4647.
56. Lebreton F, Depardieu F, Bourdon N, Fines-Guyon M, Berger P, Camiade S, Leclercq R, Courvalin P, Cattoir V: D-Ala-d-Ser VanN-type transferable vancomycin resistance in *Enterococcus faecium*. *Antimicrob Agents Chemother* 2011, 55(10):4606-4612.

57. Coque TM, Tomayko JF, Ricke SC, Okhyusen PC, Murray BE: Vancomycin-resistant enterococci from nosocomial, community, and animal sources in the United States. *Antimicrob Agents Chemother* 1996, 40(11):2605-2609.
58. Rice LB: Antimicrobial resistance in gram-positive bacteria. *Am J Infect Control* 2006, 34(5 Suppl 1):S11-19; discussion S64-73.
59. Sahm DF, Marsilio MK, Piazza G: Antimicrobial resistance in key bloodstream bacterial isolates: electronic surveillance with the Surveillance Network Database--USA. *Clin Infect Dis* 1999, 29(2):259-263.
60. Bugg TD, Wright GD, Dutka-Malen S, Arthur M, Courvalin P, Walsh CT: Molecular basis for vancomycin resistance in *Enterococcus faecium* BM4147: biosynthesis of a depsipeptide peptidoglycan precursor by vancomycin resistance proteins VanH and VanA. *Biochemistry* 1991, 30(43):10408-10415.
61. Arthur M, Molinas C, Depardieu F, Courvalin P: Characterization of Tn*1546*, a Tn*3*-related transposon conferring glycopeptide resistance by synthesis of depsipeptide peptidoglycan precursors in *Enterococcus faecium* BM4147. *J Bacteriol* 1993, 175(1):117-127.
62. Brisson-Noel A, Dutka-Malen S, Molinas C, Leclercq R, Courvalin P: Cloning and heterospecific expression of the resistance determinant *vanA* encoding high-level resistance to glycopeptides in *Enterococcus faecium* BM4147. *Antimicrob Agents Chemother* 1990, 34(5):924-927.
63. Rice LB, Carias LL, Donskey CL, Rudin SD: Transferable, plasmid-mediated *vanB*-type glycopeptide resistance in *Enterococcus faecium*. *Antimicrob Agents Chemother* 1998, 42(4):963-964.
64. Quintiliani R, Jr., Courvalin P: Characterization of Tn*1547*, a composite transposon flanked by the IS*16* and IS*256*-like elements, that confers vancomycin resistance in *Enterococcus faecalis* BM4281. *Gene* 1996, 172(1):1-8.
65. Quintiliani R, Jr., Courvalin P: Conjugal transfer of the vancomycin resistance determinant *vanB* between enterococci involves the movement of large genetic elements from chromosome to chromosome. *FEMS Microbiol Lett* 1994, 119(3):359-363.
66. Carias LL, Rudin SD, Donskey CJ, Rice LB: Genetic linkage and cotransfer of a novel, *vanB*-containing transposon (Tn*5382*) and a low-affinity penicillin-binding protein 5 gene in a clinical vancomycin-resistant *Enterococcus faecium* isolate. *J Bacteriol* 1998, 180(17):4426-4434.
67. Relman DA: The human microbiome: ecosystem resilience and health. *Nutr Rev* 2012, 70 Suppl 1:S2-9.
68. Willing BP, Russell SL, Finlay BB: Shifting the balance: antibiotic effects on host-microbiota mutualism. *Nat Rev Microbiol* 2011, 9(4):233-243.
69. Littman DR, Pamer EG: Role of the commensal microbiota in normal and pathogenic host immune responses. *Cell Host Microbe* 2011, 10(4):311-323.
70. Gilmore MS, Ferretti JJ: Microbiology. The thin line between gut commensal and pathogen. *Science* 2003, 299(5615):1999-2002.
71. Ubeda C, Taur Y, Jenq RR, Equinda MJ, Son T, Samstein M, Viale A, Socci ND, van den Brink MR, Kamboj M *et al*: Vancomycin-resistant *Enterococcus* domination of intestinal microbiota is enabled by antibiotic treatment in mice and precedes bloodstream invasion in humans. *J Clin Invest* 2010, 120(12):4332-4341.
72. Rice LB, Hutton-Thomas R, Lakticova V, Helfand MS, Donskey CJ: Beta-lactam antibiotics and gastrointestinal colonization with vancomycin-resistant enterococci. *J Infect Dis* 2004, 189(6):1113-1118.
73. Donskey CJ, Hanrahan JA, Hutton RA, Rice LB: Effect of parenteral antibiotic administration on persistence of vancomycin-resistant *Enterococcus faecium* in the mouse gastrointestinal tract. *J Infect Dis* 1999, 180(2):384-390.
74. Donskey CJ, Chowdhry TK, Hecker MT, Hoyen CK, Hanrahan JA, Hujer AM, Hutton-Thomas RA, Whalen CC, Bonomo RA, Rice LB: Effect of antibiotic therapy on the density of

vancomycin-resistant enterococci in the stool of colonized patients. *N Engl J Med* 2000, 343(26):1925-1932.

75. Brandl K, Plitas G, Mihu CN, Ubeda C, Jia T, Fleisher M, Schnabl B, DeMatteo RP, Pamer EG: Vancomycin-resistant enterococci exploit antibiotic-induced innate immune deficits. *Nature* 2008, 455(7214):804-807.
76. Kinnebrew MA, Ubeda C, Zenewicz LA, Smith N, Flavell RA, Pamer EG: Bacterial flagellin stimulates Toll-like receptor 5-dependent defense against vancomycin-resistant *Enterococcus infection*. *J Infect Dis* 2010, 201(4):534-543.
77. Heikens E, Bonten MJ, Willems RJ: Enterococcal surface protein Esp is important for biofilm formation of *Enterococcus faecium* E1162. *J Bacteriol* 2007, 189(22):8233-8240.
78. Leendertse M, Heikens E, Wijnands LM, van Luit-Asbroek M, Teske GJ, Roelofs JJ, Bonten MJ, van der Poll T, Willems RJ: Enterococcal surface protein transiently aggravates *Enterococcus faecium*-induced urinary tract infection in mice. *J Infect Dis* 2009, 200(7):1162-1165.
79. Heikens E, Leendertse M, Wijnands LM, van Luit-Asbroek M, Bonten MJ, van der Poll T, Willems RJ: Enterococcal surface protein Esp is not essential for cell adhesion and intestinal colonization of *Enterococcus faecium* in mice. *BMC Microbiol* 2009, 9:19.
80. Rice LB, Lakticova V, Carias LL, Rudin S, Hutton R, Marshall SH: Transferable capacity for gastrointestinal colonization in *Enterococcus faecium* in a mouse model. *J Infect Dis* 2009, 199(3):342-349.
81. Sillanpaa J, Nallapareddy SR, Singh KV, Prakash VP, Fothergill T, Ton-That H, Murray BE: Characterization of the ebp_{fm} pilus-encoding operon of *Enterococcus faecium* and its role in biofilm formation and virulence in a murine model of urinary tract infection. *Virulence* 2010, 1(4):236-246.
82. Nallapareddy SR, Singh KV, Murray BE: Contribution of the collagen adhesin Acm to pathogenesis of *Enterococcus faecium* in experimental endocarditis. *Infect Immun* 2008, 76(9):4120-4128.
83. Nallapareddy SR, Weinstock GM, Murray BE: Clinical isolates of *Enterococcus faecium* exhibit strain-specific collagen binding mediated by Acm, a new member of the MSCRAMM family. *Mol Microbiol* 2003, 47(6):1733-1747.
84. Choudhury T, Singh KV, Sillanpaa J, Nallapareddy SR, Murray BE: Importance of two *Enterococcus faecium* loci encoding Gls-like proteins for in vitro bile salts stress response and virulence. *J Infect Dis* 2011, 203(8):1147-1154.

Chapter 2

Genome-wide identification of ampicillin resistance determinants in *Enterococcus faecium*

Xinglin Zhang, Fernanda L. Paganelli, Damien Bierschenk, Annemarie Kuipers, Marc J. M. Bonten, Rob J. L. Willems, and Willem van Schaik

Department of Medical Microbiology, University Medical Center Utrecht, Utrecht. The Netherlands

Published in PLOS Genetics, 2012, 8(6): e1002804

Abstract

Enterococcus faecium has become a nosocomial pathogen of major importance, causing infections that are difficult to treat owing to its multi-drug resistance. In particular, resistance to the β-lactam antibiotic ampicillin has become ubiquitous among clinical isolates. Mutations in the low-affinity penicillin binding protein PBP5 have previously been shown to be important for ampicillin resistance in *E. faecium*, but the existence of additional resistance determinants has been suggested. Here, we constructed a high-density transposon mutant library in *E. faecium* and developed a transposon mutant tracking approach termed Microarray-based Transposon Mapping (M-TraM), leading to the identification of a compendium of *E. faecium* genes that contribute to ampicillin resistance. These genes are part of the core genome of *E. faecium*, indicating a high potential for *E. faecium* to evolve towards β-lactam resistance. To validate the M-TraM results, we adapted a Cre-*lox* recombination system to construct targeted, markerless mutants in *E. faecium*. We confirmed the role of four genes in ampicillin resistance by the generation of targeted mutants and further characterized these mutants regarding their resistance to lysozyme. The results revealed that *ddcP*, a gene predicted to encode a low-molecular-weight penicillin binding protein with D-alanyl-D-alanine carboxypeptidase activity, was essential for high-level ampicillin resistance. Furthermore, deletion of *ddcP* sensitized *E. faecium* to lysozyme and abolished membrane-associated D,D-carboxypeptidase activity. This study has led to the development of a broadly applicable platform for functional genomic-based studies in *E. faecium*, and it provides a new perspective on the genetic basis of ampicillin resistance in this organism.

Introduction

Enterococci rank third overall as causative agents of healthcare-associated infections [1], [2]. Up to the late 1980s, *Enterococcus faecalis* was responsible for practically all enterococcal infections, but starting from the 1990s nosocomial infections with *E. faecium* became more frequent. Currently *E. faecium* causes approximately 40% of all enterococcal infections that are acquired during hospital stay [2]–[4]. Clinical isolates of *E. faecium* have rapidly accumulated antibiotic resistance genes, including those for clinically important antibiotics such as ampicillin and vancomycin, which leads to treatment failure and increased mortality rates [2], [5]–[7]. In the USA, nosocomial infections caused by ampicillin-resistant *E. faecium* (ARE) were first detected in the 1980s and the resistance rates were steadily increasing up to 80% of *E. faecium* isolates in the 1990s [8], [9]. Vancomycin-resistant *E. faecium* (VRE) also emerged in the late 1980s and increased rapidly during the 1990s [9], [10]. Currently, VRE is widespread among clinical *E. faecium* strains in North America, but less common in hospital-acquired infections in Europe [11]. Ampicillin resistance has spread much further and it is currently being reported in over 80% of clinical *E. faecium* isolates from all over the world [1], [2] (European Antimicrobial Resistance Surveillance Network: http://www.ecdc.europa.eu/en/activities/surveillance/EARS-Net/Pages/index.aspx). In addition to ARE and VRE, the emergence of *E. faecium* strains that are resistant to new classes of antibiotics is challenging the few remaining therapeutic options [12]–[14]. Thus, the development of new anti-enterococcal agents may become critical for the successful treatment of infections caused by this multi-drug resistant organism in the future.

The intrinsic resistance to β-lactam antibiotics of enterococci was reported 60 years ago, soon after the introduction of penicillin in the early 1940s, when enterococci were found to be considerably less susceptible to β-lactams than streptococci [15]. Mutations in the high-molecular weight class B penicillin-binding protein 5 (PBP5) have been considered the main cause for the resistance to β-lactams in *E. faecium*. Upregulated expression of *pbp5* and/or mutations in the 3' end of the gene lead to a further reduced susceptibility to ampicillin [16]–[18]. However, several studies have suggested that the high minimum inhibitory concentration (MIC) of ampicillin against *E. faecium* is not exclusively due to the presence of low-affinity PBP5 but also to other genes or mechanisms that remain to be identified [19], [20]. Recently, Mainardi *et al.* [21], [22] showed in a spontaneous mutant that was obtained in the laboratory by selection on agar media containing ampicillin, that the D,D-transpeptidase activity of the PBPs could be bypassed by a β-lactam resistant L,D-transpeptidase (Ldt_{fm}) that catalyses the formation of 3→3 cross-links between peptidoglycan side chains instead of the classical 4→3 cross-links. A D,D-carboxypeptidase, termed DdcY, is an important component in the L,D-transpeptidase mediated pathway of peptidoglycan cross-linking. However, the *ddcY* gene is only present in a small proportion of

E. faecium isolates [23], again suggesting that additional ampicillin resistance determinants in *E. faecium* remained to be identified and characterized.

Genome-wide studies of clinical *E. faecium* isolates have long been hampered by a lack of appropriate genetic tools. In this study, we describe the construction of a high density *mariner* transposon mutant library and the development of a powerful tool for functional genomics, termed Microarray-based Transposon Mapping (M-TraM), in *E. faecium*. By comparing the mutant library following growth in the presence or absence of ampicillin, we identified a compendium of genes affecting the sensitivity to ampicillin. Targeted mutants of the identified genes with predicted roles in cell wall synthesis were generated for further characterization, which resulted in the identification of several intrinsic ampicillin resistance determinants in *E. faecium*. These ampicillin resistance determinants may serve as targets for the development of novel antimicrobial therapeutics.

Results

Construction of a high-density transposon mutant library in E. faecium

To attempt genome-wide transposon mutagenesis of the *E. faecium* genome, we constructed the transposon delivery plasmid pZXL5. As shown in Figure S1, this plasmid was composed of a Gram-positive thermo-sensitive replicon, a gentamicin resistant *mariner* transposon with two outward-facing T7 promoters, a nisin-inducible *mariner* transposase, a ColE1 replicon and a *cat* gene. The sequence of pZXL5 was determined by Sanger-sequencing of both DNA strands (Baseclear; Leiden, The Netherlands) and was deposited in GenBank (GenBank Accession Number: JQ088279).

Using pZXL5, we have produced a transposon mutant library in *E. faecium* strain E1162, an ampicillin-resistant clinical isolate from a bloodstream infection, for which a draft genome sequence has previously been determined [24]. The randomness of the transposon insertions and the absence of multiple transposon insertion events were determined by randomly selecting 17 mutants from the library and carrying out Southern blot hybridizations, using a fragment of the transposon as a probe (Figure S2A), as well as inverse PCR and sequence analysis to determine the location of the transposon insertion point (Figure S2B). The results showed that each mutant carries a single transposon inserted in the genome, and that the transposon was distributed in different loci in the 17 mutants. PCR footprinting was performed to estimate the genome-wide coverage of transposon insertions in the mutant library (Figure 1). An outward-facing primer was designed based on the *mariner* transposon sequence. The other PCR primer was designed for three target genes, *ddl* (which encodes a D-alanine:D-alanine ligase that is essential for bacterial cell wall

biosynthesis [25]), *esp* (which is non-essential and encodes a large surface protein involved in biofilm formation and infection [26]–[28]) and *nox* (which is a non-essential gene encoding a predicted NADH oxidase). Genomic DNA isolated from the pooled mutant library was used as a template. A range of products can be amplified by these primers, each corresponding to a transposon insertion mutant in the library. If a gene is essential for survival, its transposon insertion mutants should not be present in the library after overnight growth, and consequently no PCR products should be amplified in the corresponding size range. As expected, no PCR product was detected within the *ddl* gene (Figure 1) while many PCR bands were found in *esp* and *nox* at intervals of less than 100 bp, indicating the transposon insertions covered the nonessential genes of the genome at high density. Furthermore, mapping the transposon insertion sites to the complete genome sequence of *E. faecium* Aus0004 [29] revealed that transposon insertions were randomly distributed over the genome of this strain and not confined to a specific chromosomal region (data not shown). To establish whether pZXL5 has broad applicability in *E. faecium* we attempted to transform four other clinical *E. faecium* isolates from different geographic origins (Table S1) with pZXL5 using our optimized electroporation protocol. All four strains were efficiently transformed with transformation efficiencies ranging between 110 and 10^5 transformants per µg DNA (Figure S2C). We then continued to generate transposon mutant libraries in two of these strains (Figure S2D). These observations show that the transposon mutagenesis approach that we initially developed for strain E1162, can also be used for functional genomics in other clinical *E. faecium* isolates.

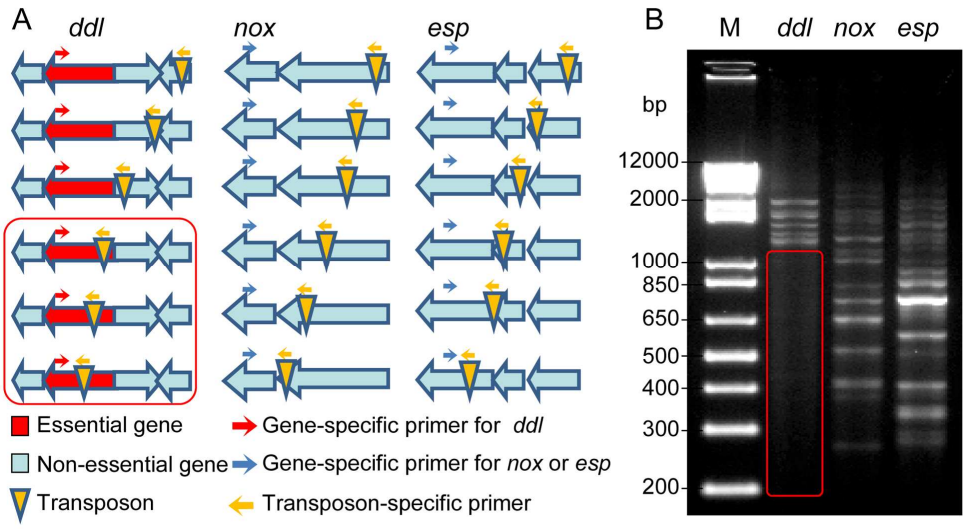


Figure 1. Footprinting analysis of the transposon mutant library. (A) Schematic overview of the transposon footprinting strategy. PCR is performed using a gene specific primer and a primer corresponding to the transposon sequence. (B) Agarose gel electrophoresis of transposon footprinting on the essential gene *ddl* (lane 1), and the non-essential genes *nox* and *esp* (lane 2 and 3, respectively). Each band represents a PCR product of a different size, corresponding to a transposon insertion in a different position. The red box represents the product size range expected for transposon insertions within the essential *ddl* gene.

M-TraM is highly reproducible

As described in the Materials and Methods and in Figure 2, we developed a technique to track the presence of all mutants in the library by simultaneously mapping the transposon insertion sites using microarray hybridization. We termed this technique M-TraM for Microarray-based Transposon Mapping. To validate the reproducibility of M-TraM, we independently grew two aliquots containing approximately 10^7 cells from the library in 20 ml of BHI broth. After 20 hours of culturing at 37°C, genomic DNA was isolated from the two replicate cultures and used for the generation of cDNA. The cDNA samples were labeled with Cy3 and Cy5 respectively and hybridized to a microarray that was designed using the *E. faecium* E1162 genome sequence. The result showed that M-TraM is highly reproducible, with a correlation coefficient of 0.94 between the two independent experiments (Figure 2C).

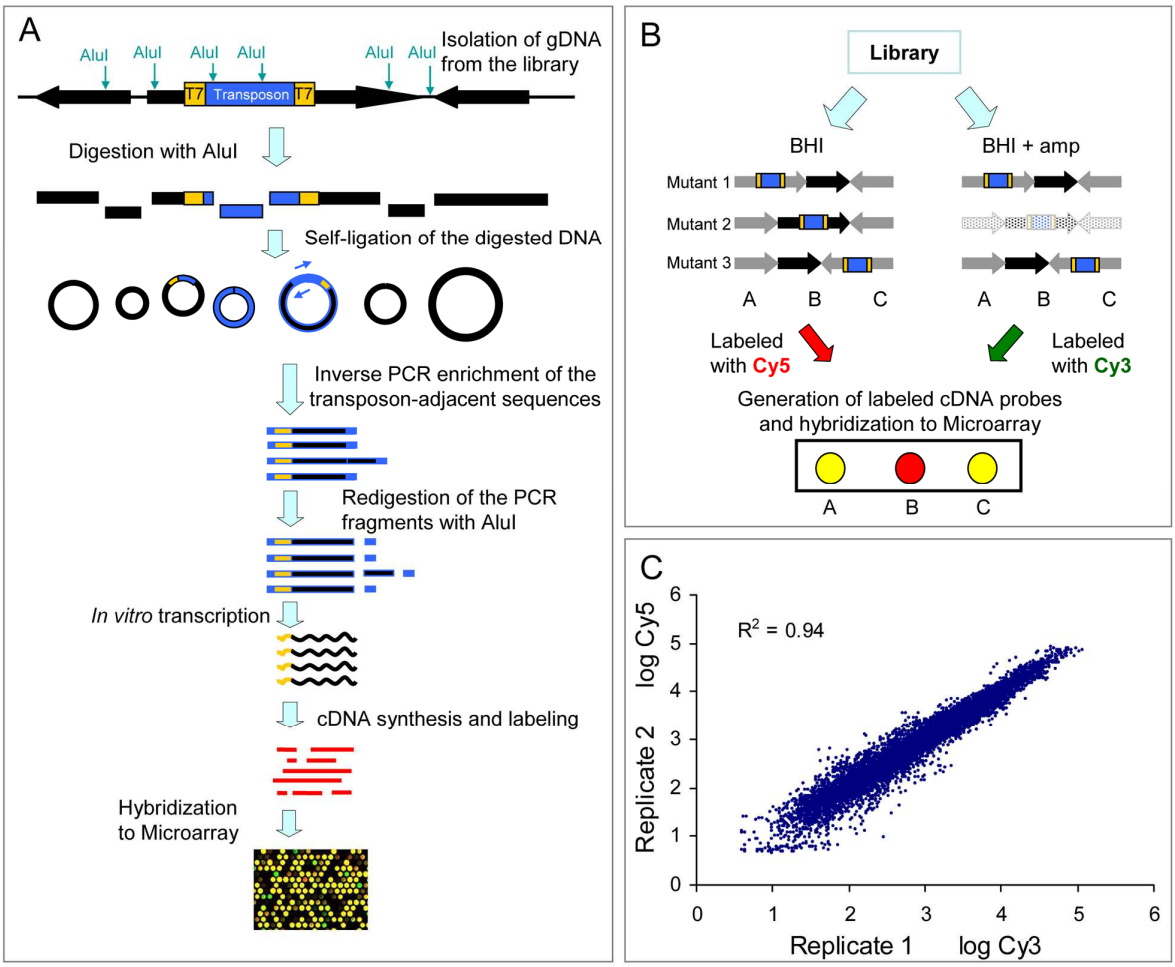


Figure 2. Schematic diagram and reproducibility of M-TraM. (A) Schematic overview of the M-TraM screening. In yellow: inverted terminal repeats (ITRs) of the *himar1* transposon with outward-facing T7 promoters; in blue: the gentamicin resistance gene in the transposon. Genomic DNA is isolated from the *E. faecium* mutant library. DNA is digested with the restriction enzyme *Alu*I, and the DNA fragments are circularized by self-ligation. The transposon-chromosome junction together with an ITR and a T7 promoter is amplified by PCR with primers (blue arrow) that hybridize to the transposon. To eliminate foreign DNA fragments that ligated into the circularized DNA of transposon-

chromosome junctions, the PCR products were re-digested with *Alu*I. The purified DNA fragments are used as template in the *in vitro* transcription reaction. The resulting RNA products are reverse transcribed into cDNA. After labelling, the cDNA is used for microarray hybridization. (B) Schematic overview of the screening strategy to identify conditionally essential genes by M-TraM. A chromosomal region encompassing three genes (A, B, and C) from three different mutants (1, 2, and 3) is shown. Each mutant carries a single transposon insertion (blue) that disrupts the function of the gene. Mutant libraries are grown in a control condition (*e.g.*, BHI) and a test condition (*e.g.*, in the presence of ampicillin). All the three genes are non-essential for growth in the control condition. Gene B is required only for the test condition, so mutant 2 exhibits attenuated growth or poorer survival only in the test condition, and will consequently be reduced or be entirely lost from this library (indicated by light shading). M-TraM samples are generated from the two conditions, labelled with different dyes, and hybridized to a microarray. The DNA probes of gene A and gene C on the microarray will hybridize to the samples generated from both conditions. However, the cDNA sample of gene B will be present at reduced levels only in the test condition. By comparing the signal intensity from the two conditions for each probe, genes involved in growth or survival of the test condition can be identified. (C) Reproducibility of M-TraM. Log-log plot of the microarray signal intensities from two independent experiments of mutant libraries grown under non-selective conditions in BHI broth.

Identification of E. faecium *genes involved in ampicillin resistance by M-TraM*

To identify genes required for ampicillin resistance, we grew the pool of mutants in the presence or absence of a subinhibitory concentration (20 μg ml^{-1}) of ampicillin, and used M-TraM to determine which mutants were selectively lost during culturing in the presence of ampicillin. Eleven genes belonging to a variety of functional categories were identified to be involved in ampicillin resistance (Table 1). Four genes involved in cell wall biogenesis were identified and we decided to further focus on these genes. The EfmE1162_0447 (*ddcP*) and EfmE1162_1886 (ldt_{fm}) genes were predicted to encode a D-alanyl-D-alanine carboxypeptidase (D,D-carboxypeptidase, DdcP) and beta-lactam-insensitive peptidoglycan transpeptidase (L,D-transpeptidase, Ldt_{fm}), respectively. Previous studies have shown that a different D,D-carboxypeptidase (DdcY) and Ldt_{fm} were able to bypass the D,D-transpeptidase activity of the PBPs by forming 3→3 cross-links instead of the classical 4→3 cross-links thereby conferring resistance to β-lactams [21], [23]. DdcP and DdcY only share 13.7% amino acid identity and have a completely different protein domain architecture as DdcY is a β-lactam-insensitive VanY-type carboxypeptidase [23], [30], while DdcP belongs to the family of low-molecular-weight (LMW) PBPs [30] (Figure S3). The *ddcY* gene is absent from 24 (including E1162) of the 29 *E. faecium* genomes available (on 9 January 2012) at NCBI Genomes. The *ddcP* gene is conserved in all 29 *E. faecium* genomes. EfmE1162_0975 (*pgt*) was predicted to encode a glycosyl transferase group 2 family protein which is 63% identical to GltB, a protein that was proposed to be involved in glycosylation of cell wall teichoic acid in serotype 4b *Listeria monocytogenes* [31]. EfmE1162_2487 (*lytG*) was predicted to encode an exo-glucosaminidase that could be acting as a peptidoglycan hydrolase involved in cell wall lysis, remodeling and cell division [32]. An overview of the protein domain architecture and predicted cellular localization of DdcP, Ldt_{fm}, Pgt and LytG is provided in Figure S3. Like

the *ddcP* gene, the *ldt*$_{fm}$, *pgt* and *lytG* genes are present in all the 29 *E. faecium* genomes as well, suggesting these genes are part of the *E. faecium* core genome. Notably the *E. faecium* ampicillin resistance determinants *ddcP*, *ldt*$_{fm}$ and *pgt* do not have homologs (defined here by proteins with >30% amino acid identity) in *E. faecalis*. It should be noted that the L,D-transpeptidase from *E. faecalis* that was biochemically characterized by Magnet *et al.*, [33] has not been experimentally linked to β-lactam resistance in *E. faecalis* and is only remotely related (26% amino acid identity) to *ldt*$_{fm}$.

Table 1: *E. faecium* genes involved in ampicillin resistance as determined by M-TraM analysis.

LocusTag[a]	Accession code	Gene Name	Annotation	Fold-change[b]	Bayesian P-value
EfmE1162_0447	ZP_06676292	*ddcP*	D-alanyl-D-alanine carboxypeptidase	32.5	$4.4x10^{-7}$
EfmE1162_2490	ZP_06678189		oxidoreductase, Gfo/Idh/MocA family	11.8	$8.6x10^{-4}$
EfmE1162_1886	ZP_06677646	*ldt*$_{fm}$	beta-lactam-insensitive peptidoglycan transpeptidase	11.1	$3.0x10^{-5}$
EfmE1162_0975	ZP_06676820	*pgt*	glycosyl transferase, group 2 family protein	10.6	$4.8x10^{-5}$
EfmE1162_0256	ZP_06676101		metallo-beta-lactamase superfamily protein	10.0	$4.8x10^{-4}$
EfmE1162_2260	ZP_06678020		dihydrodipicolinate synthase	7.3	$3.0x10^{-4}$
EfmE1162_2058	ZP_06677818		hydrolase, alpha/beta hydrolase fold family	7.1	$1.7x10^{-4}$
EfmE1162_0669	ZP_06676514		nitroreductase family protein	6.8	$2.8x10^{-4}$
EfmE1162_1943	ZP_06677703		chromosomal replication initiator protein DnaA	5.5	$7.6x10^{-4}$
EfmE1162_0064	ZP_06675909		ribosomal protein L33	5.2	$2.8x10^{-4}$
EfmE1162_2487	ZP_06678186	*lytG*	Exo-glucosaminidase	-7.5	$6.3x10^{-4}$

[a] Indicates the gene containing the transposon insertion.
[b] Indicates the fold-change derived from the ratio of the unselected control library to the ampicillin competitively selected library. *e.g.* the value 32.5 means that the relative quantity of mutants of EfmE1162_0447 in the ampicillin-selected library was 32.5-fold less than in the control library grown without selective pressure. This indicates that mutants in EfmE1162_0447 have a lower relative fitness in the presence of ampicillin than wild type cells. The value of -7.5 for EfmE1162_2487 indicates that mutants in this gene outgrow the other mutants in the ampicillin-selected library by 7.5-fold, indicating that mutants of EfmE1162_2487 have higher relative fitness in the presence of ampicillin.

Comparison of ampicillin sensitivities of targeted mutants and wild-type E1162

To validate the results of the M-TraM screen and to further characterize the role of the identified genes in ampicillin resistance, we constructed targeted mutants in the *ddcP*, *ldt*$_{fm}$ and *pgt* genes, which were identified with the most significant P-values (Table 1) and have predicted functions in cell wall biogenesis. We also generated a targeted mutant in the *lytG* gene of which the inactivation could confer hyper-resistance to ampicillin, as suggested by the M-TraM data. Targeted deletion mutants of *ddcP*, *pgt*, and *lytG* were generated using a novel Cre-*lox*-based system for the generation of markerless mutants in *E. faecium* that we developed as part of this study (Figure S4). For *ldt*$_{fm}$ no double cross-over mutant could be constructed and instead a single cross-over mutant was constructed using the pWS3 vector

[34]. Mutants of *ddcP*, *ldt*$_{fm}$, and *lytG* were also complemented *in trans* and the ampicillin resistance of E1162 (wild-type), the mutants and the complemented strains was determined. The *pgt* mutant could not be complemented as constructs containing the *pgt* gene could not be transformed to either *Escherichia coli* or *E. faecium*, presumably due to toxicity of the gene product. In the absence of ampicillin we did not detect significant differences in growth speed or cell density upon entry into stationary phase between wild-type and the mutant strains (Figure S5). When these strains were grown in BHI with 20 µg ml^{-1} ampicillin, the Δ*ddcP* mutant was dramatically affected in its growth (Figure 3A). Growth of the *ldt*$_{fm}$::pWS3 and Δ*pgt* mutants was also poorer than wild-type (Figure 3B and 3C). The *in trans* complemented strains of the Δ*ddcP* and *ldt*$_{fm}$::pWS3 strains could fully or partially restore the ampicillin resistance to wild-type levels (Figure 3A–3C). The Δ*lytG* mutant had a growth rate that was similar to the parental strain's (0.998±0.007 h^{-1} for Δ*lytG* vs. 0.989±0.018 h^{-1} for E1162) but could grow to slightly higher optical densities (Figure 3D). The *in trans* complemented *lytG* mutant exhibited a significantly lower growth rate (0.859±0.017 h^{-1}) in exponential phase (Figure 3D). The empty vector had no effect on the growth of the mutants in BHI supplemented with 20 µg ml^{-1} ampicillin (data not shown). MICs of ampicillin against the wild-type E1162 and Δ*ddcP*, *ldt*$_{fm}$::pWS3, Δ*pgt* and Δ*lytG* strains were determined by microdilution in cation-adjusted Muller-Hinton broth as 43, 8, 16, 27 and 43 µg ml^{-1}, respectively, which is in accordance with the growth performance of the mutants in BHI with 20 µg ml^{-1} ampicillin (Figure 3). The seemingly contradictory observation that the *ldt*$_{fm}$::pWS3 mutant can grow in BHI medium supplemented with ampicillin at a concentration above the MIC can be explained by the differences in growth media used and the approximately 4-fold larger inoculum size used in the growth experiments in BHI compared to the inoculum size used in the MIC determinations. The MIC of vancomycin was determined to be 0.5 µg ml^{-1} for all strains.

Comparative analysis of the transcriptome of E. faecium *E1162 during exponential growth in the absence and presence of ampicillin*

We used microarray-based transcriptome analysis on exponentially growing (OD_{660} = 0.3) *E. faecium* E1162 cultures in BHI medium with or without 20 µg ml^{-1} ampicillin to identify genes that are regulated by the exposure to sub-MIC levels of ampicillin. Compared to the untreated control, only sixteen genes were identified to be differentially regulated between the two conditions and none of these genes were upregulated more than 2.1-fold in the presence of ampicillin (Table S2), indicating that *E. faecium* does not require major transcriptional rearrangements to cope with the presence of sub-inhibitory levels of ampicillin. None of the genes that were identified by M-TraM were identified to be differentially regulated by the presence of ampicillin, which indicates that the identified ampicillin resistance determinants are constitutively expressed, even in the absence of ampicillin.

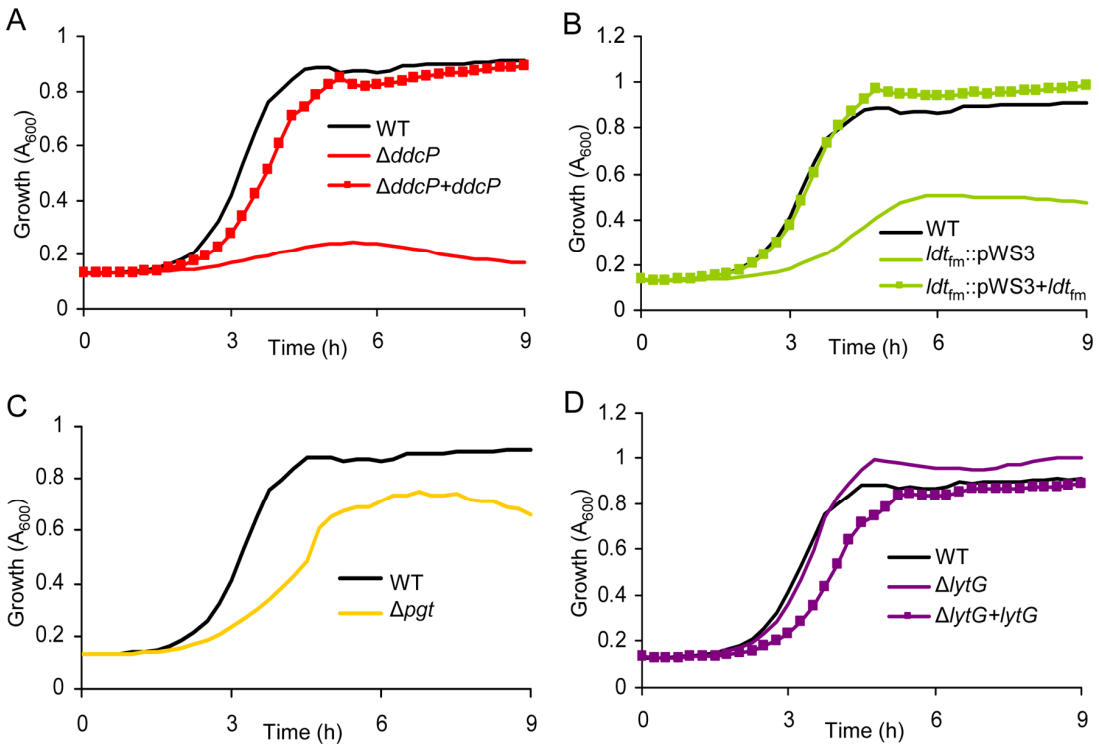


Figure 3. The effect of targeted mutations of *ddcP*, ldt_{fm}, *pgt*, and *lytG* on growth of *E. faecium* in the presence of ampicillin. Overnight cultures of mutants and wild-type *E. faecium* were inoculated at an initial cell density of OD_{660} 0.0025 in BHI with 20 µg ml^{-1} ampicillin. Growth curves of strain E1162, the different mutants (panel A: Δ*ddcP*; panel B: ldt_{fm}::pWS3; panel C: Δ*pgt*; panel D: Δ*lytG*) and *in trans* complemented strains are shown. Growth curves are mean data of three independent experiments.

Mutations in ddcP *and* ldt_{fm} *increase sensitivity to lysozyme*

After the identification of *ddcP*, ldt_{fm} and *pgt* as ampicillin resistance determinants, we studied the susceptibility of the wild-type strain E1162 and its mutants to another compound targeting the cell wall, i.e. lysozyme, which is one of the most important antimicrobial enzymes of the host innate immune system. Lysozyme kills Gram-positive bacteria by enzymatic lysis of the bacterial cell wall [35]. The results demonstrated that deletion of *pgt* had no significant effect on lysozyme resistance, while Δ*ddcP* and ldt_{fm}::pWS3 mutants were significantly more sensitive to lysozyme challenge than the wild-type strain (Figure 4). The *in trans* complemented strains of Δ*ddcP* and ldt_{fm}::pWS3 could restore the resistance to lysozyme. Hence, *ddcP* and ldt_{fm} contribute not only to β-lactam resistance but also to the resistance against the peptidoglycan-hydrolyzing enzyme lysozyme.

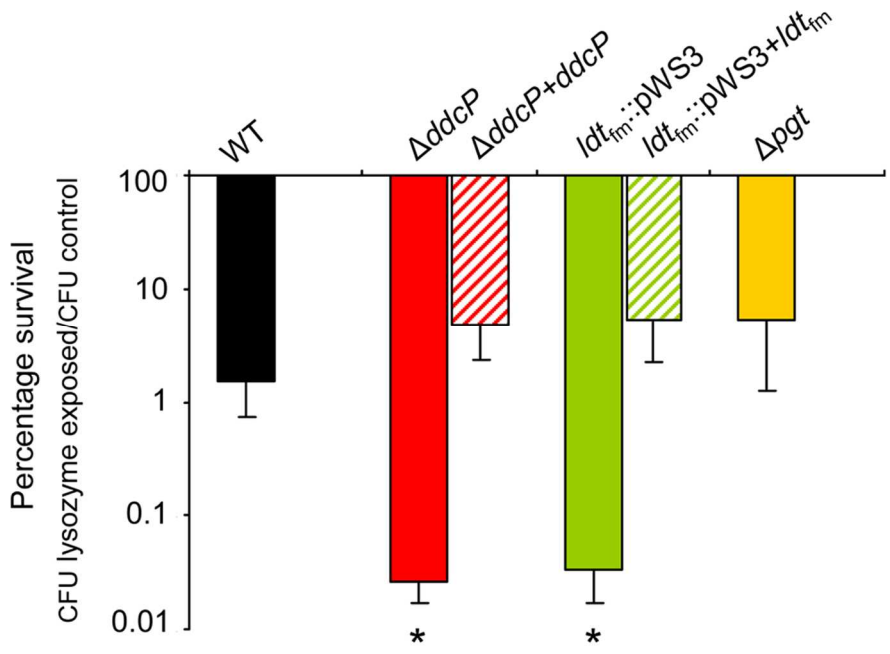


Figure 4. Percentage survival of *E. faecium* cells following a lysozyme challenge. Survival of the indicated wild-type, mutant strains and *in trans* complemented strains following a 30-minute incubation in PBS containing 0.5 mg ml^{-1} lysozyme relative to the survival of the strains after a 30-minute incubation in PBS without lysozyme. Bars represent the standard deviation of the mean of three independent experiments. Asterisks represent significant differences (P<0.005 as determined by a two-tailed Student's *t*-test) between the indicated mutants and the wild-type strain.

Disruption of ddcP *abolishes membrane-associated D,D-carboxypeptidase activity of* E. faecium *E1162*

Of all novel ampicillin resistance determinants identified in this study, the *ddcP* gene contributes most to ampicillin resistance in *E. faecium* E1162. The DdcP protein was annotated as a D-alanyl-D-alanine carboxypeptidase but a functional study confirming this activity has not been performed. To confirm its predicted function, we determined the D-alanyl-D-alanine carboxypeptidase activity in cellular extracts of E1162, the Δ*ddcP* mutant and the *in trans* complemented strain Δ*ddcP+ddcP*. As shown in Figure 5, the D-alanyl-D-alanine-carboxypeptidase activity of Δ*ddcP* membrane extracts was completely abolished. In the complemented Δ*ddcP+ddcP* strain enzymatic activity was restored in the membrane fraction, revealing that DdcP is responsible for D-alanyl-D-alanine carboxypeptidase activity in *E. faecium* E1162. The D,D-carboxypeptidase activity was approximately 5 fold lower in the cytoplasmic fractions than in the membrane fractions (data not shown), strongly suggesting that the DdcP protein is associated with the membrane. When *E. faecium* E1162 was grown in the presence of 20 µg ml^{-1} ampicillin, all D-alanyl-D-alanine carboxypeptidase activity was undetectable, which is in accordance with the designation of DdcP as a LMW-PBP.

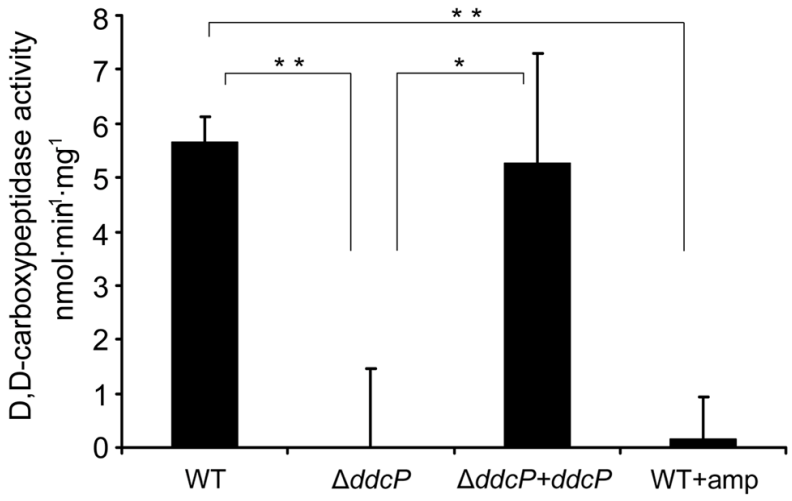


Figure 5. D,D-carboxypeptidase activity in *E. faecium* membrane fractions. Membrane extracts were isolated from the indicated strains grown in BHI medium (with or without the addition 20 µg ml^{-1} ampicillin) until an OD_{600} of 0.7. The D,D-carboxypeptidase activity (nmol min^{-1} mg^{-1}) was defined as the number of nmoles of D-Ala released from pentapeptide (7.5 mM) per min and per mg of protein in the membrane fractions. D-Ala was assayed using D-amino acid oxidase coupled to peroxidase. Bars represent the standard deviation of the mean of three independent experiments. Asterisks represent significant difference ($^{*}P$<0.05, $^{**}P$<0.005 as determined by a two-tailed Student's *t*-test) between the different strains and conditions.

Confirmation of pbp5 *as an important ampicillin resistance determinant in* E. faecium *E1162*

Notably, *pbp5* was not identified in this M-TraM screen even though this gene has been implicated in high-level ampicillin resistance in *E. faecium* [16]–[18]. Our failure to identify *pbp5* by M-TraM may partially be explained by the presence of an *Alu*I restriction site on one of the two microarray probes for *pbp5*. The other microarray probe maps to an internal region of the *pbp5* gene and is located on a small *Alu*I restriction fragment with a window of 158 nucleotides for the transposon insertion. Previous genome-wide studies using transposon mutagenesis have shown that transposon insertions might fail to fully inactivate the target genes. The failure is commonly found with insertions situated near either end of a gene, but has also been observed with internal insertions [36], [37], after which genes might be capable of intracistronic complementation. We constructed targeted insertional mutants of *pbp5*, *ddcP* and *ldt*$_{fm}$ by the insertion of the plasmid pWS3 5′ to the gene probes and close to the central region of the gene. Consistent with the M-TraM results and the phenotypes of the markerless mutants, the insertional mutants of *ddcP* (data not shown) and *ldt*$_{fm}$ (Figure 3B) were sensitized to ampicillin. Surprisingly the insertional mutation in *pbp5* did not significantly affect the sensitivity to ampicillin (Figure 6), implying that insertional mutation (by either a plasmid or a transposon) could not fully disrupt *pbp5*. Therefore a markerless deletion mutant for *pbp5* was generated to completely abolish the function of this gene using the Cre-*lox* system described above. To our knowledge, a targeted mutant of the *pbp5* gene has not been generated previously in *E. faecium*. The *pbp5* mutant was unable to grow in cultures containing 20 µg ml^{-1} ampicillin and the ampicillin resistant phenotype could be partially be restored by *in trans* complementation of Δ*pbp5* with the *pbp5* gene (Figure 6). The MIC of ampicillin for Δ*pbp5* was determined by broth microdilution to be only 0.2 µg/ml.

Our results revealed that the relative contribution of *pbp5*, *ddcP*, ldt_{fm} and *pgt* to ampicillin resistance in *E. faecium* E1162 can be summarized as *pbp5*□*ddcp*>ldt_{fm}>*pgt*. The *pbp5* mutant was also assayed for its survival in the presence of lysozyme, but no significant difference was found with the parental strain E1162 (data not shown).

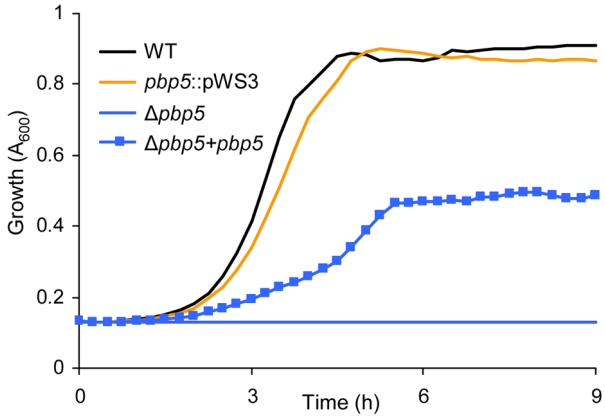


Figure 6. Growth curves of *E. faecium* E1162 and *pbp5* mutants in BHI with 20 µg ml^{-1} ampicillin. Overnight cultures of wild-type *E. faecium* E1162, the insertional *pbp5* mutant (*pbp5*::pWS3), the markerless deletion mutant Δ*pbp5* and the *in trans* complemented deletion mutant (Δ*pbp5*+*pbp5*) were inoculated at an initial cell density of OD_{660} 0.0025 in BHI with 20 µg ml^{-1} ampicillin. Growth curves are mean data of three independent experiments.

Discussion

Ampicillin resistance in *E. faecium* has emerged in the late 1970s and has spread rapidly since [38]. Practically all clinical isolates are currently resistant to ampicillin. The resistance to β-lactams of *E. faecium* complicates the treatment of infections with this organism, particularly when resistance to other antibiotics has also been acquired. The goal of the research described here was to identify genes involved in ampicillin resistance in *E. faecium* in a high-throughput fashion. We developed a system for the generation of a large random transposon mutant library in *E. faecium*, coupled to a microarray-based screening approach (termed M-TraM) to simultaneously monitor the relative fitness of individual mutants undergoing selection by growth in the presence of ampicillin.

The lack of appropriate genetic tools has long been a bottleneck for the studies of *E. faecium*. In this study, we constructed a random high-density transposon mutant library in *E. faecium*, developed a powerful screening technique to track transposon mutants and adapted the Cre-*lox* [39] recombination system to construct targeted, markerless mutants in *E. faecium*, which enabled us to perform high-throughput genome-wide analysis and specific targeted investigations in a clinical *E. faecium* isolate.

When *E. faecium* is exposed to ampicillin the D,D-transpeptidase activity of PBPs is inhibited, with the exception of the low-affinity PBP5, which can catalyze the last cross-linking step of the D,D-transpeptidation pathway of cell wall assembly [17]. This implies that any D,D-

transpeptidase activity conferred by genes other than *pbp5*, is not essential for ampicillin resistance. Consequently, no genes encoding D,D-transpeptidases were identified in this study. However, in our transposon mutant screen, we identified a novel D-alanyl-D-alanine carboxypeptidase (DdcP) that plays an important role in resistance to ampicillin. DdcP is the only gene that is responsible for D-alanyl-D-alanine carboxypeptidase activity during exponential growth of strain E1162. The observation that the enzymatic activity of DdcP is abolished in the presence of ampicillin is in accordance with the prediction that DdcP is a LMW-PBP. The *in vivo* functional roles of LMW-PBPs are relatively poorly understood but they are generally not essential for survival and are thought to contribute to peptidoglycan-remodelling in both Gram-positive and Gram-negative bacteria [40]. Deletion of the gene encoding a LMW-PBP can lead to an increased sensitivity towards β-lactam antibiotics [41]–[42]. Although a mechanistic understanding for this sensitive phenotype is currently lacking, it has been proposed that these LMW-PBPs may enzymatically inactivate β-lactams or, alternatively, save other PBPs from inactivation by sequestration of the β-lactams to the LMW-PBPs [41]. Further functional characterization of DdcP in *E. faecium*, also involving strains that have higher and lower resistance towards ampicillin than strain E1162, is needed to identify the exact mechanism by which DdcP contributes to ampicillin resistance in *E. faecium*. Previous biochemical studies have indicated that Ldt_{fm} is a crucial component of the β-lactam-insensitive L,D-transpeptidation pathway, which catalyzes the cross-links of tetrapeptides [21]–[23]. However, genetic evidence for the role of Ldt_{fm} in ampicillin resistance was so far lacking. Here, we have constructed a mutant, confirming the role of this pathway in ampicillin resistance in *E. faecium*. The identification of *pgt* as an ampicillin resistance determinant in *E. faecium* suggests that wall teichoic acid is involved in β-lactam resistance which is in line with a similar observation in methicillin-resistant *Staphylococcus aureus* (MRSA) strains [43]. None of the mutations had an effect on vancomycin resistance of *E. faecium*, presumably because tetrapeptide precursors for peptidoglycan crosslinks are present at levels that are too low to confer resistance to this antibiotic in wild-type and the mutant strains.

We did not generate mutants or performed functional analyses to confirm the function of the other genes that were identified in our M-TraM screening (Table 1). However, it seems likely that at least some of these genes also contribute to ampicillin resistance in *E. faecium*, in particular EfmE1162_2490 (predicted to function as an NAD- or NADP-dependent oxidoreductase) and EfmE1162_0256 (possibly acting as a zinc-dependent β-lactamase), because the transposon mutants in these genes appear to have a similar loss in fitness in the presence of ampicillin as the transposon mutants in ldt_{fm} and *pgt* (Table 1). The *lytG* gene was the only gene that was identified in our screen which, upon inactivation by a transposon, appeared to result in a hyper-resistant phenotype. However, this phenotype could not be confirmed in a markerless deletion mutant of *lytG*, which had the same MIC for

ampicillin as the wild-type strain. The observed slightly higher optical density reached by Δ*lytG* as compared to wild-type and the slower growth rate of the *in trans* complemented *lytG* mutant may suggest that there is a subtle role for *lytG* in ampicillin resistance in *E. faecium* but further experiments would be required to exactly determine its role.

The recent emergence of *E. faecium* as a major nosocomial pathogen can be explained by the acquisition of genes that contribute to colonization or infection [26], [44] and by the acquisition of resistance to antibiotics, particularly to ampicillin and vancomycin. Interestingly, resistance against ampicillin and vancomycin emerged predominantly in *E. faecium* while both resistances are virtually absent in *E. faecalis* [8]. This study provides insights into the genetic basis of intrinsic β-lactam resistance in *E. faecium* and identified the ampicillin resistance determinants DdcP, Ldt_{fm} and Pgt in this organism. All three genes are conserved in *E. faecium* but absent from *E. faecalis*, indicating that *E. faecium* possesses more innate β-lactam resistance determinants than *E. faecalis*. This observation supports the concept that *E. faecium* has a higher potential to develop high-level β-lactam resistance than *E. faecalis*, thereby explaining the faster emergence of ampicillin resistance in *E. faecium* than in *E. faecalis* [1], [2], [45].

We have identified several novel mechanisms, besides the low-affinity penicillin-binding protein PBP5, that contribute to ampicillin resistance in *E. faecium*. These proteins could serve as targets for the development of novel therapeutics against this multi-resistant organism. Our study showed that DdcP and Ldt_{fm} also contribute to resistance to lysozyme of *E. faecium*. An inhibitor of these proteins may thus provide the dual benefit of compromising resistance to the innate immune system as well as enhancing antibiotic susceptibility. In the course of this study we have also developed a number of novel genetic tools for *E. faecium* allowing for genome-wide analysis of this bacterium. Further functional genomic-based studies to understand the mechanisms involved in colonization, infection and antibiotic resistance of this important nosocomial pathogen are now a realistic opportunity for future research.

Materials and Methods

Bacterial strains, plasmids, and growth conditions

E. faecium and *E. coli* strains used in this study are listed in Table S1. The ampicillin-resistant *E. faecium* strain E1162 was used throughout this study. This strain was isolated from a bloodstream infection in France in 1996 and its genome has recently been sequenced [24]. Unless otherwise mentioned, *E. faecium* was grown in brain heart infusion broth (BHI; Oxoid) at 37°C. The *E. coli* strains DH5*α* (Invitrogen) and EC1000 [46] were grown in Luria-Bertani (LB) medium. Where necessary, antibiotics were used at the following

concentrations: chloramphenicol 4 μg ml^{-1} for *E. faecium* and 10 μg ml^{-1} for *E. coli*, gentamicin 300 μg ml^{-1} for *E. faecium* and 25 μg ml^{-1} for *E. coli*, spectinomycin 300 μg ml^{-1} for *E. faecium* and 100 μg ml^{-1} for *E. coli*, and erythromycin 50 μg ml^{-1} for *E. faecium* (with added lincomycin at 50 μg ml^{-1}) and 150 μg ml^{-1} for *E. coli*. All antibiotics were obtained from Sigma-Aldrich (Saint Louis, MO). Growth was determined by measuring the optical density at 660 nm (OD_{660}).

Construction of an E. faecium in vivo *transposon mutagenesis system*

The transposon delivery plasmid, pZXL5, was constructed in several steps. The Gram-positive *lacZ* gene of pCJK72 [47] was PCR (Accuprime DNA polymerase, Invitrogen) amplified using the primers pCJK72_PstI_lacZ_F and pCJK72_KpnI_lacZ_R (primer sequences are listed in Table S3) and cloned into pTEX5500ts [48] between *Pst*I and *Kpn*I sites to create pZXL1. A fragment containing the chloramphenicol resistance (Chl^r) cassette, the *lacZ* gene and a gram-positive thermosensitive origin of replication (*repA*ts, functional at 30°C, but not at 37°C) from pZXL1 was PCR amplified using the pZXL1_EcoRI _cm_ori_F and pZXL1_EcoRI _cm_ori_R. Meanwhile, another fragment containing a nisin inducible *mariner* transposase C9 and the *nisRK* genes (encoding a two-component system required for the transcriptional activation of the transposase gene in the presence of nisin), and the ColE1 origin of replication was PCR amplified from pCJK55 [47] using the primers pCJK55_EcoRI_tps_F and pCJK55_EcoRI_tps_R. These two fragments were digested with *EcoR*I and then ligated together to generate pZXL2. To construct a *mariner* transposon [49], the 5′ and 3′ ITRs of *Himar1-mariner* were amplified from pMMOrf [50] using the primer pMMOrf_SacII_ITR that resulted in two *Sac*II recognition sites at both ends of the amplified DNA. The amplified fragment was cloned into pGEM-T Easy (Promega) forming pGEM-ITR. A gentamicin-resistance cassette was PCR amplified from pAT392 [51] using primers pAT392_EheI_T7_genta_F and pAT392_EheI_T7_genta_R, resulting in a gentamicin-resistance cassette with outward-facing T7 promoters on both ends of the cassette, which allows for the generation of RNA products corresponding to the regions flanking the site of transposon integration in genomic DNA. This fragment was digested with EheI and cloned into a SmaI site present between the 5′ and 3′ ITRs in pGEM-ITR, thereby forming the transposon cassette. This transposon cassette was then cut out with *Sac*II and subcloned into the *Sac*II site of pZXL2 producing pZXL3. This vector was electroporated to *E. faecium* E1162 but pZXL3 was found to be able to replicate at 37°C in *E. faecium* E1162 and blue/white screening using *lacZ* proved to be ineffective (data not shown). We therefore replaced the *repA*ts and *lacZ* gene of pZXL3 by the *repA*ts from pAW068 [52]. To this aim, a fragment of pZXL3 containing the nisin inducible *mariner* transposase and the transposon cassette was PCR amplified by primers pZXL3_BfrI_tn_F and pZXL3_BfrI_tn_R. Another fragment carrying the *repA*ts and Chl^r cassette was amplified from pAW068 by PCR using primers pAW068_BfrI_cm_ori_F and pAW068_BfrI_cm_ori_R. After digestion with BfrI these

two PCR products was ligated together, resulting in the generation of pZXL5. All restriction enzymes were obtained from New England Biolabs.

Transposon mutant library construction and evaluation

Electrotransformation of the different *E. faecium* strains (Table S1) with the plasmid pZXL5 was performed according to previously described methodologies [26], [34] with optimizations in preparing electrocompetent cells and the cell-plasmid mixture. To obtain the electrocompetent cells, overnight cell culture from BHI was diluted 1000 fold in 25 ml BHI supplemented with 1% of glycine and 200 mM sucrose and again grown overnight at 37°C. Cells were then resuspended in same volume of pre-warmed BHI supplemented with 1% glycine and 200 mM sucrose and incubated at 37°C for 1 hour. Cells were washed three times with ice-cold wash buffer (500 mM of sucrose and 10% glycerol), and resuspended with 1.25 ml ice-cold wash buffer. A 100 µl aliquot of the cell suspension was mixed with 0.1–1 µg of plasmid and transferred into an ice-cooled electroporation cuvette (2-mm gap) and kept on ice for 20 minutes before electroporation. Gentamicin-resistant transformants were grown overnight in BHI broth supplemented with 300 µg ml^{-1} gentamicin and 10 µg ml^{-1} chloramphenicol at the permissive temperature of 28°C, after which 100 µl (approximately 10^8 viable cells) of this overnight culture were inoculated in 200 ml of pre-warmed BHI broth supplemented with gentamicin and 25 ng ml^{-1} nisin and grown overnight at the non-permissive temperature of 37°C with shaking at 150 rpm. Subsequently, 100 µl of this culture was transferred to 200 ml of fresh pre-warmed BHI broth and similarly grown overnight without nisin. Cultures were then stored at −80°C in BHI broth containing 50% (v/v) glycerol in 1 ml aliquots as mutant library stocks.

To evaluate the randomness and coverage of transposition, we performed Southern blot analysis, identified the sites of transposon insertion and used PCR footprinting. Southern blot analysis was performed as described previously [34]. Genomic DNA of 17 arbitrarily picked gentamicin-resistant colonies from the library was isolated using the Wizard Genomic DNA Purification kit (Promega), digested with HaeIII and BamHI. The probe consisting of a 414 bp fragment within the gentamicin-resistance gene was amplified from pZXL5 by PCR, using the primer pair genta_probe_F/R. To map the sites of transposon insertion, genomic DNA of 17 mutants from the library was digested with HaeIII and then self-ligated, forming circular DNA. Loci in which the transposon had inserted were amplified using the transposon-specific primer pair IPCR_HaeIII_R/F with AccuPrime DNA polymerases (Invitrogen) with the following conditions: 94°C for 1 min; 32 cycles of 94°C for 18 sec, 53°C for 30 sec, 68°C for 10 min; and 68°C for 7 min. Sequencing of the PCR product was performed using the primer IPCR_HaeIII_R and/or IPCR_HaeIII_F. PCR footprinting was conducted on genomic DNA of mutant library as described elsewhere [53] with a

transposon-specific primer, ftp_tn, and gene-specific primers, ftp_ddl, ftp_nox or ftp_esp, respectively.

Simultaneously mapping of transposon insertion sites by M-TraM

Transposon insertion mapping was based on the previously published method of Genomic Array Footprinting (GAF) [54]. Because we observed that T7 polymerase will transcribe *E. faecium* genomic DNA aspecifically (data not shown), it was required to modify GAF to specifically enrich for the junction sites of the transposon and the flanking *E. faecium* DNA. Genomic DNA from mutant libraries was isolated using the Wizard Genomic DNA Purification kit (Promega), digested with *Alu*I (New England Biolabs) and then purified on a Qiagen QIAquick PCR Purification column (Qiagen). 200 ng of the digested DNA was self-ligated by the Quick Ligation Kit (New England Biolabs) in a reaction volume of 20 µl. This ligation reaction was directly used as template for PCR amplification of the transposon–chromosome junctions with primer pair IPCR_*Alu*I_F and IPCR_*Alu*I_R in a reaction volume of 200 µl, using AccuPrime Taq DNA polymerases High Fidelity (Invitrogen) with the following conditions: 94°C for 1 min; 26 cycles of 94°C for 18 sec, 56.5°C for 30 sec, 68°C for 50 sec; and 68°C for 7 min. PCR products were purified using Qiagen QIAquick PCR Purification Kit. After purification, 200–500 ng DNA was redigested by *Alu*I and used for *in vitro* transcription (IVT) in a volume of 20 µl using the T7 MEGAshortscript kit (Ambion) at 37°C for 6 hours. The RNA was first treated with DNase (Ambion) and then purified with the MEGAclear Kit (Ambion). 5–10 µg of the purified RNA was used for generating labeled cDNA using the FairPlay III Microarray Labeling Kit (Agilent Technologies) as described in the manufacturer's protocol. Samples of both conditions (grown in BHI and BHI with 20 µg ml^{-1} ampicillin) were labeled with Cy3 or Cy5. Dyes were swapped between samples to minimize the effect of dye bias. Microarray hybridizations were carried out using the Gene Expression Hybridization Kit (Agilent) following the manufacturer's instructions, using 60 ng of labeled cDNA. The experiment was performed with four biologically independent replicates.

The microarrays used in this study were custom-made *E. faecium* E1162 microarrays using Agilent's 8×15 K platform. Probes were designed by Agilent's eArray server. As probes 60-mer oligonucleotides were designed on coding sequences (CDS) only. A total of 2650 CDS are covered by 2 probes (98.4% of the total number of CDS in the *E. faecium* E1162 genome sequence; NCBI accession number NZ_ABQJ00000000), which were spotted in duplicate. A total of 11 CDS are covered by a single probe (0.4% of the total number of CDS) and these probes were spotted in quadruplicate. For 33 CDS no probes could be designed.

Microarray data were extracted and normalized using Agilent Feature Extraction Software Version 10.7.1.1 (FE 10.7.1). Statistical differences in hybridization signals between the conditions were analyzed using Cyber-T [55] (http://cybert.microarray.ics.uci.edu/). Probes

exhibiting Bayesian P-value<0.001 were deemed statistically significant. A gene with two identical probes or all four probes meeting this criterion were classified as significantly selected during exposure to ampicillin.

Screening for genes involved in ampicillin resistance

To carry out identification of genes required for ampicillin resistance, aliquots containing approximately 10^7 CFU from the mutant pool stored at −80°C were diluted 1 to 1000 in 20 ml of BHI broth or BHI broth with 20 μg ml^{-1} ampicillin. Cells were grown at 37°C for 8 hours, after which 1 ml of the bacteria cultures were spun down and used for the extraction of genomic DNA, which was then further processed as described above. The same protocol (except that the cells were grown for 20 hours instead of 8 hours) was used to determine the reproducibility of the M-TraM screening in which two independent mutant libraries cultures were mapped using the approach described above.

Construction of targeted, markerless deletion mutants and in trans *complementation*

For this study, we developed a new method to construct markerless mutants in *E. faecium* based on the Cre-*lox* recombination system [39]. The 5′ and 3′ flanking regions (approximately 500 bp each) of the target genes were PCR amplified with the primers in Table S3. The two flanking regions were then fused together by fusion PCR (generating an *EcoR*I site between both fragments) and cloned into pWS3. Then a gentamicin-resistant cassette was PCR amplified from pAT392 using primers pAT392_EcoRI_lox66_genta_F and pAT392_EcoRI_lox71_genta_R, resulting in a gentamicin-resistant cassette flanked by *lox66* and *lox71*, which allows for the deletion of the gentamicin-resistant cassette in the presence of Cre recombinase. This fragment was digested with *EcoR*I and cloned into the *EcoR*I site that was generated between the 5′ and 3′ flanking regions in the pWS3 construct and then electrotransformed into *E. faecium* as previously described [26], [48]. A transformant containing the plasmid was grown overnight in BHI broth at 30°C supplemented with gentamicin. The cell culture was then diluted 10,000-fold in prewarmed BHI broth and grown at 37°C overnight without antibiotics. The cells were then plated on BHI agar plates with gentamicin and incubated at 37°C. Colonies were then restreaked on BHI agar plates with spectinomycin and BHI agar plates with gentamicin, respectively. The gentamicin-resistant but spectinomycin-susceptible colonies were supposed to be marked deletion mutants and checked by PCR (Table S3). To remove the marker and obtain the markerless mutants a Cre cassette was cut from pRAB1 [56] by digestion with *Pst*I and *Sac*I, blunted by Quick Ligation Kit (New England Biolabs) and cloned into the *EcoR*V site of pWS3 producing pWS3-Cre, which was subsequently electrotransformed into the marked mutants. Spectinomycin-resistant transformants containing pWS3-Cre were then grown overnight in BHI broth at 30°C supplemented with spectinomycin and then diluted 10000 fold in pre-warmed BHI broth and grown at 37°C overnight without antibiotics. These cultures were plated on BHI

agar plates and incubated at 37°C for 18–24 h. Single colonies were then restreaked on BHI agar with spectinomycin, BHI agar with gentamicin and BHI agar without antibiotics. The colonies that were susceptible to both gentamicin and spectinomycin resulted from a recombination event catalyzed by Cre and subsequent loss of the thermosensitive plasmid, resulting in a markerless deletion mutant of the gene of interest. This was verified by PCR and sequencing.

Insertional mutagenesis was performed as previously described [34]. Internal DNA fragments of target genes were PCR amplified using primers listed in Table S3, cloned to a Gram-positive thermosensitive plasmid and electrotransformed into *E. faecium* as previously described [26], [48]. After electrotransformation, the cells were recovered for 2 hours at 30°C, after which the cells were plated on BHI plates supplemented with 300 µg ml^{-1} spectinomycin at 30°C to select for transformants. Spectinomycin-resistant colonies were picked and grown overnight in 200 ml of BHI broth at an elevated temperature (37°C) to cure the plasmid. The cells were then plated on BHI agar plates with spectinomycin at 37°C. Single-cross-over integrations into the target genes were verified by PCR with a pWS3-specific primer, check_pWS3, and a gene-specific primer (Table S3).

Plasmids for the *in trans* complementation of the *ddcP*, ldt_{fm}, *lytG* and *pbp5* mutants were produced by PCR amplification of the genes using the primers listed in Table S3. PCR products were ligated into the downstream region of P_{nisA} promoter of pMSP3535 [57]. The resulting plasmids were introduced into the appropriate host strains by electroporation as described above.

Determination of growth curves and MIC

A BioScreen C instrument (Oy Growth Curves AB, Helsinki, Finland) was used to monitor effects of ampicillin on bacterial growth. Wild-type *E. faecium*, mutants and *in trans* complemented strains were grown overnight in BHI and BHI containing appropriate antibiotics. Cells were inoculated at an initial OD_{660} of 0.0025 into 300 µl BHI and BHI with ampicillin 20 µg ml^{-1} and 1 µg ml^{-1}.The cultures were incubated in the Bioscreen C system at 37°C with continuous shaking, and absorbance of 600 nm (A_{600}) was recorded every 15 min for 9 hours. Each experiment was performed in triplicate.

MIC of ampicillin of the wild-type and mutants were determined in triplicate by broth microdilution in cation-adjusted Muller-Hinton broth as previously described [58].

Transcriptome profiling

E. faecium E1162 was incubated in BHI broth and BHI broth supplemented with 20 µg ml^{-1} ampicillin for 18 hours. Cultures were then diluted to OD_{660} 0.025 in 20 ml of prewarmed BHI

broth and BHI broth containing 20 µg ml^{-1} ampicillin respectively, and grown until OD_{660} 0.3. Cells were centrifuged for 12 seconds at 169000 g at room temperature, and pellets were flash frozen in liquid N_2 prior to RNA extraction. RNA was isolated using TRI Reagent (Ambion) according to the manufacturer's protocol. RNA quantity and quality was determined by spectrophotometry (Nanodrop 1000, Thermo Scientific, Wilmington DE, USA) and by Bioanalyzer 2100 analysis (Agilent). Labeling of 5 µg of total RNA, hybridization and data analysis were performed as described above. Genes for which all four probes exhibited a Bayesian P<0.001 in Cyber-T [55] were deemed differentially expressed.

Assay for lysozyme sensitivity

To compare the lysozyme sensitivity of the parental strain E1162, the mutant strains and *in trans* complemented strains, overnight cell cultures were diluted 100 fold in fresh BHI and grown to OD_{660} 0.5. Two ml of the cell cultures were harvested by centrifugation. The pellets were resuspended in 1 ml phosphate buffered saline (PBS; NaCl 137 mM; 2.7 mM KCl; 10 mM Na_2HPO_4; 2 mM KH_2PO_4; pH 7.4) as negative control and in 1 ml PBS containing 0.5 mg ml^{-1} lysozyme. After a 30-minute incubation at 37°C, cells were washed with PBS and resuspended in 1 ml of PBS. Survival of the strains was determined following serial dilution and plating on BHI agar plates. The experiment was performed in triplicate and statistical analysis of the data was performed using a two-tailed Student's *t*-test.

Determination of D,D-carboxypeptidase activity in enterococcal extracts

The enzymatic activities in the enterococcal extracts of wild-type, Δ*ddcP* and Δ*ddcP+ddcP* were assayed as described previously with slight modifications [59], [60]. In short, strains were grown until an OD_{600} of 0.7. Bacteria were then harvested by centrifugation and lysed by treatment with lysozyme at 37°C for 1 hour followed by sonication. The membrane fraction was then pelleted by ultracentrifugation (100,000 *g*, 45 min). The supernatant (cytoplasmic fractions) was collected and the pellet (membrane fractions) was resuspended in 0.1 M phosphate buffer (pH 7.0) and both fractions were assayed for D,D-carboxypeptidase activity [59], [60]. The amounts of D-Ala released from the pentapeptide (Ala-D-γ-Glu-Lys-D-Ala-D-Ala, Sigma-Aldrich) by D,D-carboxypeptidases were determined by using D-amino acid oxidase and horseradish peroxidase in a colorimetric assay.

Microarray data accession numbers

The microarray data generated in this study have been deposited in the ArrayExpress database (http://www.ebi.ac.uk/arrayexpress) under accession numbers E-MEXP-3501 for the M-TraM screening for ampicillin resistance determinants, E-MEXP-3502 for the assay of the reproducibility of the M-TraM procedure and E-MEXP-3564 for the transcriptome analysis data.

Acknowledgments

We thank the following researchers for kindly providing us with plasmids: David Lampe for pMMorf, Marta Perego for pAW068, Ralph Bertram for pRAB1, and Chris J. Kristich for pCJK55 and pCJK72.

Author Contributions

Conceived and designed the experiments: XZ WvS. Performed the experiments: XZ DB AK FLP. Analyzed the data: XZ WvS. Wrote the paper: XZ RJLW MJMB WvS.

References

1. Hidron AI, Edwards JR, Patel J, Horan TC, Sievert DM, et al. (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.
2. Willems RJ, van Schaik W (2009) Transition of *Enterococcus faecium* from commensal organism to nosocomial pathogen. Future Microbiol 4: 1125–1135.
3. Ammerlaan HS, Troelstra A, Kruitwagen CL, Kluytmans JA, Bonten MJ (2009) Quantifying changes in incidences of nosocomial bacteraemia caused by antibiotic-susceptible and antibiotic-resistant pathogens. J Antimicrob Chemother 63: 1064–1070.
4. Top J, Willems R, Blok H, de Regt M, Jalink K, et al. (2007) Ecological replacement of *Enterococcus faecalis* by multiresistant clonal complex 17 *Enterococcus faecium*. Clin Microbiol Infect 13: 316–319.
5. Leavis HL, Bonten MJ, Willems RJ (2006) Identification of high-risk enterococcal clonal complexes: global dispersion and antibiotic resistance. Curr Opin Microbiol 9: 454–460.
6. van Schaik W, Willems RJ (2010) Genome-based insights into the evolution of enterococci. Clin Microbiol Infect 16: 527–532.
7. Arias CA, Contreras GA, Murray BE (2010) Management of multidrug-resistant enterococcal infections. Clin Microbiol Infect 16: 555–562.
8. Murdoch DR, Mirrett S, Harrell LJ, Monahan JS, Reller LB (2002) Sequential emergence of antibiotic resistance in enterococcal bloodstream isolates over 25 years. Antimicrob Agents Chemother 46: 3676–3678.
9. Grayson ML, Eliopoulos GM, Wennersten CB, Ruoff KL, De Girolami PC, et al. (1991) Increasing resistance to beta-lactam antibiotics among clinical isolates of *Enterococcus faecium*: a 22-year review at one institution. Antimicrob Agents Chemother 35: 2180–2184.
10. Murray BE (2000) Vancomycin-resistant enterococcal infections. N Engl J Med 342: 710–721.
11. Werner G, Coque TM, Hammerum AM, Hope R, Hryniewicz W, et al. (2008) Emergence and spread of vancomycin resistance among enterococci in Europe. Euro Surveill 13: pii: 19046.
12. Montero CI, Stock F, Murray PR (2008) Mechanisms of resistance to daptomycin in *Enterococcus faecium*. Antimicrob Agents Chemother 52: 1167–1170.
13. Scheetz MH, Knechtel SA, Malczynski M, Postelnick MJ, Qi C (2008) Increasing incidence of linezolid-intermediate or -resistant, vancomycin-resistant *Enterococcus faecium* strains parallels increasing linezolid consumption. Antimicrob Agents Chemother 52: 2256–2259.
14. Arias CA, Panesso D, McGrath DM, Qin X, Mojica MF, et al. (2011) Genetic basis for in vivo daptomycin resistance in enterococci. N Engl J Med 365: 892–900.
15. Murray BE (1990) The life and times of the Enterococcus. Clin Microbiol Rev 3: 46–65.
16. Fontana R, Grossato A, Rossi L, Cheng YR, Satta G (1985) Transition from resistance to hypersusceptibility to beta-lactam antibiotics associated with loss of a low-affinity penicillin-binding protein in a *Streptococcus faecium* mutant highly resistant to penicillin. Antimicrob Agents Chemother 28: 678–683.
17. Williamson R, le Bouguenec C, Gutmann L, Horaud T (1985) One or two low affinity penicillin-binding proteins may be responsible for the range of susceptibility of *Enterococcus faecium* to benzylpenicillin. J Gen Microbiol 131: 1933–1940.
18. Fontana R, Aldegheri M, Ligozzi M, Lopez H, Sucari A, et al. (1994) Overproduction of a low-affinity penicillin-binding protein and high-level ampicillin resistance in *Enterococcus faecium*. Antimicrob Agents Chemother 38: 1980–1983.
19. Galloway-Pena JR, Rice LB, Murray BE (2011) Analysis of PBP5 of Early U.S. Isolates of *Enterococcus faecium*: Sequence Variation Alone Does Not Explain Increasing Ampicillin Resistance over Time. Antimicrob Agents Chemother 55: 3272–3277.

20. Rice LB, Bellais S, Carias LL, Hutton-Thomas R, Bonomo RA, et al. (2004) Impact of specific *pbp5* mutations on expression of beta-lactam resistance in *Enterococcus faecium*. Antimicrob Agents Chemother 48: 3028–3032.
21. Mainardi JL, Legrand R, Arthur M, Schoot B, van Heijenoort J, et al. (2000) Novel mechanism of beta-lactam resistance due to bypass of DD-transpeptidation in *Enterococcus faecium*. J Biol Chem 275: 16490–16496.
22. Mainardi JL, Morel V, Fourgeaud M, Cremniter J, Blanot D, et al. (2002) Balance between two transpeptidation mechanisms determines the expression of beta-lactam resistance in *Enterococcus faecium*. J Biol Chem 277: 35801–35807.
23. Sacco E, Hugonnet JE, Josseaume N, Cremniter J, Dubost L, et al. (2010) Activation of the L,D-transpeptidation peptidoglycan cross-linking pathway by a metallo-D,D-carboxypeptidase in *Enterococcus faecium*. Mol Microbiol 75: 874–885.
24. van Schaik W, Top J, Riley DR, Boekhorst J, Vrijenhoek JE, et al. (2010) Pyrosequencing-based comparative genome analysis of the nosocomial pathogen *Enterococcus faecium* and identification of a large transferable pathogenicity island. BMC Genomics 11: 239.
25. Walsh CT (1989) Enzymes in the D-alanine branch of bacterial cell wall peptidoglycan assembly. J Biol Chem 264: 2393–2396.
26. Heikens E, Bonten MJ, Willems RJ (2007) Enterococcal surface protein Esp is important for biofilm formation of *Enterococcus faecium* E1162. J Bacteriol 189: 8233–8240.
27. Heikens E, Singh KV, Jacques-Palaz KD, van Luit-Asbroek M, Oostdijk EA, et al. (2011) Contribution of the enterococcal surface protein Esp to pathogenesis of *Enterococcus faecium* endocarditis. Microbes Infect 13: 1185–1190.
28. Leendertse M, Heikens E, Wijnands LM, van Luit-Asbroek M, Teske GJ, et al. (2009) Enterococcal surface protein transiently aggravates *Enterococcus faecium*-induced urinary tract infection in mice. J Infect Dis 200: 1162–1165.
29. Lam MM, Seemann T, Bulach DM, Gladman SL, Chen H, et al. (2012) Comparative Analysis of the First Complete *Enterococcus faecium* Genome. J Bacteriol 194: 2334–2341.
30. Sauvage E, Kerff F, Terrak M, Ayala JA, Charlier P (2008) The penicillin-binding proteins: structure and role in peptidoglycan biosynthesis. FEMS Microbiol Rev 32: 234–258.
31. Lei XH, Fiedler F, Lan Z, Kathariou S (2001) A novel serotype-specific gene cassette (*gltA-gltB*) is required for expression of teichoic acid-associated surface antigens in *Listeria monocytogenes* of serotype 4b. J Bacteriol 183: 1133–1139.
32. Horsburgh GJ, Atrih A, Williamson MP, Foster SJ (2003) LytG of *Bacillus subtilis* is a novel peptidoglycan hydrolase: the major active glucosaminidase. Biochemistry 42: 257–264.
33. Magnet S, Arbeloa A, Mainardi JL, Hugonnet JE, Fourgeaud M, et al. (2007) Specificity of L,D-transpeptidases from gram-positive bacteria producing different peptidoglycan chemotypes. J Biol Chem 282: 13151–13159.
34. Zhang X, Vrijenhoek JE, Bonten MJ, Willems RJ, Van Schaik W (2011) A genetic element present on megaplasmids allows *Enterococcus faecium* to use raffinose as carbon source. Environ Microbiol 13: 518–528.
35. Callewaert L, Michiels CW (2010) Lysozymes in the animal kingdom. J Biosci 35: 127–160.
36. Kang Y, Durfee T, Glasner JD, Qiu Y, Frisch D, et al. (2004) Systematic mutagenesis of the *Escherichia coli* genome. J Bacteriol 186: 4921–4930.
37. Christen B, Abeliuk E, Collier JM, Kalogeraki VS, Passarelli B, et al. (2011) The essential genome of a bacterium. Mol Syst Biol 7: 528.
38. Galloway-Pena JR, Nallapareddy SR, Arias CA, Eliopoulos GM, Murray BE (2009) Analysis of clonality and antibiotic resistance among early clinical isolates of *Enterococcus faecium* in the United States. J Infect Dis 200: 1566–1573.
39. Sauer B (1987) Functional expression of the *cre-lox* site-specific recombination system in the yeast *Saccharomyces cerevisiae*. Mol Cell Biol 7: 2087–2096.
40. Ghosh AS, Chowdhury C, Nelson DE (2008) Physiological functions of D-alanine carboxypeptidases in *Escherichia coli*. Trends Microbiol 16: 309–317.

41. Sarkar SK, Chowdhury C, Ghosh AS (2010) Deletion of penicillin-binding protein 5 (PBP5) sensitises *Escherichia coli* cells to beta-lactam agents. Int J Antimicrob Agents 35: 244–249.
42. Memmi G, Filipe SR, Pinho MG, Fu Z, Cheung A (2008) *Staphylococcus aureus* PBP4 is essential for beta-lactam resistance in community-acquired methicillin-resistant strains. Antimicrob Agents Chemother 52: 3955–3966.
43. Campbell J, Singh AK, Santa Maria JP Jr, Kim Y, Brown S, et al. (2011) Synthetic lethal compound combinations reveal a fundamental connection between wall teichoic acid and peptidoglycan biosyntheses in *Staphylococcus aureus*. ACS Chem Biol 6: 106–116.
44. Leavis HL, Willems RJ, van Wamel WJ, Schuren FH, Caspers MP, et al. (2007) Insertion sequence-driven diversification creates a globally dispersed emerging multiresistant subspecies of *E. faecium*. PLoS Pathog 3: e7.
45. Bush LM, Calmon J, Cherney CL, Wendeler M, Pitsakis P, et al. (1989) High-level penicillin resistance among isolates of enterococci. Implications for treatment of enterococcal infections. Ann Intern Med 110: 515–520.
46. Leenhouts K, Buist G, Bolhuis A, ten Berge A, Kiel J, et al. (1996) A general system for generating unlabelled gene replacements in bacterial chromosomes. Mol Gen Genet 253: 217–224.
47. Kristich CJ, Nguyen VT, Le T, Barnes AM, Grindle S, et al. (2008) Development and use of an efficient system for random *mariner* transposon mutagenesis to identify novel genetic determinants of biofilm formation in the core *Enterococcus faecalis* genome. Appl Environ Microbiol 74: 3377–3386.
48. Nallapareddy SR, Singh KV, Murray BE (2006) Construction of improved temperature-sensitive and mobilizable vectors and their use for constructing mutations in the adhesin-encoding acm gene of poorly transformable clinical *Enterococcus faecium* strains. Appl Environ Microbiol 72: 334–345.
49. Akerley BJ, Lampe DJ (2002) Analysis of gene function in bacterial pathogens by GAMBIT. Methods Enzymol 358: 100–108.
50. Lampe DJ, Akerley BJ, Rubin EJ, Mekalanos JJ, Robertson HM (1999) Hyperactive transposase mutants of the *Himar1 mariner* transposon. Proc Natl Acad Sci U S A 96: 11428–11433.
51. Arthur M, Depardieu F, Snaith HA, Reynolds PE, Courvalin P (1994) Contribution of VanY D,D-carboxypeptidase to glycopeptide resistance in *Enterococcus faecalis* by hydrolysis of peptidoglycan precursors. Antimicrob Agents Chemother 38: 1899–1903.
52. Wilson AC, Perego M, Hoch JA (2007) New transposon delivery plasmids for insertional mutagenesis in *Bacillus anthracis*. J Microbiol Methods 71: 332–335.
53. Chaudhuri RR, Allen AG, Owen PJ, Shalom G, Stone K, et al. (2009) Comprehensive identification of essential *Staphylococcus aureus* genes using Transposon-Mediated Differential Hybridisation (TMDH). BMC Genomics 10: 291.
54. Bijlsma JJ, Burghout P, Kloosterman TG, Bootsma HJ, de Jong A, et al. (2007) Development of genomic array footprinting for identification of conditionally essential genes in *Streptococcus pneumoniae*. Appl Environ Microbiol 73: 1514–1524.
55. Baldi P, Long AD (2001) A Bayesian framework for the analysis of microarray expression data: regularized *t* -test and statistical inferences of gene changes. Bioinformatics 17: 509–519.
56. Leibig M, Krismer B, Kolb M, Friede A, Gotz F, et al. (2008) Marker removal in staphylococci via Cre recombinase and different *lox* sites. Appl Environ Microbiol 74: 1316–1323.
57. Bryan EM, Bae T, Kleerebezem M, Dunny GM (2000) Improved vectors for nisin-controlled expression in gram-positive bacteria. Plasmid 44: 183–190.
58. Andrews JM (2001) Determination of minimum inhibitory concentrations. J Antimicrob Chemother 48: Suppl 15–16.
59. Lebreton F, Depardieu F, Bourdon N, Fines-Guyon M, Berger P, et al. (2011) D-Ala-D-Ser VanN-Type Transferable Vancomycin Resistance in *Enterococcus faecium*. Antimicrob Agents Chemother 55: 4606–4612.

60. Arthur M, Depardieu F, Reynolds P, Courvalin P (1996) Quantitative analysis of the metabolism of soluble cytoplasmic peptidoglycan precursors of glycopeptide-resistant enterococci. Mol Microbiol 21: 33–44.

Supplementary data

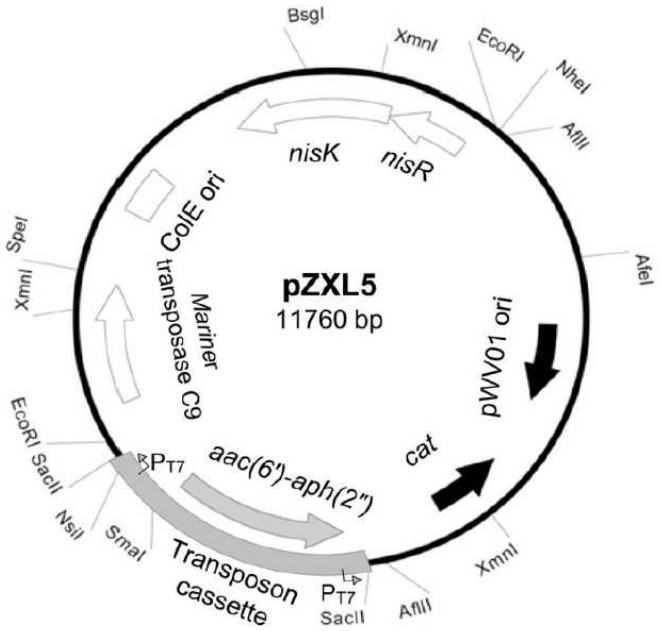


Figure S1. Map of pZXL5. This plasmid contains a Gram-positive thermo-sensitive pWVO1 replicon and the chloramphenicol acetyltransferase (*cat*) gene from pAW068, a nisin inducible *mariner* transposase (including *nisA* promoter, the transposase, *nisK* and *nisR*) and a ColE1 replicon from pCJK55, and a *mariner* transposon carrying the gentamicin resistance gene *acc(6')-aph(2")* with two outward-facing T7 promoters. Arrows indicate the direction of transcription. The T7 promoters (P_{T7}) and unique or relevant restriction sites are shown.

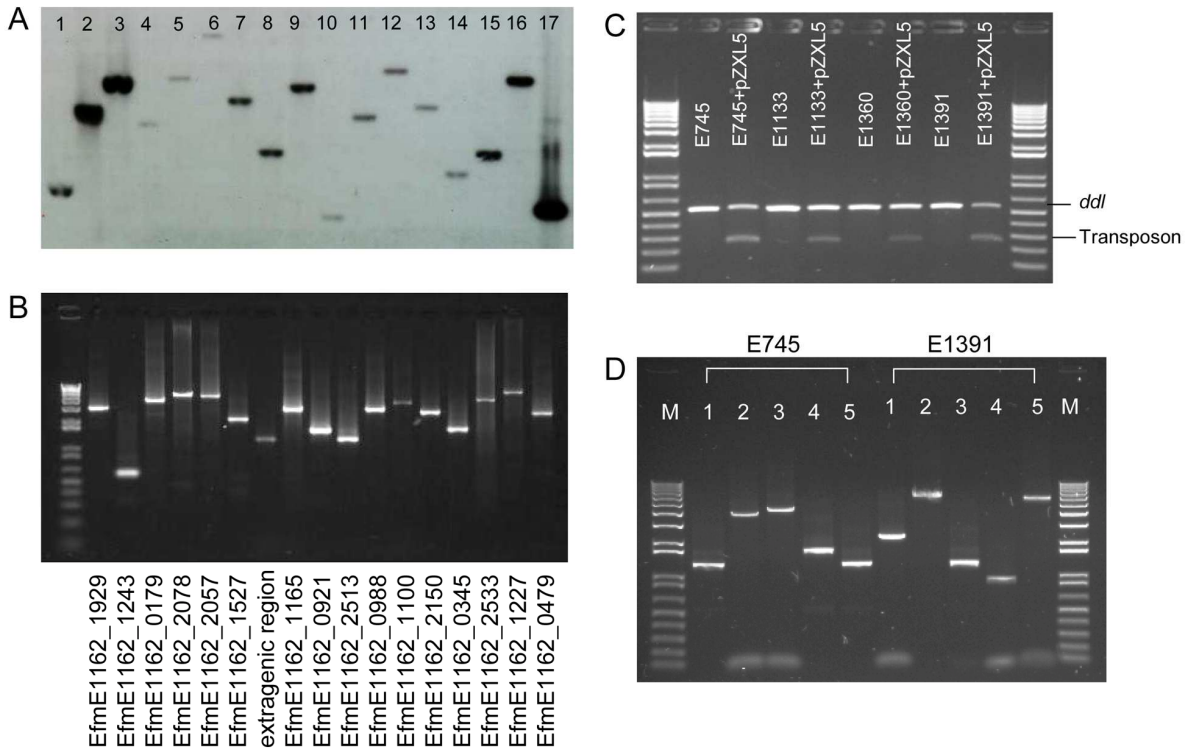


Figure S2. **Evaluation of the transposon mutant library in *E. faecium* E1162 and electroporation of pZXL5 to four other clinical *E. faecium* isolates and subsequent generation of transposon mutant libraries in strains E745 and E1391.** (A) Southern blot analysis of 17 randomly selected *E. faecium* transposon insertion mutants (lane 1 to 17) from the mutant library. Genomic DNA was digested with HaeIII and BamHI, and hybridized to an ECL-labeled probe specific for the transposon. (B) Inverse PCR and sequencing analysis of 17 randomly selected *E. faecium* transposon insertion mutants (lane 2 to 18). Genomic DNA was digested with HaeIII, self-ligated and the transposon-chromosome junction was PCR amplified using transposon-specific primers. The PCR products were sequenced and matched to the E1162 genome sequence. The transposon insertion loci of the mutants are indicated. (C) Multiplex PCR verification of the electroporation of pZXL5 into four clinical *E. faecium* strains. The expected sizes of the PCR products of *ddl* (housekeeping gene in the *E. faecium* genome) and the gentamicin resistance cassette on the *mariner* transposon (in pZXL5) are indicated. The primers used for the multiplex PCR are listed in Table S3: ftp_ddl and ddl_1 were used for *ddl*, genta_in_F and genta_in_R were used for pZXL5. (D) Inverse PCR and sequencing analysis of randomly selected transposon insertion mutants from the libraries generated with E745 and E1391. Five mutants were selected from each library. Inverse PCR was performed as described in (B).

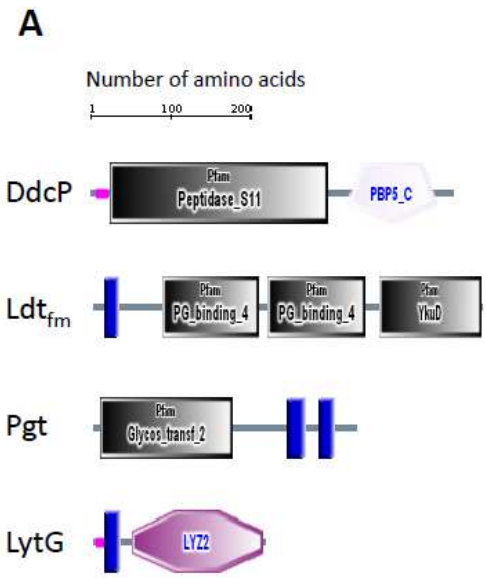
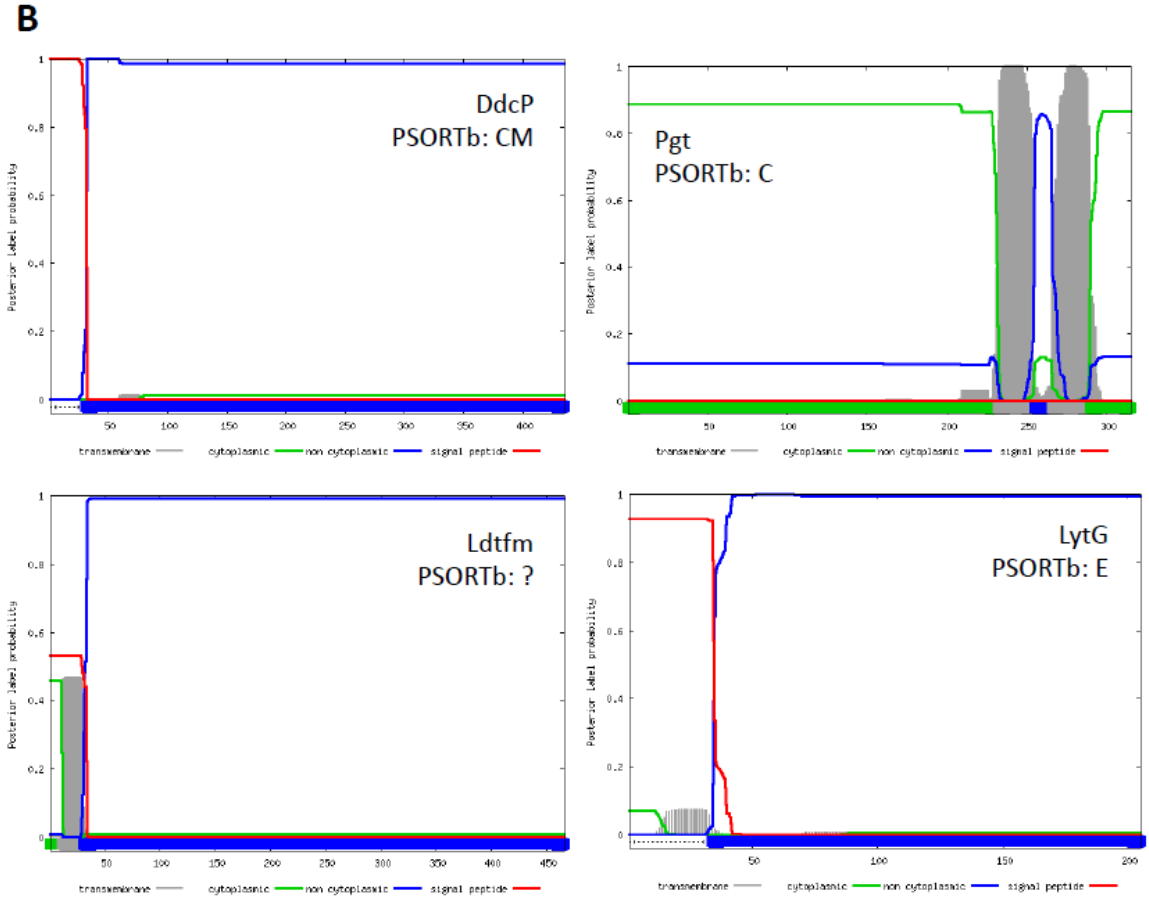


Figure S3. Predicted protein domain architecture and cellular localization of Ddcp, Ldt_{fm}, Pgt, and LytG. (A) Protein domain visualizations and the annotations of protein domains were made using SMART at http://smart.embl.de/. Blue horizontal bars indicate transmembrane regions. Pink stretches indicate regions of low complexity. Pfam domain Peptidase_S11 in DdcP is predicted to function as a serine peptidase with D-Ala-D-Ala carboxypeptidases activity. The PBP5_C domain in DdcP is homologous to the C-terminal domain of *E. coli* low-molecular weight penicillin-binding protein Pbp5, which has no known catalytic function. It could be involved in mediating interactions with other cell wall-synthesizing enzymes, thereby allowing the protein to be recruited to areas of active cell wall synthesis. Alternatively, it could function as a linker domain that positions the active site in the catalytic domain closer to the peptidoglycan layer. The two Pfam PG_binding_4 domains in Ldt_{fm} are predicted to act as a L,D-transpeptidase domain which can cross-link two peptidoglycan side-chains. The Pfam YkuD domain in Ldt_{fm} is frequently encountered in proteins with peptidoglycan-binding domains, but its function is unknown. The Pfam Glycos_transf_2 domain that is present in Pgt is also found in a diverse family of glycosyl transferases that transfer a sugar moiety from an activated nucleotide substrate to a range of substrates including teichoic acids. The Lyz2 domain that was identified in LytG is present in eubacterial enzymes that are distantly related to eukaryotic lysozymes.



(B) Cellular localization of the proteins was predicted by Phobius (http://phobius.sbc.su.se/) and PSORTb (http://www.psort.org/psortb/; CM: cytoplasmic membrane; C: cytoplasm; E: extracellular)

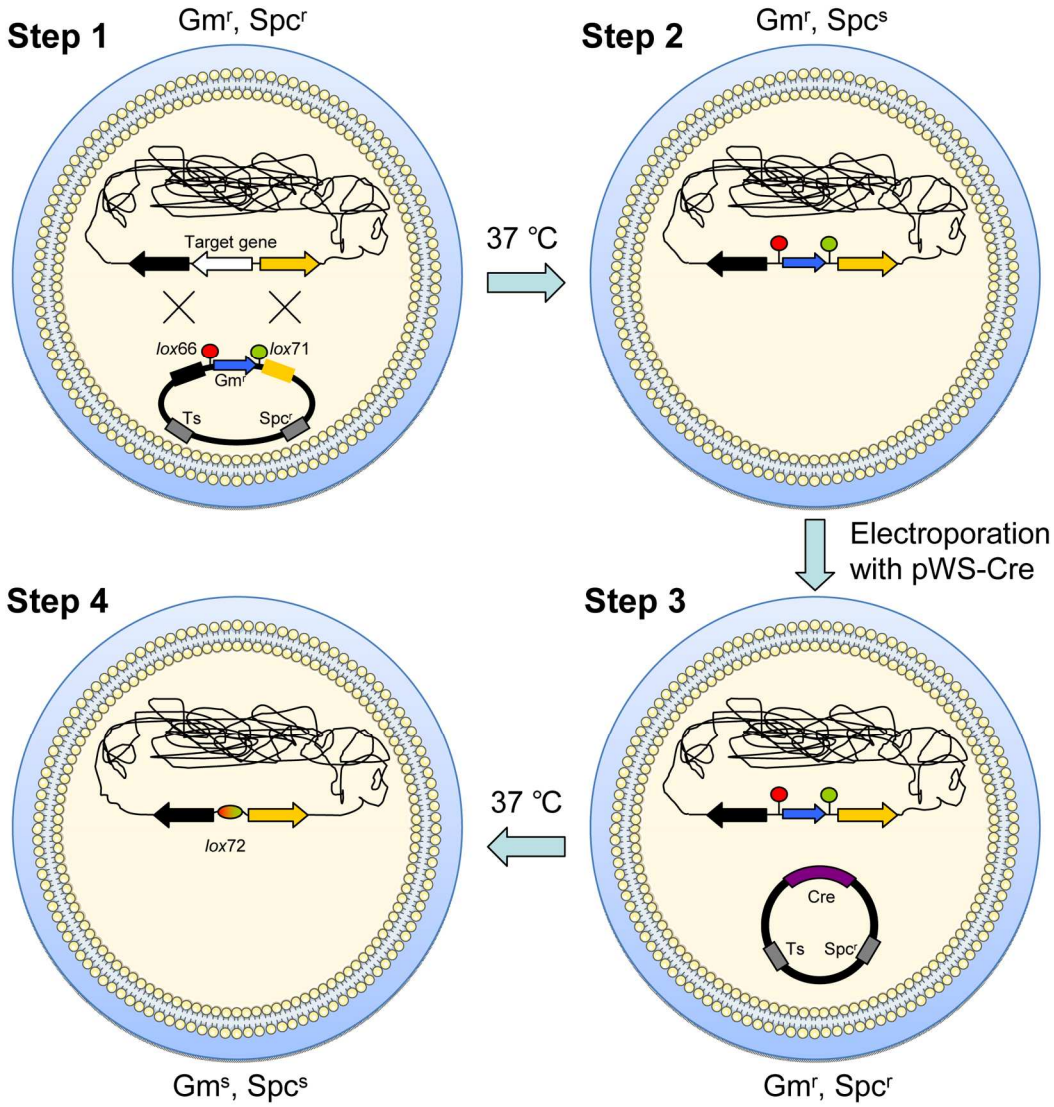


Figure S4. Schematic diagram of the Cre-*lox* recombination system for the construction of markerless mutant in *E. faecium*. A detailed description of the procedure is provided in the Materials and Methods section. Step 1: The gene replacement construct carrying *in vitro*-altered sequences (UpFlankingRegion-*lox*66-Gm$_r$-*lox*71-DownFlankingRegion) is introduced into E1162 by electroporation, and transformants are incubated at permissive temperature (30°C) for double crossover events. Step 2: The cells are passaged at a non-permissive temperature (37°C) for plasmid replication. Double-crossover integrants, are screened using agar plates supplemented with appropriate antibiotics. Step 3: Subsequently, the thermosensitive plasmid pWS3-Cre is electrotransformed into the marked mutants, and the *lox*66-Gm$_r$-*lox*71 cassette is removed from the chromosome by the Cre-mediated excision during overnight culture of the transformants at 30°C. Step 4: Subsequent overnight culturing of the cells at 37°C leads to the loss of pWS3-Cre, resulting in a markerless double crossover mutant in which the gene is replaced by a lox72 site. (Gm^r: gentamicin resistant; Gm^s: gentamicin susceptible; Spc^r: spectinomycin resistant; Spc^s: spectinomycin susceptible; Ts: thermosensitive replicon).

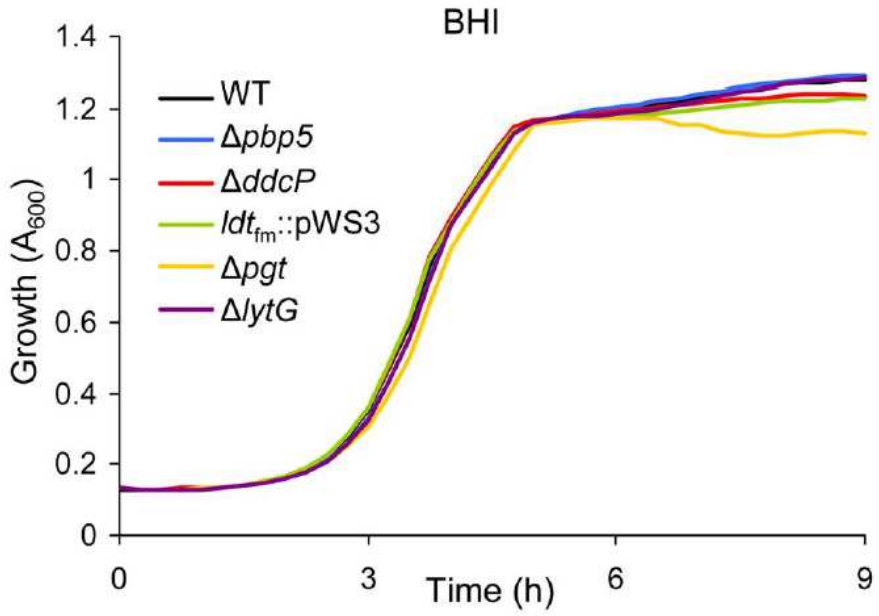


Figure S5. Growth curves of targeted mutants and wild-type *E. faecium* E1162 in BHI medium without added antibiotics. Overnight cultures of mutants and wild-type strain were inoculated at an initial cell density of OD_{660} 0.0025 in BHI and grown at 37°C with shaking in the Bioscreen C instrument. Growth curves represent mean data from three independent experiments.

Table S1: Strains and plasmids used in this study.

Strain or plasmid	Relevant characteristic(s)	Source or reference
E. faecium		
E1162	Clinical isolate (bloodstream infection), isolated in France, 1996	[1]
E745	Strain isolated in 2000 in The Netherlands from patient feces during a VRE outbreak; resistant to vancomycin	[2]
E1133	Strain isolated in 2002 in the United States of America from patient feces during a VRE outbreak; resistant to vancomycin	[2]
E1360	Clinical isolate (2001, United States of America); resistant to vancomycin	[2]
E1391	Clinical isolate (2000, United Kingdom)	[2]
Δ*ddcP*	Markerless deletion mutant of *ddcP* of E1162	This study
ldt_{fm}::pWS3	Single-crossover insertional mutant of ldt_{fm} of E1162	This study
Δ*pgt*	Markerless deletion mutant of *pgt* of E1162	This study
Δ*lytG*	Markerless deletion mutant of *lytG* of E1162	This study
Δ*pbp5*	Markerless deletion mutant of *pbp5* of E1162	This study
Δ*ddcP+ddcP*	Complementation strain of Δ*ddcP*; Δ*ddcP* harboring pMSP3535-*ddcP*	This study
ldt_{fm}::pWS3+ldt_{fm}	Complementation strain of ldt_{fm}::pWS3; ldt_{fm}::pWS3 harboring pMSP3535-ldt_{fm}	This study
Δ*lytG+lytG*	Complementation strain of Δ*lytG*; Δ*lytG* harboring pMSP3535-*lytG*	This study
Δ*pbp5+pbp5*	Complementation strain of Δ*pbp5*; Δ*pbp5* harboring pMSP3535-*pbp5*	This study
E. coli		
DH5α	*E. coli* host strain for routine cloning	Invitrogen
EC1000	MC1000 *glgB*::*repA*	[3]
Plasmids		
pMMOrf	Contains 5' and 3' ITR from *Himar1*	[4]
pAT392	Shuttle expression vector (Gen^r Spc^r $oriR_{pUC}$ $oriR_{pAMβ1}$ $oriT_{RK2}$ P_2)	[5]
pCJK55	Contains *mariner* transposase C9 induced by nisin	[6]
pTEX5500ts	Shuttle plasmid, gram-positive thermosensitive origin of replication; Chl^r, Gent^r	[7]
pCJK72	Mobilizable delivery plasmid carrying EfaMarTn transposable element; gram-positive *lacZ* gene	[6]
pAW068	Transposon delivery plasmid carrying *Himar1* transpose; Gram-positive thermosensitive origin; Chl^r, Spc^r	[8]
pWS3	Gram-positive thermosensitive origin; Spc^r	[9]
pZXL1	Gram-positive *lacZ* gene form pCJK72 cloned into pTEX5500ts	This study
pZXL2	Contains nisin induced *mariner* transposase C9 from pCJK55 and gram-positive thermosensitive origin of replication, Chl^r cassette from pZXL1	This study
pZXL3	A Gent^r *mariner* transposon with two outward-facing T7 promoters cloned into pZXL2	This study
pZXL5	Transposon delivery plasmid carrying nisin induced *mariner* transposase C9 and Gent^r *mariner* transposon with two outward-facing T7 promoters from pZXL3; Gram-positive thermosensitive origin of replication and Chl^r cassette from pAW068	This study
pGEM-T Easy	Cloning vector	Promega
pGEM-ITR	ITR of Himar1 *from pMMorf* ligated into pGEM-T Easy	This study
pGEM-Tn1	pGEM-T Easy carrying a Gent^r *mariner* transposon with two outward-facing T7 promoters	This study
pRAB1	*cat bla*, P_{pagA}-*cre*; expression of *cre*	[10]
pWS3-Cre	pWS3 derivative expressing Cre in *E. faecium*	This study
pMSP3535	pAMβ1 (from pIL252), ColE1 replicon, Ermr, nisRK, PnisA	[11]
pMSP3535-*ddcP*	Complementation plasmid for *ddcP*; pMSP3535 carrying gene *ddcP*	This study
pMSP3535-ldt_{fm}	Complementation plasmid for ldt_{fm}; pMSP3535 carrying gene ldt_{fm}	This study
pMSP3535-*lytG*	Complementation plasmid for *lytG*; pMSP3535 carrying gene *lytG*	This study
pMSP3535-*pbp5*	Complementation plasmid for *pbp5*; pMSP3535 carrying gene *pbp5*	This study

Table S2: Comparative transcriptome analysis of *E. faecium*E1162 during mid-exponential growth in the presence (20 µg ml^{-1}) and absence of ampicillin.

LocusTag[a]	Accession code	Annotation	Expression ratio BHI+amp/BHI
EfmE1162_0046	ZP_06675891	PyrRbifunctionalprotein	1.9
EfmE1162_0047	ZP_06675892	uracilpermease	2.0
EfmE1162_0048	ZP_06675893	aspartatecarbamoyltransferase	1.9
EfmE1162_0049	ZP_06675894	Dihydroorotase	2.1
EfmE1162_0050	ZP_06675895	carbamoyl-phosphate synthase, small subunit	1.9
EfmE1162_0051	ZP_06675896	carbamoyl-phosphate synthase, large subunit	1.9
EfmE1162_0052	ZP_06675897	dihydroorotate dehydrogenase electron transfer subunit	2.0
EfmE1162_0053	ZP_06675898	dihydroorotate dehydrogenase B, catalytic subunit	1.9
EfmE1162_0054	ZP_06675899	orotidine 5'-phosphate decarboxylase	2.0
EfmE1162_0055	ZP_06675900	orotatephosphoribosyltransferase	2.0
EfmE1162_0229	ZP_06676074	conservedhypotheticalprotein	1.8
EfmE1162_0230	ZP_06676075	accessory gene regulator protein B, putative	1.7
EfmE1162_0231	ZP_06676076	hypotheticalprotein	1.8
EfmE1162_0443	ZP_06676288	hypotheticalprotein	1.7
EfmE1162_0683	ZP_06676528	peptidoglycan-binding LysM	0.5
EfmE1162_2616	ZP_06678315	aggregation promoting protein	0.4

[a] All genes exhibiting significantly different expression during mid-exponential growth (OD_{660} = 0.3) in BHI and BHI with 20 µg ml^{-1} ampicillin for all four probes per gene (Bayesian P-value <0.001) in microarray hybridizations are included.

Table S3: Primers used in this study[a].

Primer	Sequence
pAT392_EheI_T7_genta_F	5′-GGTGGGCGCC**CTCCCTATAGTGAGTCGTATTA**AAACCCAGCGAACCATTTGAGGTGATAGGTAAG-3′
pAT392_EheI_T7_genta_R	5′-GCTCGGCGCC**CTCCCTATAGTGAGTCGTATTA**CAGGAAACAGCTATGACCATGATTACGCCAAGC-3′
pMMOrf_SacII_ITR	5′-TCCCCGCGGTAACAGGTTGGCTGATAAGTCCCCGGTCT-3′
pZXL3_BfrI_tn_F	5′-ACGCCTTAAGAAAGTCGAAGGGGGTTTTT-3′
pZXL3_tn_R	5′-TTCACGAACGAAAATCAAGC-3′
pAW068_BfrI_cm_ori_F	5′-ACGCCTTAAGCTGGCGTAATAGCGAAGAGG-3′
pAW068_BfrI_cm_ori_R	5′-ACGCCTTAAGACGCTCAGTGGAACGAAAAC-3′
pCJK72_PstI_lacZ_F	5′-AACTGCAGGTGAGCTTGGACTAGAAAAAAACTTCACAA-3′
pCJK72_KpnI_lacZ_R	5′-GGGGTACCTTATTTTTGACACCAGACCAACTGGTAATGG-3′
pZXL1_EcoRI_cm_ori_F	5′-GGAATTCTGTACAATTGCTAGCGTACGG-3′
pZXL1_EcoRI_cm_ori_R	5′-GGAATTCGACTTCTCTACTCCTTTTAATC-3′
pCJK55_EcoRI_tps_F	5′-GGAATTCATGCATGCGTCGACTCTAG-3′
pCJK55_EcoRI_tps_R	5′-GGAATTCAAGCTTCGGCCGATCGATG-3′
genta_probe_F	5′-ACAGAGCCTTGGGAAGATGA-3′
genta_probe_R	5′-TGCCTTAACATTTGTGGCATT-3′
IPCR_HaeIII_R	5′-CCCCCTGAAATCCTTACAT-3′
IPCR_HaeIII_F	5′-AAACAGGAATTTATCGAAAATGGT-3′
ftp_tn	5′-AAACAGGAATTTATCGAAAATGGT-3′
ftp_ddl	5′-AAAAAGAAATCGCACCG-3′
ftp_nox	5′-TACGGTTGTAGTGACTATTTTTTC-3′
ftp_esp	5′-GGACTTGCATTAGCAAAATC-3′
IPCR_AluI_F	5′-GTTCGCTGGGTTTTAATACGACTCACT-3′
IPCR_AluI_R	5′-CAGCGAACCATTTGAGGTGATAGGT-3′
pAT392_EcoRI_lox66_genta_F	5′-GAGGGAATTC**TACCGTTCGTATAGCATACATTATACGAAGTTAT**GATAAACCCAGCGAACCATTTGAGG-3′
pAT392_EcoRI_lox71_genta_R	5′-CTCCGAATTC**TACCGTTCGTATAATGTATGCTATACGAAGTTAT**TCAATCTTTATAAGTCCTTTTATAA-3′
Inst_EcoRI_pbp5_R	5′-GGAATTCACCTGTTGTCCCACGAAGAT-3′
Inst_XhoI_pbp5_F	5′-CCGCTCGAGATCAAAGCGATTGCTTCCTC-3′
pbp5_check	5′-ACCGTCTGCATCTGTAATGCT -3′
Inst_EcoRI_ddcP_R	5′-GGAATTCAGGCTCCTGCAAGTTCAGTC-3′
Inst_XhoI_ddcP_F	5′-CCGCTCGAGAAAAATCTCTGGCAGCGAAG-3′
ddcP_check	5′-GTGATCGTACCGACGAAACA -3′
Inst_EcoRI_ldt_{fm}_R	5′-GGAATTCTGTTTGGATCGTCCAACTGT-3′
Inst_XhoI_ldt_{fm}_F	5′-CCGCTCGAGCATCTGAAGACAGCTCACTGAAA-3′
ldt_{fm}_check	5′-CCACATATGCTGGTTCTCCA-3′
check_pWS3	5′-GGGGATTTTATGCGTGAGAA-3′
dele_XhoI_pbp5_dn_R	5′-CCGCTCGAGTGACGCTTGTAGCGATTTTG-3′
dele_EcoRI_pbp5_dn_F	5′-GACCACACAAGAAGCAGGAATTCTGATTCAGCAACCAAACGAG-3′
dele_EcoRI_pbp5_up_R	5′-CAGAATTCCTGCTTCTTGTGTGGTCAGG-3′
dele_XmaI_pbp5_up_F	5′-CCCCCCGGGAAAAATCGAACAGGCGCTTA-3′
check_pbp5_dn	5′-GCTGGGATAGCTGTCAGTCA-3′
check_pbp5_up	5′-GGAATGACAAGCAAGAGAAGG-3′
dele_XhoI_ddcP_up_F	5′-CCGCTCGAGCCACAACACTATTTTTCCATACAA-3′
dele_EcoRI_ddcP_up_R	5′-GATCGAATTCGGTAGAAAAGAAGCAAGGCAAA-3′
dele_EcoRI_ddcP_dn_F	5′-GCTTCTTTTCTACCGAATTCGATCACCTTGGCAGAAGAT-3′
dele_XmaI_ddcP_dn_R	5′-CCCCCCGGGTGAAAAATAGGAAAAGGCAAAGA-3′
check_ddcP_up	5′-GCCATAAACGTACCCCTCCT-3′
check_ddcP_dn	5′-GATCGTGATCAAGGCAATCA-3′
dele_XhoI_pgt_dn_R	5′-CCGCTCGAGCCACATTACGTACCAATGGATG-3′
dele_EcoRI_pgt_dn_F	5′-GAGGAAGCAACAATTTCAGAATTCCTTCTGCGATTTTAAGTACTGG-3′
dele_EcoRI_pgt_up_R	5′-CGCAGAAGGAATTCTGAAATTGTTGCTTCCTCATTG-3′
dele_XmaI_pgt up_F	5′-CCCCCCGGGTGCAATCGAAGGTTCTTG-3′
check_pgt_dn	5′-ACAAACAGTGGGGCAAGAAG-3′
check_pgt_up	5′-GGGTAGCTTCAACGATTTGG-3′
dele_XhoI_lytG_dn_R	5′-CCGCTCGAGTCAAAACGACCCCTAACGAA-3′
dele_EcoRI_lytG_dn_F	5′-GGAAAAAGCAAAAGCGAATTCTTGCAAGATGCAGGTTATGC-3′
dele_EcoRI_lytG_up_R	5′-GCAAGAATTCGCTTTTGCTTTTTCCACTT-3′
dele_XmaI_lytG_up_F	5′-CCCCCCGGGTGGTCTCAGCTTACCGGTCT-3′
check_lytG_dn	5′-TTTCGACGTTACCGTACTGAT-3′
check_lytG_up	5′-CGCATTGCTCCAGAATATGA-3′
comp_BamHI_ddcP_F	5′-CGGGATCCCGCTTCCCGCTTATTAGATG-3′
comp_PstI_ddcP_R	5′-AACTGCAGTTTTCGCCAATAAACGATGG-3′

comp_BamHI_pbp5_F	5′-CGGGATCCGGAATGACAAGCAAGAGAAGG-3′
comp_XhoI_pbp5_R	5′-CCGCTCGAGTCATTTTACAAATTGGACAGCAA-3′
comp_BamHI_ldt$_{fm}$_F	5′-CGGGATCCAAGTGACAGCGAGCAAGACC-3′
comp_PstI_ldt$_{fm}$_R	5′-AACTGCAGGCTGCTGAAATGTAAGTAGCACA-3′
comp_BamHI_pgt_F	5′-CGGGATCCAGAACGAATACGCACCCACT-3′
comp_PstI_pgt_R	5′-AACTGCAGTGCATCAAGCAAACGAATTT-3′
comp_BamHI_lytG_F	5′-CGGGATCCTGAACCTTGGCACTTACGTT-3′
comp_PstI_lytG_R	5′-AACTGCAGCGAACGGATCAATTTTCCAA-3′
genta_in_F	5′-CGGGAATTCCAGATTGCCTTGAATATATTGAC-3′
genta_in_R	5′-GACGGATCCTTATTTATCACCTTTTTCATAATCA-3′
ddl_1	5′-GAGACATTGAATATGCCTTATG-3′

[a] The restriction sites are underlined, and the sequences of the T7 promoter, the *lox66* and *lox71* sites are in bold.

Chapter 3

Identification of a genetic determinant in clinical *Enterococcus faecium* strains which contributes to intestinal colonization during antibiotic treatment

Xinglin Zhang[1], Janetta Top[1], Mark de Been[1], Damien Bierschenk[1], Malbert Rogers[1], Masja Leendertse[2], Marc J. M. Bonten[1], Tom van der Poll[2], Rob J. L. Willems[1], Willem van Schaik[1]

[1] Department of Medical Microbiology, University Medical Center Utrecht, Utrecht. The Netherlands
[2] Center for Infection and Immunity Amsterdam, Academic Medical Center, Amsterdam. The Netherlands

Published in The Journal of Infectious Diseases, 2013, doi: 10.1093/infdis/jit076

Abstract

Intestinal colonization by antibiotic-resistant *E. faecium* is the first step in a process that can lead to infections in hospitalized patients. By comparative genome analysis and subsequent PCR screening we identified a locus, encoding a putative phosphotransferase system (PTS), that is widespread in isolates from hospital outbreaks (84.2%) and non-outbreak clinical infections (66.0%), but is absent from human commensal isolates. A deletion in this PTS significantly impaired the ability of *E. faecium* to colonize the murine intestinal tract during antibiotic treatment. This is the first description of a determinant that contributes to intestinal colonization in clinical *E. faecium* strains.

Introduction

Enterococcus faecium is a Gram-positive bacterium that inhabits the intestinal tracts of humans and animals as a commensal. In the last two decades *E. faecium* has emerged as a multi-resistant nosocomial pathogen causing bacteremia, endocarditis, surgical wound, urinary tract, and device-related infections in debilitated patients [1]. *E. faecium* is intrinsically resistant to several clinically important antibiotics and has recently acquired resistance to many more antibiotics, including vancomycin, daptomycin and linezolid [1].
The establishment of high-level intestinal colonization by enterococci is a crucial step in a process that can finally lead towards nosocomial infections [1]. Once *E. faecium* has dominantly colonized the gastrointestinal tract, it may cross the intestinal lining and access the bloodstream causing bacteremia. It can also serve as reservoir for fecal contamination of the urinary tract and the skin, ultimately leading to catheter-related infections [1]. 16S rRNA-based studies have shown that enterococci are present at relatively low levels in the colon and feces of healthy humans. In hospitalized patients, however, enterococci can become one of the most prominent bacterial groups in the gut microbiota, presumably as a consequence of antibiotic therapy [2]. Antibiotic treatment alters intestinal microbial diversity by eliminating susceptible organisms, which provides an opportunity for multi-resistant enterococci to colonize these niches [3, 4]. Antibiotic therapy also indirectly enables colonization by multi-resistant enterococci, because antibiotic-mediated depletion of the microbiota affects the host intestinal immune defenses [1, 5].

However, neither the antibiotic resistance of enterococci nor the reduced immune status of the host can explain the notable feature that *E. faecium* isolates that are responsible for the majority of nosocomial infections are genetically distinct from isolates that are indigenous to humans and animals [6-8]. This observation implies that nosocomial *E. faecium* strains possess unique traits that enable them to benefit from the antibiotic-induced perturbations of the intestinal tract in hospitalized patients, thereby facilitating high-level intestinal colonization preceding infection. The specific genetic determinants of nosocomial *E. faecium* that promote gastrointestinal colonization during antibiotic therapy remain to be identified.

In this study, we describe a comparative genomic analysis which led to the identification of a gene cluster encoding a carbohydrate phosphotransferase system (PTS; here termed PTS^{clin}) that is highly enriched in clinical *E. faecium* isolates but is absent from commensal isolates. Deletion of the *pstD* gene, which is predicted to encode the enzyme IID subunit of PTS^{clin}, significantly impaired the ability of *E. faecium* to colonize the murine intestinal tract during antibiotic treatment.

Materials and Methods

Bacterial strains and growth conditions

Enterococcus faecium isolates (Supplementary Table 1) were routinely grown at 37°C on Trypticase soy agar II plates supplemented with 5% sheep blood (TSAB; Becton Dickinson, Alphen aan den Rijn, the Netherlands) or in brain–heart infusion (BHI) broth.

Comparative genomics analysis

Genome sequence information of all publicly available *E. faecium* strains (1 complete and 29 draft genomes, downloaded on May 2012) was retrieved from GenBank (ftp://ftp.ncbi.nih.gov/genomes/Bacteria; ftp://ftp.ncbi.nih.gov/genomes/Bacteria_DRAFT). To determine groups of orthologous proteins in *E. faecium*, we extracted all 86,257 annotated protein sequences from the 30 *E. faecium* genomes and used these as input for an all-vs-all sequence similarity search, using BLAST v2.2.24 [9] (default settings, except for: -F 'm S'; -e 1e-100; -z 86,257 [i.e. the number of sequences in the database]).Groups of orthologous proteins were predicted from the Blast output using OrthoMCL v2.0.2 [10, 11]. The *E. faecium* core proteome was defined by those orthologous groups that contained exactly 30 proteins, including exactly one protein from each of the 30 *E. faecium* genomes.

The core proteome was used to assess *E. faecium* phylogeny. Associated DNA sequences were extracted for each core protein, after which nucleotide alignments were built per orthologous group using MUSCLE v3.7 [12]. Prior to concatenating the individual alignments, alignment gaps were removed with TrimAl v1.2 [13]. The alignment contains 1,082,099 bp, including 75012 variable SNP positions. A bootstrapped Neighbour-Joining (NJ) tree was constructed from the concatenated alignment of the SNPs only, using CLUSTAL W v2.0.10 [14] (running 1000 bootstraps) (Figure 1*A*). Proteins that were exclusively present in ≥80% isolates of clade A were identified. TX0133B, TX0133C, TX0133a01 and TX0133a04 were excluded in this analysis because they represent the same strain as TX0133A [15].

Multiple genome alignment was performed using the comparative genomics tool VistaPoint [16] to further analyze the PTS^{clin} gene cluster. Conservation is represented as a curve that is calculated as a windowed-average (window size of 100 bp) identity score for the alignment. Peaks in the curve represent conservation between aligned sequences.

Multiplex PCR screening of PTS^{clin} and genotypic clustering of the E. faecium *isolates*

A total of 140 *E. faecium* isolates (Supplementary Table 1) were screened for the presence of PTS^{clin} gene cluster. The presence of this gene cluster was determined by multiplex PCR, in which a pair of primers specific for the *E. faecium* housekeeping gene *ddl* (5'-AAAAAGAAATCGCACCG-3' and 5'-GAGACATTGAATATGCCTTATG-3') and a pair of primers

specific for *ptsD* (5′-TATCAACGCGATCAAAACGA-3′ and 5′-CGTTCGCATACAGCTTTTCA-3′) were mixed together for each PCR reaction. The PCR products of *ddl* and *ptsD* were 560 bp and 242 bp, respectively. PCR amplification conditions using HotStarTaq and performed in a volume of 25-µl were as follows: initial denaturation at 95°C for 15 min, followed by 30 cycles of 30 s at 94°C, 30 s at 55°C and 72°C (the time depending on the size of the PCR product. All tested strains were positive for *ddl*. Analysis of MLST data was performed using the goeBURST algorithm [17].

Construction of ptsD deletion mutant

The construction of a double cross-over mutant in the *ptsD* gene (NCBI accession number: ZP_06677678) was performed as described previously [18]. *E. faecium* E1162 was used as the parental strain for the mutant construction. E1162 was isolated from a hospital-acquired bloodstream infection and its genome sequence has previously been determined [19]. Briefly, the 5′ flanking regions of *ptsD* was PCR amplified using primers 5′-CCGCTCGAGCGCAGTTGGAACGATTATTG-3′ and 5′-CCAAAAGAATTCAAGGAATGGTCGTCGAA-AAAT-3′, and the 3′ flanking region was PCR amplified using primers 5′-CGACGACCA-TTCCTTGAATTCTTTTGGAATGGTCGTCTTGG-3′ and 5′- CCCCCCGGGCGATGATCCGATTCTGAC-AA-3′. The two flanking regions were then fused together by fusion PCR (generating an EcoRI site between both fragments) and cloned into pWS3 [20], resulting in pDEL5a. Then, a gentamicin-resistance cassette which was flanked by *lox66*- and *lox71*-sites [18] was cloned into the EcoRI site that was generated between the 5′ and 3′ flanking regions in pDEL5a, resulting plasmid pDEL5b which were then electrotransformed into *E. faecium* E1162. The *ptsD* mutant, in which the *ptsD* gene was replaced by a gentamicin-resistance cassette, were obtained by growing the gentamicin-resistant transformants at appropriate temperatures supplemented with appropriate antibiotics as described previously [18]. The mutant was verified by PCR using primers 5′-GCCGTTGTTATGGCGTATCT-3′ and 5′-GCGATCGTGTCCTTATTTCC-3′.

Mice intestinal colonization

Intestinal colonization of mice by *E. faecium* strains was performed as described elsewhere [21-23] with some modifications. Specific pathogen-free 7-week-old male BALB/c mice (*n* = 8, Harlan Sprague-Dawley, Horst, The Netherlands) were decolonized for two days with subcutaneous injections of ceftriaxone (12 mg/ml, 100 µl per injection) two times a day, before inoculation of bacteria. Two days after the initiation of the antibiotic treatment, no *E. faecium* were identified from the fecal pellets and mice were inoculated orogastrically using a feeding tube with 1×10^7 CFUs in 300 µl PBS containing an equal proportion of wild-type E1162 and the *ptsD* mutant, which were individually grown overnight in BHI. Dilutions of the inoculum were plated on Slanetz Bartley agar (SBA; both wild-type E1162 and the mutant strain will grow on these plates) and SBA supplemented with 100 µg/ml gentamicin, to

quantify viable counts for the mutant. After inoculation the mice were allowed to drink *ad libitum* from water supplemented with 0.2 g/l cefoxitin. Stool samples were collected at 1, 3, 6 and 10 days of colonization and used to determine viable counts of wild-type and mutant *E. faecium* as described below. Mice were sacrificed after 10 days of colonization. The small intestines, cecum, and colon were collected, weighed and homogenized in 10 volumes PBS. Viable counts were determined by serially diluting the homogenates and plating onto SBA and SBA supplemented with 100 µg/ml gentamicin. After incubation at 37°C for two days, colonies were counted, and the strain identity was confirmed by PCR detection of the *esp* gene in E1162 wild-type and mutant. The PCR amplification condition was described above. The competitive index (CI) of mutant versus wild-type was determined as previously described [24] by the following formula: CI = (CFU mutant output/ CFU wild-type output)/(CFU mutant inoculum/CFU wild-type inoculum). Statistical significance between wild-type and mutant was assessed by the paired two-tailed Student's *t*-test of log-transformed data. $P < 0.05$ was considered statistically significant. The Animal Care and Use Committee of the University of Amsterdam (Amsterdam, The Netherlands) approved all experiments.

Results

Comparative genomic analysis identified genes specifically present in clinical E. faecium isolates

To determine whether specific traits distinguish commensal and clinical *E. faecium* isolates, we performed genome-wide comparisons with publicly available *E. faecium* genomes. Two clearly separated clades were identified, similar to previous analyses [7, 25] (Figure 1*A*). Most of the clinical isolates are members of clade A, while most of the human commensal isolates belong to clade B (Figure 1*A*). Comparative genomic analysis identified 93 genes (including 9 gene clusters) that were exclusively present in ≥80% isolates (four of the five TX0133-derived strains were excluded in this analysis) of clade A (Supplementary Table 2). Since gene clusters are more likely to confer complete functions than single genes, further analysis was performed on the gene clusters. We found that the largest gene cluster (cluster 7, Supplementary Table 2) was also present in a human commensal strain E1039 that clustered inside clade A, while the second largest gene cluster (cluster 5) was absent in all the commensal strains. We therefore decided to focus on the latter gene cluster, which is 15 kb in size (Figure 2) and specifically enriched in clinical isolates. It contains four genes that collectively encode a mannose family PTS, which is involved in the active transport of carbohydrates over the cytoplasmic membrane [26]. This gene cluster is chromosomally located in the complete genome sequence of *E. faecium* Aus0004 and the highly similar gene

synteny of the genes flanking this cluster in other *E. faecium* genomes strongly suggests a chromosomal location in the other genomes as well.

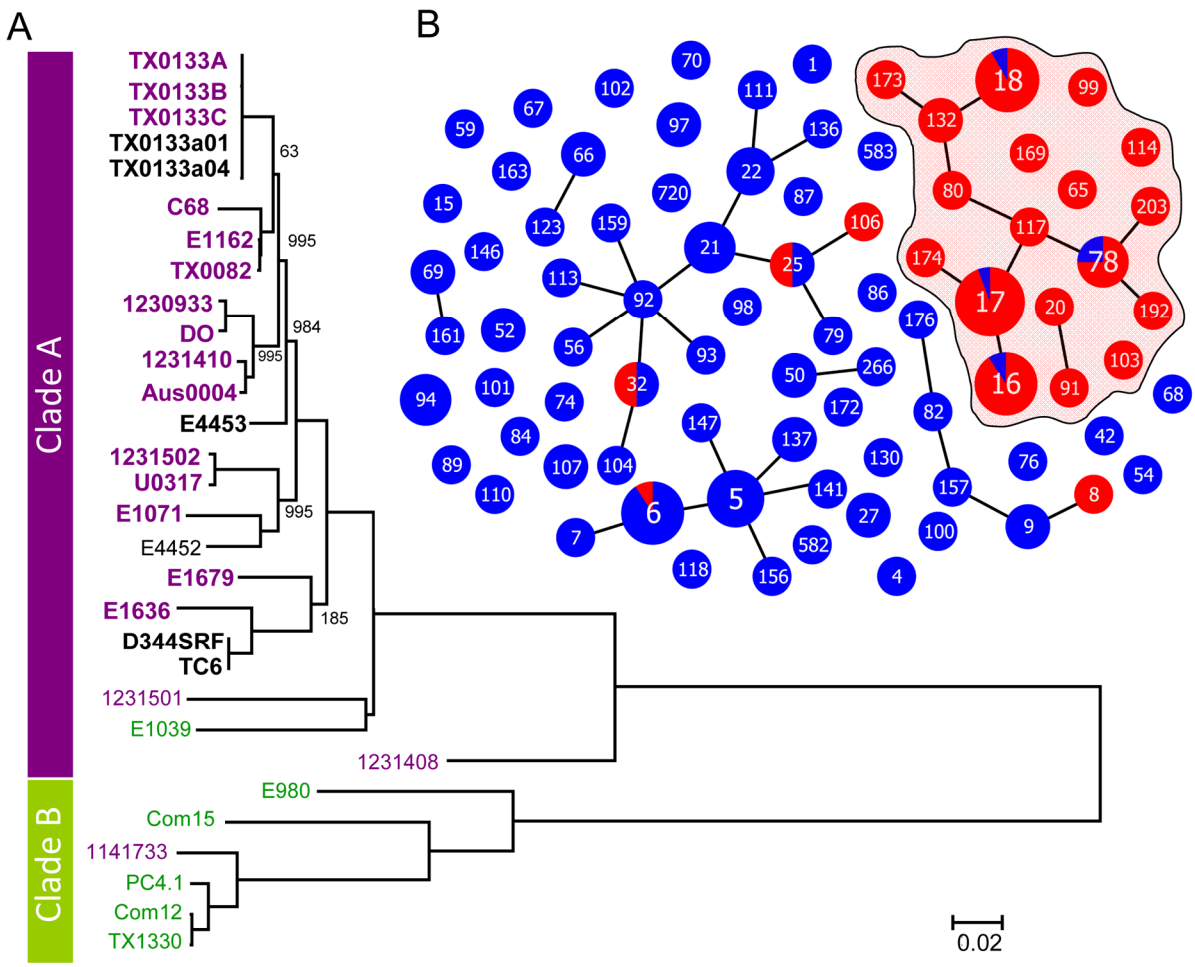


Figure 1. Phylogenomic and genotypic analysis of *E. faecium*. (A) A bootstrapped Neighbour-Joining tree was constructed based on the SNPs of the concatenated sequences of 1046 genes that are present in all 30 *E. faecium* genome sequences. Clades A and B are indicated. The hospital-associated isolates are displayed in purple and the commensal isolates are displayed in green. TX0133a01, TX0133a04, D344SRF and TC6 are laboratory strains, and E4453 and E4452 are isolates from dog stools. The isolates carrying the PTS^{clin} gene cluster are indicated in bold. Unless otherwise indicated, the bootstrap value for all nodes is 1000. (B) A population snapshot of 140 *E. faecium* isolates based on MLST data using the goeBURST algorithm [17]. Each circle represents a different ST, and the ST is indicated by the number in the circle. Circle size corresponds to the number of strains tested for each ST. STs are connected to their single locus variant with black lines. Pie charts indicate the distribution of presence (red) and absence (blue) of the PTS^{clin} gene cluster. STs that are commonly associated with clinical infections [8] are surrounded by pink shading.

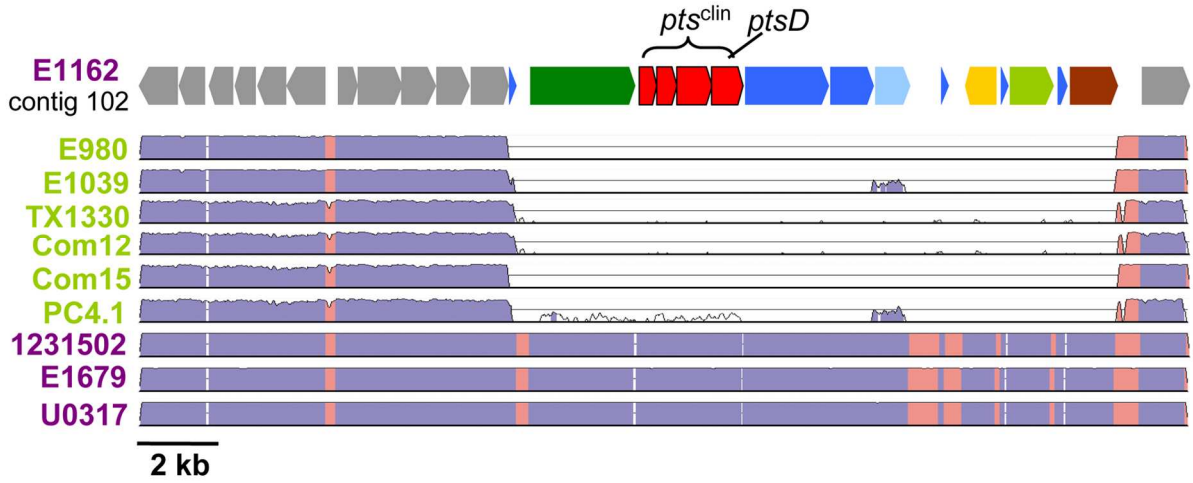


Figure 2. Overview and multiple genome alignment of the PTS^{clin} gene cluster in *E. faecium*. Arrows indicate the coding sequences of contig102 of E1162. Blue: hypothetical protein; dark green: AAA superfamily ATPase with a possible regulatory function; red: PTS system of the mannose/fructose/sorbose family; light blue: fructokinase; yellow: transcriptional regulator; green: replication initiation factor family; brown: phage integrase family protein; grey: genes flanking the PTS^{clin} gene cluster. Contig102 of E1162 was aligned to the other 29 *E. faecium* genomes, in which the alignments with 6 commensal strains (green) and 3 clinical strains (purple) are shown. The VISTA curve is calculated as a windowed-average (window size of 100 bp) identity score for the alignment. Peaks represent conservation between aligned sequences. Regions of high conservation are colored as coding (blue) or non-coding (pink).

The PTS^{clin} cluster is highly prevalent in clinical E. faecium *isolates, absent from commensal* E. faecium *isolates*

The presence and absence of this gene cluster was determined by PCR among 140 *E. faecium* isolates from diverse sources and genetic backgrounds (Supplementary Table 1). Together with 20 additional non-laboratory *E. faecium* strains of which the genomes have been sequenced, we analyzed the prevalence of the PTS^{clin} cluster in 160 *E. faecium* strains in total. This gene cluster is enriched in the isolates of hospital outbreaks (16/19, 84.2%) and non-outbreak clinical infections (33/50, 66.0%), but is absent (0/33; 0%) in human commensal isolates (Table 1). Notably the PTS^{clin} cluster was also identified in 3 out of the 17 strains isolated from animals (from pig, poultry and dog), suggesting a shared reservoir of genetic information for animal and clinical *E. faecium* strains.

Table 1. The presence of PTS^{clin} in *E. faecium* strains from diverse sources

Origin of isolates	Number of tested strains	PTS^{clin} positive	PTS^{clin} positive (%)
Human commensal	33	0	0
Clinical infections	50	33	66.0
Hospital outbreaks[a]	19	16	84.2
Hospital surveillance[b]	33	9	27.3
Animal	17	3	17.6
Environment	8	0	0

An overview of the strains used is provided in Supplementary Table 1 and Figure 1.
[a] Strains isolated from stools or infections of hospitalized patients during a hospital outbreak of *E. faecium* infections.
[b] *E. faecium* strains isolated from stools of hospitalized patients without an enterococcal infection.

Analysis of multilocus sequence typing (MLST) data of the 160 *E. faecium* isolates showed that 60 isolates belong to sequence types (STs) that are commonly associated with hospital infections [8]. Among these 60 strains, 56 (93%) isolates carry PTS^{clin}, which is significantly ($P < 0.0001$) more frequent than among the other 100 strains, of which 5 isolates (5%) contain PTS^{clin} (Figure 1*B*).

PTS^{clin} contributes to intestinal colonization of E. faecium in a mouse model

A mannose family PTS is composed of general cytoplasmic enzymes (EI and HPr), carbohydrate-specific cytoplasmic proteins (EIIA and EIIB), and carbohydrate-specific proteins (EIIC and EIID) which together form a membrane-associated complex [26]. In this study we constructed a deletion mutant of the *ptsD* gene , which encodes a sugar-specific membrane-associated EIID subunit that is required for carbohydrate transport [26]. The mutant strain and wild-type E1162 were tested on 65 different carbohydrates to identify the substrates of PTS^{clin}, but no significant difference in growth was observed between the two strains (data not shown). Possibly PTS^{clin} is only expressed under the complex conditions that are prevalent in the intestinal tract and which we were unable to replicate in the laboratory.
The mutant strain and wild-type E1162 exhibited same levels of minimum inhibitory concentration (MIC) of cefoxitin and ceftriaxone, which are 256 and 512 µg/ml respectively. These two strains had same fitness during competitive growth in BHI, and showed identical growth ratio in BHI supplemented with 20 µg/ml cefoxitin (Supplementary Figure 1).

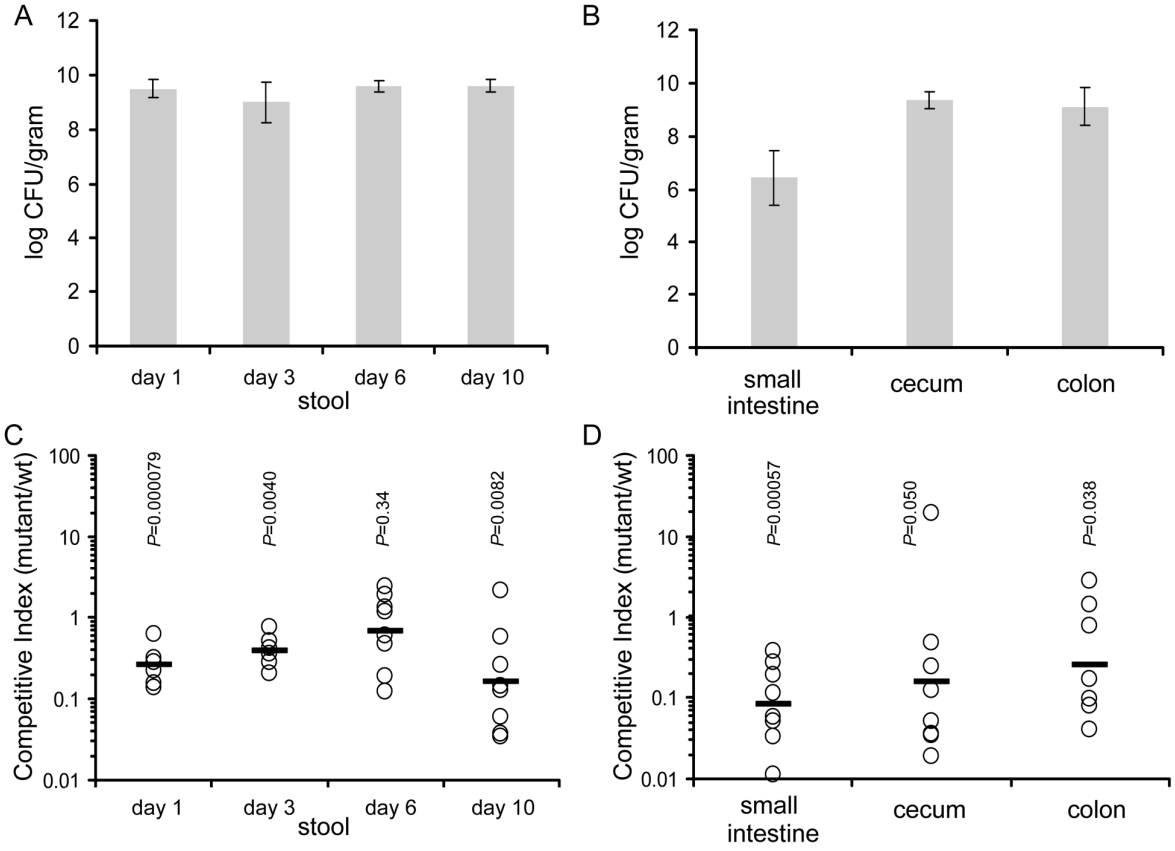


Figure 3. Intestinal colonization by *E. faecium* in a murine model. Eight mice were orally inoculated with a 1:1 mix of mixed wild-type E1162 and the *ptsD* mutant strain. Stool samples were collected at 1, 3, 6 and 10 days of colonization, and samples of small intestine, cecum and colon were collected at 10 days of colonization. (A and B) Total numbers of *E. faecium* were determined in stool, small intestine, cecum and colon. Data are expressed as geometric means of 8 mice. Error bars indicate the standard deviation. (C and D) Horizontal bars indicate the geometric means of CI values, and individual CI values are illustrated with circles. *P*-values (paired two-tailed Student's *t*-test) for comparisons between wild-type and the mutant are indicated.

To investigate the role of PTS^{clin} in intestinal colonization, cell suspensions of the E1162 strain and its isogenic *ptsD*-deficient mutant were mixed in a 1:1 ratio (10^7 colony forming units (CFUs) of each strain) and used to orogastrically inoculate mice. As shown in Figure 3*AB* *E. faecium* was able to successfully colonize the intestinal tract of mice at high level. The CUF per gram stool samples were stabilized over 10^9 at all time points. The data presented in Figure 3*C* showed that the mutant had significant lower viable counts compared to the wild-type, in stool samples collected in all the time points except day 6. Figure 3*D* demonstrated that the mutant colonized the small intestines and the colon at significantly lower levels than the parental strain. The difference between colonization of the cecum by E1162 and the mutant approached statistical significance ($P = 0.05$). We conclude that the deletion of *ptsD* in *E. faecium* E1162 resulted in a lower intestinal colonization in this mouse model.

Discussion

In this work, we have described the identification of a PTS that is enriched in clinical *E. faecium* isolates and which contributes to intestinal colonization of *E. faecium* in a mouse model. PTS is crucially important in the uptake and metabolism of carbohydrates in both Gram-negative and Gram-positive bacteria [26]. Microbes in the intestinal tract interact to form complex food webs for the degradation and metabolism of carbohydrates [27]. Consequently, genes involved in carbohydrate metabolism in gut-dwelling commensals and opportunistic pathogens are of crucial importance for intestinal colonization [27, 28]. In this study, we identify PTS^{clin} as a determinant for intestinal colonization of a clinical *E. faecium* isolate, indicating the importance of adaptations in carbohydrate metabolism to intestinal colonization of clinical *E. faecium* strains.

Several homologues of PTS^{clin} can be identified in genomes of anaerobic gut commensals. Particularly the IIC and IID-subunits have homologs with a high amino acid identity (>65%) in *Catenibacterium mitsuokai*, *Ruminococcus torques* and *R. gnavus*, suggesting that the element in *E. faecium* may have originated from anaerobes that colonize the mammalian gut. We propose that PTS^{clin} allows *E. faecium* to benefit from the disturbed metabolic network after the administration of antibiotics.

The observation that commensal *E. faecium* strains do not carry PTS^{clin} indicates that this element does not contribute to colonization of the unperturbed intestinal tract, possibly because of the competition by better adapted, highly abundant strictly anaerobic gut commensals that carry similar PTS. Another possibility is that the PTS^{clin} could be targeted by certain bacteriocins produced by the gut microflora in healthy humans, since mannose PTSs are common targets of bacteriocins [29].

The data described in this study implicate the significant, and previously unrecognized, contribution of a PTS in gut colonization of clinical isolates of *E. faecium*. This is the first description of a genetic determinant that contributes to intestinal colonization in the presence of antibiotics and which is specific to clinical isolates of *E. faecium*.

Acknowledgments

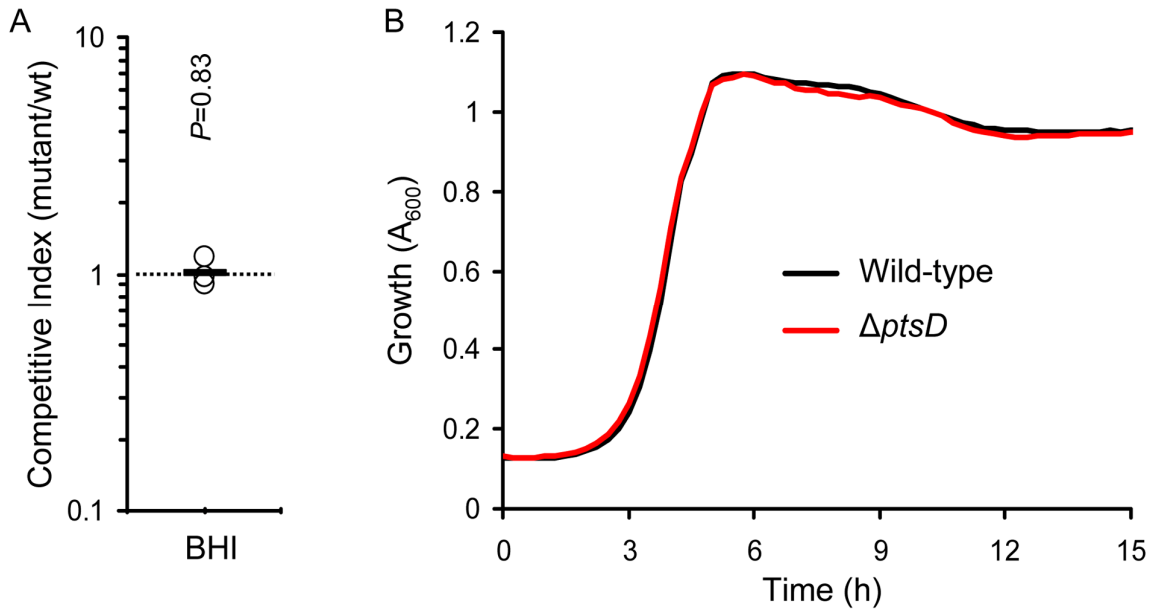
The authors wish to thank Joost Daalhuisen and Marieke ten Brink for their expert technical assistance during the animal experiments. This work was supported by the Nederlandse Organisatie voor Wetenschappelijk Onderzoek [NWO-VENI grant 916.86.044] and the European Union Seventh Framework Programme (FP7-HEALTH-2011-single-stage) Evolution and Transfer of Antibiotic Resisance (EvoTAR) [grant agreement number 282004].

References

1. Arias CA, Murray BE: The rise of the *Enterococcus*: beyond vancomycin resistance. *Nat Rev Microbiol* 2012, 10(4):266-278.
2. Ubeda C, Taur Y, Jenq RR, Equinda MJ, Son T, Samstein M, Viale A, Socci ND, van den Brink MR, Kamboj M *et al*: Vancomycin-resistant *Enterococcus* domination of intestinal microbiota is enabled by antibiotic treatment in mice and precedes bloodstream invasion in humans. *J Clin Invest* 2010, 120(12):4332-4341.
3. Palmer KL, Gilmore MS: Multidrug-resistant enterococci lack CRISPR-*cas*. *MBio* 2010, 1(4).
4. Gilmore MS, Ferretti JJ: Microbiology. The thin line between gut commensal and pathogen. *Science* 2003, 299(5615):1999-2002.
5. Brandl K, Plitas G, Mihu CN, Ubeda C, Jia T, Fleisher M, Schnabl B, DeMatteo RP, Pamer EG: Vancomycin-resistant enterococci exploit antibiotic-induced innate immune deficits. *Nature* 2008, 455(7214):804-807.
6. Willems RJ, Top J, van Santen M, Robinson DA, Coque TM, Baquero F, Grundmann H, Bonten MJ: Global spread of vancomycin-resistant *Enterococcus faecium* from distinct nosocomial genetic complex. *Emerg Infect Dis* 2005, 11(6):821-828.
7. Galloway-Pena J, Roh JH, Latorre M, Qin X, Murray BE: Genomic and SNP analyses demonstrate a distant separation of the hospital and community-associated clades of *Enterococcus faecium*. *PLoS One* 2012, 7(1):e30187.
8. Willems RJ, van Schaik W: Transition of *Enterococcus faecium* from commensal organism to nosocomial pathogen. *Future Microbiol* 2009, 4(9):1125-1135.
9. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ: Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 1997, 25(17):3389-3402.
10. Li L, Stoeckert CJ, Jr., Roos DS: OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Res* 2003, 13(9):2178-2189.
11. van Dongen S: Graph clustering by flow simulation. *PhD thesis* 2000.
12. Edgar RC: MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 2004, 32(5):1792-1797.
13. Capella-Gutierrez S, Silla-Martinez JM, Gabaldon T: trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 2009, 25(15):1972-1973.
14. Thompson JD, Higgins DG, Gibson TJ: CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 1994, 22(22):4673-4680.
15. Arias CA, Torres HA, Singh KV, Panesso D, Moore J, Wanger A, Murray BE: Failure of daptomycin monotherapy for endocarditis caused by an *Enterococcus faecium* strain with vancomycin-resistant and vancomycin-susceptible subpopulations and evidence of in vivo loss of the *vanA* gene cluster. *Clin Infect Dis* 2007, 45(10):1343-1346.
16. Frazer KA, Pachter L, Poliakov A, Rubin EM, Dubchak I: VISTA: computational tools for comparative genomics. *Nucleic Acids Res* 2004, 32(Web Server issue):W273-279.
17. Feil EJ, Li BC, Aanensen DM, Hanage WP, Spratt BG: eBURST: inferring patterns of evolutionary descent among clusters of related bacterial genotypes from multilocus sequence typing data. *J Bacteriol* 2004, 186(5):1518-1530.
18. Zhang X, Paganelli FL, Bierschenk D, Kuipers A, Bonten MJ, Willems RJ, van Schaik W: Genome-Wide Identification of Ampicillin Resistance Determinants in *Enterococcus faecium*. *PLoS Genet* 2012, 8(6):e1002804.
19. van Schaik W, Top J, Riley DR, Boekhorst J, Vrijenhoek JE, Schapendonk CM, Hendrickx AP, Nijman IJ, Bonten MJ, Tettelin H *et al*: Pyrosequencing-based comparative genome analysis of the nosocomial pathogen *Enterococcus faecium* and identification of a large transferable pathogenicity island. *BMC Genomics* 2010, 11:239.

20. Zhang X, Vrijenhoek JE, Bonten MJ, Willems RJ, van Schaik W: A genetic element present on megaplasmids allows *Enterococcus faecium* to use raffinose as carbon source. *Environ Microbiol* 2011, 13(2):518-528.
21. Heikens E, Leendertse M, Wijnands LM, van Luit-Asbroek M, Bonten MJ, van der Poll T, Willems RJ: Enterococcal surface protein Esp is not essential for cell adhesion and intestinal colonization of *Enterococcus faecium* in mice. *BMC Microbiol* 2009, 9:19.
22. Krueger WA, Krueger-Rameck S, Koch S, Carey V, Pier GB, Huebner J: Assessment of the role of antibiotics and enterococcal virulence factors in a mouse model of extraintestinal translocation. *Crit Care Med* 2004, 32(2):467-471.
23. Rice LB: Antibiotics and gastrointestinal colonization by vancomycin-resistant enterococci. *Eur J Clin Microbiol Infect Dis* 2005, 24(12):804-814.
24. Lawley TD, Chan K, Thompson LJ, Kim CC, Govoni GR, Monack DM: Genome-wide screen for *Salmonella* genes required for long-term systemic infection of the mouse. *PLoS Pathog* 2006, 2(2):e11.
25. Palmer KL, Godfrey P, Griggs A, Kos VN, Zucker J, Desjardins C, Cerqueira G, Gevers D, Walker S, Wortman J *et al*: Comparative genomics of enterococci: variation in *Enterococcus faecalis*, clade structure in *E. faecium*, and defining characteristics of *E. gallinarum* and *E. casseliflavus*. *MBio* 2012, 3(1):e00318-00311.
26. Deutscher J, Francke C, Postma PW: How phosphotransferase system-related protein phosphorylation regulates carbohydrate metabolism in bacteria. *Microbiol Mol Biol Rev* 2006, 70(4):939-1031.
27. Koropatkin NM, Cameron EA, Martens EC: How glycan metabolism shapes the human gut microbiota. *Nat Rev Microbiol* 2012, 10(5):323-335.
28. Chang DE, Smalley DJ, Tucker DL, Leatham MP, Norris WE, Stevenson SJ, Anderson AB, Grissom JE, Laux DC, Cohen PS *et al*: Carbon nutrition of *Escherichia coli* in the mouse intestine. *Proc Natl Acad Sci U S A* 2004, 101(19):7427-7432.
29. Opsata M, Nes IF, Holo H: Class IIa bacteriocin resistance in *Enterococcus faecalis* V583: the mannose PTS operon mediates global transcriptional responses. *BMC Microbiol* 2010, 10:224.

Supplementary data



Supplementary Figure 1. *In vitro* analysis of growth fitness of wild-type E1162 and the *ptsD* mutant. (*A*) A 1:1 mix of mixed wild-type E1162 and the *ptsD* mutant strain was inoculated into BHI and grown for 20 generations. Competitive index (CI) of three independent experiments were determined. Horizontal bar indicates the geometric means of CI value, and individual CI values of each experiment are illustrated with a circle. The *P*-value (paired two-tailed Student's *t*-test) for the difference between wild-type and the mutant is indicated. (*B*) Overnight cultures of wild-type E1162 (black) and the *ptsD* mutant (red) were inoculated at an initial cell density of OD_{660} 0.0025 in BHI with 20 µg/ml cefoxitin. Growth curves are mean data of three independent experiments.

Supplementary Table 1. *E. faecium* isolates used for Multiplex PCR screening of PTS^{clin}

Category	Strains[a]	ST	Country[b]	Year	Isolation site	PTS^{clin}
Human commensal	E0060	147	NLD	1996	Faeces	-
	E0092	6	NLD	1997	Faeces	-
	E0128	82	NLD	1996	Faeces	-
	E0135	6	NLD	1996	Faeces	-
	E0980	**94**	**NLD**	**1998**	**Faeces**	**-**
	E1002	54	NLD	1998	Faeces	-
	E1039	**42**	**NLD**	**1998**	**Faeces**	**-**
	E1485	101	ESP	2000	Faeces	-
	E1590	163	IRL	2001	Faeces	-
	E1764	6	BEL	1996	Faeces	-
	E1766	136	BEL	1996	Faeces	-
	E0129	6	NLD	1996	Faeces	-
	E0130	6	NLD	1996	Faeces	-
	E0131	5	NLD	1996	Faeces	-
	E0138	6	NLD	1996	Faeces	-
	E0139	6	NLD	1996	Faeces	-
	E0278	9	NLD	1996	Faeces	-
	E0400	15	NLD	1997	Faeces	-
	E0403	7	NLD	1997	Faeces	-
	E1028	56	NLD	1998	Faeces	-
	E1037	97	NLD	1998	Faeces	-
	E1046	98	NLD	1998	Faeces	-
	E1050	92	NLD	1998	Faeces	-
	E1327	118	NLD	1998	Faeces	-
	E1488	32	ESP	2000	Faeces	-
	E1489	102	ESP	2000	Faeces	-
	E1582	59	BEL	2001	Faeces	-
	E1583	161	BEL	2001	Faeces	-
	E1768	137	BEL	1996	Faeces	-
Hospital outbreaks	E0510	17	AUS	1998	Blood	-
	E1760	173	AUS	2000	ND	+
	E1679	**114**	**BRA**	**1998**	**Catheter**	**+**
	E1644	78	DEU	2002	Catheter urine	+
	E1716	17	DNK	ND	ND	+
	E1717	18	DNK	ND	ND	+
	E1441	16	GRC	2000	Peritoneal fluid	+
	E1435	65	GRC	1999	Blood	+
	E0470	16	NLD	2000	ND	+
	E0734	16	NLD	2000	Environment	+
	E1652	18	NLD	2002	Faeces	+
	E0745	16	NLD	2000	Faeces	+
	E0772	50	NLD	2000	Faeces	-
	E1340	17	NOR	1999	WouND	+
	E0013	18	GBR	1992	Urine	+
	E0300	20	USA	1994	Urine	+
	E0155	17	USA	ND	ND	+
	E0161	16	USA	1995	Faeces	-
	E1132	16	USA	2001	Faeces	+
Clinical infections	E0073	22	NLD	1995	Faeces	-

Category	Strains[a]	ST	Country[b]	Year	Isolation site	PTS[clin]
	E0125	21	NLD	1995	Bile	-
	E0318	91	USA	1994	Peritoneal fluid	+
	E0333	80	ISR	1997	Blood	+
	E0380	17	GBR	1997	Blood	+
	E0699	157	GBR	ND	ND	-
	E1162	**17**	**FRA**	**ND**	**Blood**	**+**
	E1172	99	POL	ND	Urine	+
	E1250	25	CHE	ND	Blood	-
	E1263	78	AUT	ND	Blood	-
	E1282	17	DEU	ND	ND	+
	E1283	130	DEU	ND	Blood	-
	E1292	17	ITA	ND	Blood	+
	E1302	17	POL	ND	Blood	+
	E1307	132	PRT	ND	Blood	+
	E1308	17	PRT	ND	Blood	+
	E1360	16	USA	2001	ND	+
	E1391	16	GBR	2000	ND	+
	E1403	84	GBR	2000	ND	-
	E1421	100	GBR	2000	ND	-
	E1423	94	GBR	2000	ND	-
	E1463	17	ESP	1998	Blood	+
	E1467	18	ESP	1997	Blood	+
	E1499	74	ESP	1999	Urine	-
	E1500	18	ESP	1997	Blood	+
	E1620	67	NLD	1950	ND	-
	E1621	86	NLD	1959	Blood	-
	E1623	22	NLD	1960	Pus	-
	E1625	22	NLD	1961	Cerebrospinal fluid	-
	E1636	**106**	**NLD**	**1961**	**Blood**	**+**
	E1721	169	TZA	ND	Blood	+
	E1728	132	TZA	ND	Blood	+
	E1731	18	TZA	ND	Blood	+
	E1734	18	ESP	1995	Blood	+
	E1735	16	ESP	2001	Blood	+
	E1737	18	ESP	1997	Blood	+
	E1794	18	USA	1991	Blood	+
Hospital surveillance	E0005	25	FRA	1986	Faeces	+
	E0027	146	GBR	1992	Faeces	-
	E0321	17	FRA	1997	Faeces	+
	E0322	79	FRA	1997	Faeces	-
	E0802	5	NLD	2000	Faeces	-
	E0849	21	NLD	2000	Faeces	-
	E1071	**32**	**NLD**	**ND**	**ND**	**+**
	E1133	117	USA	2001	Faeces	+
	E1147	16	NLD	1998	Faeces	+
	E1149	6	NLD	1995	Faeces	-
	E1316	5	NLD	2000	Blood	-
	E1554	50	NLD	2002	Faeces	-
	E1643	78	DEU	2002	Faeces	+
	E1674	110	BRA	ND	Faeces	-
	E1675	111	BRA	ND	Faeces	-
	E1850	18	ESP	ND	Faeces	-
	E0075	87	NLD	1995	Faeces	-

Category	Strains[a]	ST	Country[b]	Year	Isolation site	PTS[clin]
	E0492	89	NLD	1999	ND	-
	E0729	5	NLD	2000	Faeces	-
	E0767	6	NLD	2000	Faeces	-
	E0808	6	NLD	2000	Faeces	-
	E1141	123	NLD	1995	Faeces	-
	E1493	103	ESP	2001	Faeces	+
	E1552	5	NLD	2002	Faeces	-
	E1627	66	NLD	1979	Faeces	-
	E1638	5	NLD	ND	ND	-
	E1640	5	NLD	ND	ND	-
	E1669	94	BRA	ND	Faeces	-
	E1670	94	BRA	ND	Faeces	-
	E1677	113	BRA	ND	Faeces	-
	E1695	97	BRA	1998	Faeces	-
	E1762	174	AUS	2000	ND	+
Animal	E1573	21	BEL	1994	Faeces	-
	E0172	1	NLD	1996	Faeces	-
	E0211	4	NLD	1996	Faeces	-
	E0466	21	NLD	ND	ND	-
	E0463	27	NLD	1996	Faeces	-
	E1574	27	BEL	1995	Faeces	-
	E1607	76	NOR	1956	Food	-
	E1619	70	NOR	1964	Food	-
	E1576	159	ZAF	2001	Faeces	-
	E0144	6	NLD	1996	Faeces	+
	E0685	137	ESP	ND	Faeces	-
	E1781	141	BEL	2001	Carcasses	-
	E0045	9	GBR	1992	Faeces	-
	E0429	8	NLD	1997	Faeces	+
	E1622	104	NLD	1959	Faeces	-
Environment	E0695	176	GBR	ND	Faeces	-
	E0696	156	GBR	ND	Faeces	-
	E1628	68	NLD	1981	Environment	-
	E1629	93	NLD	1981	Environment	-
	E1630	69	NLD	1981	Environment	-
	E1631	69	NLD	1981	Environment	-
	E1634	66	NLD	1982	Environment	-
	E1759	172	FRA	1985	ND	-

[a] Strains of which genome sequences have been published are indicated in bold.
[b] Country codes are the three letter codes as defined in ISO 3166-1.
ND: not determined.

Supplementary Table 2. Genes and gene clusters specifically present in clade A *E. faecium* isolates

Gene cluster	LocusTag	Genbank Accession code	Annotation
Cluster 1	EfmE1162_0138	293559445	DNA-binding response regulator
	EfmE1162_0139	293559446	histidine kinase-related ATPase, putative
	EfmE1162_0140	293559447	hypothetical protein
	EfmE1162_0141	293559448	hypothetical protein
	EfmE1162_0142	293559449	hypothetical protein
Cluster 2	EfmE1162_0285	293559596	transcriptional regulator, DeoR family
	EfmE1162_0286	293559597	tagatose-1,6-bisphosphate aldolase gaty
	EfmE1162_0287	293559598	pts system fructose-specific eiibbc component
Cluster 3	EfmE1162_1160	293560515	transcriptional regulator
	EfmE1162_1161	293560516	6-phosphogluconate dehydrogenase
	EfmE1162_1162	293560517	D-3-phosphoglycerate dehydrogenase
Cluster 4	EfmE1162_1402	293560771	beta-glucosidase
	EfmE1162_1403	293560772	hypothetical protein
	EfmE1162_1404	293560773	pts system beta-glucoside-specific eiibca component
	EfmE1162_1405	293560774	beta-glucosidase
Cluster 5	EfmE1162_1917	293563224	PTS system mannose/fructose/sorbose family IIC component
	EfmE1162_1918	293563225	PTS system mannose/fructose/sorbose family IID component
	EfmE1162_1919	293563226	hypothetical protein
	EfmE1162_1920	293563227	conserved hypothetical protein
	EfmE1162_1921	293563228	fructokinase
	EfmE1162_1923	293563230	transcriptional regulator, putative
	EfmE1162_1925	293563232	replication initiation factor family
	EfmE1162_1927	293563234	phage integrase family protein
Cluster 6	EfmE1162_2549	293563811	pts system N-acetylglucosamine-specific eiicba component
	EfmE1162_2550	293563812	YdjC-like protein
	EfmE1162_2551	293563813	transcription antiterminator
Cluster 7	EfmE1162_2309	293564004	putative HTH-type transcriptional regulator YcjW
	EfmE1162_2310	293564005	putative sucrose phosphorylase
	EfmE1162_2311	293564006	ABC transporter, substrate-binding protein
	EfmE1162_2312	293564007	ABC transporter, permease protein
	EfmE1162_2313	293564008	inner membrane ABC transporter permease protein YcjP
	EfmE1162_2314	293564009	oxidoreductase, zinc-binding dehydrogenase family
	EfmE1162_2315	293564010	AP endonuclease, family 2
	EfmE1162_2316	293564011	putative oxidoreductase YcjS

Gene cluster	LocusTag	Genbank Accession code	Annotation
	EfmE1162_2317	**293564012**	**kojibiose phosphorylase**
	EfmE1162_2318	**293564013**	**glycerate kinase**
	EfmE1162_2319	**293564014**	**beta-phosphoglucomutase**
	EfmE1162_1537	**293564056**	**lipoprotein, NLP/P60 family**
Cluster 8	**EfmE1162_1536**	**293564057**	**conjugative transposon membrane protein**
	EfmE1162_1535	**293564058**	**conjugative transposon protein**
	EfmE1162_1804	**293564101**	**transcriptional regulator, LysR family**
Cluster 9	**EfmE1162_1803**	**293564102**	**hexapeptide-repeat containing-acetyltransferase**
	EfmE1162_1802	**293564103**	**Short-chain dehydrogenase/reductase SDR**
Genes not part of gene cluster	EfmE1162_0055	293559358	orotate phosphoribosyltransferase
	EfmE1162_0070	293559374	5-formyltetrahydrofolate cyclo-ligase
	EfmE1162_0093	293559399	general stress protein, putative
	EfmE1162_0300	293559612	PrgW
	EfmE1162_0302	293559614	ATPase, ParA family protein
	EfmE1162_0331	293559644	lipoprotein, NLP/P60 family
	EfmE1162_0332	293559645	conjugative transposon membrane protein
	EfmE1162_0356	293559673	guanylate kinase
	EfmE1162_0364	293559681	lipoprotein
	EfmE1162_0407	293559731	conserved hypothetical protein
	EfmE1162_0416	293559740	Cof family protein
	EfmE1162_0476	293559803	hypothetical protein
	EfmE1162_0537	293559870	integral membrane protein
	EfmE1162_0701	293560042	integral membrane protein
	EfmE1162_0788	293560129	ABC transporter, permease protein
	EfmE1162_0833	293560176	co-chaperone GrpE
	EfmE1162_0876	293560220	hypothetical protein
	EfmE1162_0997	293560347	conserved hypothetical protein
	EfmE1162_1022	293560374	ribosomal subunit interface protein
	EfmE1162_1190	293560545	GapA
	EfmE1162_1228	293560584	hypothetical protein
	EfmE1162_1269	293560626	conserved hypothetical protein
	EfmE1162_1314	293560674	lipoprotein, putative
	EfmE1162_1356	293560722	nucleotide-binding protein ExpZ
	EfmE1162_1389	293560757	transposase
	EfmE1162_1417	293560786	cardiolipin synthetase
	EfmE1162_1499	293562872	conserved hypothetical protein
	EfmE1162_1527	293562903	transposase, Mutator family
	EfmE1162_1550	293562916	putative phosphoglucomutase
	EfmE1162_1551	293562917	lactose phosphotransferase system repressor
	EfmE1162_1555	293562921	arabinogalactan endo-1,4-beta-galactosidase
	EfmE1162_1648	293563015	hypothetical protein

Gene cluster	LocusTag	Genbank Accession code	Annotation
	EfmE1162_1701	293563068	hypothetical protein
	EfmE1162_1702	293563069	conserved hypothetical protein
	EfmE1162_1728	293563095	hypothetical protein
	EfmE1162_1852	293563152	phosphoglycerol transferase
	EfmE1162_1881	293563184	transposase
	EfmE1162_2083	293563394	LysM domain protein
	EfmE1162_2099	293563410	dihydrofolate reductase
	EfmE1162_2116	293563427	heat shock protein HslVU, ATP-dependent protease subunit HslV
	EfmE1162_2145	293563457	hypothetical protein
	EfmE1162_2147	293563459	hypothetical protein
	EfmE1162_2148	293563460	site-specific recombinase, phage integrase family
	EfmE1162_2482	293563741	hypothetical protein
	EfmE1162_2514	293563774	septum formation protein Maf
	EfmE1162_2602	293563872	conjugative transposon protein
	EfmE1162_2603	293563873	conjugative transposon membrane protein
	EfmE1162_2609	293563880	ATP/GTP-binding conjugative transposon protein
	EfmE1162_2684	293563975	cell wall surface anchor family protein
	EfmE1162_2688	293563979	BNR/Asp-box repeat domain protein

The 9 gene clusters are indicated in bold and each gene cluster is separated by blank lines.

Chapter 4

Functional genomic analysis of bile salt resistance in *Enterococcus faecium*

Xinglin Zhang, Damien Bierschenk, Janetta Top, Iacovos Anastasiou, Marc J. M. Bonten, Rob J. L. Willems, Willem van Schaik

Department of Medical Microbiology, University Medical Center Utrecht, Utrecht, The Netherlands

Submitted for publication

Abstract

Enterococcus faecium is a Gram-positive commensal bacterium of the mammalian intestinal tract. In the last two decades it has also emerged as a multi-resistant nosocomial pathogen. In order to survive in and colonize the human intestinal tract *E. faecium* must resist the deleterious actions of bile. The molecular mechanisms exploited by this bacterium to tolerate bile are as yet unexplored. In this study we used a high-throughput quantitative screening approach of transposon mutant libraries, termed Microarray-based Transposon Mapping (M-TraM), to identify the genetic determinants required for resistance to bile salts in *E. faecium* E1162. The gene *gltK*, which is predicted to encode a glutamate/aspartate transport system permease protein, was identified by M-TraM to be involved in bile resistance. The role of GltK in bile salt resistance was confirmed by the subsequent observation that the deletion of *gltK* significantly sensitized *E. faecium* E1162 to bile salts. To further characterize the response of *E. faecium* E1162 to bile salts, we performed a transcriptome analysis to identify genes that are regulated by exposure to 0.02% bile salts. Exposure to bile salts resulted in major transcriptional rearrangements, predominately in genes involved in carbohydrate, nucleotide and coenzyme transport and metabolism. These findings add to a better understanding of the molecular mechanisms by which *E. faecium* responds and resists the antimicrobial action of bile salts.

Introduction

Enterococcus faecium is a common inhabitant of the intestines of humans and animals and is present in many different environments [1, 2]. However, during the past two decades *E. faecium* has rapidly emerged as an important multi-drug resistant nosocomial pathogen around the world and is now frequently responsible for hospital-acquired bloodstream, urinary tract and surgical wound infections [3-5]. The establishment of high-level intestinal colonization by enterococci is a crucial step in a process that can finally lead towards nosocomial infections [5].

Enterococci are known as being highly tolerant to hostile environments including high temperature conditions and high salt concentrations [6]. Enterococci are also relatively resistant to chemical disinfectants like chlorine, glutaraldehyde and alcohol [7-9]. In order to survive in and colonize the intestinal tract, a bacterium must be able to cope with stresses encountered during adaptation to the conditions that occur in the human intestinal tract. Bile represents a major challenge to the intestinal microflora. The human liver daily secretes up to one liter of bile which is stored in the gall bladder and exported into the intestine [10]. Bile is a complex mixture composed mainly of bile salts, phospholipids, cholesterol, proteins and bilirubin [11]. Bile salts are amphipathic molecules that act as detergents, aiding in lipid solubilization and digestion, but they also play a role in host defenses, as bile salts have potent antimicrobial properties that can cause damage to the DNA, proteins and membranes of enteric bacteria [12, 13]. In both Gram-positive and Gram-negative bacteria the disruption of bile tolerance loci often leads to impaired intestinal survival [14-16], while a mutation resulting in high-level bile resistance of *Escherichia coli* displays a fitness advantage in intestinal colonization [17].

As a successful colonizer of the intestinal tract, *E. faecium* must have developed mechanisms to sense, respond to and tolerate bile during its evolution as a gut commensal. Previously, two genetic loci *gls33-glsB* and *gls20-glsB1* that encode Gls-like proteins in *E. faecalis* and *E. faecium* were identified to be involved in bile resistance and pathogenicity in a mouse peritonitis model [18, 19]. *E. faecium* was also shown to possess bile salt hydrolase (BSH) activity [20], which was conferred by the protein encoded by the *bsh* gene (accession no. AY260046) [21]. In this study, we performed a genome-wide identification of the genetic loci required for bile salt resistance in *E*. *faecium*, using a high-throughput quantitative screening approach of transposon mutant libraries, termed Microarray-based Transposon Mapping (M-TraM) [22]. We also studied the transcriptional response of *E*. *faecium* to bile salts-induced stress.

Material and Methods

Bacterial strains, plasmids and growth conditions

E. faecium and *E. coli* strains used in this study are listed in Table 1. The *E. faecium* strain E1162 was used throughout this study. This strain was isolated from a bloodstream infection in France in 1996 and its genome has previously been sequenced [23]. Unless otherwise mentioned, *E. faecium* was grown in brain heart infusion broth (BHI; Oxoid) at 37°C. The *E. coli* strains DH5α (Invitrogen) and EC1000 [24] were grown in Luria-Bertani medium. Where necessary, antibiotics were used at the following concentrations: gentamicin at 300 µg ml^{-1} for *E. faecium* and 25 µg ml^{-1} for *E. coli*, spectinomycin at 300 µg ml^{-1} for *E. faecium* and 100 µg ml^{-1} for *E. coli*. All antibiotics were obtained from Sigma-Aldrich (Saint Louis, MO). Growth of cultures was determined by measuring the optical density at 660 nm (OD_{660}).

Table 1. Strains and plasmids used in this study.

Strain or plasmid	Relevant characteristic(s)	Source or reference
E. faecium		
E1162	Clinical isolate (bloodstream infection), isolated in France, 1996	[23]
Δ*gltK*	Markerless deletion mutant of *gltK* gene of E1162	This study
Δ*gspA*	Markerless deletion mutant of *gspA* gene of E1162	This study
Δ*gltK+gltK*	Complementation strain of Δ*gltK*; Δ*gltK* harboring pEF25-*gltK*	This study
Δ*gspA+gspA*	Complementation strain of Δ*gspA*; Δ*gspA* harboring pEF25-*gspA*	This study
E. coli strains		
DH5α	*E. coli* host strain for routine cloning	Invitrogen
EC1000	MC1000 *glgB*::*repA*; host strain for pWS3 derived vectors	[24]
Plasmids		
pWS3	Gram-positive thermosensitive origin of replication; Spc^r	[25]
pDEL3a	pWS3 carrying the 5′ and 3′ flanking regions of gene *gltK* for mutant construction	This study
pDEL4a	pWS3 carrying the 5′ and 3′ flanking regions of *gspA* gene cluster for mutant construction	This study
pDEL3b	pDEL3a with a Gen^r cassette which was flanked by *lox66*- and *lox71*-sites cloned between the 5′ and 3′ flanking regions	This study
pDEL4b	pDEL4a with a Gen^r cassette which was flanked by *lox66*- and *lox71*-sites cloned between the 5′ and 3′ flanking regions	This study
pWS3-Cre	pWS3 derivative expressing Cre in *E. faecium*	[22]
pEF25	Shuttle plasmid pAT18 with spectinomycin resistance cassette cloned in the EcoRI site; Spc^r Ery^r	[26]
pEF25-*gltK*	Complementation plasmid for Δ*gltK*; pEF25 carrying *gltK*	This study
pEF25-*gspA*	Complementation plasmid for Δ*gspA*; pEF25 carrying *gspA*	This study

Screening for genes involved in bile salt resistance using M-TraM

M-TraM, a high throughput screening technique of transposon mutant libraries has previously been described in detail [22]. Here we use this technique to perform a genome-wide identification of genes involved in bile salt resistance in *E. faecium*. Briefly, aliquots containing approximately 10^7 colony-forming units (CFU) from the mutant pool were used to inoculate 20 ml of BHI broth or BHI broth supplemented with 0.02% bile salts (cholic acid sodium salt:deoxycholic acid sodium salt 1:1, Sigma-Aldrich). Cells were grown at 37°C for 20 hours, after which 1 ml of the cultures were spun down and used for the extraction of genomic DNA, which was then further processed as described previously [22]. Statistical differences in hybridization signals between the conditions were analyzed using Cyber-T [27] (http://cybert.microarray.ics.uci.edu/). Probes exhibiting a Bayesian P-value <0.005 were deemed statistically significant. A gene with two identical probes or when all four probes passed this threshold were classified as significantly selected during exposure to bile salts. In an addition, genes which were selected between 0.5- and 2-fold were deemed biologically insignificant and were filtered out. This experiment was performed with four biological replicates.

The microarray data generated in the M-TraM screening have been deposited in the ArrayExpress database (http://www.ebi.ac.uk/arrayexpress) under accession number E-MEXP-3797.

Transcriptome profiling

E. faecium E1162 was grown in 3 ml BHI broth for 18 hours. Cultures were then diluted 100 fold in 20 ml of prewarmed BHI broth (in a 50-ml Falcon tube) and grown until OD_{660} 0.3. Two ml aliquots of the cultures were centrifuged for 12 seconds at 16900 *g* at room temperature, and pellets were flash frozen in liquid N_2 prior to RNA extraction. This sample served as the negative control (t = 0 min) prior to the addition of bile salts. Bile salts (final concentration 0.02%, w/v) were added into the remaining 18 ml of culture. After 5 and 15 minutes of incubation at 37°C, 2 ml aliquots of the cultures were centrifuged and flash frozen as described above. RNA isolation, cDNA synthesis and hybridization were performed as described in our previous work [22]. In this experiment, the expression of genes at t = 5 min. and t = 15 min. were compared to t = 0 min. Analysis for statistical significance was performed using the Web-based VAMPIRE microarray suite (http://sasquatch.ucsd.edu/vampire/) as described previously [28, 29]. A gene of which all four probes on the microarray were identified as differentially expressed with a false discovery rate <0.001, were classified as significantly differentially expressed between samples. Genes with an expression between 0.5- and 2-fold different were deemed biologically insignificant and were filtered out. This experiment was performed with two biological replicates.

The microarray data generated in the transcriptome analysis have been deposited in the ArrayExpress database (http://www.ebi.ac.uk/arrayexpress) under accession number E-MEXP-3796.

Construction of markerless deletion mutants and in trans complementation

Markerless gene deletion mutants in the *gltK* gene (locustag: EfmE1162_1760) and the *gspA* gene (locustag: EfmE1162_1186) were created via the Cre-*lox* recombination system as previously described [22, 30]. Briefly, the 5′ and 3′ flanking regions (approximately 500 bp each) of the target genes were PCR amplified with the primers in Table 2. The two flanking regions were then fused together by fusion PCR (generating an EcoRI site between both fragments) and cloned into pWS3 [25], resulting in pDEL3a and pDEL4a (plasmids used or generated in this study are listed in Table 1). Then, a gentamicin-resistance cassette which was flanked by *lox66*- and *lox71*-sites [22] was cloned into the EcoRI site that was generated between the 5′ and 3′ flanking regions in pDEL3a and pDEL4a, respectively. The resulting plasmids pDEL3b and pDEL4b were then electrotransformed into *E. faecium* E1162. Marked deletion mutants were obtained by growing the gentamicin-resistant transformants as described previously [22]. The plasmid pWS3-Cre [22], carrying a gene encoding Cre recombinase, was introduced into the marked mutant by electroporation. Further culturing for the removal of the gentamicin resistance cassette and subsequent loss of pWS-Cre was performed as described previously [22]. Excision of the gentamicin resistance cassette and loss of pWS3-Cre was verified by PCR using primers listed in Table 2.

Table 2: Primers used in this study.

Primer	Sequence [a]
delete_XmaI_gltK_up_F	5′-CCCCCCGGGCCAAGCAGGTACGATTGGAT-3′
delete_EcoRI_gltK_up_R	5′-AACCGGAAAGCAGAGAATTCTCGAAAACAATGAAACTTCAACA-3′
delete_EcoRI_gltK_dn_F	5′-TCGAGAATTCTCTGCTTTCCGGTTACTTGG-3′
delete_XhoI_gltK_dn_R	5′-CCGCTCGAGGGAAGGATCACACCGATGAC-3′
check_gltK_up	5′-CGGAACGTTAATGGCAATCT-3′
check_gltK_dn	5′-CCGTACCAATCGTACCGATAA-3′
delete_XmaI_gspA_up_F	5′-CCCCCCGGGCCTCCTTTTTGGACTTTCTCG-3′
delete_EcoRI_gspA_up_R	5′-ACCACATTTAGCTGCAGAATTCGACGGCTTTCCGTTGTGTAG-3′
delete_EcoRI_gspA_dn_F	5′-CGAATTCTGCAGCTAAATGTGGTACGAA-3′
delete_XhoI_gspA_dn_R	5′-CCGCTCGAGGCCAAGTGAAAGCTTTGGAA-3′
check_gspA_up	5′-GCTCGAATTCTTCGATTGCT-3′
check_gspA_dn	5′-TGATGAGCCGTTAAATGGAA-3′
complement_BamHI_gltK_F	5′-ACGGGATCCTTTTAGCAATCGTAGCTGGTTT-3′
complement_XhoI_gltK_R	5′-ACCGCTCGAGCGTGAATTTTCAAGTGCTC-3′
complement_BamHI_gspA_F	5′-ACGGGATCCTGAAAAACCTTCGATCGTTCA-3′
complement_XhoI_gspA_R	5′-ACCGCTCGAGTCCATTCCTACTCCCCCTCT-3′
pAT392_EcoRI_lox66_genta_F	5′-AGGGAATTCTACCGTTCGTATAGCATACATTATACGAAGTTATG ATAAACCCAGCGAACCATTTGAGG-3′
pAT392_EcoRI_lox71_genta_R	5′-CTCCGAATTCTACCGTTCGTATAATGTATGCTATACGAAGTTATT CAATCTTTATAAGTCCTTTTATAA-3′

[a] Restriction sites are underlined.

In trans complemented strains of *gltK* and *gspA* gene deletion mutants were generated as described previously [22, 26]. The *gltK* and *gspA* were PCR amplified, respectively, from the genomic DNA of E1162 using the primers listed in Table 2. The PCR products were cloned into the shuttle vector pEF25 [26]. The resulting plasmids, pEF25-*gltK* and pEF25-*gspA*, were introduced into the corresponding mutant strains by electroporation as described above.

Determination of growth curves

A BioScreen C instrument (Oy Growth Curves AB, Helsinki, Finland) was used to determine the effects of bile salts on bacterial growth. Wild type, mutants and the *in trans* complemented strains were grown overnight in BHI (containing appropriate antibiotics for the *in trans* complemented strains). Cells were inoculated at an initial OD_{660} of 0.0025 into 300 µl BHI and BHI with 0.02%, 0.04%, 0.08% and 0.16% of bile salts, respectively. The cultures were incubated in the Bioscreen C system at 37°C with continuous shaking, and absorbance of 600nm (A_{600}) was recorded every 15 min for 12 hours. Each experiment was performed in triplicate.

Bile salt adaptation and challenge assays

To compare the sensitivity to bile salts of the parental strain E1162, the mutant strains and *in trans* complemented strains, overnight cultures were diluted 100 fold in fresh BHI and grown to OD_{660} 0.3. One ml of the cell cultures were harvested by centrifugation at 12500 g for 1 minute and adapted to bile salts by resuspending the cells in BHI containing sub-lethal levels of bile salts (0.02%) or in BHI without any additions. After a 15-minute adaptation period, viable counts were determined by serial dilution and plating on BHI agar plates (time point 0). Adapted and non-adapted cells were spun down as described above and resuspended in BHI containing 0.3% bile salts, a concentration that is physiologically relevant [31]. After 5, 30 and 60 minutes of incubation at 37°C, aliquots of cells were washed with PBS and viable counts were determined following serial dilution and plating on BHI agar plates. The experiment was performed in triplicate and statistical analysis of the data was performed using an unpaired two-tailed Student's *t*-test.

Results and Discussion

Identification of genetic determinants involved in bile salt resistance in E. faecium *by M-TraM*

To identify genes that are required for bile salt resistance in *E. faecium* E1162, we grew the pool of mutants in the presence or absence of a sub-lethal concentration (0.02%) of bile salts for 20 hours, and used M-TraM to determine which mutants were less resistant to bile salts and therefore selectively lost during culturing in the presence of bile salts. Seventy-five

genes belonging to a variety of functional categories were identified to be involved in bile resistance (Supplementary Table 1), in which a single gene, *gltK* (locus tag EfmE1162_1760), encoding a putative glutamate/aspartate transport system permease protein, was identified by M-TraM with the highest fold change (11.5 fold, which was saliently higher than the other identified genes), indicating that this gene may contribute considerably to bile resistance in *E. faecium*. Consequently, we decided to further study the function of this gene in bile resistance (further described below). The function of GltK and its homologues in other microorganisms in bile resistance has not been experimentally verified. BLAST analysis showed that GltK is conserved (with amino acid identities >97%) in all of the 69 *E. faecium* genomes available (on 30 October 2012) at NCBI Genomes, indicating that the *gltK* gene is part of the *E. faecium* core genome. A gene (locus tag: EfmE1162_2043) encoding a putative cardiolipin synthetase, which functions as an enzyme in phospholipid metabolism and is involved in enterococcal daptomycin resistance [32, 33], was identified as contributing to bile resistance in the M-TraM screening and possibly acts by protecting the cells from membrane-associated damage induced by bile. In *E. faecalis*, the *sagA* gene was previously shown to be important in maintaining cell wall integrity and resistance to bile [34]. The *E. faecium* homolog (locus tag: EfmE1162_2437) of the *sagA* gene was also identified by M-TraM as potentially contributing to bile resistance.

The *bsh* gene which encodes BSH that hydrolyses bile salts [21] is conserved in all the 69 publicly available *E. faecium* genomes, including E1162 (locus tag: EfmE1162_2656). However this gene was not identified by M-TraM screening, presumably because BSH does not provide protection despite its predicted activity in the hydrolysis of bile salts. It is also possible that in the M-TraM screening, during which many different transposon insertion mutants are pooled together, the minor proportion of BSH-deficient mutants could be compensated by the extracellular bile salt hydrolase activity that is produced by cells that carry other mutations. We did not identify the two Gls-like protein-encoding loci which were shown to be involved in bile resistance in a previous study [19]. However, single deletions of either locus only resulted in a minor effect on bile salt resistance possibly due to mutual compensation of the two loci [19], which may also explain why we did not identify these loci in the M-TraM screening, as the mutant library only contains mutants that are inactivated in a single locus by transposon insertion [22].

Transcriptional responses of E. faecium *to bile salt-induced stress*

A microarray-based transcriptome analysis was used to identify genes that are regulated by exposure to bile salts. Compared to the untreated control, 214 (175 up-regulated and 39 down-regulated) and 190 (119 up-regulated and 71 down-regulated) genes were identified to be differentially expressed at 5 min and 15 min incubation with bile salts, respectively (Table S2). Genes identified at either time points were grouped by COG functional categories

and the percentage abundance of each group was compared to the overall COG-based composition of the E1162 genome [23] (Figure 1). Genes in COG categories F (nucleotide transport and metabolism) and H (coenzyme transport and metabolism) were overrepresented among the down-regulated genes during exposure to bile salts, possibly reflecting a general energy-saving strategy by the cell during its rapid adaptation to bile stress. No genes from these categories exhibited up-regulated expression during bile salt exposure, indicating that the decreased expression of genes in these two functional categories is a significant transcriptional response of *E. faecium* to bile salts. Among the up-regulated genes, genes in COG category G (carbohydrate transport and metabolism) were overrepresented at 5 min. after exposure to bile salts. These included genes that are predicted to encode proteins involved in the utilization of a variety of sugars including maltose, maltodextrin, cellobiose, galactose, fructose, mannose and lactose. Interestingly, no genes from COG category G were down-regulated after 5 min. This observation indicates that bile salts positively impact on expression of genes involved in carbohydrate transport and metabolism immediately after exposure to bile salts, but this response becomes less prominent after longer periods of time. We also found a number of categories that were enriched in up-regulated genes, including C (energy production and conversion), E (amino acid transport and metabolism), O (posttranslational modification, protein turnover; chaperones) and Q (secondary metabolites biosynthesis, transport and catabolism). These data suggest an involvement of these functional categories in the *E. faecium* response to bile salts.

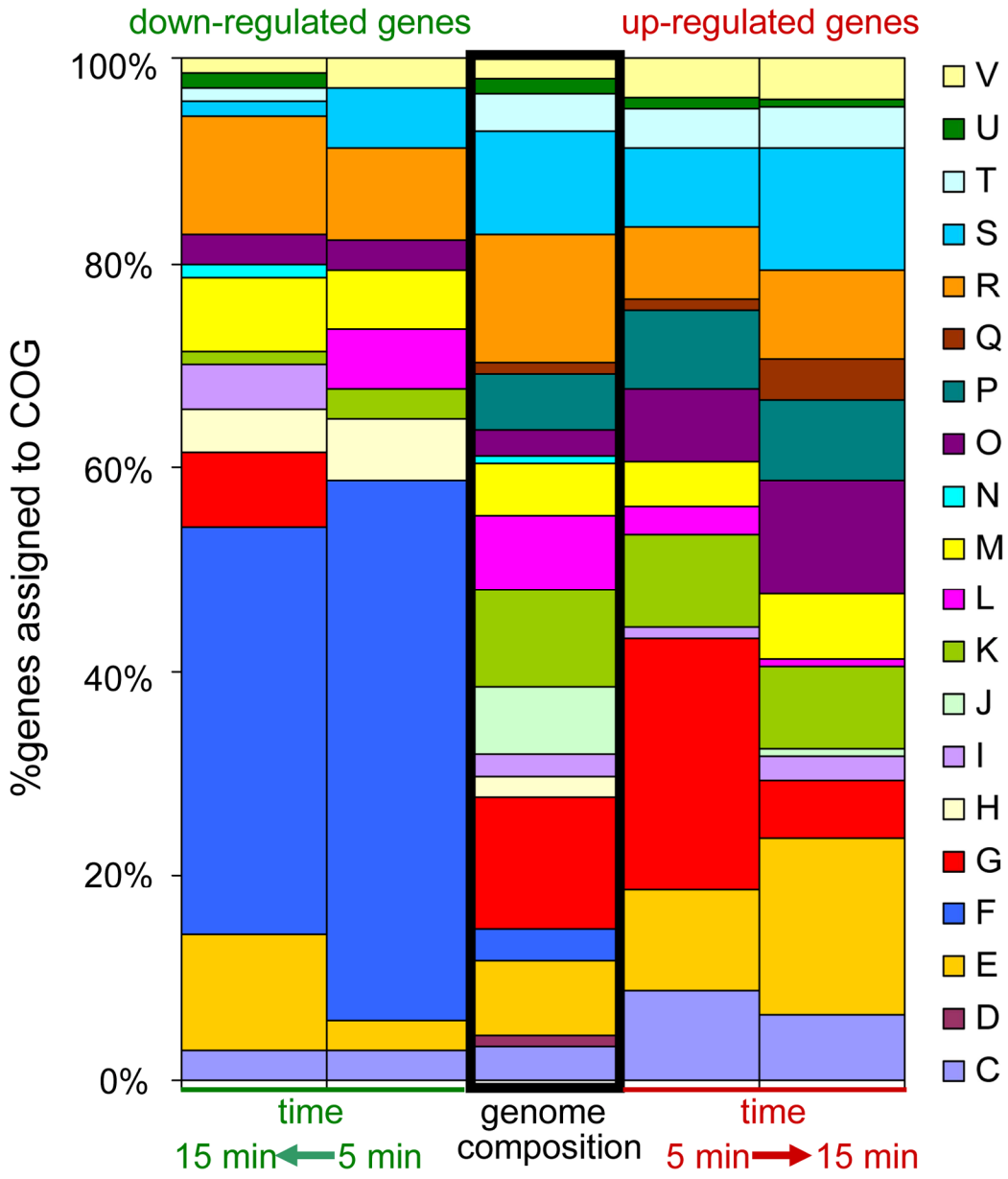


Figure 1. COG classification of differentially expressed genes during exposure to bile salts. The down-regulated (left panel) and up-regulated (right panel) genes during bile treatment were grouped by COG functional categories, respectively, and the percentage abundance of each group was compared to abundance of each COG in the E1162 genome. The one-letter codes represent the following COG functional categories: C: energy production and conversion; D: cell cycle control, cell division, chromosome partitioning; E: amino acid transport and metabolism; F: nucleotide transport and metabolism; G: carbohydrate transport and metabolism; H: coenzyme transport and metabolism; I: lipid transport and metabolism; J: translation, ribosomal structure and biogenesis; K: transcription; L: replication, recombination and repair; M: cell wall/membrane/envelope biogenesis; N: cell motility; O: posttranslational modification, protein turnover; chaperones; P: inorganic ion transport and metabolism; Q: secondary metabolites biosynthesis, transport and catabolism; R: general function prediction only; S: function unknown; T: signal transduction mechanisms; U: intracellular trafficking, secretion, and vesicular transport; V: defense mechanisms.

We further focused on a gene EfmE1162_1186 (*gspA*) which is predicted to encode a general stress protein A. This gene was identified by both transcriptome analysis (4.6 and 47.0 fold up-regulated at 5 min and 15 min of bile salts treatment, respectively) and M-TraM (2.8 fold less signal in bile-exposed library compared to the control). GspA is also highly conserved (with amino acid identities >98%) in 66 of the 69 *E. faecium* genomes. We observed that both of the two Gls-like protein-encoding loci (EfmE1162_1192-EfmE1162_1193 and EfmE1162_1201-EfmE1162_1202) were induced over eight-fold during exposure to bile salts, although they were not identified by M-TraM screening. However, the *bsh* gene was not identified to be differentially expressed in BHI with bile salts, indicating that the expression of this gene is not regulated by bile salts despite its role in bile salt hydrolysis.

The transcriptional responses of *E. faecalis* to bovine bile has been investigated in a previous study [35]. A striking common finding of this study and our work is that a large gene cluster (locus tags EfmE1162_0724-EfmE1162_0731 in *E. faecium* E1162 and EF1492-EF1500 in *E. faecalis* V583), which putatively encodes a V-type ATPase, exhibits upregulated expression during exposure to bile salts. V-type ATPases are membrane proteins that function as proton- or sodium ion pumps that build up ion gradients at the expense of ATP [36]. Induction of this gene cluster suggested that *E. faecium* may generate a proton gradient to respond to bile mediated stress. The link between bile mediated stress and maintenance of the proton motive force (PMF) was previously demonstrated in *Lactobacillus plantarum* [37], *Bifidobacterium longum* [38] and *B. animalis* [39]. Bile salts can induce DNA damage in bacteria, and consequently DNA mismatch repair proteins are important for bacterial bile resistance [12, 40, 41]. In this study we identified a gene (locus tag: EfmE1162_1335), encoding the DNA mismatch repair protein MutS, that was higher expressed (23.0 fold at 5 min and 9.5 fold at 15 min) after addition of bile salts to the culture medium.

Effect of bile salts on growth of E. faecium *E1162 wild-type and the mutant strains.*

To determine the role of GltK and GspA in bile salt resistance, markerless deletion mutants in *gltK* and *gspA* were constructed in *E. faecium* E1162, and the mutants were complemented *in trans*. The growth of *E. faecium* E1162 wild type (WT), the isogenic mutants and the complemented strains in BHI and BHI supplemented with bile salts were determined. In the absence of bile salts the wild-type strain and its isogenic mutants grew identically (data not shown). When these strains were grown in BHI with 0.02%, 0.04%, 0.08% and 0.16% bile salts, the growth rate of the Δ*gltK* mutant was decreased compared to WT and this difference was most salient in BHI with 0.08% bile salts. The growth rate of the Δ*gltK* mutant could be restored to WT levels upon *in trans* complementation (Figure 2), indicating that GltK contributes to bile resistance of *E. faecium*. Only very minor effects on growth rate and optical density in stationary phase were observed due to the deletion of *gspA,* indicating that this gene has an insignificant role in bile resistance of *E. faecium*. In

addition, the sensitivity of both mutants to other stresses, including different pHs, oxidative and osmotic stress, were examined and no significant difference between the mutants and wild-type strain was detected (data not shown).

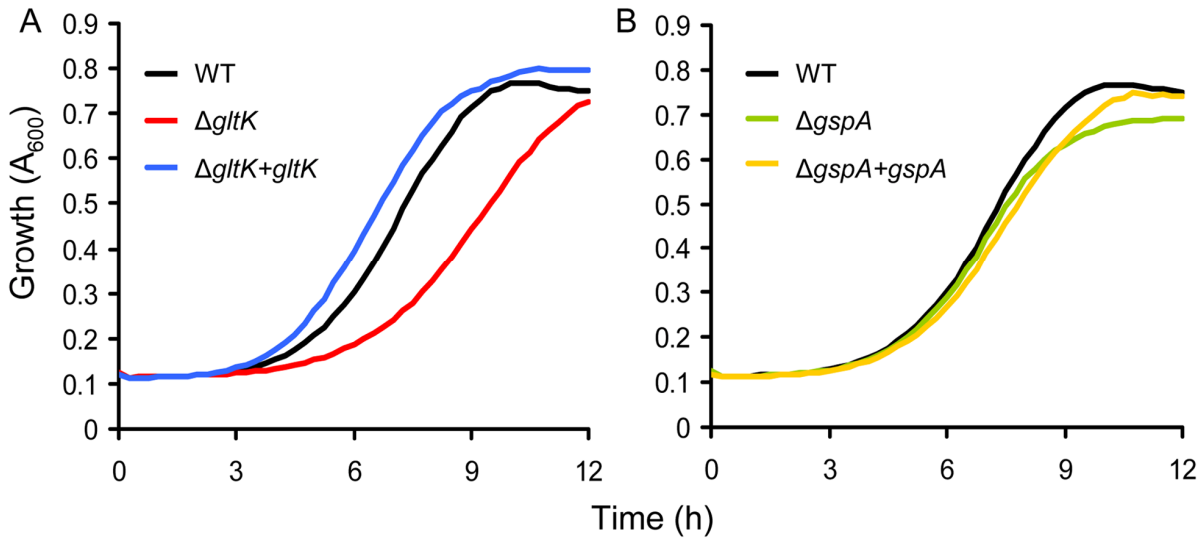


Figure 2. Growth of *E. faecium* in BHI with 0.08% bile salts. Overnight cultures of *E. faecium* strains were inoculated at an initial cell density of OD_{660} 0.0025 in BHI with 0.08% bile salts. Growth curves of wild-type E1162, the two mutants (panel A: Δ*gltK*, and panel B: Δ*gspA*) and *in trans* complemented strains are shown. Growth curves are the means of three independent experiments.

Bile salt adaptation and challenge assays

Exponential-phase cells of E1162 wild-type, Δ*gltK* mutant and *in trans* complemented strain were adapted to 0.02% bile salts for 15 min or were left unadapted, and then challenged with 0.3% bile salts (Figure 3). Viable counts for the unadapted cells dropped below the detection limit (<50 CFU/ml) after this challenge, indicating that these cells were sensitive to this high concentration of bile salts. In contrast, the adapted cells were more tolerant to 0.3% bile salts, with 10^4 CFU/ml surviving after 5 min of being exposed to bile salts and no significant further killing occurring during the remainder of the 1-hour experiment. These results showed that adaptation to low levels of bile salts provided *E. faecium* substantial protection to levels of bile salts that are lethal to non-adapted cells. The deletion of *gltK* reduced the protection provided by the adaptation to a sub-lethal concentration of bile salts, leading to an approximately 1-log lower survival of pre-adapted *gltK* cells than the survival of the wild type E1162 cells. Survival of the *in trans* complemented strain upon pre-adaptation to bile salts was similar to that of the wild type. The Δ*gspA* mutant was also included in this assay, but no significant difference was observed compared to wild-type E1162 (data not shown), again indicating that *gspA* was not required for bile resistance although its expression was highly induced by bile salts. These results suggest that the bile

salt-regulated genes do not necessarily contribute to bile resistance. Previous studies indicated that the protective adaptation to bile salts mainly results from changes in membrane composition and architecture that are independent of *de novo* protein synthesis [42, 43]. Flahaut *et al*. showed that a 5 second-adaptation of *E. faecalis* to low level bile salts could provide substantial protection against challenge with lethal bile salt concentrations, and the addition of chloramphenicol during the adaptation period did not prevent development of acquired tolerance [43]. A similar result was also observed in *L. monocytogenes* [42]. However, the bile salt-regulated genes, rather than directly contributing to bile resistance, could be involved in other functions including virulence and carbohydrate metabolism [10]. It has previously been established in *Salmonella* [44, 45] and *Vibrio* [46-48] that bile can be used as an environmental cue to influence the regulation genes involved in intestinal colonization and virulence. We identified many genes involved in carbohydrate metabolism that exhibited upregulated expression upon exposure to bile salts, e.g. a gene cluster (locus tags: EfmE1162_1484 - EfmE1162_1489) putatively involved in maltose utilization (Zhang *et al*., unpublished work). This may suggest that *E. faecium* senses bile as an environmental signal indicating that it has entered the host gut, leading to a rapid adjustment of the cell's carbohydrate metabolism.

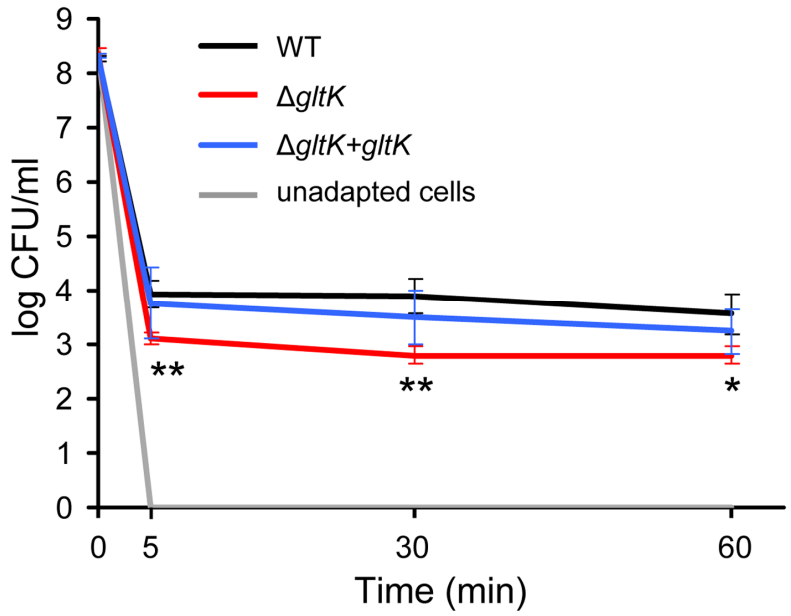


Figure 3. Survival of *E. faecium* cells following exposure to 0.3% bile salts with and without pre-adaptation. Exponential-phase cells of the E1162 wild type, the Δ*gltK* mutant and its *in trans* complemented strain Δ*gltK+gltK* were adapted to 0.02% bile salts for 15 min or were left unadapted, and then challenged with 0.3% bile salts. Viable cells were determined immediately before exposure to 0.3% bile salts (t = 0 min) and after 5, 30 and 60 min of challenge. The grey line represents unadapted cultures of these strains, in which viable counts dropped to undetectable levels (<50 CFU/ml) after being challenged by bile salts. Bars represent the standard deviation of the mean of three independent experiments. Asterisks represent significant differences (*: P<0.05, **: P<0.01, as determined by an unpaired two-tailed Student's *t*-test) between the Δ*gltK* mutant and the wild type strain.

Responding and being resistant to bile are important features of bacteria that inhabit the gut [10]. In the present work, we have identified a genetic determinant in *E. faecium* that

contributes to bile salts resistance, and studied the transcriptional response of *E. faecium* to bile salts. These findings add to a better understanding of the molecular mechanisms that lead to bile resistance in *E. faecium*.

References

1. Top J, Willems R, Bonten M: Emergence of CC17 *Enterococcus faecium*: from commensal to hospital-adapted pathogen. *FEMS Immunol Med Microbiol* 2008, 52(3):297-308.
2. Sghir A, Gramet G, Suau A, Rochet V, Pochart P, Dore J: Quantification of bacterial groups within human fecal flora by oligonucleotide probe hybridization. *Appl Environ Microbiol* 2000, 66(5):2263-2266.
3. Willems RJ, Top J, van Schaik W, Leavis H, Bonten M, Siren J, Hanage WP, Corander J: Restricted Gene Flow among Hospital Subpopulations of *Enterococcus faecium*. *MBio* 2012, 3(4).
4. Willems RJ, van Schaik W: Transition of *Enterococcus faecium* from commensal organism to nosocomial pathogen. *Future Microbiol* 2009, 4(9):1125-1135.
5. Arias CA, Murray BE: The rise of the *Enterococcus*: beyond vancomycin resistance. *Nat Rev Microbiol* 2012, 10(4):266-278.
6. Facklam RR, Collins MD: Identification of *Enterococcus* species isolated from human infections by a conventional test scheme. *J Clin Microbiol* 1989, 27(4):731-734.
7. Bradley CR, Fraise AP: Heat and chemical resistance of enterococci. *J Hosp Infect* 1996, 34(3):191-196.
8. Freeman R, Kearns AM, Lightfoot NF: Heat resistance of nosocomial enterococci. *Lancet* 1994, 344(8914):64-65.
9. Kearns AM, Freeman R, Lightfoot NF: Nosocomial enterococci: resistance to heat and sodium hypochlorite. *J Hosp Infect* 1995, 30(3):193-199.
10. Begley M, Gahan CG, Hill C: The interaction between bacteria and bile. *FEMS Microbiol Rev* 2005, 29(4):625-651.
11. Esteller A: Physiology of bile secretion. *World J Gastroenterol* 2008, 14(37):5641-5649.
12. Merritt ME, Donaldson JR: Effect of bile salts on the DNA and membrane integrity of enteric bacteria. *J Med Microbiol* 2009, 58(Pt 12):1533-1541.
13. Hofmann AF, Hagey LR: Bile acids: chemistry, pathochemistry, biology, pathobiology, and therapeutics. *Cell Mol Life Sci* 2008, 65(16):2461-2483.
14. Begley M, Sleator RD, Gahan CG, Hill C: Contribution of three bile-associated loci, *bsh*, *pva*, and *btlB*, to gastrointestinal persistence and bile tolerance of *Listeria monocytogenes*. *Infect Immun* 2005, 73(2):894-904.
15. Kristich CJ, Wells CL, Dunny GM: A eukaryotic-type Ser/Thr kinase in *Enterococcus faecalis* mediates antimicrobial resistance and intestinal persistence. *Proc Natl Acad Sci U S A* 2007, 104(9):3508-3513.
16. Reynolds MM, Bogomolnaya L, Guo J, Aldrich L, Bokhari D, Santiviago CA, McClelland M, Andrews-Polymenis H: Abrogation of the twin arginine transport system in *Salmonella enterica* serovar Typhimurium leads to colonization defects during infection. *PLoS One* 2011, 6(1):e15800.
17. De Paepe M, Gaboriau-Routhiau V, Rainteau D, Rakotobe S, Taddei F, Cerf-Bensussan N: Trade-off between bile resistance and nutritional competence drives *Escherichia coli* diversification in the mouse gut. *PLoS Genet* 2011, 7(6):e1002107.
18. Teng F, Nannini EC, Murray BE: Importance of *gls24* in virulence and stress response of *Enterococcus faecalis* and use of the Gls24 protein as a possible immunotherapy target. *J Infect Dis* 2005, 191(3):472-480.
19. Choudhury T, Singh KV, Sillanpaa J, Nallapareddy SR, Murray BE: Importance of two *Enterococcus faecium* loci encoding Gls-like proteins for in vitro bile salts stress response and virulence. *J Infect Dis* 2011, 203(8):1147-1154.
20. Franz CM, Specht I, Haberer P, Holzapfel WH: Bile salt hydrolase activity of Enterococci isolated from food: screening and quantitative determination. *J Food Prot* 2001, 64(5):725-729.

21. Wijaya A, Hermann A, Abriouel H, Specht I, Yousif NM, Holzapfel WH, Franz CM: Cloning of the bile salt hydrolase (*bsh*) gene from *Enterococcus faecium* FAIR-E 345 and chromosomal location of *bsh* genes in food enterococci. *J Food Prot* 2004, 67(12):2772-2778.
22. Zhang X, Paganelli FL, Bierschenk D, Kuipers A, Bonten MJ, Willems RJ, van Schaik W: Genome-Wide Identification of Ampicillin Resistance Determinants in *Enterococcus faecium*. *PLoS Genet* 2012, 8(6):e1002804.
23. van Schaik W, Top J, Riley DR, Boekhorst J, Vrijenhoek JE, Schapendonk CM, Hendrickx AP, Nijman IJ, Bonten MJ, Tettelin H *et al*: Pyrosequencing-based comparative genome analysis of the nosocomial pathogen *Enterococcus faecium* and identification of a large transferable pathogenicity island. *BMC Genomics* 2010, 11:239.
24. Leenhouts K, Buist G, Bolhuis A, ten Berge A, Kiel J, Mierau I, Dabrowska M, Venema G, Kok J: A general system for generating unlabelled gene replacements in bacterial chromosomes. *Mol Gen Genet* 1996, 253(1-2):217-224.
25. Zhang X, Vrijenhoek JE, Bonten MJ, Willems RJ, van Schaik W: A genetic element present on megaplasmids allows *Enterococcus faecium* to use raffinose as carbon source. *Environ Microbiol* 2011, 13(2):518-528.
26. Top J, Sinnige JC, Majoor EA, Bonten MJ, Willems RJ, van Schaik W: The recombinase IntA is required for excision of *esp*-containing ICE*Efm1* in *Enterococcus faecium*. *J Bacteriol* 2011, 193(4):1003-1006.
27. Baldi P, Long AD: A Bayesian framework for the analysis of microarray expression data: regularized *t*-test and statistical inferences of gene changes. *Bioinformatics* 2001, 17(6):509-519.
28. Lebreton F, van Schaik W, Sanguinetti M, Posteraro B, Torelli R, Le Bras F, Verneuil N, Zhang X, Giard JC, Dhalluin A *et al*: AsrR Is an Oxidative Stress Sensing Regulator Modulating *Enterococcus faecium* Opportunistic Traits, Antimicrobial Resistance, and Pathogenicity. *PLoS Pathog* 2012, 8(8):e1002834.
29. Hsiao A, Ideker T, Olefsky JM, Subramaniam S: VAMPIRE microarray suite: a web-based platform for the interpretation of gene expression data. *Nucleic Acids Res* 2005, 33(Web Server issue):W627-632.
30. Sauer B: Functional expression of the *cre-lox* site-specific recombination system in the yeast *Saccharomyces cerevisiae*. *Mol Cell Biol* 1987, 7(6):2087-2096.
31. Van Deest BW, Fordtran JS, Morawski SG, Wilson JD: Bile salt and micellar fat concentration in proximal small bowel contents of ileectomy patients. *J Clin Invest* 1968, 47(6):1314-1324.
32. Palmer KL, Daniel A, Hardy C, Silverman J, Gilmore MS: Genetic basis for daptomycin resistance in enterococci. *Antimicrob Agents Chemother* 2011, 55(7):3345-3356.
33. Arias CA, Panesso D, McGrath DM, Qin X, Mojica MF, Miller C, Diaz L, Tran TT, Rincon S, Barbu EM *et al*: Genetic basis for *in vivo* daptomycin resistance in enterococci. *N Engl J Med* 2011, 365(10):892-900.
34. Breton YL, Maze A, Hartke A, Lemarinier S, Auffray Y, Rince A: Isolation and characterization of bile salts-sensitive mutants of *Enterococcus faecalis*. *Curr Microbiol* 2002, 45(6):434-439.
35. Solheim M, Aakra A, Vebo H, Snipen L, Nes IF: Transcriptional responses of *Enterococcus faecalis* V583 to bovine bile and sodium dodecyl sulfate. *Appl Environ Microbiol* 2007, 73(18):5767-5774.
36. Senior AE: The proton-translocating ATPase of *Escherichia coli*. *Annu Rev Biophys Biophys Chem* 1990, 19:7-41.
37. Bron PA, Molenaar D, de Vos WM, Kleerebezem M: DNA micro-array-based identification of bile-responsive genes in *Lactobacillus plantarum*. *J Appl Microbiol* 2006, 100(4):728-738.
38. Sanchez B, Champomier-Verges MC, Anglade P, Baraige F, de Los Reyes-Gavilan CG, Margolles A, Zagorec M: Proteomic analysis of global changes in protein expression during bile salt exposure of *Bifidobacterium longum* NCIMB 8809. *J Bacteriol* 2005, 187(16):5799-5808.

39. Sanchez B, de los Reyes-Gavilan CG, Margolles A: The F1F0-ATPase of *Bifidobacterium animalis* is involved in bile tolerance. *Environ Microbiol* 2006, 8(10):1825-1833.
40. Payne A, Schmidt TB, Nanduri B, Pendarvis K, Pittman JR, Thornton JA, Grissett J, Donaldson JR: Proteomic analysis of the response of *Listeria monocytogenes* to bile salts under anaerobic conditions. *J Med Microbiol* 2012.
41. Prieto AI, Ramos-Morales F, Casadesus J: Bile-induced DNA damage in *Salmonella enterica*. *Genetics* 2004, 168(4):1787-1794.
42. Begley M, Gahan CG, Hill C: Bile stress response in *Listeria monocytogenes* LO28: adaptation, cross-protection, and identification of genetic loci involved in bile resistance. *Appl Environ Microbiol* 2002, 68(12):6005-6012.
43. Flahaut S, Hartke A, Giard JC, Benachour A, Boutibonnes P, Auffray Y: Relationship between stress response toward bile salts, acid and heat treatment in *Enterococcus faecalis*. *FEMS Microbiol Lett* 1996, 138(1):49-54.
44. Prouty AM, Brodsky IE, Manos J, Belas R, Falkow S, Gunn JS: Transcriptional regulation of *Salmonella enterica* serovar Typhimurium genes by bile. *FEMS Immunol Med Microbiol* 2004, 41(2):177-185.
45. Prouty AM, Gunn JS: *Salmonella enterica* serovar typhimurium invasion is repressed in the presence of bile. *Infect Immun* 2000, 68(12):6763-6769.
46. Krukonis ES, DiRita VJ: From motility to virulence: Sensing and responding to environmental signals in *Vibrio cholerae*. *Curr Opin Microbiol* 2003, 6(2):186-190.
47. Schuhmacher DA, Klose KE: Environmental signals modulate ToxT-dependent virulence factor expression in *Vibrio cholerae*. *J Bacteriol* 1999, 181(5):1508-1514.
48. Gupta S, Chowdhury R: Bile affects production of virulence factors and motility of *Vibrio cholerae*. *Infect Immun* 1997, 65(3):1131-1134.

Supplementary data

Supplementary Table 1. *E. faecium* genes involved in bile salt resistance as determined by M-TraM analysis

LocusTag	Annotation	M-TraM	COG	Expression	
				5 min	15 min
EfmE1162_1760	glutamate/aspartate transport system permease protein GltK	11.54	E	-	-
EfmE1162_1642	NAD+ synthetase	7.98	H	-	-
EfmE1162_0824	conserved hypothetical protein	4.91	-	-	-
EfmE1162_1905	phosphopantothenoylcysteine decarboxylase	4.65	H	-	-
EfmE1162_1575	ribosomal protein S12	4.25	J	-	-
EfmE1162_2043	cardiolipin synthetase	4.02	I	-	-
EfmE1162_1851	conserved hypothetical protein	3.93	NT	-	-
EfmE1162_0041	formate--tetrahydrofolate ligase	3.91	F	-	-
EfmE1162_2174	permease	3.82	P	-	-
EfmE1162_1873	glutaredoxin	3.76	-	0.47	0.39
EfmE1162_0252	O-methyltransferase family protein	3.74	R	-	-
EfmE1162_1632	ABC transporter, iron chelate uptake transporter	3.64	P	-	-
EfmE1162_0116	GatB/Yqey domain protein	3.63	S	-	-
EfmE1162_1025	cell division ATP-binding protein FtsE	3.60	D	-	-
EfmE1162_1999	protein YjjG	3.54	J	-	-
EfmE1162_1957	hypothetical protein	3.49	-	-	-
EfmE1162_1853	inner membrane symporter YgjU	3.37	E	-	-
EfmE1162_0323	DNA topoisomerase III	3.37	L	-	-
EfmE1162_0276	conserved hypothetical protein	3.28	S	-	-
EfmE1162_0620	UDP-N-acetylmuramate--alanine ligase	3.21	M	-	-
EfmE1162_1073	hypothetical protein	3.18	-	-	-
EfmE1162_0770	conserved hypothetical protein	3.18	-	-	-
EfmE1162_0615	dephospho-CoA kinase	2.82	HH	-	-
EfmE1162_1186	general stress protein A	2.79	M	4.64	47.02
EfmE1162_1277	metal-dependent hydrolase	2.77	R	-	-
EfmE1162_2515	DNA mismatch repair protein MutL	2.71	L	-	-
EfmE1162_0509	YycH protein	2.62	S	-	-
EfmE1162_2437	SagA	2.59	S	0.19	-
EfmE1162_1460	conserved hypothetical protein	2.51	S	5.44	-
EfmE1162_1505	glycosyltransferase	2.50	M	-	-
EfmE1162_0302	ATPase, ParA family protein	2.43	D	-	-
EfmE1162_0731	V-type Na-ATPase	2.33	C	3.39	-
EfmE1162_1100	PTS system, IIC component	0.50	G	-	-
EfmE1162_2034	transcriptional regulator, GntR family	0.39	K	-	-
EfmE1162_0167	N-acetylmuramic acid 6-phosphate etherase	0.39	R	-	-
EfmE1162_0795	ABC transporter, ATP-binding protein	0.37	R	-	-
EfmE1162_2036	beta-glucosidase A	0.35	G	3.72	-
EfmE1162_0640	two-component response regulator	0.35	T	-	-
EfmE1162_1639	transcriptional regulator, MerR family	0.35	K	-	-
EfmE1162_0792	membrnae protein, puative	0.34	R	-	-
EfmE1162_1271	von Willebrand factor type A domain protein	0.33	PO	-	-
EfmE1162_2152	helix-turn-helix motif	0.33	K	-	-
EfmE1162_1190	GapA	0.33	-	-	13.58
EfmE1162_0449	voltage-gated chloride channel family protein	0.33	P	-	-

LocusTag	Annotation	M-TraM	COG	Expression	
				5 min	15 min
EfmE1162_0410	aminopeptidase C	0.32	E	-	3.74
EfmE1162_2282	beta-lactamase	0.32	V	-	-
EfmE1162_2234	LysM domain protein	0.31	-	0.15	-
EfmE1162_2687	putative multidrug export ATP-binding/permease protein	0.31	V	-	-
EfmE1162_0310	addiction module toxin, Txe/YoeB family	0.30	S	-	-
EfmE1162_1509	coenzyme A disulfide reductase	0.30	R	-	-
EfmE1162_0477	thioredoxin family protein	0.30	O	-	-
EfmE1162_1160	transcriptional regulator	0.30	K	-	-
EfmE1162_1818	tRNA-dihydrouridine synthase	0.30	J	-	-
EfmE1162_2672	ribosomal large subunit pseudouridine synthase B	0.30	J	-	-
EfmE1162_1756	CBS domain protein	0.29	R	-	-
EfmE1162_1483	hypothetical protein	0.29	-	-	-
EfmE1162_1465	transcriptional regulator, PadR family	0.29	K	-	-
EfmE1162_2361	NAD-dependent deacetylase	0.29	K	-	-
EfmE1162_0918	hypothetical protein	0.28	-	-	-
EfmE1162_2626	cadmium-translocating P-type ATPase	0.28	P	-	-
EfmE1162_0552	altronate hydrolase	0.27	G	7.41	-
EfmE1162_1199	GapA	0.27	-	-	9.00
EfmE1162_0776	phage antirepressor protein	0.27	K	-	-
EfmE1162_2399	hypothetical protein	0.27	-	-	-
EfmE1162_0850	D-xylulose 5-phosphate/D-fructose 6-phosphate phosphoketolase	0.26	G	-	-
EfmE1162_2053	2,5-diketo-d-gluconic acid reductase a	0.26	R	-	-
EfmE1162_0817	arginine deiminase pathway transcriptional regulator, Crp family	0.26	T	-	-
EfmE1162_0558	hypothetical protein	0.24	S	-	-
EfmE1162_1550	putative phosphoglucomutase	0.23	R	-	-
EfmE1162_2483	inner membrane protein	0.23	G	-	-
EfmE1162_1958	phosphosugar-binding transcriptional regulator, RpiR family	0.23	K	-	-
EfmE1162_0873	magnesium chelatase, subunit ChII family protein	0.22	R	-	-
EfmE1162_0703	phosphopyruvate hydratase	0.21	G	-	-
EfmE1162_1508	coenzyme A disulfide reductase	0.17	P	-	-
EfmE1162_1033	phosphate transport system regulatory protein PhoU	0.17	P	-	-

Supplementary Table 2. Expression ratios of *E. faecium* E1162 genes that exhibit significant differences in expression at 5 min and/or 15 min of exposure to 0.02% bile salts

LocusTag	Annotation	COG	Expression	
			5 min	15 min
EfmE1162_0005	hypothetical protein	-	6.31	17.17
EfmE1162_0009	cysteine synthase A	E	-	3.66
EfmE1162_0051	carbamoyl-phosphate synthase, large subunit	EF	0.43	0.37
EfmE1162_0052	dihydroorotate dehydrogenase electron transfer subunit	HC	0.38	0.34
EfmE1162_0053	dihydroorotate dehydrogenase B, catalytic subunit	F	0.33	0.30
EfmE1162_0054	orotidine 5'-phosphate decarboxylase	F	0.30	0.29
EfmE1162_0055	orotate phosphoribosyltransferase	F	0.36	0.26
EfmE1162_0078	UDP-glucose 4-epimerase	M	2.77	-
EfmE1162_0079	galactose-1-phosphate uridylyltransferase	G	3.17	-
EfmE1162_0080	galactose operon repressor GalR	K	4.25	-
EfmE1162_0143	diacylglycerol kinase catalytic domain protein	IR	3.50	-
EfmE1162_0153	hypothetical protein	L	2.98	-
EfmE1162_0154	hypothetical protein	-	3.49	-
EfmE1162_0156	integral membrane protein	H	0.37	0.18
EfmE1162_0162	ABC transporter protein	V	4.65	3.90
EfmE1162_0163	membrane spanning protein	R	4.19	3.82
EfmE1162_0188	beta-glucosidase A	G	4.83	-
EfmE1162_0189	alanine racemase	K	3.43	-
EfmE1162_0191	deoxyguanosIne kinase	F	-	0.12
EfmE1162_0192	nucleoside 2-deoxyribosyltransferase	F	-	0.11
EfmE1162_0193	conserved hypothetical protein	R	0.33	0.07
EfmE1162_0194	hypothetical protein	-	0.35	-
EfmE1162_0223	transcriptional regulator Spx	P	4.32	10.39
EfmE1162_0224	regulatory protein spx	P	4.97	13.29
EfmE1162_0228	deoxyguanosIne kinase	F	-	0.45
EfmE1162_0230	accessory gene regulator protein B, putative	OTK	2.43	6.09
EfmE1162_0231	hypothetical protein	-	-	4.79
EfmE1162_0242	high affinity ribose transport protein RbsD	G	14.96	-
EfmE1162_0243	pts system beta-glucoside-specific eiibca component	GG	14.27	-
EfmE1162_0259	DltD protein	M	4.44	3.31
EfmE1162_0260	D-alanyl carrier protein	IQ	5.09	3.69
EfmE1162_0261	protein DltB	M	5.86	4.21
EfmE1162_0262	D-alanine-activating enzyme	Q	6.69	3.85
EfmE1162_0263	putative D-Ala-teichoic acid biosynthesis protein	-	6.94	4.04
EfmE1162_0266	anaerobic ribonucleoside-triphosphate reductase	F	0.47	0.18
EfmE1162_0267	anaerobic ribonucleoside-triphosphate reductase activating protein	O	-	0.21
EfmE1162_0278	conserved hypothetical protein	S	6.43	8.84
EfmE1162_0279	conserved hypothetical protein	-	6.13	9.16
EfmE1162_0280	conserved hypothetical protein	R	6.16	7.77
EfmE1162_0293	CTP synthase	F	0.47	0.31
EfmE1162_0295	glutamine synthetase, type I	E	-	0.21
EfmE1162_0296	HTH-type transcriptional regulator GlnR	K	-	0.29
EfmE1162_0355	pyrimidine nucleoside transport protein	F	0.28	0.25
EfmE1162_0356	guanylate kinase	F	0.38	-
EfmE1162_0367	maltose/maltodextrin ABC transporter, permease protein	G	3.44	-
EfmE1162_0371	YitT family protein	S	3.29	4.97
EfmE1162_0372	transcriptional regulator, MerR family	S	3.76	-
EfmE1162_0386	penicillin-binding protein 3	M	-	0.46
EfmE1162_0389	enoyl-[acyl-carrier-protein] reductase	I	-	0.39

LocusTag	Annotation	COG	Expression	
			5 min	15 min
EfmE1162_0410	aminopeptidase C	E	-	3.74
EfmE1162_0433	ATP-dependent Clp protease, proteolytic subunit ClpP	OU	4.36	8.04
EfmE1162_0446	hypothetical protein	-	9.61	7.97
EfmE1162_0455	membrane protein OxaA 2	U	2.30	-
EfmE1162_0472	negative regulator of genetic competence ClpC/mecB	O	2.81	4.50
EfmE1162_0473	transcriptional regulator CtsR	K	3.32	4.33
EfmE1162_0482	glutamine-fructose-6-phosphate transaminase	M	-	0.47
EfmE1162_0491	hypothetical protein	-	4.66	23.40
EfmE1162_0492	hypothetical protein	-	3.95	21.15
EfmE1162_0499	conserved hypothetical protein	-	13.02	15.04
EfmE1162_0502	chaperonin GroL	O	2.38	4.21
EfmE1162_0503	chaperonin GroS	O	2.76	3.81
EfmE1162_0514	CBS domain protein	R	2.24	-
EfmE1162_0543	glycosyl hydrolase, family 38	G	3.16	-
EfmE1162_0544	fructokinase	KG	2.88	-
EfmE1162_0546	penicillin-binding protein 2A	M	7.11	5.16
EfmE1162_0547	hypothetical protein	S	3.07	-
EfmE1162_0552	altronate hydrolase	G	7.41	-
EfmE1162_0553	altronate oxidoreductase	G	8.85	-
EfmE1162_0554	transcriptional regulator	K	2.66	-
EfmE1162_0665	PTS system, fructose-specific family, IIABC components	GGTG	15.50	3.62
EfmE1162_0666	1-phosphofructokinase	G	18.28	3.09
EfmE1162_0667	lactose phosphotransferase system repressor	KG	16.68	-
EfmE1162_0674	conserved hypothetical protein	S	31.60	17.66
EfmE1162_0724	V-type ATPase, D subunit	C	2.12	-
EfmE1162_0725	V-type ATP synthase beta chain	C	2.82	3.42
EfmE1162_0726	V-type ATP synthase alpha chain	C	3.04	3.45
EfmE1162_0727	ATP synthase (F/14-kDa) subunit	C	3.26	3.22
EfmE1162_0728	ATP synthase (C/AC39) subunit	C	3.50	3.12
EfmE1162_0729	V-type sodium ATP synthase subunit E	-	3.52	-
EfmE1162_0730	V-type sodium ATP synthase subunit K	C	3.45	2.93
EfmE1162_0731	V-type Na-ATPase	C	3.39	-
EfmE1162_0732	hypothetical protein	-	3.50	-
EfmE1162_0745	transcriptional regulator	K	2.77	4.06
EfmE1162_0746	lipoprotein-releasing system ATP-binding protein LolD	V	3.28	5.59
EfmE1162_0747	antimicrobial peptide ABC transporter permease	V	3.69	6.21
EfmE1162_0814	carbamate kinase	E	4.88	7.63
EfmE1162_0815	ornithine carbamoyltransferase	E	5.17	7.47
EfmE1162_0816	arginine deiminase	E	4.94	8.33
EfmE1162_0827	glutamine ABC transporter, permease/substrate-binding protein	ETE	-	0.21
EfmE1162_0828	glutamine transport ATP-binding protein GlnQ	E	-	0.23
EfmE1162_0831	chaperone protein DnaJ	O	-	3.47
EfmE1162_0832	chaperone protein DnaK	O	2.79	4.57
EfmE1162_0833	co-chaperone GrpE	O	3.82	5.24
EfmE1162_0834	heat-inducible transcription repressor HrcA	K	3.86	5.07
EfmE1162_0844	glucose 1-dehydrogenase	IQR	-	4.14
EfmE1162_0862	phosphatidate cytidylyltransferase	I	-	0.38
EfmE1162_0863	di-trans,poly-cis-decaprenylcistransferase	I	-	0.41
EfmE1162_0901	hypothetical protein	-	2.12	-
EfmE1162_0932	hypothetical protein	-	3.20	-
EfmE1162_0959	deoxyguanosIne kinase	F	0.47	0.22
EfmE1162_0980	emap domain protein	R	-	0.36

LocusTag	Annotation	COG	Expression	
			5 min	15 min
EfmE1162_0981	hypothetical serine protease YyxA	O	14.00	8.60
EfmE1162_1000	glyceraldehyde-3-phosphate dehydrogenase, type I	G	-	3.47
EfmE1162_1002	amino acid permease family protein	E	-	0.32
EfmE1162_1019	cellobiose-specific phosphotransferase enzyme iib component	G	3.17	-
EfmE1162_1020	Helicase-like:Type III restriction enzyme, res subunit:DEAD/DEAH box helicase-like	L	3.14	-
EfmE1162_1021	competence protein F	R	3.71	-
EfmE1162_1029	phosphate ABC transporter, permease protein PstC	P	2.51	-
EfmE1162_1030	phosphate ABC transporter, permease protein PstA	P	2.37	-
EfmE1162_1031	phosphate ABC transporter, ATP-binding protein	P	2.29	-
EfmE1162_1032	phosphate ABC transporter, ATP-binding protein	P	2.12	-
EfmE1162_1034	hypothetical protein	-	2.41	-
EfmE1162_1037	conserved hypothetical protein	S	14.83	30.02
EfmE1162_1038	PspC domain family	KT	14.70	23.75
EfmE1162_1039	integral membrane protein	S	17.63	27.37
EfmE1162_1108	conserved hypothetical protein	-	20.17	14.34
EfmE1162_1120	undecaprenyldiphospho-muramoylpentapeptide beta-N-acetylglucosaminyltransferase	M	-	0.45
EfmE1162_1142	PTS system mannose IID component	G	2.19	-
EfmE1162_1143	PTS system, mannose-specific IIC component	G	2.11	0.49
EfmE1162_1145	fructose-specific phosphotransferase enzyme iib component	G	2.16	0.45
EfmE1162_1159	hypothetical protein	-	2.56	-
EfmE1162_1174	conserved hypothetical protein	-	2.25	-
EfmE1162_1183	hypothetical protein	Q	-	22.88
EfmE1162_1184	phosphatase YidA	R	2.98	19.04
EfmE1162_1185	general stress protein A	M	4.77	54.78
EfmE1162_1186	general stress protein A	M	4.64	47.02
EfmE1162_1187	hypothetical protein	-	-	10.45
EfmE1162_1188	Mg2+ and Co2+ transporter	P	2.76	47.19
EfmE1162_1189	oxidoreductase	IQR	-	34.59
EfmE1162_1190	GapA	-	-	13.58
EfmE1162_1191	GapB	S	-	14.20
EfmE1162_1192	Gls24	S	-	11.07
EfmE1162_1193	GlsB	S	-	8.53
EfmE1162_1194	hypothetical protein	-	0.44	0.19
EfmE1162_1199	GapA	-	-	9.00
EfmE1162_1200	conserved hypothetical protein	S	-	9.60
EfmE1162_1201	Gls24	S	-	8.73
EfmE1162_1202	GlsB	S	-	8.37
EfmE1162_1210	putative N-acetylmannosamine-6-phosphate 2-epimerase	G	3.84	-
EfmE1162_1251	putative oxidoreductase YhhX	R	-	0.25
EfmE1162_1263	hypothetical protein	S	2.26	18.50
EfmE1162_1303	oxidoreductase, aldo/keto reductase family	C	3.18	-
EfmE1162_1305	PTS system, cellobiose-specific IIc component	G	3.62	-
EfmE1162_1327	xanthine/uracil permeases family protein	R	0.33	0.15
EfmE1162_1333	Ser/Thr protein phosphatase family protein	R	-	2.94
EfmE1162_1335	DNA mismatch repair protein MutS, putative	-	22.98	9.54
EfmE1162_1358	adenine-specific methyltransferase	L	2.35	-
EfmE1162_1366	adenine deaminase	F	0.22	0.17
EfmE1162_1373	foldase protein PrsA	O	3.97	2.95
EfmE1162_1401	oligo-1,6-glucosidase	G	3.95	-
EfmE1162_1402	beta-glucosidase	G	2.37	-
EfmE1162_1412	glycerol kinase	C	3.45	-

LocusTag	Annotation	COG	Expression	
			5 min	15 min
EfmE1162_1413	aerobic glycerol-3-phosphate dehydrogenase	C	7.40	-
EfmE1162_1414	glycerol uptake facilitator protein	G	7.54	-
EfmE1162_1442	transcriptional regulator, MarR family	K	0.42	-
EfmE1162_1443	transposase	L	0.41	-
EfmE1162_1444	transposase	L	0.44	-
EfmE1162_1457	aldose 1-epimerase	G	4.35	-
EfmE1162_1458	ABC-type antimicrobial peptide transport system, permease component	V	5.38	-
EfmE1162_1459	macrolide export ATP-binding/permease protein MacB	V	5.49	-
EfmE1162_1460	conserved hypothetical protein	S	5.44	-
EfmE1162_1476	conserved hypothetical protein	-	7.27	15.63
EfmE1162_1477	lincomycin resistance protein LmrB	GEPR	8.74	15.54
EfmE1162_1478	transcriptional regulator, MerR family	K	9.17	6.08
EfmE1162_1484	endonuclease/exonuclease/phosphatase family protein	R	8.25	-
EfmE1162_1485	pts system, iibc component	GGG	11.00	-
EfmE1162_1486	maltose phosphorylase	G	22.46	-
EfmE1162_1487	beta-phosphoglucomutase	R	19.56	-
EfmE1162_1488	aldose 1-epimerase	G	9.76	-
EfmE1162_1489	sugar-binding transcriptional regulator, LacI family	K	2.57	-
EfmE1162_1499	conserved hypothetical protein	-	0.47	-
EfmE1162_1500	dihydroxyacetone kinase, L subunit	G	4.20	-
EfmE1162_1501	dihydroxyacetone kinase, DhaK subunit	G	5.15	-
EfmE1162_1502	dihydroxyacetone kinase, phosphotransfer subunit	S	5.63	-
EfmE1162_1521	hypothetical protein	-	18.66	16.34
EfmE1162_1522	conserved hypothetical protein	-	16.09	21.71
EfmE1162_1561	pyrimidine-nucleoside phosphorylase	F	-	0.43
EfmE1162_1562	deoxyribose-phosphate aldolase	F	-	0.36
EfmE1162_1563	cytidine deaminase	F	-	0.39
EfmE1162_1564	basic membrane protein family	R	-	0.22
EfmE1162_1565	ribose import ATP-binding protein RbsA	R	-	0.29
EfmE1162_1566	ABC transporter, permease protein	R	-	0.35
EfmE1162_1567	ABC transporter, permease protein	R	-	0.41
EfmE1162_1568	phosphopentomutase	G	-	0.44
EfmE1162_1569	purine nucleoside phosphorylase I, inosine and guanosine-specific	F	-	0.44
EfmE1162_1570	purine nucleoside phosphorylase	F	-	0.44
EfmE1162_1619	oligopeptide transport ATP-binding protein AppD	EP	7.85	7.32
EfmE1162_1620	oligopeptide transport ATP-binding protein AppF	E	7.91	7.51
EfmE1162_1621	oligopeptide transport system permease protein AppB	EP	7.56	7.30
EfmE1162_1622	peptide ABC transporter, permease protein	EP	5.05	-
EfmE1162_1623	peptide ABC transporter, peptide-binding protein	E	4.19	6.93
EfmE1162_1665	transcriptional regulator, TetR family	K	6.08	4.01
EfmE1162_1666	ABC transporter, ATP-binding/permease protein	V	4.83	4.39
EfmE1162_1667	ABC transporter, transmembrane region	V	4.03	4.11
EfmE1162_1701	hypothetical protein	R	0.30	-
EfmE1162_1702	conserved hypothetical protein	S	0.40	0.23
EfmE1162_1728	hypothetical protein	GEPR	22.77	22.32
EfmE1162_1798	YolG	C	4.58	3.63
EfmE1162_1799	transcriptional regulator, MarR family	K	5.06	3.28
EfmE1162_1800	muramidase-2	NU	-	0.37
EfmE1162_1865	glycine betaine/L-proline ABC transporter, glycine betaine/L-proline-binding/permease protein	EE	3.89	7.89
EfmE1162_1866	glycine betaine transport ATP-binding protein opuAA	E	5.07	8.59

LocusTag	Annotation	COG	Expression	
			5 min	15 min
EfmE1162_1872	hypothetical protein	O	0.25	-
EfmE1162_1873	glutaredoxin	-	0.47	0.39
EfmE1162_1874	ribonucleoside-diphosphate reductase, alpha subunit	F	-	0.35
EfmE1162_1875	ribonucleoside-diphosphate reductase, beta subunit	F	-	0.37
EfmE1162_1882	SagA	M	0.18	-
EfmE1162_1906	phosphopantothenate--cysteine ligase	H	-	0.47
EfmE1162_1929	inosine-5'-monophosphate dehydrogenase	FF	0.32	0.24
EfmE1162_1955	adenylosuccinate synthetase	F	-	0.32
EfmE1162_1961	cadmium-translocating P-type ATPase	P	-	4.75
EfmE1162_1988	glycosyl transferase, group 1 family protein	M	8.10	4.79
EfmE1162_1989	glycosyl transferase, group 1 family protein	M	6.09	4.93
EfmE1162_1992	ClpE	O	3.46	9.07
EfmE1162_2019	phosphotransferase system	GGTG	2.51	-
EfmE1162_2026	lactose phosphotransferase system repressor	KG	2.32	-
EfmE1162_2035	PTS system, cellobiose-specific IIC component	G	9.60	-
EfmE1162_2036	beta-glucosidase A	G	3.72	-
EfmE1162_2037	universal stress protein family	T	3.03	-
EfmE1162_2115	tyrosine recombinase XerC	L	5.85	3.87
EfmE1162_2116	heat shock protein HslVU, ATP-dependent protease subunit HslV	O	6.09	4.54
EfmE1162_2117	heat shock protein HslVU, ATPase subunit HslU	O	5.60	4.56
EfmE1162_2118	GTP-sensing transcriptional pleiotropic repressor CodY	K	4.13	3.83
EfmE1162_2119	lacx protein	G	3.02	-
EfmE1162_2124	formate acetyltransferase	C	-	0.38
EfmE1162_2165	FeS assembly protein SufD	O	-	0.49
EfmE1162_2196	conserved hypothetical protein	-	2.25	-
EfmE1162_2197	adp-ribose pyrophosphatase	LR	2.66	-
EfmE1162_2198	5-bromo-4-chloroindolyl phosphate hydrolysis protein	R	18.24	17.37
EfmE1162_2199	tellurite resistance protein	P	8.45	10.53
EfmE1162_2201	glucose uptake protein GlcU	G	-	0.45
EfmE1162_2211	sugar ABC transporter permease	G	2.87	-
EfmE1162_2212	inner membrane protein	G	3.02	-
EfmE1162_2213	sugar ABC transporter substrate-binding protein, putative	G	4.00	-
EfmE1162_2214	manganese catalase	P	-	16.38
EfmE1162_2228	phosphoglucomutase/phosphomannomutase family protein	G	2.47	-
EfmE1162_2234	LysM domain protein	-	0.15	-
EfmE1162_2253	NADH peroxidase	R	-	4.11
EfmE1162_2254	aspartate aminotransferase	E	-	3.69
EfmE1162_2255	amidohydrolase	R	2.14	6.79
EfmE1162_2256	aspartokinase	E	-	5.68
EfmE1162_2257	aspartate-semialdehyde dehydrogenase	E	2.90	7.74
EfmE1162_2258	2,3,4,5-tetrahydropyridine-2-carboxylate N-succinyltransferase	E	3.08	7.51
EfmE1162_2259	dihydrodipicolinate reductase	E	3.53	7.61
EfmE1162_2260	dihydrodipicolinate synthase	EM	-	5.41
EfmE1162_2261	diaminopimelate epimerase	E	3.23	5.54
EfmE1162_2262	diaminopimelate decarboxylase	E	3.81	6.82
EfmE1162_2265	cysteine synthase	E	-	0.39
EfmE1162_2266	methionine gamma-lyase	E	-	0.45
EfmE1162_2285	conserved hypothetical protein	S	5.90	4.01
EfmE1162_2286	sensor protein LiaS	T	7.37	4.78
EfmE1162_2287	DNA-binding response regulator, LuxR family	TK	7.27	4.71
EfmE1162_2288	TrkA-N domain family	P	3.68	-
EfmE1162_2289	HesB/YadR/YfhF-family protein	S	3.22	-

LocusTag	Annotation	COG	Expression	
			5 min	15 min
EfmE1162_2295	pheromone cAD1 lipoprotein	-	-	0.27
EfmE1162_2300	cell surface protein	-	-	0.21
EfmE1162_2353	Esp	M	-	0.47
EfmE1162_2370	Na+-driven multidrug efflux pump	V	0.35	0.33
EfmE1162_2372	ribosomal protein L25, Ctc-form	J	-	3.43
EfmE1162_2382	hypothetical protein	-	0.39	0.36
EfmE1162_2413	pur operon repressor	F	-	0.40
EfmE1162_2437	SagA	S	0.19	-
EfmE1162_2453	dihydrolipoyllysine-residue acetyltransferase component of pyruvatedehydrogenase complex	C	2.43	-
EfmE1162_2454	pyruvate dehydrogenase E1 component subunit beta	C	2.48	-
EfmE1162_2455	pyruvate dehydrogenase	C	2.72	-
EfmE1162_2457	guanosine monophosphate reductase	F	-	0.25
EfmE1162_2471	phosphoribosylformylglycinamidine synthase II	F	0.48	-
EfmE1162_2472	phosphoribosylformylglycinamidine synthase I	F	0.44	-
EfmE1162_2474	phosphoribosylaminoimidazole-succinocarboxamide synthase	F	0.45	-
EfmE1162_2477	phosphoribosylaminoimidazole carboxylase, ATPase subunit	F	-	0.42
EfmE1162_2478	phosphoribosylaminoimidazole carboxylase, catalytic subunit	F	0.47	0.31
EfmE1162_2479	xanthine permease	F	0.35	0.42
EfmE1162_2480	xanthine phosphoribosyltransferase	F	0.35	0.39
EfmE1162_2530	LrgB family protein	M	0.46	-
EfmE1162_2532	glycerol facilitator-aquaporin gla	G	-	0.39
EfmE1162_2588	ATP-dependent chaperone ClpB	O	8.42	11.80
EfmE1162_2611	hypothetical protein	-	-	0.31
EfmE1162_2612	hypothetical protein	-	0.28	0.19
EfmE1162_2616	aggregation promoting protein	-	0.46	-
EfmE1162_2639	hypothetical protein	S	12.04	9.72
EfmE1162_2640	conserved hypothetical protein	S	12.05	11.20
EfmE1162_2641	glycerophosphoryl diester phosphodiesterase family protein	CC	4.98	3.59
EfmE1162_2643	hypothetical protein	-	2.12	-
EfmE1162_2682	hypothetical protein	-	3.13	3.84
EfmE1162_2684	cell wall surface anchor family protein	M	-	0.29

Chapter 5

A genetic element present on megaplasmids allows *Enterococcus faecium* to use raffinose as carbon source

Xinglin Zhang, Fernanda L. Paganelli, Damien Bierschenk, Annemarie Kuipers, Marc J. M. Bonten, Rob J. L. Willems, and Willem van Schaik

Department of Medical Microbiology, University Medical Center Utrecht, Utrecht. The Netherlands

Published in Environmental Microbiology, 2011, 13: 518–528

Abstract

Enterococcus faecium is a commensal of the gastrointestinal tract of humans and animals. Since the 1990s, it has also emerged as a nosocomial pathogen. Little is known about carbon metabolism of *E. faecium* even though the ability to utilize different sugars could be an important factor in adapting to different ecological niches. In this study we identify an *E. faecium* gene cluster that is responsible for the metabolism of the α-galactoside sugar raffinose. Phenotypic testing of seven *E. faecium* isolates of which the genomes were previously sequenced, showed that one isolate (strain E980) could grow on raffinose. Genome analysis identified a gene cluster containing two genes encoding α-galactosidases (termed *agaA* and *agaB*) that was uniquely present in E980. The *agaA* and *agaB* genes were significantly more frequently found in strains that are phylogenetically related to E980 and were more prevalent in surveillance isolates from hospital and community sources than in isolates from clinical infections. Disruption of the α-galactosidase gene *agaB*, but not of *agaA*, disabled growth on raffinose in strain E980. In all strains *agaA* and *agaB* are carried on megaplasmids that are between 150 and 300 kb in size. Filter-mating experiments showed that the megaplasmid of E980 can be transferred to a plasmidless recipient which then gains the ability to grow on raffinose. The observation that raffinose utilization by *E. faecium* is a trait carried by megaplasmids, indicates that these megaplasmids can have important roles in shaping the competitive fitness of *E. faecium* in the environment, for example by expanding the metabolic repertoire of this organism.

Introduction

Enterococcus faecium is a facultatively anaerobic Gram-positive bacterium commonly found in the alimentary tracts of humans and other animals as well as in soil, water and food [1]. In humans, *E. faecium* has long been considered a harmless member of the gastrointestinal flora. However, during the past two decades it has rapidly emerged as an important nosocomial pathogen around the world, where it is responsible for hospital-acquired bloodstream, urinary tract and surgical wound infections [2-4] These infections are often difficult to treat as clinical isolates of *E. faecium* have acquired resistance against several clinically important antibiotics.

Previous studies based on multi-locus sequence typing (MLST), microarray-based comparative genomic hybridization (CGH) and genome sequencing of *E. faecium* demonstrated the considerable genome plasticity of this bacterium [5-7]. The acquisition of a variety of metabolic pathways, resistance determinants, virulence factors, and other traits through horizontal gene transfer may allow *E. faecium* to adapt to different ecological niches and survive host defenses.

Few studies have addressed the basic metabolism of *E. faecium* despite its importance in colonization and adaptation to particular environments [8-10]. In this work, we focused on a genetic element involved in the metabolism of raffinose in *E. faecium*. Raffinose is an α-galactoside sugar, which consists of an α-1−6-linked galactopyranosyl unit linked to C-6 of the glucose moiety of sucrose. Raffinose is a common constituent of beans, vegetables and other plants [11]. Raffinose cannot be digested by the human gastrointestinal tract but can be degraded and utilized as carbon source by a number of commensal bacteria. The metabolism of raffinose depends on the activity of α-galactosidases [12-18] These enzymes function as glycoside hydrolases that cleave the galactose unit from the sucrose moiety of raffinose.

In several Gram-positive bacteria the α-galactosidases that confer the ability to utilize raffinose as carbon source have been functionally characterized [15, 16, 19]. An illustrative example is provided by *Streptococcus mutans* and *Streptococcus pneumoniae* where the α-galactosidase genes are located on a well-characterized multiple-sugar metabolism (*msm*) operon [15, 16]. These gene clusters contain eight genes, products of which are involved in the uptake of raffinose and the metabolism of sucrose. In these gene clusters, a gene encoding an α-galactosidase is essential for raffinose utilization and its expression is regulated by gene products belonging to the AraC/XylS family [15, 16, 19].

The ability to utilize raffinose has long been known to be variably present among enterococci [20] and has been used as a phenotypic marker for the identification of different *Enterococcus* species [21, 22]. Among *E. faecium* strains the ability to utilize raffinose is variably present [21, 23] but the genetic determinants for the metabolism of this sugar have so far not been identified. In this study, we set out to identify the genes responsible for the phenotype of raffinose utilization in *E. faecium* and to determine if this trait is linked to *E. faecium* strains isolated from specific environmental niches or genetic backgrounds.

Results

Identification of an α-galactosidase gene cluster in E. faecium *E980*

In a recent study draft genome sequences of seven *E. faecium* isolates were analyzed to quantify the considerable diversity among *E. faecium* genomes [6]. To further study the phenotypic consequences of the genetic diversity in *E. faecium*, high-throughput phenotypic characterization of these seven isolates was performed using Phenotype MicroArray technology [24]. This analysis showed considerable differences in the repertoire of carbohydrate sources that can be metabolized by the seven sequenced strains (Supporting Information – Figure 1). One of the most obvious differences among the strains was the ability of strain E980 to effectively utilize raffinose as a carbon source, while the other six strains could not grow on this sugar. Strain E980 was also uniquely able to grow on a related α-galactoside sugar (stachyose), suggesting the presence of a dedicated system for the uptake and metabolism of raffinose and other α-galactosides in strain E980.

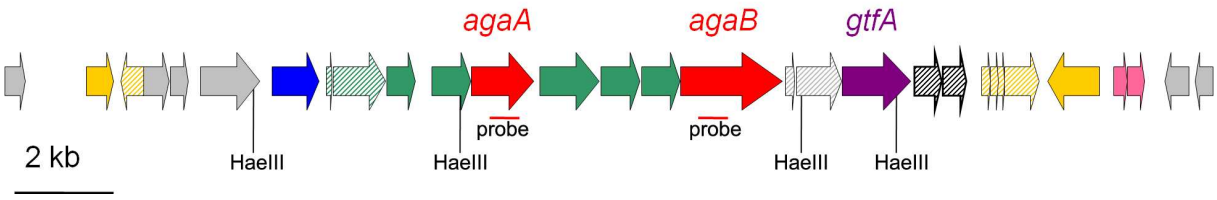


Figure 1. Overview of the gene cluster involved in raffinose metabolism in E980. Red: α-galactosidases *agaA* and *agaB*; purple: sucrose phosphorylase *gtfA*; dark blue: transcriptional regulator of the LacI/GalR family; green: sugar ABC transporters; pink: toxin-antitoxin (TA) module; yellow: putative integrases/recombinase; grey: hypothetical protein; black: transposase. The dashed arrows indicate pseudogenes with the color indicating the predicted function of the intact original gene. The location of HaeIII restricton sites and probes used for Southern hybridization are indicated.

To identify candidate genes responsible for the ability to metabolize raffinose in *E. faecium*, the genome of strain E980 was analyzed for the presence of proteins that belong to Cluster of Orthologous Groups (COG; [25]) COG1486 and COG3345, which are both predicted to act as α-galactosidases. A total of five proteins (three proteins belonging to COG1486, two proteins belonging to COG3345) were identified in the genome of strain E980. Three of these five putative α-galactosidases were also present in the other sequenced *E. faecium* isolates and so are unlikely to be involved in the metabolism of raffinose. Two genes, that we renamed *agaA* and *agaB* (locus tags EfmE980_0223 and EfmE980_0219, which encode proteins of 436 and 727 amino acids respectively; AgaA and AgaB were assigned to COG1486 and COG3345, respectively), were found to be uniquely present in E980 and were located on a genetic element that also carried genes with putative roles in the uptake and subsequent phosphorylation of carbohydrates (Figure 1).

Both AgaA and AgaB appear to be relatively closely related to two predicted α-galactosidase of *E. casseliflavus* (83% and 68% amino acid identity to EGAG_03044 and EGAG_00458, respectively, from *E. casseliflavus* EC30). It should be noted that these *E. casseliflavus* genes are not located in close proximity to each other as are *agaA* and *agaB* in *E. faecium* E980. The GtfA protein (locus tag EfmE980_0216) is predicted to convert the sucrose moiety, which results from the action of α-galactosidase on raffinose, into fructose and glucose-1-phosphate [16]. These sugars can then be further metabolized through other pathways in *E. faecium*. Homologs of the GtfA protein are not present in other enterococcal genomes but there is a clear homolog (82% amino acid identity) encoded in the genomes of the Gram-positive human commensals *Granulicatella elegans* and *Granulicatella adiacens*.

The proteins encoded by the genes located between *agaA* and *agaB* have predicted roles as sugar ABC transporters (two permease proteins and one substrate binding protein). These proteins are homologous (58 – 75% amino acid identity) to the *Streptococcus pneumoniae* raffinose uptake system [15]. Upstream of *agaA*, a transcriptional regulator was identified. This protein belongs to the LacI/GalR family of transcriptional regulators, which is widespread family of regulatory proteins that prominently feature in regulating the expression of genes encoding metabolic pathways [26].
Several pseudogenes are also present in the α-galactosidase gene cluster. All mutations responsible for the gene truncations were confirmed by PCR and Sanger sequencing. Based on sequence similarities these pseudogenes may represent remnants of sugar ABC transporters, transposase and integrases/recombinase-encoding genes. Interestingly, the gene cluster is flanked by two integrase genes and directly adjacent to one of these integrase genes is a prokaryote toxin-antitoxin (TA) module, which belongs to the *phd/doc*

family [27]. This suggests that the *agaAB* gene cluster is mobilizable and/or that this gene cluster is part of a larger mobile genetic element.

Growth on raffinose and detection of agaA *and* agaB *in a large panel of* E. faecium *strains*
To determine the frequency of raffinose metabolism among *E. faecium*, growth experiments were conducted with 122 *E. faecium* isolates (Table 1 and Supporting Information - Table 1) in M1 (a medium developed in this study to analyze the metabolism of carbon sources in *E. faecium*) and M1+R (M1 supplemented with 0.5% (w/v) raffinose), followed by Southern hybridizations with probes for *agaA* and *agaB* (Table 1 and Figure 2). Out of all 122 isolates, 20 (16%) isolates could utilize raffinose and 14 of these isolates were found to contain both *agaA* and *agaB* by subsequent Southern hybridization. In addition, in none of the 13 randomly selected isolates that were unable to grow in M1+R *agaA* or *agaB* could be detected. These findings demonstrate that the presence of these two genes and the ability to utilize raffinose is highly correlated. In all cases *agaA* (data not shown) and *agaB* (Figure 2B) were found on the same HaeIII restriction fragment of 7039 bp indicating that in all of the tested strains these genes are part of a conserved gene cluster. Among the 26 strains with MLST sequence type related to E980, 9 (35%) isolates contain *agaA* and *agaB*, which is significantly ($P < 0.001$) more frequent than among the other 96 strains, in which 5 isolates (5%) are positive for *agaA* and *agaB* (Figure 2C). These two genes are also significantly ($P < 0.01$) more common among strains that were isolated from hospital and community surveillance programs (9 out of 32 isolates; 28%) than in strains isolated from other ecological niches (5 out of 90 isolates; 6%). This observation suggests that the ability to utilize raffinose can give a competitive advantage predominantly to human commensal strains.

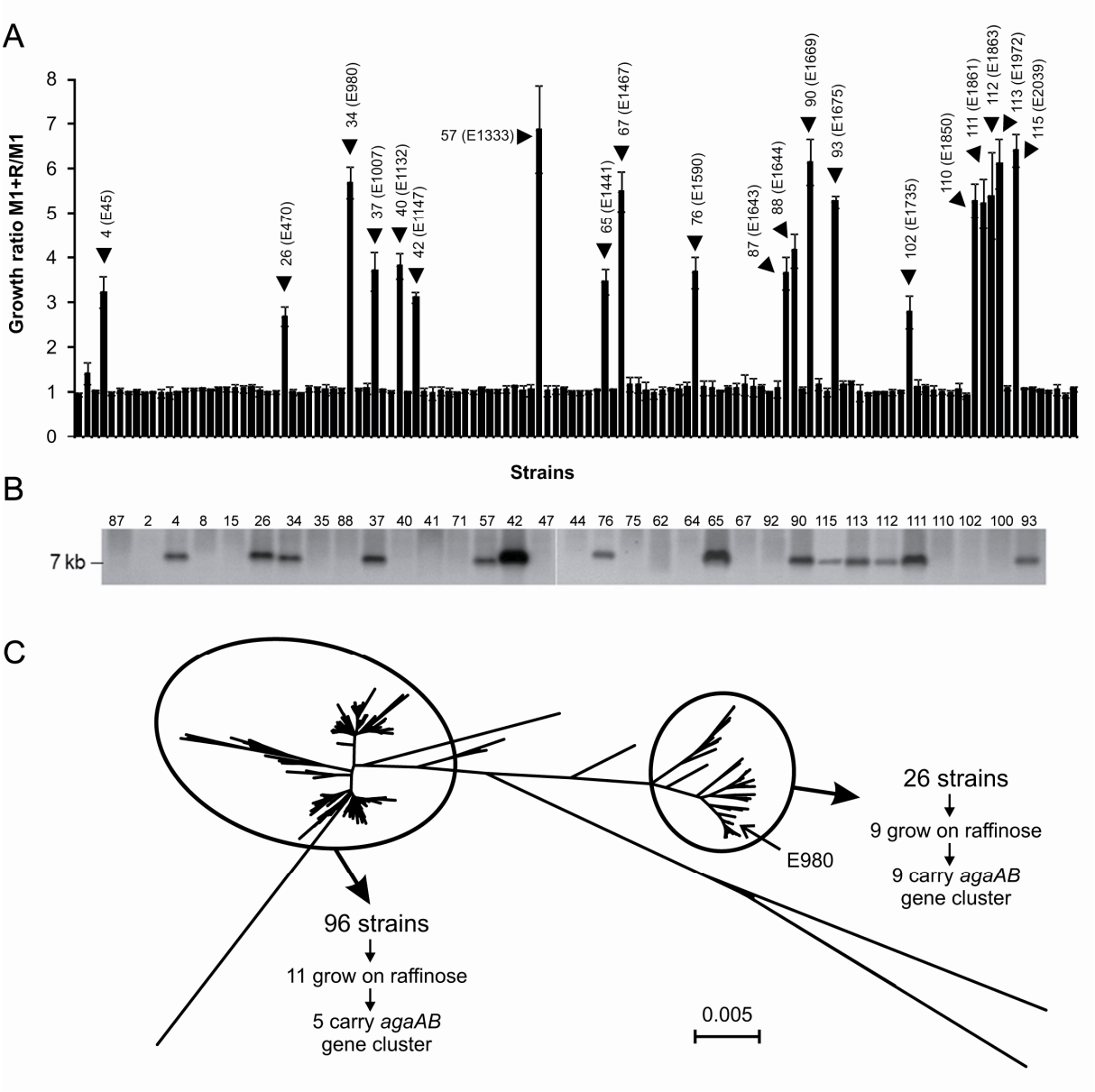


Figure 2. Analysis of the growth of *E. faecium* on raffinose. Panel A: 122 *E. faecium* isolates were cultured in M1 and M1+R respectively for 20 hours. The OD was measured at 655 nm. The growth of the isolates on M1+R was expressed as a ratio relative to growth in M1. Isolates with a growth ratio > 2 were regarded as positive for growth on raffinose. This experiment was repeated three times in duplicate and error bars indicate standard deviation. Strains that can grow on raffinose are indicated with their strain code (Table 1). Panel B: Southern analysis of the presence of *agaB* in *E. faecium*. Genomic DNA of 33 *E. faecium* isolates was digested with HaeIII and hybridized to ECL-labeled probes specific for *agaA* and *agaB*, respectively. Strain 34 is E980, which was used as a positive control; strain 44 is E1162, which was used as negative control. The HaeIII restricton sites in the α-galactosidase gene cluster and the PCR products used as probes are indicated in Figure 1. Strain codes refer to Table 1. Panel C: Phylogenetic tree inferred using the Neighbor-Joining method of *E. faecium* based on MLST data. For clarity a number of STs (105, 164, 325, 326, 331, 332, 335, 353, 357, 362, 419 and 470) that are distantly related to the main *E. faecium* population were omitted from this tree. The number of isolates that are able to use raffinose as carbon source and that carry the *agaAB* gene cluster is indicated. The position of strain E980 in this phylogenetic tree is also shown.

Table 1. ***E. faecium*** **strains used in this study.**

Category	Strain code in this study	Strain	ST	Growth on raffinose	*agaAB* gene cluster	Category	Strain code in this study	Strain	ST	Growth on raffinose	*agaAB* gene cluster
Epidemic	11	E161	16	-		Clinical	98	E1728	132	-	
	26	E470	16	+	+		97	E1721	169	-	
	29	E734	16	-			**100**	**E1733**	**178**	-	-
	30	E745	16	-			**115**	**E2039**	**296**	+	+
	40	E1132	16	+	-		**117**	**E3569**	**296**	-	
	65	E1441	16	+	+		**113**	**E1972**	**299**	+	+
	10	E155	17	-			**118**	**E3596**	**329**	-	
	27	E510	17	-			**119**	**E4135**	**345**	-	
	58	E1340	17	-			**120**	**E4136**	**346**	-	
	95	E1716	17	-			**121**	**E4168**	**352**	-	
	2	E13	18	-	-	Hospital	55	E1316	5	-	
	89	E1652	18	-		surveillance	43	E1149	6	-	
	96	E1717	18	-			72	E1554	10	-	
	14	E300	20	-			42	E1147	16	+	+
	15	E302	20	-	-		19	E321	17	-	
	16	E303	20	-			110	E1850	18	+	-
	17	E304	20	-			33	E849	21	-	
	64	E1435	65	-	-		1	E5	25	-	
	88	E1644	78	+	-		39	E1071	32	-	-[a]
	94	E1679	114	-	-[a]		87	E1643	78	+	-
	105	E1760	173	-			20	E322	79	-	
Clinical	55	E1316	5	-			**90**	**E1669**	**94**	+	+
	59	E1360	16	-			**91**	**E1670**	**94**	-	
	60	E1391	16	-			92	E1674	110	-	-
	102	E1735	16	+	-		93	E1675	111	+	+
	22	E380	17	-			41	E1133	117	-	-
	44	E1162	17	-	-[a]		3	E27	146	-	
	50	E1284	17	-			**111**	**E1861**	**289**	+	+
	51	E1292	17	-			**112**	**E1863**	**289**	+	+
	52	E1302	17	-		Community	8	E135	6	-	-
	54	E1308	17	-		surveillance	106	E1764	6	-	
	66	E1463	17	-			38	E1039	42	-	-[a]
	67	E1467	18	+	-		36	E1002	54	-	
	70	E1500	18	-			**37**	**E1007**	**61**	+	+
	99	E1731	18	-			**57**	**E1333**	**61**	+	+
	101	E1734	18	-			7	E128	82	-	
	103	E1737	18	-			**34**	**E980**	**94**	+	+[a]
	109	E1794	18	-			**35**	**E986**	**94**	-	-
	6	E125	21	-			68	E1485	101	-	
	5	E73	22	-			**56**	**E1327**	**118**	-	
	82	E1623	22	-			107	E1766	136	-	
	83	E1625	22	-			**76**	**E1590**	**163**	+	+
	46	E1250	25	-		Environment	84	E1628	68	-	
	31	E772	50	-			85	E1630	69	-	
	79	E1620	67	-			**104**	**E1759**	**172**	-	
	69	**E1499**	**74**	-		Animal	12	E172	1	-	
	48	E1263	78	-			13	E211	4	-	
	122	U0317	78	-	-[a]		9	E144	6	-	
	21	E333	80	-			23	E429	8	-	
	61	E1403	84	-			4	E45	9	+	+
	47	**E1262**	**85**	-	-		25	E466	21	-	
	71	**E1503**	**85**	-	-		73	E1573	21	-	
	80	E1621	86	-			24	E463	27	-	
	18	E318	91	-			74	E1574	27	-	
	63	E1423	94	-			78	E1619	70	-	
	114	**E1974**	**94**	-			**77**	**E1607**	**76**	-	
	45	E1172	99	-			81	E1622	104	-	
	62	**E1421**	**100**	-	-		28	E685	137	-	
	86	E1636	106	-	-[a]		108	E1781	141	-	
	49	E1283	130	-			75	E1576	159	-	-
	53	E1307	132	-		Probiotic	**116**	**E2363**	**296**	-	

Strains with STs that belong to the same cluster as strain E980 (Figure 2C) are indicated in bold.
[a] For these strains the absence or presence of the *agaAB* gene cluster was determined by analysis of the draft genome sequences [6].
A more detailed version of this table is included as Supporting Information – Table 1.

Disruption of the agaB *gene abolishes growth on raffinose*

To determine which of the two α-galactosidases in the identified gene cluster is primarily responsible for the ability to grow on raffinose, two disruption mutants (*agaA*::pWS3 and *agaB*::pWS3) were generated and their growth in M1, M1+G (M1 supplemented with 0.5% (w/v) glucose) and M1+R was determined (Figure 3). The results showed that both mutants had similar growth curves as the parental strain E980 in M1 and M1+G (data not shown). However, in M1+R, the growth of mutant *agaB*::pWS3 was significantly decreased while the growth of mutant *agaA*::pWS3 was not affected compared to the wild-type, indicating that *agaB* is the main α-galactosidase responsible for the metabolism of raffinose in strain E980.

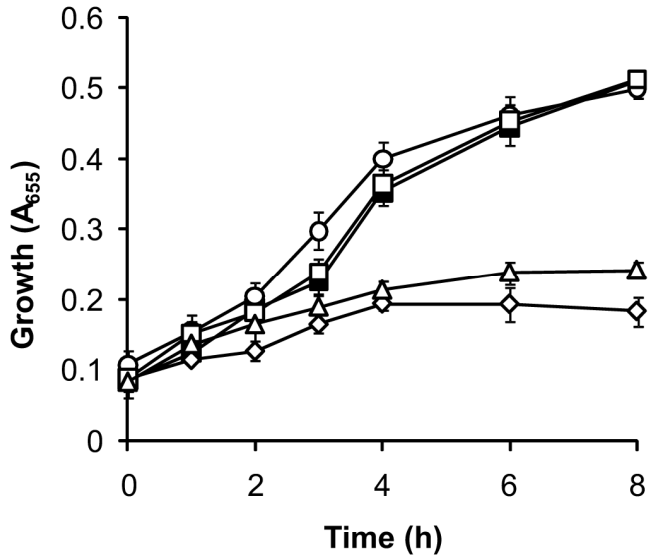


Figure 3. Growth of *E. faecium* E980, BM4105RF, a transconjugant and the *agaA* and *agaB* disruption mutants in M1+R. Overnight cell cultures of E980 (closed squares), the transconjugant (open circles), BM4105RF (open diamonds), *agaA*::pWS3 (open squares) and *agaB*::pWS3 (open triangles) were used to inoculate M1+R at A_{655} = 0.1. Growth was measured at regular intervals. This experiment was performed in triplicate and error bars indicate standard deviation.

The raffinose gene cluster is carried on a conjugative megaplasmid

We performed filter mating experiments with strain E980 as donor and BM4105RF as recipient to study if the raffinose metabolism gene cluster is located on a mobile genetic element. Because the genetic element is not linked to an antibiotic resistance marker, we selected for the transfer of the raffinose gene cluster by plating the cells after filter-mating on M1+R agar plates supplemented with rifampicin and fusidic acid, which inhibit the growth of the donor strain (E980) but not of the recipient BM4105RF strain. Three independent experiments showed that the raffinose gene cluster could be transferred by conjugation between E980 and BM4105RF with frequencies between 10^{-8} and 10^{-10} per donor cell. Pulsed-Field Gel Electrophoresis (PFGE) of the transconjugants confirmed that they had the

same genetic background as BM4105RF and were not E980 cells that had spontaneously acquired resistance to rifampicin and fusidic acid (Figure 4A). The subsequent Southern hybridization showed the transfer of *agaB* from donor to recipient. Intense hybridization to the DNA remaining in the well, which is indicative for extra-chromosomal supercoiled DNA, suggested that this mobilizable gene cluster is located on a plasmid (Figure 4B). Subsequent PFGE using S1 nuclease treated DNA (S1-PFGE) and Southern hybridization analysis confirmed that this mobilizable gene cluster is located on a megaplasmid of approximately 200 kb in size (Figure 4C and 4D). Growth experiments showed that the growth of E980, BM4105RF and the transconjugant is essentially identical in M1 and M1+G (data not shown), but that the transconjugant, unlike the recipient BM4105RF, could grow on raffinose as efficiently as E980. This indicates that the horizontally transferred genetic information was sufficient for efficient metabolism of raffinose (Figure 3).

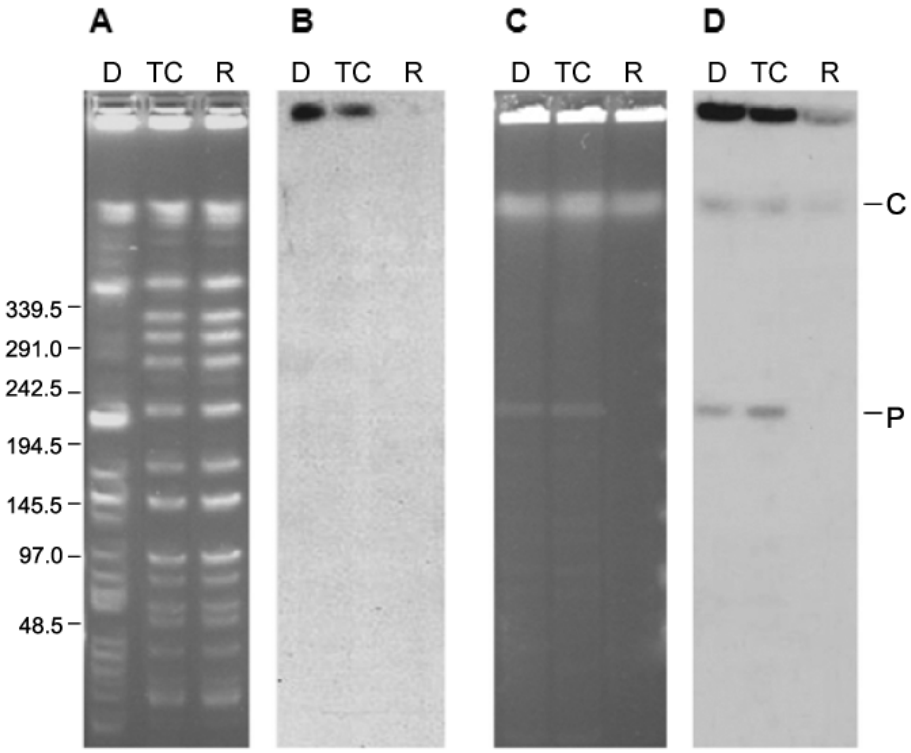


Figure 4. Location and transfer of the *agaAB* gene cluster. Panel A: PFGE analysis of SmaI digested DNA of *E. faecium* isolates. Panel B: Corresponding Southern blot with an ECL-labeled probe hybridizing to *agaB*. Panel C: S1-PFGE analysis of S1 nuclease digested DNA of *E. faecium* isolates. Panel D: Southern blot of S1-PFGE with an ECL-labeled probe hybridizing to *agaB*. In all panels D denotes the donor strain E980; TC denotes the transconjugant and R denotes the recipient strain BM4105RF. Marker sizes (in kb) are indicated. The letters C and P indicate chromosomal and plasmid DNA, respectively.

Location of this raffinose gene cluster in other E. faecium *isolates*

S1-PFGE and Southern hybridization were subsequently performed in the other 13 *E. faecium* isolates harboring *agaA* and *agaB*, which showed that the raffinose gene cluster was carried on megaplasmids that were present in all of the 13 isolates but absent in the negative control E1162 (Figure 5). The sizes of the plasmids ranged from approximately 150 kb to 300 kb. Most isolates had only one megaplasmid, except strain E1441, which carries two different megaplasmids of which only the larger one (approximately 195 kb in size) contained the α-galactosidase gene cluster.

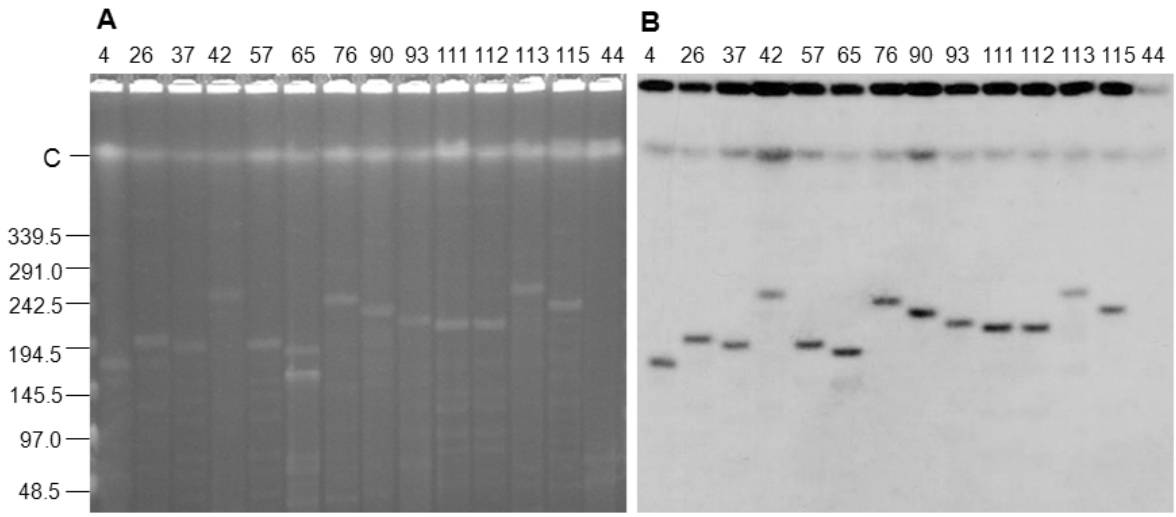


Figure 5. Location of the raffinose gene cluster in the 13 *E. faecium* isolates containing *agaA* and *agaB*. Panel A: S1-PFGE of the 13 isolates containing *agaA* and *agaB*. Strain 44 was E1162, which serves as a negative control. C indicates chromosomal DNA. Panel B: The corresponding Southern hybridization with a labeled probe specific for *agaB*. Strain codes in both panels refer to Table 1. Marker sizes (in kb) are indicated.

Discussion

In recent years, research on *E. faecium* has focused mainly on antibiotic resistance, epidemiological typing and the identification and characterization of proteins involved in virulence and colonization. Consequently, certain areas including the metabolism of carbohydrates by *E. faecium* remain poorly understood despite their potential importance in the colonization and adaptation to particular environmental niches.

This study provides an analysis into the metabolism of the α-galactoside sugar raffinose by *E. faecium* on both the phenotypic and genetic level. We show that a genetic element identified in the genome of E980 confers the ability to grow on raffinose to *E. faecium* and

that this element is only present in raffinose metabolizing strains. The raffinose utilization gene clusters of *S. pneumoniae*, *S. mutans*, *Lactobacillus plantarum*and *Lactococcusraffinolactis* contain only one α-galactosidase gene [13, 15, 16, 18]. In order to determine if both the *agaA* and *agaB* genes are indispensable for raffinose metabolism in *E. faecium* E980, we constructed insertion mutants for both genes. The *agaA* mutant is not affected in its ability to grow on raffinose, which indicates that *agaA* is not involved in the metabolism of raffinose in *E. faecium*. Interestingly, the AgaA protein of *E. faecium* shares 49% sequence identity with the MelA protein from *Escherichia coli* K12. This protein is involved in the metabolism of the α-galactosides melibiose, melibiitol and galactinol, but not of raffinose, which can only be metabolized by *E. coli* due to the presence of specific plasmids, that carry genes for raffinose utilization [12, 28]. In *E. faecium* AgaA may fulfill a similar role in the metabolism of α-galactosides other than raffinose. The disruption of *agaB* completely abolishes the metabolism of raffinose, indicating that the α-galactosidase encoded by this gene is crucial for the breakdown of raffinose into sugars that can be further metabolized.

We also show that the α-galactosidase gene cluster is located on megaplasmids and that the plasmid from strain E980 can be transferred to a plasmidless recipient which then gains the ability to grow on raffinose. This finding indicates that this gene cluster can spread by horizontal transfer through the *E. faecium* population. Horizontal gene transfer of metabolic pathways is an important factor in microbial evolution and contributes to the adaptation to new niches [29]. Large (>150 kb) conjugative plasmids are common in *E. faecium* and they can carry genes involved in antibiotic resistance and colonization of the gastrointestinal tract[30-33] In addition, a putative virulence gene of *E. faecium*, hyl(Efm), is carried on megaplasmids and is commonly found in the hospital-adapted *E. faecium* polyclonal subgroup CC17 [34, 35]. However, this virulence gene was not present on the megaplasmid of the non-CC17 strain E980 [6]. It appears unlikely that hyl(Efm) and the *agaAB* gene cluster are commonly co-localized on the same megaplasmid, because hyl(Efm) is highly associated with isolates from infections of hospitalized patients [34, 35] while the raffinose gene cluster is more frequently present in human commensal isolates. In this study we show that megaplasmids can also carry complete metabolic pathways that notably contribute to the phenotype of *E. faecium*.

Growth experiments demonstrated that out of the 122 isolates representing the whole *E. faecium* population, 20 isolates could utilize raffinose. Of these strains, 6 did not have the *agaA* or *agaB* gene, suggesting that besides the genetic element that we describe here, there are other mechanisms contributing to raffinose metabolism in *E. faecium*.

Based on the strain collection used in this study, we found that the genetic element conferring raffinose metabolism is more prevalent in hospital and community surveillance isolates (Table 1). These isolates can thus be considered to be representative of human commensal *E. faecium* strains. Consequently, it appears that the acquisition of the metabolic pathway that allows growth on raffinose may provides a competitive advantage to *E. faecium* strains that colonize the human gastrointestinal tract. In the strains that were isolated from infections of hospitalized patients the ability to grow on raffinose is less common, indicating that the metabolism of this sugar does not contribute to an infectious phenotype of *E. faecium.*

Our study illustrates that invasive and colonizing *E. faecium* strains differ genetically and phenotypically. Most studies so far identified putative virulence genes acquired by and enriched in invasive and hospital outbreak isolates [5, 7, 36-39]. Utilization of raffinose, as described in this study, may represent a trait specific for the commensal *E. faecium* flora of the human gastrointestinal tract. It is interesting to note that raffinose is a plant sugar and that *E. faecium* has many genetic determinants for the metabolism of plant sugars (such as arabinose and pectin), which are absent from the related organism *E. faecalis* [6]. The acquisition of genes that confer the ability to grow on raffinose may be a further specialization of *E. faecium* to grow on plants (from which *E. faecium* can frequently be isolated [40]) or their breakdown products.

Raffinose metabolism is more frequently found in *E. faecium* strains with MLST sequence types related to E980 (Figure 2C), and in these strains raffinose metabolism was always linked to the presence of *agaA* and *agaB*. Presumably strains from this genetic background are more adapt in integrating this element into their gene pool. However, the element is also, but less frequently, found in *E. faecium* strains that are not closely related to E980, indicating that gene flow exists between strains from different genetic backgrounds. In a number of STs (e.g. ST16 and ST94) the *agaAB* gene cluster is present in a sub-set of isolates in our strain collection, while it is absent in other isolates with the same ST. This indicates that the raffinose gene cluster can be acquired and/or lost at relatively high frequencies.

The data described in this study add to the growing recognition of the important role that plasmid-mediated traits have to *E. faecium* biology and illustrate the considerable metabolic flexibility of this organism. The interplay between metabolic adaptations and the ability of *E. faecium* to colonize a wide variety of environmental niches may be key to understanding the ubiquitous spread of this organism.

Materials and Methods

Bacterial isolates and growth conditions

In this study, 122 *E. faecium* isolates from different environmental niches and genetic backgrounds were selected for growth experiments on raffinose (Table 1 and Supporting Information - Table 1). These strains, which originate from different epidemiological sources (isolates from epidemic infections in hospitals (n = 21), isolated from non-epidemic clinical infections (n = 50), hospital (n = 19) and community (n = 13) surveillance isolates , environmental isolates (n = 3), animal isolates (n = 15) and a probiotic strain (n = 1)) and phylogenetic backgrounds, were selected as a representation of the global *E. faecium* population. BM4105RF is a rifampicin and fusidic acid resistant derivative of the plasmid-free strain *E. faecium* BM4105 [41] and this strain was used as a recipient in filter mating experiments. *E. faecium* isolates were grown at 37°C on Trypticase soy agar II plates supplemented with 5% sheep blood (TSAB; Becton Dickinson, Alphen aan den Rijn, The Netherlands) or in brain heart infusion (BHI) broth, and when appropriate, the antibiotics fusidic acid and rifampicin were used at concentrations of 20 μg/ml. Antibiotics were obtained from Sigma-Aldrich (St. Louis, MO).

Phenotype MicroArray and genome sequence analysis

High-throughput phenotypic characterization of the seven *E. faecium* isolates that we recently used in a genome sequencing-based study [6]was performed using Phenotype MicroArrays (Biolog Inc, Hayward CA, USA). All assays were performed in duplicate and at 37°C. Growth in the Phenotype MicroArray plates was measured colorimetrically over a period of 24 h. at 15 min. intervals using a tetrazolium dye that is present in the growth medium. α-Galactosidases with a potential role in raffinose metabolism were identified by using the COG-based annotation of the proteins encoded by the draft genome sequences of the seven *E. faecium* isolates that were analyzed previously [6].

Growth experiment on raffinose

To test the ability of *E. faecium* to utilize raffinose as carbon source, we developed a medium that minimizes growth of *E. faecium* when no carbon source is added. This medium, termed M1, consists of 10 g tryptone and 0.5 g yeast extract in 1 L phosphate buffered saline (PBS; 138 mM NaCl, 2.7 mM KCl, 140 mM Na_2HPO_4, 1.8 mM KH_2PO_4, adjusted to pH 7.4 with HCl). To supplement this medium with a carbon source, filter-sterilized solutions of 10% (w/v) raffinose or glucose were added to a final concentration of 0.5%, resulting in M1+R and M1+G respectively. All of the 122 *E. faecium* isolates were grown on TSAB plates overnight at 37°C. Then a single colony for each strain was picked up and inoculated in BHI broth. After

an over night incubation at 37°C with shaking, 2 µl of the overnight culture were added to 198 µl of M1, M1+R and BHI broth, respectively, and cultured in 96-well micro-titer plates at 37°C for 20 hours. The absorbance of the cultures was then measured at 655 nm. To quantify the growth of the *E. faecium* strains the ratio was calculated between the absorbance of the cultures in M1+R medium and M1. Isolates that had a growth ratio >2 were regarded as being able to utilize raffinose as carbon source for growth.

Southern blot analysis

Southern blot analysis was performed on genomic DNA isolated from 33 *E. faecium* isolates (including all the 20 strains that can grow on raffinose and 13 randomly chosen strains that are unable to grow on raffinose). Genomic DNA was isolated using the Wizard Genomic DNA Purification kit (Promega, Leiden, The Netherlands), digested with HaeIII (Promega), and fragments were separated by gel electrophoresis on a 0.8% agarose gel and blotted onto a Hybond N^+ nylon membrane (GE Healthcare, Diegem, Belgium). DNA was fixed onto the membrane by incubation for 30 min in 0.5 M NaOH followed by a neutralization in neutralizing solution (1.5 M NaCl, 0.5 M Tris-HCl and 1 mM EDTA) for 15 min. The membrane was hybridized overnight at 42°C with 100 ng probe. Probes were generated by PCR on genomic DNA of E980 using the primers *agaA*_F and *agaA*_R for *agaA*, *agaB*_F and *agaB*_R for *agaB* (Supporting Information - Table 2). Amplified DNA probes were purified with a PCR purification kit (Qiagen) and labeled according to the ECL nucleic acid labeling kit (GE Healthcare). Hybridization and washing was performed as described previously [36]. The membranes were exposed to Hyperfilm ECL (GE Healthcare).

Phylogenetic and statistical analyses

The evolutionary history of *E. faecium* was inferred with the Neighbor-Joining method and the concatenated sequences of the seven alleles in the *E. faecium* MLST database (http:// efaecium.mlst.net) using MEGA4 [42]. Comparisons between groups were made using Fisher's exact test (two-tailed).

Gene disruptions

In order to generate disruption mutants in *E. faecium* E980, we constructed a thermosensitive vector, pWS3, which is a derivative of the widely used Gram-positive thermosensitive plasmid pG^+host9 [43]. pWS3 was created by PCR amplifying the spectinomycin-resistance cassette from plasmid pAW068 [44] and cloning this amplified fragment into Bst1107I and ScaI digested pG^+host9, which results in the substitution of the erythromycin-resistance cassette from pG^+host9 with a spectinomycin-resistance cassette in pWS3. Subsequently internal DNA fragments of 436 bp and 439 bp of *agaA* and *agaB*, respectively, were PCR amplified using primers ZXL140 and ZXL141 for *agaA*, and ZXL138 and ZXL139 for *agaB* (Supporting Information - Table 2), and cloned into pWS3 between EcoRI

and XhoI restriction sites. The resulting vectors were then electrotransformed into E980 as previously described [45, 46]. After electrotransformation, the cells were allowed to recover for 2 h at the permissive temperature of 30°C, after which the cells were plated on BHI plates supplemented with 300 µg/ml of spectinomycin at 30°C to select for transformants. Spectinomycin-resistant colonies were picked and grown overnight in 200 ml BHI broth supplemented with 300 µg/ml spectinomycin at an elevated temperature (37°C) to cure the plasmid. The cells were then plated on BHI agar plates with spectinomycin at 37°C. Single-crossover integrations into *agaA* and *agaB* were verified by PCR with primers ZXL146 (based on pWS3) and ZXL144 for *agaA,* primers ZXL146 and ZXL142 for *agaB* (Supporting Information - Table 2).

Filter matings

The filter mating conjugation experiments were performed as previously described with modifications [47]. Donor (E980) and recipient (BM4105RF) were grown overnight in BHI broth at 37°C followed by 200-fold dilution in fresh BHI broth and grown to A_{660} = 1. Then 1 ml of donor and 1 ml of recipient were mixed together and collected onto a 0.45 µm nitrocellulose filter. The filter was placed on TSAB and incubated overnight at 37°C. Cells were removed from the filters by vortexing in 1 ml PBS and plated on M1+R agar plates supplemented with fusidic acid and rifampicin for selecting transconjugants, on BHI agar without antibiotic for counting the total number of colony forming units (CFUs) of donor and recipient cells, and on BHI agar with fusidic acid and rifampicin for counting the of CFUs of the recipient. Colonies of transconjugants from each mating were then grown on BHI with fusidic acid and rifampicin and verified by PCR and PFGE, as described below. The conjugation frequency was calculated as the number of transconjugate divided by the number of CFUs of the donor. The conjugation experiment was performed in triplicate. Growth curves of E980, BM4105RF, the transconjugant, *agaA*::pWS3 and *agaB*::pWS3 were measured in M1, M1+R and M1+G, respectively. Briefly, cells were grown over night in BHI, subsequently they were washed with PBS and diluted to an A_{655} of 0.1 with 5 ml of M1, M1+R and M1+G, respectively, and then incubated at 37°C with shaking at 200 rpm. Absorbance was measured at 1, 2, 3, 4, 6 and 8 hours post inoculation.

PFGE, S1-PFGE and hybridizations

PFGE was performed as described previously with minor modifications [48]. S1-PFGE was performed similarly to PFGE [33]. The only difference was that each plug was digested 15 min at 37°C with 14 U of S1 nuclease (Takara) instead of SmaI. The subsequent Southern hybridizations of PFGE and S1-PFGE were performed as described above with probe hybridizing to *agaB*.

Acknowledgments

We would like to thank Dr. Emmanuelle Maguin (INRA, Jouy-en-Josa, France) and Dr. Marta Perego (The Scripps Research Institute, La Jolla, USA) for kindly proving us with pG$^+$host9 and pAW068, respectively and Jenny Laverde-Gomez (Robert Koch Institute, Wernigerode, Germany) for helpful discussions on S1-PFGE. This research was funded through a grant from the European Union Sixth Framework Programme under contract LSHE-CT-2007-037410 and a NWO-Veni grant (916.86.044).

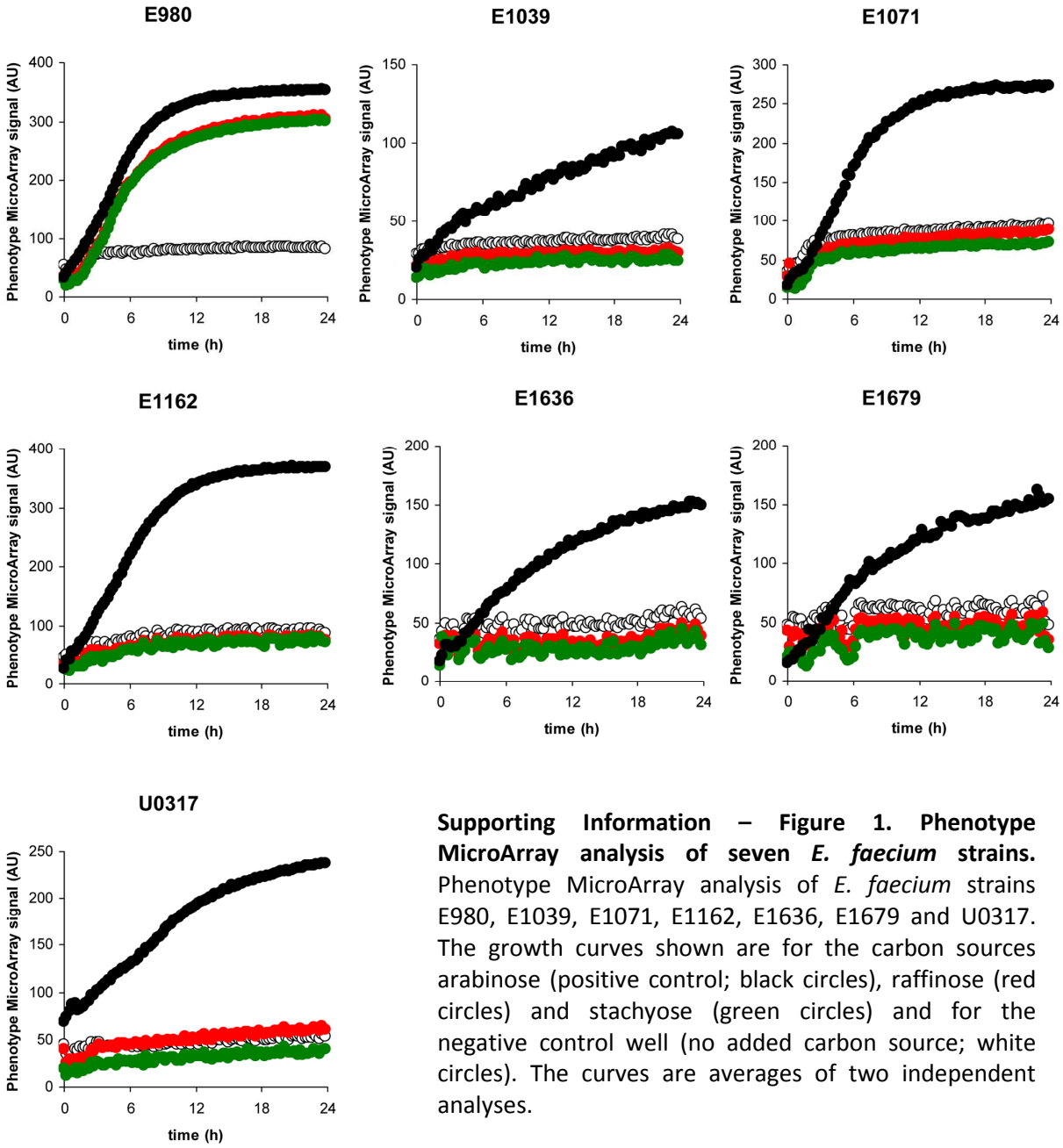
References

1. Top J, Willems R, Bonten M: Emergence of CC17 *Enterococcus faecium*: from commensal to hospital-adapted pathogen. *FEMS Immunol Med Microbiol* 2008, 52(3):297-308.
2. Leavis HL, Bonten MJ, Willems RJ: Identification of high-risk enterococcal clonal complexes: global dispersion and antibiotic resistance. *Curr Opin Microbiol* 2006, 9(5):454-460.
3. Murray BE: Vancomycin-resistant enterococcal infections. *N Engl J Med* 2000, 342(10):710-721.
4. Werner G, Coque TM, Hammerum AM, Hope R, Hryniewicz W, Johnson A, Klare I, Kristinsson KG, Leclercq R, Lester CH *et al*: Emergence and spread of vancomycin resistance among enterococci in Europe. *Euro Surveill* 2008, 13(47):19046.
5. Leavis HL, Willems RJ, van Wamel WJ, Schuren FH, Caspers MP, Bonten MJ: Insertion sequence-driven diversification creates a globally dispersed emerging multiresistant subspecies of *E. faecium*. *PLoS Pathog* 2007, 3(1):e7.
6. van Schaik W, Top J, Riley DR, Boekhorst J, Vrijenhoek JE, Schapendonk CM, Hendrickx AP, Nijman IJ, Bonten MJ, Tettelin H *et al*: Pyrosequencing-based comparative genome analysis of the nosocomial pathogen *Enterococcus faecium* and identification of a large transferable pathogenicity island. *BMC Genomics* 2010, 11:239.
7. Willems RJ, Top J, van Santen M, Robinson DA, Coque TM, Baquero F, Grundmann H, Bonten MJ: Global spread of vancomycin-resistant *Enterococcus faecium* from distinct nosocomial genetic complex. *Emerg Infect Dis* 2005, 11(6):821-828.
8. Ford M, Perry JD, Gould FK: Use of cephalexin-aztreonam-arabinose agar for selective isolation of *Enterococcus faecium*. *J Clin Microbiol* 1994, 32(12):2999-3001.
9. Deibel RH, Lake DE, Niven CF, Jr.: Physiology Of The Enterococci As Related To Their Taxonomy. *J Bacteriol* 1963, 86:1275-1282.
10. Kalcheva EO, Shanskaya VO, Maliuta SS: Activities and regulation of the enzymes involved in the first and the third steps of the aspartate biosynthetic pathway in *Enterococcus faecium*. *Arch Microbiol* 1994, 161(4):359-362.
11. Andersen KE, Bjergegaard C, Moller P, Sorensen JC, Sorensen H: Compositional variations for alpha-galactosides in different species of leguminosae, brassicaceae, and barley: a chemotaxonomic study based on chemometrics and high-performance capillary electrophoresis. *J Agric Food Chem* 2005, 53(14):5809-5817.
12. Aslanidis C, Schmid K, Schmitt R: Nucleotide sequences and operon structure of plasmid-borne genes mediating uptake and utilization of raffinose in *Escherichia coli*. *J Bacteriol* 1989, 171(12):6753-6763.
13. Boucher I, Vadeboncoeur C, Moineau S: Characterization of genes involved in the metabolism of alpha-galactosides by *Lactococcus raffinolactis*. *Appl Environ Microbiol* 2003, 69(7):4049-4056.
14. Burkardt HJ, Mattes R, Schmid K, Schmitt R: Properties of two conjugative plasmids mediating tetracycline resistance, raffinose catabolism and hydrogen sulfide production in *Escherichia coli*. *Mol Gen Genet* 1978, 166(1):75-84.
15. Rosenow C, Maniar M, Trias J: Regulation of the alpha-galactosidase activity in *Streptococcus pneumoniae*: characterization of the raffinose utilization system. *Genome Res* 1999, 9(12):1189-1197.
16. Russell RR, Aduse-Opoku J, Sutcliffe IC, Tao L, Ferretti JJ: A binding protein-dependent transport system in *Streptococcus mutans* responsible for multiple sugar metabolism. *J Biol Chem* 1992, 267(7):4631-4637.
17. Schmid K, Schmitt R: Raffinose metabolism in *Escherichia coli* K12. Purification and properties of a new alpha-galactosidase specified by a transmissible plasmid. *Eur J Biochem* 1976, 67(1):95-104.

18. Silvestroni A, Connes C, Sesma F, De Giori GS, Piard JC: Characterization of the *melA* locus for alpha-galactosidase in *Lactobacillus plantarum*. *Appl Environ Microbiol* 2002, 68(11):5464-5471.
19. Gallegos MT, Schleif R, Bairoch A, Hofmann K, Ramos JL: Arac/XylS family of transcriptional regulators. *Microbiol Mol Biol Rev* 1997, 61(4):393-410.
20. Orla-Jensen S: (1919) The lactic acid bacteria. Mem Acad Roy Sci Danemark Sect Sci 8 Ser 5: 81–197.
21. Manero A, Blanch AR: Identification of *Enterococcus spp*. with a biochemical key. *Appl Environ Microbiol* 1999, 65(10):4425-4430.
22. Facklam RR, Collins MD: Identification of *Enterococcus species* isolated from human infections by a conventional test scheme. *J Clin Microbiol* 1989, 27(4):731-734.
23. Day AM, Sandoe JA, Cove JH, Phillips-Jones MK: Evaluation of a biochemical test scheme for identifying clinical isolates of *Enterococcus faecalis* and *Enterococcus faecium*. *Lett Appl Microbiol* 2001, 33(5):392-396.
24. Bochner BR: Global phenotypic characterization of bacteria. *FEMS Microbiol Rev* 2009, 33(1):191-205.
25. Tatusov RL, Natale DA, Garkavtsev IV, Tatusova TA, Shankavaram UT, Rao BS, Kiryutin B, Galperin MY, Fedorova ND, Koonin EV: The COG database: new developments in phylogenetic classification of proteins from complete genomes. *Nucleic Acids Res* 2001, 29(1):22-28.
26. Weickert MJ, Adhya S: A family of bacterial regulators homologous to Gal and Lac repressors. *J Biol Chem* 1992, 267(22):15869-15874.
27. Gerdes K, Christensen SK, Lobner-Olesen A: Prokaryotic toxin-antitoxin stress response loci. *Nat Rev Microbiol* 2005, 3(5):371-382.
28. Schmitt R: Analysis of melibiose mutants deficient in alpha-galactosidase and thiomethylgalactoside permease II in *Escherichia coli* K-12. *J Bacteriol* 1968, 96(2):462-471.
29. Pal C, Papp B, Lercher MJ: Adaptive evolution of bacterial metabolic networks by horizontal gene transfer. *Nat Genet* 2005, 37(12):1372-1375.
30. Garcia-Migura L, Liebana E, Jensen LB: Transposon characterization of vancomycin-resistant *Enterococcus faecium* (VREF) and dissemination of resistance associated with transferable plasmids. *J Antimicrob Chemother* 2007, 60(2):263-268.
31. Rice LB, Lakticova V, Carias LL, Rudin S, Hutton R, Marshall SH: Transferable capacity for gastrointestinal colonization in *Enterococcus faecium* in a mouse model. *J Infect Dis* 2009, 199(3):342-349.
32. Zhu X, Zheng B, Wang S, Willems RJ, Xue F, Cao X, Li Y, Bo S, Liu J: Molecular characterisation of outbreak-related strains of vancomycin-resistant *Enterococcus faecium* from an intensive care unit in Beijing, China. *J Hosp Infect* 2009, 72(2):147-154.
33. Freitas AR, Novais C, Ruiz-Garbajosa P, Coque TM, Peixe L: Clonal expansion within clonal complex 2 and spread of vancomycin-resistant plasmids among different genetic lineages of *Enterococcus faecalis* from Portugal. *J Antimicrob Chemother* 2009, 63(6):1104-1111.
34. Arias CA, Panesso D, Singh KV, Rice LB, Murray BE: Cotransfer of Antibiotic Resistance Genes and a hyl_{Efm}-Containing Virulence Plasmid in *Enterococcus faecium*. *Antimicrob Agents Chemother* 2009, 53(10):4240-4246.
35. Freitas AR, Tedim AP, Novais C, Ruiz-Garbajosa P, Werner G, Laverde-Gomez JA, Canton R, Peixe L, Baquero F, Coque TM: Global spread of the *hyl(Efm)* colonization-virulence gene in megaplasmids of the *Enterococcus faecium* CC17 polyclonal subcluster. *Antimicrob Agents Chemother* 2010, 54(6):2660-2665.
36. Hendrickx AP, van Wamel WJ, Posthuma G, Bonten MJ, Willems RJ: Five genes encoding surface-exposed LPXTG proteins are enriched in hospital-adapted *Enterococcus faecium* clonal complex 17 isolates. *J Bacteriol* 2007, 189(22):8321-8332.

37. Leavis H, Top J, Shankar N, Borgen K, Bonten M, van Embden J, Willems RJ: A novel putative enterococcal pathogenicity island linked to the esp virulence gene of *Enterococcus faecium* and associated with epidemicity. *J Bacteriol* 2004, 186(3):672-682.
38. Nallapareddy SR, Singh KV, Okhuysen PC, Murray BE: A functional collagen adhesin gene, *acm*, in clinical isolates of *Enterococcus faecium* correlates with the recent success of this emerging nosocomial pathogen. *Infect Immun* 2008, 76(9):4110-4119.
39. Willems RJ, Homan W, Top J, van Santen-Verheuvel M, Tribe D, Manzioros X, Gaillard C, Vandenbroucke-Grauls CM, Mascini EM, van Kregten E *et al*: Variant *esp* gene as a marker of a distinct genetic lineage of vancomycin-resistant *Enterococcus faecium* spreading in hospitals. *Lancet* 2001, 357(9259):853-855.
40. Muller T, Ulrich A, Ott EM, Muller M: Identification of plant-associated enterococci. *J Appl Microbiol* 2001, 91(2):268-278.
41. Carlier C, Courvalin P: Emergence of 4',4"-aminoglycoside nucleotidyltransferase in enterococci. *Antimicrob Agents Chemother* 1990, 34(8):1565-1569.
42. Tamura K, Dudley J, Nei M, Kumar S: MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* 2007, 24(8):1596-1599.
43. Maguin E, Prevost H, Ehrlich SD, Gruss A: Efficient insertional mutagenesis in lactococci and other gram-positive bacteria. *J Bacteriol* 1996, 178(3):931-935.
44. Wilson AC, Perego M, Hoch JA: New transposon delivery plasmids for insertional mutagenesis in *Bacillus anthracis*. *J Microbiol Methods* 2007, 71(3):332-335.
45. Nallapareddy SR, Singh KV, Murray BE: Construction of improved temperature-sensitive and mobilizable vectors and their use for constructing mutations in the adhesin-encoding *acm* gene of poorly transformable clinical *Enterococcus faecium* strains. *Appl Environ Microbiol* 2006, 72(1):334-345.
46. Heikens E, Bonten MJ, Willems RJ: Enterococcal surface protein Esp is important for biofilm formation of *Enterococcus faecium* E1162. *J Bacteriol* 2007, 189(22):8233-8240.
47. Morton TM, Eaton DM, Johnston JL, Archer GL: DNA sequence and units of transcription of the conjugative transfer gene complex (trs) of *Staphylococcus aureus* plasmid pGO1. *J Bacteriol* 1993, 175(14):4436-4447.
48. Murray BE, Singh KV, Heath JD, Sharma BR, Weinstock GM: Comparison of genomic DNAs of different enterococcal isolates using restriction endonucleases with infrequent recognition sites. *J Clin Microbiol* 1990, 28(9):2059-2063.

Supplementary data



Supporting Information – Figure 1. Phenotype MicroArray analysis of seven *E. faecium* strains. Phenotype MicroArray analysis of *E. faecium* strains E980, E1039, E1071, E1162, E1636, E1679 and U0317. The growth curves shown are for the carbon sources arabinose (positive control; black circles), raffinose (red circles) and stachyose (green circles) and for the negative control well (no added carbon source; white circles). The curves are averages of two independent analyses.

Supporting Information - Table 1. *E. faecium* strains used in this study

Category	Strain code	Strain	ST	Growth on raffinose	*agaAB*	CC17 ?[b]	Isolation site	Year	Source	Country[c]	*van* gene
Epidemic	11	E161	16	-		Y	Faeces	1995	Hospitalized patient	USA	*vanA*
	26	E470	16	+	+	Y		1999	Hospitalized patient	NLD	*vanA*
	29	E734	16	-		Y		2000	Hospital environment	NLD	*vanA*
	30	E745	16	-		Y	Faeces	2000	Hospitalized patient	NLD	*vanA*
	40	E1132	16	+	-	Y	Faeces	2001	Hospitalized patient	USA	*vanA*
	65	E1441	16	+	+	Y		2000	Hospitalized patient	GRC	*vanA*
	10	E155	17	-		Y	Faeces	1995	Hospitalized patient	USA	*vanA*
	27	E510	17	-		Y	Blood	1998	Hospitalized patient	AUS	*vanB*
	58	E1340	17	-		Y	Wound	1999	Hospitalized patient	NOR	
	95	E1716	17	-		Y			Hospitalized patient	DNK	
	2	E13	18	-	-	Y	Urine	1992	Hospitalized patient	GBR	*vanA*
	89	E1652	18	-		Y	Faeces	2002	Hospitalized patient	NLD	*vanA*
	96	E1717	18	-		Y			Hospitalized patient	DNK	
	14	E300	20	-		Y	Urine	1994	Hospitalized patient	USA	*vanA*
	15	E302	20	-	-	Y	Faeces	1994	Hospitalized patient	USA	*vanA*
	16	E303	20	-		Y	Urine	1994	Hospitalized patient	USA	*vanA*
	17	E304	20	-		Y	Vascular catheter	1994	Hospitalized patient	USA	*vanA*
	64	E1435	65	-	-	Y		1999	Hospitalized patient	GRC	*vanA*
	88	E1644	78	+	-	Y		2002	Hospitalized patient	DEU	*vanA*
	94	E1679	114	-	-[a]		Vascular catheter	1998	Hospitalized patient	BRA	*vanA*
	105	E1760	173	-		Y		2000	Hospitalized patient	AUS	*vanB*
Clinical	55	E1316	5	-			Blood	2000	Hospitalized patient	NLD	*vanA*
	59	E1360	16	-		Y		2001	Hospitalized patient	USA	
	60	E1391	16	-		Y		2000	Hospitalized patient	GBR	
	102	E1735	16	+	-	Y		2001	Hospitalized patient	ESP	
	22	E380	17	-		Y	Blood	1997	Hospitalized patient	GBR	*vanA*
	44	E1162	17	-	-[a]	Y	Blood		Hospitalized patient	FRA	
	50	E1284	17	-		Y	Blood		Hospitalized patient	DEU	

Category	Strain code	Strain	ST	Growth on raffinose	*agaAB*	CC17 ?[b]	Isolation site	Year	Source	Country[c]	*van* gene
	51	E1292	17	-		Y	Blood		Hospitalized patient	ITA	
	52	E1302	17	-		Y	Blood		Hospitalized patient	POL	
	54	E1308	17	-		Y	Blood		Hospitalized patient	PRT	
	66	E1463	17	-		Y	Blood	1998	Hospitalized patient	ESP	
	67	E1467	18	+	-	Y	Blood	1997	Hospitalized patient	ESP	
	70	E1500	18	-		Y	Blood	1997	Hospitalized patient	ESP	
	99	E1731	18	-		Y	Blood	nd	Hospitalized patient	TZA	
	101	E1734	18	-		Y		1995	Hospitalized patient	ESP	
	103	E1737	18	-		Y		1997	Hospitalized patient	ESP	
	109	E1794	18	-		Y		1991	Hospitalized patient	USA	
	6	E125	21	-			Bile	1995	Hospitalized patient	NLD	*vanA*
	5	E73	22	-			Pus	1995	Hospitalized patient	NLD	*vanA*
	82	E1623	22	-			Pus	1960	Hospitalized patient	NLD	
	83	E1625	22	-			Cerebrospinal fluid	1961	Hospitalized patient	NLD	
	46	E1250	25	-			Blood		Hospitalized patient	CHE	
	31	E772	50	-			Faeces	2000	Hospitalized patient	NLD	*vanA*
	79	E1620	67	-			Blood	1957	Hospitalized patient	NLD	
	69	**E1499**	**74**	**-**			Urine	1999	Hospitalized patient	ESP	
	48	E1263	78	-		Y	Blood		Hospitalized patient	AUT	
	122	U0317	78	-	-[a]	Y	Urine	2005	Hospitalized patient	NLD	
	21	E333	80	-		Y	Blood	1997	Hospitalized patient	ISR	*vanA*
	61	E1403	84	-				2000	Hospitalized patient	GBR	
	47	**E1262**	**85**	**-**	**-**		Blood		Hospitalized patient	ESP	
	71	**E1503**	**85**	**-**	**-**		Blood	1995	Hospitalized patient	ESP	
	80	E1621	86	-			Blood	1959	Hospitalized patient	NLD	
	18	E318	91	-		Y	Peritoneal fluid	1994	Hospitalized patient	USA	*vanA*
	63	E1423	94	-				2000	Hospitalized patient	GBR	
	114	**E1974**	**94**	**-**			Pleura	2001	Hospitalized patient	DEU	
	45	E1172	99	-			Urine		Hospitalized patient	POL	
	62	**E1421**	**100**	**-**	**-**			2000	Hospitalized	GBR	

Category	Strain code	Strain	ST	Growth on raffinose	*agaAB*	CC17 ?[b]	Isolation site	Year	Source	Country[c]	*van* gene
									patient		
	86	E1636	106	-	-[a]		Blood	1961	Hospitalized patient	NLD	
	49	E1283	130	-			Blood		Hospitalized patient	DEU	
	53	E1307	132	-		Y	Blood		Hospitalized patient	PRT	
	98	E1728	132	-		Y	Blood		Hospitalized patient	TZA	
	97	E1721	169	-			Blood		Hospitalized patient	TZA	
	100	**E1733**	**178**	**-**	**-**		Blood		Hospitalized patient	TZA	*vanA*
	115	**E2039**	**296**	**+**	**+**			2000	Hospitalized patient	DEU	
	117	**E3569**	**296**	**-**			Blood	2005	Hospitalized patient	NLD	
	113	**E1972**	**299**	**+**	**+**		Blood	2000	Hospitalized patient	DEU	
	118	**E3596**	**329**	**-**			Blood	1999	Hospitalized patient	NLD	
	119	**E4135**	**345**	**-**			Blood	1988	Hospitalized patient	GBR	
	120	**E4136**	**346**	**-**			Blood	1989	Hospitalized patient	GBR	
	121	**E4168**	**352**	**-**			Blood	2002	Hospitalized patient	GBR	
Hospital	32	E802	5	-			Faeces	2000	Hospitalized patient	NLD	*vanA*
surveillance	43	E1149	6	-				1995	Hospitalized patient	NLD	*vanA*
	72	E1554	10	-				2002	Hospitalized patient	NLD	*vanA*
	42	E1147	16	+	+	Y		1998	Hospitalized patient	NLD	*vanA*
	19	E321	17	-		Y	Faeces	1997	Hospitalized patient	FRA	*vanA*
	110	E1850	18	+	-	Y	Faeces	2001	Hospitalized patient	ESP	
	33	E849	21	-			Faeces	2000	Hospitalized patient	NLD	*vanA*
	1	E5	25	-			Faeces	1986	Hospitalized patient	FRA	*vanA*
	39	E1071	32	-	-[a]		Faeces	2000	Hospitalized patient	NLD	*vanA*
	87	E1643	78	+	-	Y		2002	Hospitalized patient	DEU	*vanA*
	20	E322	79	-			Faeces	1997	Hospitalized patient	FRA	*vanA*
	90	**E1669**	**94**	**+**	**+**				Hospitalized patient	BRA	
	91	**E1670**	**94**	**-**					Hospitalized patient	BRA	
	92	E1674	110	-	-				Hospitalized patient	BRA	
	93	E1675	111	+	+				Hospitalized patient	BRA	
	41	E1133	117	-	-	Y	Faeces	2001	Hospitalized patient	USA	*vanA*

Category	Strain code	Strain	ST	Growth on raffinose	*agaAB*	CC17 ?[b]	Isolation site	Year	Source	Country[c]	*van* gene
	3	E27	146	-			Faeces	1992	Hospitalized patient	GBR	*vanA*
	111	**E1861**	**289**	**+**	**+**		Faeces	2001	Hospitalized patient	ESP	
	112	**E1863**	**289**	**+**	**+**		Faeces	2001	Hospitalized patient	ESP	
Community	8	E135	6	-	-		Faeces	1996	Non-hospitalized person	NLD	*vanA*
surveillance	106	E1764	6	-			Faeces	1996	Non-hospitalized person	BEL	
	38	E1039	42	-	-[a]		Faeces	1998	Non-hospitalized person	NLD	
	36	E1002	54	-			Faeces	1998	Non-hospitalized person	NLD	
	37	**E1007**	**61**	**+**	**+**		Faeces	1998	Non-hospitalized person	NLD	
	57	**E1333**	**61**	**+**	**+**		Faeces		Non-hospitalized person	NLD	
	7	E128	82	-			Faeces	1996	Non-hospitalized person	NLD	*vanA*
	34	**E980**	**94**	**+**	**+**[a]		Faeces	1998	Non-hospitalized person	NLD	
	35	**E986**	**94**	**-**	**-**		Faeces	1998	Non-hospitalized person	NLD	
	68	E1485	101	-			Faeces	2000	Non-hospitalized person	ESP	
	56	**E1327**	**118**	**-**			Faeces	1998	Non-hospitalized person	NLD	
	107	E1766	136	-			Faeces	1996	Non-hospitalized person	BEL	
	76	**E1590**	**163**	**+**	**+**		Faeces	2001	Non-hospitalized person	IRL	
Environment	84	E1628	68	-				1981	River water	NLD	
	85	E1630	69	-				1981	River water	NLD	
	104	**E1759**	**172**	**-**				1985	Environment	FRA	
Animal	12	E172	1	-			Faeces	1996	Calf	NLD	*vanA*
	13	E211	4	-			Faeces	1996	Calf	NLD	*vanA*
	9	E144	6	-			Faeces	1996	Pig	NLD	*vanA*
	23	E429	8	-			Faeces	1997	Chicken	NLD	*vanA*
	4	E45	9	+	+		Faeces	1992	Chicken	GBR	*vanA*
	25	E466	21	-			Faeces	1996	Cat	NLD	*vanA*
	73	E1573	21	-			Rumen	1994	Bison	BEL	

Category	Strain code	Strain	ST	Growth on raffinose	*agaAB*	CC17 ?[b]	Isolation site	Year	Source	Country[c]	*van* gene
	24	E463	27	-			Faeces	1996	Dog	NLD	*vanA*
	74	E1574	27	-				1995	Dog	BEL	
	78	E1619	70	-				1964	Fish burger	NOR	
	77	**E1607**	**76**	**-**				1956	Cheese	NOR	
	81	E1622	104	-				1959	Mouse	NLD	
	28	E685	137	-					Pig	ESP	
	108	E1781	141	-			Carcass	2001	Pig	BEL	
	75	E1576	159	-	-		Caecum	2001	Ostrich	ZAF	
Probiotic	**116**	**E2363**	**296**	**-**					Probiotic	NLD	

Strains with STs that belong to the same cluster as strain E980 (Fig. 2C) are indicated in bold.

[a] For these strains the absence or presence of the agaAB gene cluster was determined by analysis of draft genome sequences.

[b] "Y" indicates that the strain could be assigned to the poylclonal subcluster CC17 based on MLST.

[c] Country codes are the three letter codes as defined in ISO 3166-1.

Supporting Information - Table 2. Sequences of primers used in this study.

Primer name	Sequence (5' to 3')[a]	Product
ZXL138	CCACTCGAGCATGCACTTGCTTGGCTAAA	*agaB* disruption mutant
ZXL139	GGAATTCTCACGAAACGGTGATTCAAA	*agaB* disruption mutant
ZXL140	CCGCTCGAGATGCCTGCAATTTTCCATTG	*agaA* disruption mutant
ZXL141	GGAATTCTCGAATGAAAATCGTGCTGA	*agaA* disruption mutant
ZXL142	TCCCAAATTTCATCCCGTTA	Verification of *agaB* mutant
ZXL144	TGGATGTGGATTTGGTTTTTC	Verification of *agaA* mutant
ZXL146	GGGGATTTTATGCGTGAGAA	Verification of mutants
*agaA*_F	GGGTGTTTTAAAAGCGACGA	Probe for *agaA*
*agaA*_R	TTCACGGCTTCTCGTATGTG	Probe for *agaA*
*agaB*_F	AGCAGCCATCAACAAAATCC	Probe for *agaB*
*agaB*_R	CATGCACTTGCTTGGCTAAA	Probe for *agaB*

[a] Introduced restriction sites are underlined.

Chapter 6

A LacI-family regulator activates maltodextrin metabolism of *Enterococcus faecium*

Xinglin Zhang, Malbert Rogers, Damien Bierschenk, Marc J. M. Bonten, Rob J. L. Willems, Willem van Schaik

Department of Medical Microbiology, University Medical Center Utrecht, Utrecht. The Netherlands

Manuscript in preparation

Abstract

Enterococcus faecium is a gut commensal of humans and animals. In the intestinal tract, *E. faecium* will have access to a wide variety of carbohydrates, including maltodextrins and maltose, which are the sugars that result from the enzymatic digestion of starch by host-derived and microbial amylases. In this study, we identified the genetic determinants for maltodextrin utilization of *E*. *faecium* E1162. We generated deletions in the *mdxABCD-pulA* gene cluster that is homologous to maltodextrin uptake genes in other Gram-positive bacteria and in the *mdxR* gene, which is predicted to encode a LacI family regulator of *mdxABCD-pulA*. Both these mutations impaired growth on maltodextrin but had no effect on the growth on maltose and glucose. Comparative transcriptome analysis showed that eight genes (including *mdxABCD-pulA*) were significantly less expressed in the isogenic Δ*mdxR* mutant strain compared to the parental strain when grown on maltose. Quantitative real-time RT-PCR confirmed the results of transcriptome analysis and showed that the transcription of a putative maltose utilization gene cluster is induced in M1 medium supplemented with maltose and is not regulated by MdxR. The *mdxABCD-pulA* gene cluster is conserved in 66 of 68 publicly available *E. faecium* genome sequences. Understanding the maltodextrin metabolism of *E. faecium* could yield novel insights into the underlying mechanisms that contribute to the gut commensal lifestyle of *E. faecium*.

Introduction

Enterococci are facultatively anaerobic Gram-positive bacteria commonly found in the gastrointestinal tracts of humans and animals [1]. In the last twenty years, *E. faecium* has emerged as a clinical pathogen of major importance. This development has been linked to its ability to efficiently acquire antibiotic resistance genes and genomic islands that may contribute to virulence [2, 3].

The ability of both commensal and clinical *E. faecium* strains to effectively colonize the intestinal tract determines the ecological success of this species. Therefore, understanding the mechanisms of successful host colonization is important for the development of novel strategies to prevent or treat infections with these opportunistic pathogens. The metabolism of carbohydrates in complicated food webs of the mammalian intestinal tract is crucially important for gut colonization of commensals and opportunistic pathogens [4-8]. Carbohydrate metabolism of *E. faecium* remains poorly understood despite its potential importance in colonization and adaptation to particular environmental niches [9].

One of the main energy and carbon sources for bacteria in the intestine originates from complex polysaccharides, such as starch [4]. Starch is a plant storage glycan that consists of glucose monomers joined via α-1,4 glycosidic linkages with additional branches introduced by α-1,6 linked glucose moieties. In the human intestinal tract, starch is digested by host-derived and microbial amylases. Its breakdown products (mainly maltose and maltodextrins) can be absorbed by the host small intestine [10], but can also reach the colon [11, 12] where they can be metabolized by bacteria from several genera [13, 14]. The metabolism of maltodextrin has been investigated in *Escherichia coli* [15, 16] and in several Gram-positive bacteria, including *Bacillus subtilis* [17, 18], *Listeria monocytogenes* [19] and *Streptococcus pyogenes* [20, 21]. The maltose/maltodextrin regulon in *E. coli* consists of ten genes encoding four glycoside hydrolases, a maltodextrin phosphorylase, a maltodextrin glucosidase, a periplasmic α-amylase, together with an ATP-binding cassette (ABC) transporter [15, 16]. In *B. subtilis*, maltose and maltodextrin are separately transported by a maltose-specific phosphotransferase system and a maltodextrin-specific ABC transporter, respectively [17], while in *L. monocytogenes* both maltose and maltodextrin are taken up by the same ABC transporter [19]. In this study, we identified the determinants of maltodextrin uptake and metabolism in *E. faecium*.

Materials and Methods

Bacterial strains, plasmids and growth conditions

E. faecium and *E. coli* strains used in this study are listed in Table 1. The *E. faecium* strain E1162 (with sequence type 17) was used throughout this study. This strain was isolated from a bloodstream infection in France in 1996 and its genome has previously been sequenced [22]. Unless otherwise mentioned, *E. faecium* was grown in brain heart infusion broth (BHI; Oxoid) at 37°C. The *E. coli* strains DH5α (Invitrogen) and EC1000 [23] were grown in Luria-Bertani medium. Where necessary, antibiotics were used at the following concentrations: gentamicin at 300 μg ml^{-1} for *E. faecium* and 25 μg ml^{-1} for *E. coli*, spectinomycin at 300 μg ml^{-1} for *E. faecium* and 100 μg ml^{-1} for *E. coli*. All antibiotics were obtained from Sigma-Aldrich (Saint Louis, MO). Growth of cultures was determined by measuring the optical density at 660 nm (OD_{660}).

Table 1: Strains and plasmids used in this study.

Strain or plasmid	Relevant characteristic(s)	Source or reference
E. faecium		
E1162	Clinical isolate (bloodstream infection), isolated in France, 1996	[22]
Δ*mdxR*	Markerless deletion mutant of *mdxR* of E1162	This study
Δ*mdxABCD-pulA*	Markerless deletion mutant of the *mdxABCD-pulA* gene cluster of E1162	This study
Δ*mdxR+mdxR*	Complementation strain of Δ*mdxR*; Δ*mdxR* harboring pEF25-*mdxR*	This study
E. coli strains		
DH5α	*E. coli* host strain for routine cloning	Invitrogen
EC1000	MC1000 *glgB*::*repA*; host strain for pWS3 derived vectors	[23]
Plasmids		
pWS3	Gram-positive thermosensitive origin; Spc^r	[9]
pDEL1a	pWS3 carrying the 5′ and 3′ flanking regions of *mdxR* for mutant construction	This study
pDEL2a	pWS3 carrying the 5′ and 3′ flanking regions of *mdxABCD-pulA* gene cluster for mutant construction	This study
pDEL1b	pDEL1a with a Gen^r cassette which was flanked by *lox66*- and *lox71*-sites cloned between the 5′ and 3′ flanking regions	This study
pDEL2b	pDEL2a with a Gen^r cassette which was flanked by *lox66*- and *lox71*-sites cloned between the 5′ and 3′ flanking regions	This study
pWS3-Cre	pWS3 derivative expressing Cre in *E. faecium*	[24]
pEF25	Shuttle plasmid pAT18 with spectinomycin resistance cassette cloned in the EcoRI site; Spc^r Ery^r	[25]
pEF25-*mdxR*	Complementation plasmid for Δ*mdxR*; pEF25 carrying gene *mdxR*	This study

Construction of deletion mutants and in trans *complementation*

Markerless gene deletion mutants in the *mdxR* gene (locustag: EfmE1162_2133) and the *mdxABCD-pulA* gene cluster (locustag: EfmE1162_0366 - EfmE1162_0370) were created via

the Cre-*lox* recombination system as previously described [24, 26]. Briefly, the 5′ and 3′ flanking regions (approximately 500 bp each) of the target genes were PCR amplified with the primers in Table 2. The two flanking regions were then fused together by fusion PCR (generating an EcoRI site between both fragments) and cloned into pWS3 [9], resulting in pDEL1a and pDEL2a (plasmids used or generated in this study are listed in Table 1). Then a gentamicin-resistance cassette which was flanked by *lox66*- and *lox71*-sites [24] was cloned into the EcoRI site that was generated between the 5′ and 3′ flanking regions in pDEL1a and pDEL2a, respectively. The resulting plasmids pDEL1b and pDEL2b were then electrotransformed into *E. faecium* E1162. Marked mutants were obtained by growing the gentamicin-resistant transformants at appropriate temperatures supplemented with appropriate antibiotics [24]. The plasmid pWS3-Cre [24], carrying a gene encoding Cre recombinase, was introduced into the marked mutant by electroporation and further culturing for the removal of the gentamicin resistance cassette and subsequent loss of pWS-Cre was performed as described previously (33). Excision of the gentamicin resistance cassette and loss of pWS3-Cre was verified by PCR using primers listed in Table 2.

Table 2: Primers used in this study.

Primer	Sequence [a]
delete_XhoI_mdxR_up_F	5′-CCGCTCGAGCCTGCACCTTTGGAATATGG-3′
delete_EcoRI_mdxR_up_R	5′-AACCTTGACTCGCCCCTTGAATTCACGTTTTGCAACATCTGCT-3′
delete_EcoRI_mdxR_dn_F	5′-GAATTCAAGGGGCGAGTCAAGGTTAT-3′
delete_XmaI_mdxR_dn_R	5′-CCCCCCGGGTGATTGGTAATGGCCGGTAT-3′
check_mdxR_up	5′-CATGATCAGCTTGCAGTTGG-3′
check_mdxR_dn	5′-GTGTCAACAGATGCGTTTCG-3′
delete_XhoI_mdxABCD-pulA_up_F	5′-CCGCTCGAGTGCTTGCTGATAAGCATCGT-3′
delete_EcoRI_mdxABCD-pulA_up_R	5′-CGACCGGAAAGTGAAGAATTCAATAGGTTCCATGGCAGCAG-3′
delete_EcoRI_mdxABCD-pulA_dn_F	5′-GGAACCTATTGAATTCTTCACTTTCCGGTCGATGAT-3′
delete_XmaI_mdxABCD-pulA_dn_R	5′-CCCCCCGGGTTCCTAGACCGCTGACACCT-3′
check_mdxABCD-pulA_up	5′-TCTTTCTTTGGCAGCCATTT-3′
check_mdxABCD-pulA_dn	5′-GACAGACAACAACCGATCTGAA-3′
qPCR_mdxR_F	5′-AGCCGACAGCAACAGTCTGA-3′
qPCR_mdxR_R	5′-GCTCGCCGTTCAAGCATTAT-3′
qPCR_malP_F	5′-AGCGCAGCAAGCAGAAAAAG-3′
qPCR_malP_R	5′-CTTCCATCGTTCTGCCCAAG-3′
qPCR_malT_F	5′-GAATCGGTGCGCTTTCTTGT-3′
qPCR_malT_R	5′-CGTGGCATTGATTCTTGCTG-3′
qPCR_tuf_F	5′-TACACGCCACTACGCTCAC-3′
qPCR_tuf_R	5′-AGCTCCGTCCATTTGAGCAG-3′
pAT392_EcoRI_lox66_genta_F	5′-AGGGAATTCTACCGTTCGTATAGCATACATTATACGAAGTTATG ATAAACCCAGCGAACCATTTGAGG-3′
pAT392_EcoRI_lox71_genta_R	5′-CTCCGAATTCTACCGTTCGTATAATGTATGCTATACGAAGTTATT CAATCTTTATAAGTCCTTTTATAA-3′

[a] Restriction sites are underlined.

An *in trans* complementated strain (Δ*mdxR+mdxR*) of the *mdxR* deletion mutant (Δ*mdxR*) were produced as described previously [24, 25]. The gene *mdxR* was PCR amplified from the genomic DNA of E1162 using the primers listed in Table 2. The PCR product was cloned into

the shuttle vector pEF25 [25]. The resulting plasmid, pEF25-*mdxR*, was introduced into the Δ*mdxR* mutant strain by electroporation as described above (33).

Determination of growth curves

A BioScreen C instrument (Oy Growth Curves AB, Helsinki, Finland) was used to determine the growth of *E. faecium* strains on starch, maltodextrin, maltose and glucose. Strains were grown overnight in BHI containing appropriate antibiotics overnight at 37°C. Subsequently, cells were inoculated at an initial OD_{660} of 0.0025 in M1 medium. M1 is a semi-defined medium that minimizes growth of *E. faecium* when no carbon source is added to the medium [9]. Here, M1 medium was supplemented with starch, maltodextrin (dextrose equivalent: 4-7), maltose or glucose (2.5 g/l) as carbon sources (all carbohydrates were purchased from Sigma-Aldrich). Cultures were incubated in the Bioscreen C system at 37°C with continuous shaking. The absorbance at 600nm (A_{600}) was recorded every 15 min for 12 hours. Each experiment was performed in triplicate.

Transcriptome profiling

Transcriptome comparisons were performed between the parental strain wild-type E1162 and mutant strain Δ*mdxR* grown to mid-exponential (OD_{660} = 0.3) phase in BHI and M1 supplemented with maltose (M1+Maltose), respectively. RNA isolation, cDNA synthesis and hybridization were performed as described in our previous work [24]. Analysis for statistical significance were performed using the Web-based VAMPIRE microarray suite (http://sasquatch.ucsd.edu/vampire/) as described previously [27, 28]. A gene of which all four probes on the microarray were identified as differentially expressed with a false discovery rate <0.05, were classified as significantly differentially expressed between samples. In an addition, genes which exhibited an expression between 0.5- and 2-fold different from the wild-type were deemed biologically insignificant and were filtered out. This experiment was performed with four biological replicates.

The microarray data generated for the transcriptome analysis in this study have been deposited in the ArrayExpress database (http://www.ebi.ac.uk/arrayexpress) under accession numbers E-MEXP-3759 (M1+maltose) and E-MEXP-3760 (BHI).

Quantitative real-time RT-PCR (qRT-PCR)

The total RNA samples of the transcriptome profiling experiment were also used for qRT-PCR. The absence of genomic DNA was verified by PCR prior to reverse transcription. The cDNA was synthesized from total RNA (~1.0 µg) by using the Superscript III First-Strand Synthesis System (Invitrogen, Breda, The Netherlands) according to the manufacturer's instructions. Using synthesized cDNAs, qRT-PCR was performed using Maxima SYBR Green/ROX qPCR Master Mix (Fermentas, St. Leon-Rot, Germany) and a StepOnePlus

instrument (Applied Biosystems, Nieuwekerk a/d IJssel, The Netherlands) with the following program: 95°C for 10 min, and subsequently 40 cycles of 95°C for 15 sec, 55°C for 1 min. The expression of *tufA* was used as a housekeeping control [27]. The Ct value of each sample was normalized by the amplification efficiency. The relative transcript level (fold difference relative to *tufA*) of the tested genes were calculated by using the normalized Ct value of the *tufA* housekeeping control minus the normalized Ct value of the tested genes. The resulting value represents a log2 transformed fold difference in gene expression. Statistical significance between wild-type and mutant was assessed by the unpaired two-tailed Student's *t*-test. This experiment was performed with four biological replicates.

Results

Identification of genes putatively involved in maltodextrin utilization in E. faecium

The genome sequence of strain E1162 was analyzed to identify genes potentially responsible for the utilization of maltodextrin and maltose in *E. faecium*. A search of the E1162 genome for the orthologs of the *L*. *monocytogenes* and *B*. *subtilis* maltodextrin utilization proteins led to the identification of a gene cluster (locus tags EfmE1162_0366 - EfmE1162_0370; here termed *mdxABCD-pulA*), that is predicted to encode maltose/maltodextrin ABC transporter proteins and a neopullulanase which is predicted to hydrolyze the α-1,4 linkages in starch. The encoded proteins are homologous to the maltodextrin utilization proteins of *L*. *monocytogenes* EGD-e (amino acid identity: 31%-79%) [19] and *B*. *subtilis* 168 (amino acid identity: 25%-72%) [17] (Figure 1). In *L*. *monocytogenes* EGD-e, a regulator gene *lmo2128* is located immediately upstream of the maltodextrin utilization gene cluster. In *E. faecium* E1162, a gene (locus tag: EfmE1162_2133, here termed *mdxR*) encoding a LacI family transcriptional regulator shares the highest amino acid identity (53%) with Lmo2128. *mdxR* is located on a different contig and consequently is not in the immediate vicinity of the *mdxABCD-pulA* gene cluster. Analysis of the complete genome sequence of *E. faecium* Aus0004 [29] showed that *mdxR* and the *mdxABCD-pulA* gene cluster are located at a distance of 28 kb. A gene (EfmE1162_1486, termed *malP* here) that is homologous to *lmo2121*, which is the last gene in the maltodextrin utilization gene cluster of *L*. *monocytogenes* EGD-e and which encodes maltose phosphorylase, is also not part of the *mdxABCD-pulA* gene cluster in E1162. In *E. faecium* Aus0004, *malP* and the *mdxABCD-pulA* gene cluster are spaced 48 kb apart. Evidently, there are major differences in genomic organization between the maltodextrin utilization gene clusters of *E. faecium* and *L*. *monocytogenes*, possibly reflecting functional differences in the metabolism of maltodextrins and maltose between the two organisms.

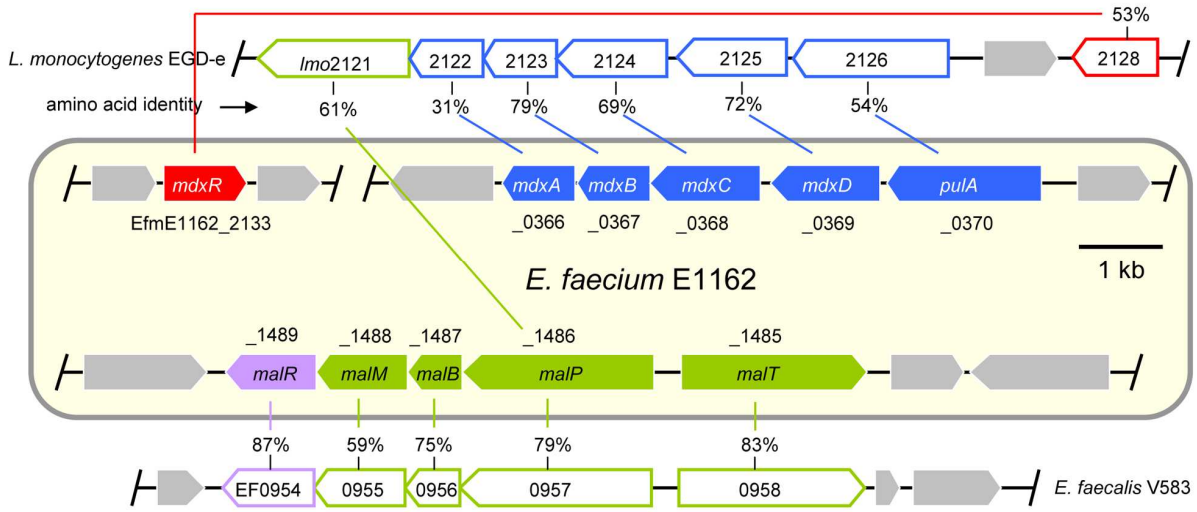


Figure 1. Schematic representation of the gene clusters involved in maltodextrin and maltose utilization of *E. faecium* E1162. Genes are represented by arrows that are drawn to scale. Genes putatively encoding proteins involved in maltodextrin utilization are indicated in blue (genes encoding transporter proteins or enzymes involved in the breakdown of maltodextrin) or in red (encoding a putative transcriptional regulator). Genes predicted to be involved in the uptake of maltose are indicated in green (genes encoding transporters or enzymes involved in maltose metabolism) or in purple (encoding a putative transcriptional regulator). The grey arrows represent the adjacent genes that are not predicted to be involved in maltodextrin or maltose utilization. Gene names are indicated in the arrows and the gene locus tags are indicated above or below the arrows. The homologues in *L. monocytogenes* EGD-e or in *E. faecalis* V583 are shown with corresponding colors above or below the gene clusters of *E. faecium*. The gene locus tags of the homologs are indicated in the arrows. Lines link the homologous genes with corresponding colors. Amino acid identities are indicated.

BLAST analysis showed that all genes of the *mdxABCD-pulA* gene cluster are conserved (with amino acid identities >82%) in 66 of the 68 *E. faecium* genomes available (on 16 October 2012) at NCBI Genomes, but no homologous gene cluster is present in *E. faecalis*. The observation that *E. faecalis* strains without this gene cluster could also grow in M1 with maltodextrin (data not shown) indicates that *E. faecalis* has different maltodextrin utilization mechanisms than *E. faecium*. The functional conservation of maltodextrin utilization in *E. faecium* and *E. faecalis* indicates that these traits have been conserved throughout the evolution of these organisms and thus likely contribute to fitness of these commensal bacteria.

Deletion of mdxR *and the* mdxABCD-pulA *gene cluster impair growth on maltodextrin*

To determine the role of the *mdxABCD-pulA* gene cluster in the ability to grow on maltodextrin, a marker-less deletion mutant (Δ*mdxABCD-pulA*) of all five genes in the *mdxABCD-pulA* gene cluster was constructed in *E. faecium* E1162. Additionally, a deletion mutant of the gene (*mdxR*) putatively encoding a transcriptional regulator involved in regulating expression of the *mdxABCD-pulA* gene cluster was also generated to characterize

its role in gene regulation and carbohydrate utilization in *E. faecium* E1162. The mutant Δ*mdxR* was complemented *in trans* (Δ*mdxR+mdxR*), but the *in trans* complemented strain for mutant Δ*mdxABCD-pulA* could not be constructed, presumably due to the large size (6.3 kb) of the DNA fragment encompassing this gene cluster. Growth of *E. faecium* E1162 wild-type (WT), the isogenic mutants and the complemented strain on M1 [9] supplemented with starch, maltodextrin, maltose or glucose were determined (Figure 2). The strains did not exhibit appreciable growth on M1 with starch, demonstrating that E1162 could not directly utilize starch. All strains showed identical growth in M1 with glucose, indicating that the introduced mutations did not cause a general growth defect. However, in M1 supplemented with maltodextrin, the growth of the Δ*mdxR* and Δ*mdxABCD–pulA* mutants were impaired, while in M1 supplemented with maltose the growth of these two mutants was comparable to wild-type and the complemented strain. This shows that the *mdxABCD-pulA* gene cluster of *E. faecium* is essential for the metabolism of maltodextrin but not for maltose. We note that E1162 wild-type exhibited a diauxic growth pattern in M1+maltodextrin, suggesting there were two types of carbon sources present in the medium. Presumably, the maltodextrin used contains residues of other sugars, such as glucose or maltose, which were utilized by *E. faecium* prior to maltodextrin. This may explain the growth of the Δ*mdxR* and Δ*mdxABCD–pulA* mutants in M1+maltodextrin during the first three hours of growth.

MdxR positively regulates the gene expression of the mdxABCD-pulA *gene cluster*

In *L. monocytogenes* and *B. subtilis*, the expression of the maltodextrin/maltose utilization systems is induced by maltose or maltodextrin [17, 19]. To identify the genes that are controlled by MdxR in *E. faecium*, the transcriptome of the Δ*mdxR* mutant was compared to its parental strain *E. faecium* 1162 grown to mid-exponential phase in M1 supplemented with maltose and BHI (which we routinely use to culture *E. faecium* and contains 2 g/l glucose), respectively (Table 3). We observed eight genes (excluding *mdxR*) that were significantly less expressed and four genes that were significantly higher expressed in the Δ*mdxR* mutant strain in comparison to the parental strain during growth in M1 with maltose. All the genes of the *mdxABCD-pulA* gene cluster were expressed at lower levels in the Δ*mdxR* mutant, which confirmed the prediction that MdxR regulates the transcription of this cluster of genes. Three other genes (EfmE1162_1270, EfmE1162_1401, and EfmE1162_1402, which are not located in the immediate vicinity of the *mdxABCD-pulA* gene cluster and, on the basis of homology to *L. monocytogenes* and *B. subtilis*, were predicted to be involved in sugar metabolism were also less expressed in the *mdxR* deletion mutant. EfmE1162_1270 and EfmE1162_1401 were annotated as encoding oligo-1,6-glucosidases that share 43% and 51% amino acid identity with the MalL protein of *B. subtilis* 168. MalL is involved in the breakdown of maltodextrin in *B. subtilis* 168 [17] which suggests that EfmE1162_1270 and EfmE1162_1401 also have a role in the metabolism of maltodextrins in *E. faecium*. EfmE1162_1402 is predicted to encode a β-glucosidase that hydrolyzes the β-1,4 bonds of

sugars like cellobiose, thereby releasing glucose. There were four genes expressed at higher levels in Δ*mdxR* than in wild-type E1162 during growth in M1+maltose. Three of these genes (EfmE1162_1412 - EfmE1162_1414) were predicted to be involved in glycerol utilization. We could not observe a positive effect of this higher expression on glycerol utilization, because both wild-type E1162 and Δ*mdxR* were not able to grow in M1 supplemented with glycerol in aerobic or anaerobic conditions (data not shown), which was consistent with the previously reported inability of *E. faecium* to grow on glycerol as carbon source [30]. When grown in BHI, only three genes encoding hypothetical proteins were expressed at slightly lower levels in Δ*mdxR* than in wild-type E1162, indicating that regulation by MdxR is unimportant in a rich, glucose-containing medium.

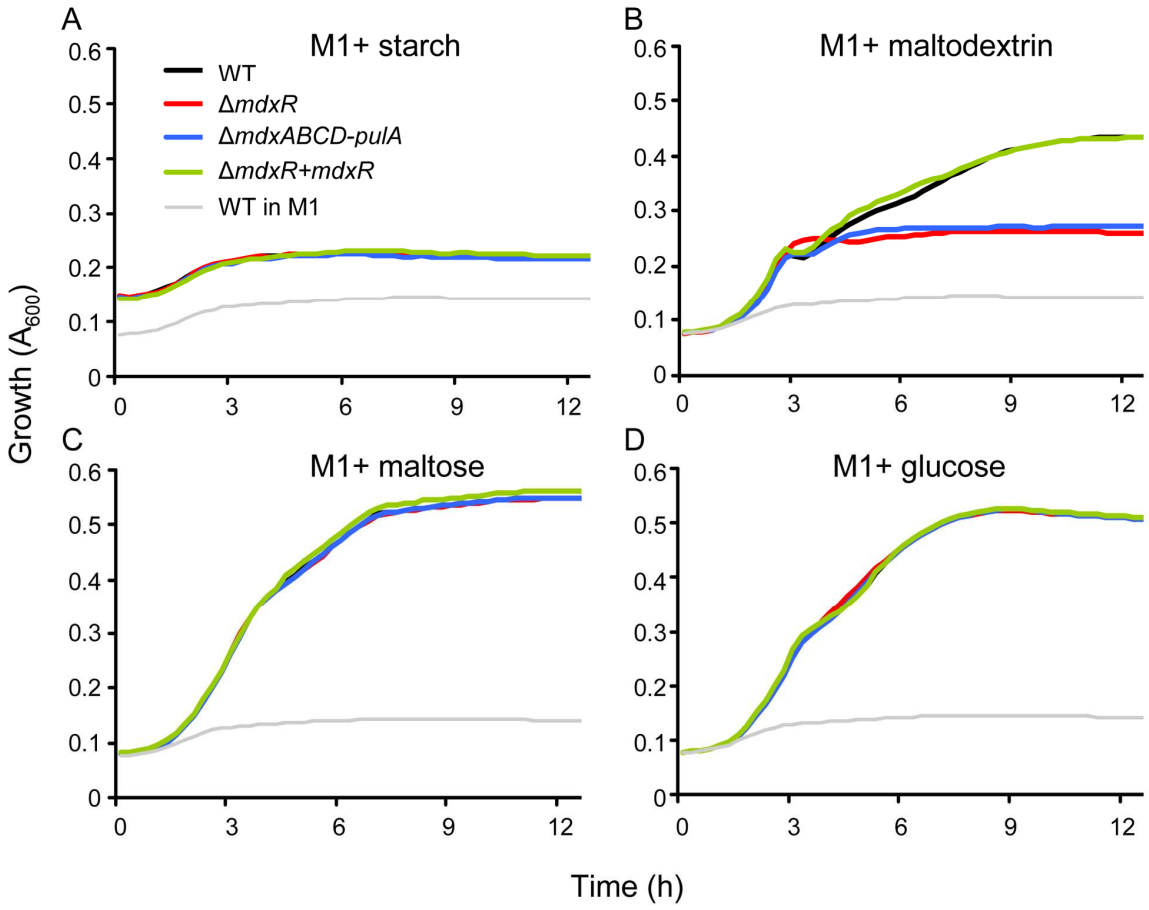


Figure 2. Growth of *E. faecium* on starch, maltodextrins, maltose and glucose. Growth curves of *E. faecium* E1162 wild-type (black), its isogenic mutants Δ*mdxR* (red) and Δ*mdxABCD-pulA* (blue), and the *in trans* complemented strain Δ*mdxR+mdxR* (green) on starch (panel A), maltodextrin (panel B), maltose (panel C) and glucose (panel D) are shown. The growth curve of E1162 wild-type in M1 was shown in grey as negative control. Note that A_{600} at the start of the experiment is higher in M1 + starch than in the other conditions, due to increased turbidity of the medium containing starch. Overnight cultures were inoculated at an initial OD_{660} of 0.0025 into 300 µl semi-defined minimal medium M1 [9], M1 supplemented with 2.5 g/l of starch maltodextrin, maltose or glucose as sole carbon source, respectively, and then incubated in the Bioscreen C system at 37°C with continuous shaking. The absorbance of 600nm (A_{600}) was recorded every 15 min for 12 hours. Growth curves are mean data of three independent experiments.

Table 3. Comparative transcriptome analysis of *E. faecium* E1162 and the Δ*mdxR* mutant during mid-exponential growth in M1+maltose and BHI

LocusTag [a]	Accession code	Gene name	Annotation	Expression ratio Δ*mdxR*/WT	
				M1+maltose	BHI
EfmE1162_0366	ZP_06676211	*mdxA*	conserved hypothetical protein	0.16	-
EfmE1162_0367	ZP_06676212	*mdxB*	maltose/maltodextrin ABC transporter, permease protein	0.12	-
EfmE1162_0368	ZP_06676213	*mdxC*	maltose/maltodextrin ABC transporter, permease protein	0.27	-
EfmE1162_0369	ZP_06676214	*mdxD*	maltose/maltodextrin ABC transporter, binding protein	0.27	-
EfmE1162_0370	ZP_06676215	*pulA*	neopullulanase	0.34	-
EfmE1162_1270	ZP_06677115		oligo-1,6-glucosidase	0.23	-
EfmE1162_1401	ZP_06677246		oligo-1,6-glucosidase	0.13	-
EfmE1162_1402	ZP_06677247		beta-glucosidase	0.16	-
EfmE1162_0373	ZP_06676218		conserved hypothetical protein	2.40	-
EfmE1162_1412	ZP_06677257		glycerol kinase	3.05	-
EfmE1162_1413	ZP_06677258		aerobic glycerol-3-phosphate dehydrogenase	3.56	-
EfmE1162_1414	ZP_06677259		glycerol uptake facilitator protein	3.94	-
EfmE1162_0318	ZP_06676163		hypothetical protein	-	0.46
EfmE1162_0408	ZP_06676253		hypothetical protein	-	0.46
EfmE1162_0932	ZP_06676777		hypothetical protein	-	0.40

[a] Genes (excluding *mdxR*) exhibiting significantly different (False Discovery Rate <0.05 and fold-change in expression <0.5 or >2) expression between E1162 wild-type and Δ*mdxR* during mid-exponential growth (OD_{660} = 0.3) in M1+maltose or in BHI are included. This experiment was performed with four biological replicates.

The transcription of a putative maltose utilization gene cluster was induced in M1+maltose and expressed independently of mdxR

Our results above showed that inactivation of the *mdxABCD-pulA* gene cluster of E1162 impaired the growth on maltodextrin but had no effect on the growth on maltose, suggesting that E1162 possesses a maltose utilization system which works independently of the maltodextrin utilization system. In *Enterococcus faecalis* a gene cluster composed of five genes has been identified as being responsible for maltose uptake and utilization [31]. A homology search in E1162 for the orthologs of this gene cluster identified a gene cluster (EfmE1162_1485 - EfmE1162_1489) which is predicted to encode five proteins with amino acid identities ranging from 59% to 87% to their homologs in *E. faecalis* V583 (Figure 1), and this gene cluster is conserved among all the available genomes of *E. faecium* (amino acid identities >96%).

We used qRT-PCR to analyze the expression of *mdxB* (putatively encoding the permease component of the maltodextrin ABC transporter) and two genes (*malT* and *malP*) of the maltose utilization gene cluster in wild-type E1162 and its isogenic mutant Δ*mdxR* grown in BHI and in M1+maltose (Figure 3). We found that, in both wild-type and mutant, the transcription of *mdxB* was at equally low level when grown in BHI, but was strongly

upregulated when grown in M1+maltose. Consistent with the observations in the transcriptome analysis, the transcription of *mdxB* was significantly lower (0.14 fold) in Δ*mdxR* than in wild-type E1162 when grown in M1+maltose. In both E1162 and the *mdxR* mutant, the transcription of *malT* and *malP* was greatly induced in M1+maltose compared to BHI. These qPCR data indicate that the EfmE1162_1485 - EfmE1162_1489 gene cluster is regulated independently of MdxR and may be involved in maltose utilization in *E. faecium*, similar to its homologous gene cluster in *E. faecalis* [31].

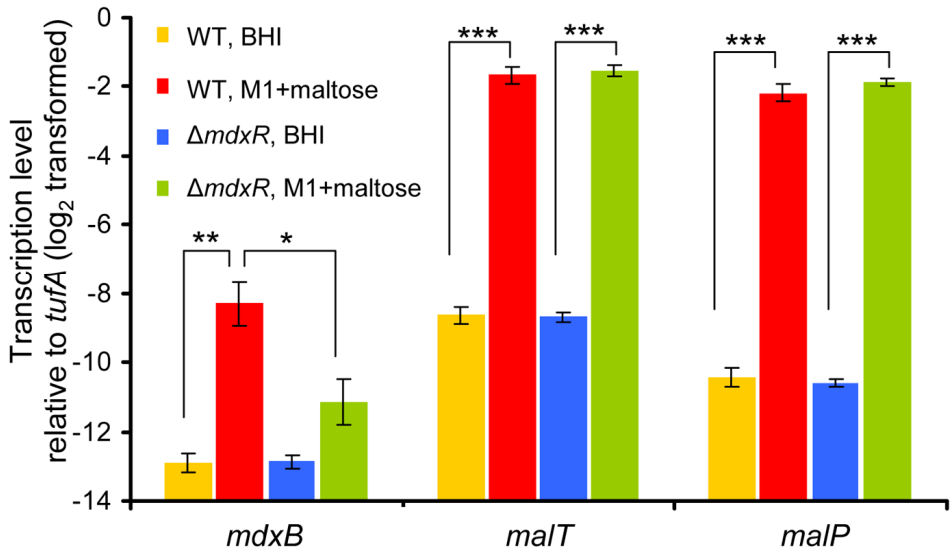


Figure 3. qRT-PCR analysis of *mdxB, malT* and *malP* expression ratios in wild-type E1162 and its isogenic mutant Δ*mdxR* grown in BHI and maltose. The data from the qRT-PCR were normalized using *tufA* as an internal standard [27]. The differences in gene expression (log_2-transformed data) relative to *tufA* are shown. Error bars represent the standard deviation of the mean of four biological replicates. Asterisks represent significant differences (*: P<0.01, **: P<0.001, ***: P<0.0001, as determined by an unpaired two-tailed Student's *t*-test) between the indicated samples.

Discussion

Carbohydrate utilization is a fundamental metabolic function in bacteria and is an important determinant for niche-adaptation by commensal bacteria [9, 32, 33]. In this work, we have identified the genetic determinants of *E. faecium* that contribute to its ability to metabolize maltodextrin which is present as a breakdown product of starch in the digestive tract of humans and animals.

Our results confirm previous data by showing that *E. faecium* cannot grow on starch [34], indicating that *E. faecium* can only benefit from the degradation products of starch due to the action of amylolytic enzymes, which are produced by the host and other gut commensals. Mutations in *mdxABCD-pulA* did not affect the growth rate on maltose,

suggesting that *E. faecium* has an independent system for the utilization of maltose. Similarly in *B. subtilis*, maltose is transported by maltose-specific phosphotransferase system, while maltodextrin is transported by a maltodextrin-specific ABC transporter [17]. In contrast, *L. monocytogenes* takes up both maltose and maltodextrin by the same ABC transporter [19]. Therefore, we propose that, similar to the situation in *B. subtilis*, the proteins encoded by the *mdxABCD-pulA* gene cluster are exclusively involved in maltodextrin transport and metabolism, and do not have a role in maltose metabolism.

We found eight genes that were less expressed in Δ*mdxR* than in wild-type E1162 when grown on M1+maltose. These include all the genes of the *mdxABCD-pulA* gene cluster and three additional genes (EfmE1162_1270, EfmE1162_1401, and EfmE1162_1402) with predicted functions in sugar metabolism. The roles of these three genes in utilizing maltodextrin remain to be determined. When grown in BHI, only three genes encoding hypothetical proteins were expressed at slightly lower levels in Δ*mdxR* than in wild-type E1162, suggesting that the expression of the *mdxABCD-pulA* gene cluster is induced by particular sugars, like maltose, and repressed in medium which contains a rapidly-metabolized sugar such as glucose. In line with catabolite repression by glucose, the expression levels of *mdxB*, *malP* and *malT* were also low when grown in BHI.

The data obtained by qRT-PCR confirmed the results of transcriptome analysis and showed that the expression of maltose utilization gene cluster is at considerably higher levels in M1+maltose than in BHI, but is unaffected by the deletion of *mdxR*. Although *mdxB* was significantly higher expressed in wild-type E1162 than in Δ*mdxR* when grown in M1+maltose, upregulation of *mdxB* expression was also observed in Δ*mdxR* grown in M1+maltose compared to growth of this strain in BHI, suggesting MdxR is not the sole regulator governing expression of the *mdxABCD-pulA* gene cluster. A possible explanation is that transcription of the *mdxABCD-pulA* gene cluster is coregulated by MdxR and HPr kinase/P-Ser-HPr phosphatase (HPrK/P) systems, which in *L. monocytogenes* [19] and *L. lactis* [35] represses the expression of genes involved in maltodextrin or maltose utilization in the presence of glucose. Consequently, upregulation of *mdxB* in M1+maltose may result from release of glucose mediated repression rather than stimulation of maltose. However, this upregulation is insufficient to support the effective growth on maltodextrin, for which MdxR is required.

E. faecium possesses a wide range of carbohydrate metabolic pathways [22] allowing utilization of a variety of sugars. In this study, we have identified an *E. faecium* gene cluster that is responsible for utilization of maltodextrin, a potential important carbon source for *E. faecium* in the gut. These data could contribute to the mechanistic understanding of the success of *E. faecium* as a gut commensal.

References

1. Top J, Willems R, Bonten M: Emergence of CC17 *Enterococcus faecium*: from commensal to hospital-adapted pathogen. *FEMS Immunol Med Microbiol* 2008, 52(3):297-308.
2. Willems RJ, Top J, van Schaik W, Leavis H, Bonten M, Siren J, Hanage WP, Corander J: Restricted Gene Flow among Hospital Subpopulations of *Enterococcus faecium*. *MBio* 2012, 3(4).
3. Willems RJ, van Schaik W: Transition of *Enterococcus faecium* from commensal organism to nosocomial pathogen. *Future Microbiol* 2009, 4(9):1125-1135.
4. Koropatkin NM, Cameron EA, Martens EC: How glycan metabolism shapes the human gut microbiota. *Nat Rev Microbiol* 2012, 10(5):323-335.
5. Chang DE, Smalley DJ, Tucker DL, Leatham MP, Norris WE, Stevenson SJ, Anderson AB, Grissom JE, Laux DC, Cohen PS *et al*: Carbon nutrition of *Escherichia coli* in the mouse intestine. *Proc Natl Acad Sci U S A* 2004, 101(19):7427-7432.
6. Marco ML, Peters TH, Bongers RS, Molenaar D, van Hemert S, Sonnenburg JL, Gordon JI, Kleerebezem M: Lifestyle of *Lactobacillus plantarum* in the mouse caecum. *Environ Microbiol* 2009, 11(10):2747-2757.
7. Thomas M, Wrzosek L, Ben-Yahia L, Noordine ML, Gitton C, Chevret D, Langella P, Mayeur C, Cherbuy C, Rul F: Carbohydrate metabolism is essential for the colonization of *Streptococcus thermophilus* in the digestive tract of gnotobiotic rats. *PLoS One* 2011, 6(12):e28789.
8. Sonnenburg ED, Zheng H, Joglekar P, Higginbottom SK, Firbank SJ, Bolam DN, Sonnenburg JL: Specificity of polysaccharide use in intestinal bacteroides species determines diet-induced microbiota alterations. *Cell* 2010, 141(7):1241-1252.
9. Zhang X, Vrijenhoek JE, Bonten MJ, Willems RJ, van Schaik W: A genetic element present on megaplasmids allows *Enterococcus faecium* to use raffinose as carbon source. *Environ Microbiol* 2011, 13(2):518-528.
10. Gray GM: Starch digestion and absorption in nonruminants. *J Nutr* 1992, 122(1):172-177.
11. Englyst HN, Cummings JH: Digestion of the carbohydrates of banana (Musa paradisiaca sapientum) in the human small intestine. *Am J Clin Nutr* 1986, 44(1):42-50.
12. Englyst HN, Cummings JH: Digestion of polysaccharides of potato in the small intestine of man. *Am J Clin Nutr* 1987, 45(2):423-431.
13. Cotta MA: Interaction of ruminal bacteria in the production and utilization of maltooligosaccharides from starch. *Appl Environ Microbiol* 1992, 58(1):48-54.
14. Wang X, Conway PL, Brown IL, Evans AJ: In vitro utilization of amylopectin and high-amylose maize (Amylomaize) starch granules by human colonic bacteria. *Appl Environ Microbiol* 1999, 65(11):4848-4854.
15. Boos W, Shuman H: Maltose/maltodextrin system of *Escherichia coli*: transport, metabolism, and regulation. *Microbiol Mol Biol Rev* 1998, 62(1):204-229.
16. Dippel R, Boos W: The maltodextrin system of *Escherichia coli*: metabolism and transport. *J Bacteriol* 2005, 187(24):8322-8331.
17. Schonert S, Seitz S, Krafft H, Feuerbaum EA, Andernach I, Witz G, Dahl MK: Maltose and maltodextrin utilization by *Bacillus subtilis*. *J Bacteriol* 2006, 188(11):3911-3922.
18. Kamionka A, Dahl MK: *Bacillus subtilis* contains a cyclodextrin-binding protein which is part of a putative ABC-transporter. *FEMS Microbiol Lett* 2001, 204(1):55-60.
19. Gopal S, Berg D, Hagen N, Schriefer EM, Stoll R, Goebel W, Kreft J: Maltose and maltodextrin utilization by *Listeria monocytogenes* depend on an inducible ABC transporter which is repressed by glucose. *PLoS One* 2010, 5(4):e10349.
20. Shelburne SA, 3rd, Keith DB, Davenport MT, Horstmann N, Brennan RG, Musser JM: Molecular characterization of group A *Streptococcus* maltodextrin catabolism and its role in pharyngitis. *Mol Microbiol* 2008, 69(2):436-452.

21. Shelburne SA, 3rd, Fang H, Okorafor N, Sumby P, Sitkiewicz I, Keith D, Patel P, Austin C, Graviss EA, Musser JM *et al*: MalE of group A *Streptococcus* participates in the rapid transport of maltotriose and longer maltodextrins. *J Bacteriol* 2007, 189(7):2610-2617.
22. van Schaik W, Top J, Riley DR, Boekhorst J, Vrijenhoek JE, Schapendonk CM, Hendrickx AP, Nijman IJ, Bonten MJ, Tettelin H *et al*: Pyrosequencing-based comparative genome analysis of the nosocomial pathogen *Enterococcus faecium* and identification of a large transferable pathogenicity island. *BMC Genomics* 2010, 11:239.
23. Leenhouts K, Buist G, Bolhuis A, ten Berge A, Kiel J, Mierau I, Dabrowska M, Venema G, Kok J: A general system for generating unlabelled gene replacements in bacterial chromosomes. *Mol Gen Genet* 1996, 253(1-2):217-224.
24. Zhang X, Paganelli FL, Bierschenk D, Kuipers A, Bonten MJ, Willems RJ, van Schaik W: Genome-Wide Identification of Ampicillin Resistance Determinants in *Enterococcus faecium*. *PLoS Genet* 2012, 8(6):e1002804.
25. Top J, Sinnige JC, Majoor EA, Bonten MJ, Willems RJ, van Schaik W: The recombinase IntA is required for excision of *esp*-containing ICE*Efm1* in *Enterococcus faecium*. *J Bacteriol* 2011, 193(4):1003-1006.
26. Sauer B: Functional expression of the *cre-lox* site-specific recombination system in the yeast *Saccharomyces cerevisiae*. *Mol Cell Biol* 1987, 7(6):2087-2096.
27. Lebreton F, van Schaik W, Sanguinetti M, Posteraro B, Torelli R, Le Bras F, Verneuil N, Zhang X, Giard JC, Dhalluin A *et al*: AsrR Is an Oxidative Stress Sensing Regulator Modulating *Enterococcus faecium* Opportunistic Traits, Antimicrobial Resistance, and Pathogenicity. *PLoS Pathog* 2012, 8(8):e1002834.
28. Hsiao A, Ideker T, Olefsky JM, Subramaniam S: VAMPIRE microarray suite: a web-based platform for the interpretation of gene expression data. *Nucleic Acids Res* 2005, 33(Web Server issue):W627-632.
29. Lam MM, Seemann T, Bulach DM, Gladman SL, Chen H, Haring V, Moore RJ, Ballard S, Grayson ML, Johnson PD *et al*: Comparative analysis of the first complete *Enterococcus faecium* genome. *J Bacteriol* 2012, 194(9):2334-2341.
30. Huycke MM: 2002. Physiology of enterococci, p 133-176. *In*: The Enterococci: Pathogenesis, Molecular Biology, and Antibiotic Resistance. Gilmore MS, Clewell, DB, Courvalin P, Dunny GM, Murray BE, Rice LB (ed), ASM Press, Washington, DC.
31. Le Breton Y, Pichereau V, Sauvageot N, Auffray Y, Rince A: Maltose utilization in *Enterococcus faecalis*. *J Appl Microbiol* 2005, 98(4):806-813.
32. Cantarel BL, Lombard V, Henrissat B: Complex carbohydrate utilization by the healthy human microbiome. *PLoS One* 2012, 7(6):e28742.
33. Abubucker S, Segata N, Goll J, Schubert AM, Izard J, Cantarel BL, Rodriguez-Mueller B, Zucker J, Thiagarajan M, Henrissat B *et al*: Metabolic reconstruction for metagenomic data and its application to the human microbiome. *PLoS Comput Biol* 2012, 8(6):e1002358.
34. Deibel RH, Lake DE, Niven CF, Jr.: Physiology Of The Enterococci As Related To Their Taxonomy. *J Bacteriol* 1963, 86:1275-1282.
35. Monedero V, Kuipers OP, Jamet E, Deutscher J: Regulatory functions of serine-46-phosphorylated HPr in *Lactococcus lactis*. *J Bacteriol* 2001, 183(11):3391-3398.

Chapter 7

Summarizing discussion

The Gram-positive bacterium *Enterococcus faecium* has long been perceived as a harmless member of the mammalian gut microbiota. However, in the past two decades it has rapidly emerged as a nosocomial pathogen causing clinical infections in hospitals all over the world. The acquisition of traits like multi-drug resistance and the elevated ability to colonize the intestinal tract of hospitalized patients have contributed to the recent ecological transition of *E. faecium* [1-3]. *E. faecium* is able to cause hospital-associated infections such as bacteraemia, endocarditis, urinary tract and surgical site infections that contribute to increased patient mortality and healthcare costs [1, 3]. Numerous studies have been performed on the molecular epidemiology and population structure of *E. faecium* [2]. Analyses based on Multi-Locus Sequence Typing (MLST) revealed the existence of three evolutionarily separate hospital-associated lineages originating from sequence types 17, 18, and 78 [4-7]. Comparative genomic analyses showed that hospital-acquired *E. faecium* isolates were genetically (based on the gene content in the genomes and on the sequences of genes in the core genome) distinct from indigenous intestinal isolates [8-11]. Although the knowledge generated from epidemiological studies and genome sequence analyses has been increasing rapidly in recent years, the functional characterization of genetic determinants of *E. faecium* that contribute to its success as both a gut commensal and a nosocomial pathogen is still in its infancy. Several important genetic determinants, such as the *van* gene cluster that confers vancomycin resistance [12-15] and the ampicillin resistance determinants *pbp5* [16-18] and *ddcY* [19, 20], have been studied in *E. faecium*, but they were mostly characterized from an epidemiological or biochemical perspective and were not genetically characterized, *i.e.* isogenic gene deletions and subsequent *in trans* complementation were lacking. A few virulence determinants, including *esp* [21, 22], ebp_{fm} [23], *acm* [24, 25], *gls33-glsB* and *gls20-glsB1* [26], have been studied in *E. faecium* by genetic approaches, but these genes were initially identified and had previously been characterized in other closely related organisms, like *E. faecalis*. In this work, we developed powerful molecular tools and techniques that allow efficient genetic manipulation of *E. faecium*. This has facilitated functional genomics-based studies to systemically identify and characterize genetic determinants of antibiotic resistance and niche adaptation that are responsible for the ecological success of this nosocomial pathogen.

New molecular tools and approaches for functional genomic study of *E. faecium*

With recent advances in sequencing technologies, the number of sequenced bacterial genomes has increased explosively. In 2008 there was only one publicly available *E. faecium* genome sequence (*E. faecium* DO), while currently (December 14, 2012) 69 *E. faecium* genome sequences can be accessed through NCBI Genomes. *In silico* analyses of these genome sequences have provided considerable information on the functional roles of *E.*

faecium genes [27, 28], based on their homology to genes of known function in other organisms. However, a large number of genotype-phenotype associations cannot be simply predicted based on studies in *Escherichia coli, Bacillus subtilis* or other model organisms and therefore need to be determined experimentally [29, 30].

Identification and functional characterization of the genetic determinants that are responsible for antibiotic resistance, colonization and virulence are necessary for our understanding of the mechanisms underlying transition of *E. faecium* from a commensal organism to a nosocomial pathogen, and may contribute to the development of novel therapeutic strategies to counter colonization and infection by *E. faecium*. However such studies have long been hampered by the lack of appropriate genetic tools for *E. faecium*.

This thesis describes many novel tools for the study of *E. faecium*. We developed a microarray based on the genome sequence of *E. faecium* E1162 using Agilent's 8×15 K platform, which, for the first time, facilitated the analysis of the *E. faecium* transcriptome [31, 32]. Transposon mutagenesis is a powerful tool for functional genomic studies and has been used extensively in many bacterial species, but not yet in *E. faecium* due to the lack of an appropriate transposon delivery vector and low electroporation efficiency. We constructed a high-density transposon mutant library in *E. faecium* and developed a transposon mutant tracking approach termed Microarray-based Transposon Mapping (M-TraM) [31]. In addition, the Cre-*lox* recombination system was adapted for use in *E. faecium* to construct targeted markerless mutants [31]. The tools developed in this study have removed many bottlenecks to the genetic study of *E. faecium*, and make it possible to perform both high-throughput genome-wide analysis and specific targeted investigations in clinical *E. faecium* isolates (Figure 1).

M-TraM provides a broadly applicable platform for functional genomics of *E. faecium* and can be potentially used for other Gram-positive organisms. This approach has been used in our laboratory for functional genomic studies in *E. faecium*. However this microarray-based mapping approach still has some technical limitations. M-TraM might fail to identify important mutants if there is an AluI restriction site present close to or inside the region that hybridizes to the microarray probes, since this may result in low or no signal on the microarray. Conversely, genes might be falsely identified when the AluI restriction fragments flank two or more genes. Recently described procedures, *e.g.* Tn-seq and INSeq, have resolved the drawbacks of microarray-based methods of screening transposon mutant libraries by using accurate, sequencing-based mapping of the transposon insertion sites [33-35]. Moreover, DNA sequencing has become increasingly cost-effective and is already less costly than microarray-based approaches for certain experiments. While M-TraM has shown itself to be extremely useful for performing the first truly functional genomic studies in *E.*

faecium, in future studies sequencing based approaches may be used to further improve the resolution of transposon mutant screening approaches in *E. faecium*.

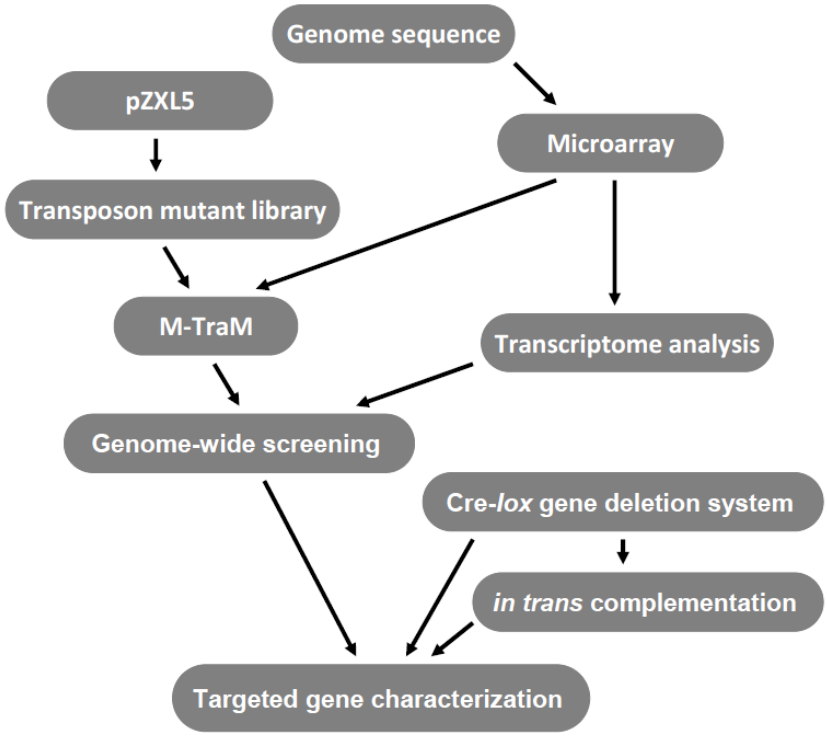


Figure 1. M**olecular tools and approaches developed in this work and their application in functional genomic study of** *E. faecium*. The plasmid pZXL5 is a *mariner* transposon delivery vector with a thermo-sensitive origin of replication [31] developed in our laboratory. Other tools and approaches are discussed in the text.

Antibiotic resistance

E. faecium is intrinsically resistant to many antibiotics and has the remarkable ability to rapidly gain resistance through the acquisition of genetic determinants or through sporadic mutations [1, 36]. Nowadays, *E. faecium* isolates from hospitals are resistant to nearly all clinically important antibiotics, which poses enormous challenges for the treatment of infections caused by this organism [1, 36]. The β-lactam antibiotic ampicillin has been used extensively to treat bacterial infections for more than fifty years. Ampicillin and its roughly equivalent successor amoxicillin are still commonly used for antibiotic therapy at present, but currently almost all clinical *E. faecium* isolates are resistant to these antibiotics [1, 3]. In this work, we performed high-throughput genome-wide analysis and targeted gene investigations in a clinical *E. faecium* isolate (strain E1162), by which several novel intrinsic ampicillin resistance determinants (*ddcP, ldt*$_{fm}$ and *pgt*), were identified [31] (Chapter 2). The results showed that although the low-affinity penicillin-binding protein PBP5 is the most important ampicillin resistance determinant of *E. faecium* E1162, other determinants such as

ddcP, ldt_{fm} and *pgt*, also contribute to PBP5-mediated high-level resistance to ampicillin. The *ddcP, ldt_{fm}* and *pgt* genes are highly conserved among *E. faecium* but absent from *E. faecalis*, indicating that *E. faecium* may possess more innate β-lactam resistance determinants than *E. faecalis* [31]. This finding may explain the fact that over 80% of clinical *E. faecium* isolates from all over the world are ampicillin resistance but less than 5% of clinical *E. faecalis* strains are resistant to ampicillin [3]. Identification of the intrinsic resistance determinants provides insights into the understanding of the genetic basis and mechanisms by which an organism is able resist the action of antibiotics, either by intrinsic mechanisms or by the acquisition of resistance determinants. Intrinsic antibiotic resistance determinants have recently been recognized as potential targets for the development of novel drugs particularly for multi-drug resistant Gram-negative bacteria such as *Acinetobacter baumannii* [37, 38] and *Pseudomonas aeruginosa* [39-44]. These studies generally use transposon mutagenesis and screening for sensitive phenotypes and have revealed that intrinsic resistance determinants belong to a variety of functional groups [37-44]. The intrinsic resistome of multidrug resistant *E. faecium* had not yet been studied despite its increasing clinical importance as a nosocomial pathogen. This work opened a door to comprehensive understanding of antibiotic resistance genetic basis of *E. faecium*, by defining the ampicillin resistome of this bacterium.

Colonization and carbohydrate metabolism

Antibiotic treatment eliminates the susceptible organisms of the gut flora, which opens up niches for multi-resistant enterococci to colonize [45] (Figure 2) This antibiotic-mediated high-level intestinal colonization by enterococci is often seen in hospitalized patients that ultimately develop enterococcal infections [46, 47]. In this work, we identified a genetic element that encodes a carbohydrate phosphotransferase system (PTS, termed PTS^{clin} here) which is enriched in clinical *E. faecium* strains and contributes to murine intestinal colonization during antibiotic treatment (Chapter 3). This is the first description of a determinant that contributes to intestinal colonization of clinical *E. faecium* strains. Previous epidemiological studies showed that most of the clinical (hospital-adapted) *E. faecium* isolates are genetically different from the indigenous *E. faecium* isolates of healthy individuals [2, 7-9]. The results obtained in this study showed that clinical *E. faecium* strains have acquired genes contributing to intestinal colonization of a microbiota that was disrupted by antibiotic treatment, indicating that clinical *E. faecium* strains have evolved traits other than antibiotic resistance to adapt to and proliferate in hospital niches. The identification of PTS^{clin} implicates a previously unrecognized role of carbohydrate utilization in intestinal colonization of clinical *E. faecium* strains and exemplifies the role of genetic

elements that have been acquired by horizontal gene transfer in determining the ecological success of hospital-adapted *E. faecium* strains.

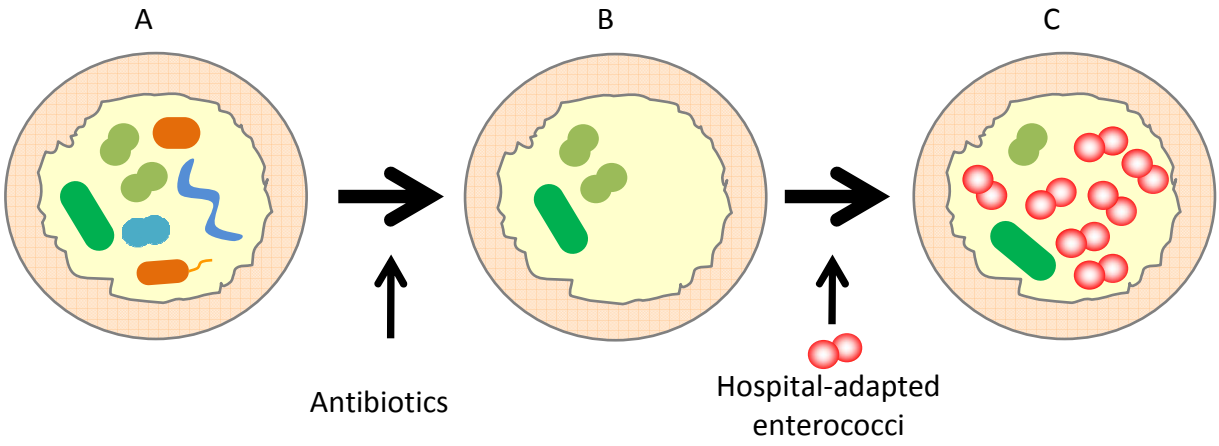


Figure 2. Antibiotic-mediated high-level intestinal colonization by hospital-adapted enterococci. (A) In the absence of antibiotics, the commensal flora, including indigenous enterococci, colonizes the human intestine. (B) Antibiotic treatment leads to a reduction in susceptible organisms and opens up new niches. (C) Hospital-adapted multi-resistant enterococci dominantly colonize the intestine, probably because they have acquired additional traits that confer the ability to colonize the intestinal tract at high levels.

In contrast to PTS^{clin}, a plasmid-carried genetic element involved in raffinose sugar utilization was identified to be enriched in *E. faecium* isolates from healthy individuals and is almost absent from clinical isolates [48] (Chapter 5). Perhaps understandably, due to its prominence as a nosocomial pathogen, most studies of *E. faecium* have so far focused on traits that are enriched in clinical isolates. Utilization of raffinose, as described in this work, may represent a trait specific for commensal *E. faecium* strains of healthy individuals. The ability to utilize the non-digestible plant sugar raffinose as carbon source may shape the ecological distribution of *E. faecium* by increasing the colonization fitness in the intestine of healthy humans.

We also identified an *E. faecium* gene cluster that is responsible for utilization of maltodextrin, a carbon source of potential importance for the growth of *E. faecium* in the gut (Chapter 6). This gene cluster is conserved in nearly all *E. faecium* genomes and therefore represents a basic biological trait of *E. faecium*. The data provides mechanistic insights on the ability of *E. faecium* to utilize maltodextrins as carbon source, which could contribute to the understanding of the success of *E. faecium* as a gut commensal.

Carbohydrate metabolism is a fundamental function in bacteria and is an important determinant for niche-adaptation by commensal bacteria [49, 50]. Its crucially important

role in gut colonization of commensals and opportunistic pathogens has been demonstrated by previous studies [51-55]. Our findings indicate that carbohydrate metabolism has a similarly important role in the ecological success of *E. faecium*, both as a gut commensal and as a nosocomial pathogen. The ability of efficiently unitizing carbohydrates in the gut is important but it also needs other traits, such as the ability to tolerate the antimicrobial effects of bile, to successfully colonize the gut. In *Listeria monocytogenes, E*. *faecalis* and *Salmonella enterica*, the disruption of bile tolerance loci leads to impaired intestinal survival [56-58], while a mutant of *E*. *coli* with elevated ability in bile tolerance displays a fitness advantage in intestinal colonization [59]. We identified a gene *gltK* which is important for bile tolerance of *E. faecium* and investigated the transcriptional response of *E*. *faecium* to bile salts (Chapter 4). Our findings contribute to a better understanding of the different mechanisms that have evolved in *E. faecium* to colonize the harsh environment of intestine.

Conclusion

This thesis describes the development and implementation of new molecular tools for both genome-wide and targeted genetic analyses in *E. faecium*. These tools provide a platform for functional genomic-based studies and allowed studies into the biology of this increasingly important nosocomial pathogen. We used these tools in studies to identify and functionally characterize genetic determinants that are involved in ampicillin resistance, bile salt resistance, carbohydrate utilization and intestinal colonization (Table 1). These determinants contribute importantly to the ecological success of *E. faecium* as a human commensal bacterium and as an emerging nosocomial pathogen. The knowledge generated by this study may be applied for development of therapeutic strategies against the colonization and infection by *E. faecium*.

Table 1. Genetic determinants identified and functionally characterized in this thesis

Gene name	LocusTag	Annotation	Function determined	Genetic approaches used in identification and characterization	Gene distribution
ddcP	EfmE1162_0447	D-alanyl-D-alanine carboxypeptidase	ampicillin and lysozyme resistance	M-TraM, transcriptome analysis, markerless gene deletion, *in trans* complementation	core genome
ldt_{fm}	EfmE1162_1886	beta-lactam-insensitive peptidoglycan transpeptidase	ampicillin and lysozyme resistance	M-TraM, transcriptome analysis, markerless gene deletion, *in trans* complementation	core genome
pgt	EfmE1162_0975	glycosyl transferase, group 2 family protein	ampicillin resistance	M-TraM, transcriptome analysis, markerless gene deletion	core genome
lytG	EfmE1162_2487	exo-glucosaminidase	not determined	M-TraM, transcriptome analysis, markerless gene deletion, *in trans* complementation	core genome
pbp5	EfmE1162_0386	low affinity penicillin-binding protein 5	ampicillin resistance	M-TraM, transcriptome analysis, markerless gene deletion, *in trans* complementation	core genome
ptsD	EfmE1162_1918	PTS system mannose/fructose /sorbose family IID componen	intestinal colonization	comparative genomics analysis, markerless gene deletion, *in trans* complementation	enriched in clinical isolates
gltK	EfmE1162_1760	glutamate/aspartate transport system permease protein	bile salt resistance	M-TraM, transcriptome analysis, markerless gene deletion, *in trans* complementation	core genome
gspA	EfmE1162_1186	general stress protein A	not determined	M-TraM, transcriptome analysis, markerless gene deletion, *in trans* complementation	conserved in 66 of 69 genomes
agaA	EfmE980_0223	alpha-galactosidase	not determined	comparative genomics analysis, insertional gene disruption, PFGE, S1-PFGE	enriched in community isolates
agaB	EfmE980_0219	alpha-galactosidase AgaN	raffinose utilization	comparative genomics analysis, insertional gene disruption, PFGE, S1-PFGE	enriched in community isolates
mdxR	EfmE1162_2133	transcriptional regulator	positive regulation of maltodextrin utilization	transcriptome analysis, markerless gene deletion, *in trans* complementation	core genome
mdxABCD-pulA	EfmE1162_0366-EfmE1162_0370	maltose/maltodextrin ABC transporter, and neopullulanase	maltodextrin utilization	transcriptome analysis, markerless gene deletion, qRT-PCR	conserved in 66 of 69 genomes

References

1. Arias CA, Murray BE: The rise of the *Enterococcus*: beyond vancomycin resistance. *Nat Rev Microbiol* 2012, 10(4):266-278.
2. Willems RJ, van Schaik W: Transition of *Enterococcus faecium* from commensal organism to nosocomial pathogen. *Future Microbiol* 2009, 4(9):1125-1135.
3. Hidron AI, Edwards JR, Patel J, Horan TC, Sievert DM, Pollock DA, Fridkin SK: 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* 2008, 29(11):996-1011.
4. Willems RJ, Top J, van Schaik W, Leavis H, Bonten M, Siren J, Hanage WP, Corander J: Restricted gene flow among hospital subpopulations of *Enterococcus faecium*. *MBio* 2012, 3(4):e00151-00112.
5. Homan WL, Tribe D, Poznanski S, Li M, Hogg G, Spalburg E, Van Embden JD, Willems RJ: Multilocus sequence typing scheme for *Enterococcus faecium*. *J Clin Microbiol* 2002, 40(6):1963-1971.
6. Willems RJ, Hanage WP, Bessen DE, Feil EJ: Population biology of Gram-positive pathogens: high-risk clones for dissemination of antibiotic resistance. *FEMS Microbiol Rev* 2011, 35(5):872-900.
7. Willems RJ, Top J, van Santen M, Robinson DA, Coque TM, Baquero F, Grundmann H, Bonten MJ: Global spread of vancomycin-resistant *Enterococcus faecium* from distinct nosocomial genetic complex. *Emerg Infect Dis* 2005, 11(6):821-828.
8. Galloway-Pena J, Roh JH, Latorre M, Qin X, Murray BE: Genomic and SNP analyses demonstrate a distant separation of the hospital and community-associated clades of *Enterococcus faecium*. *PLoS One* 2012, 7(1):e30187.
9. Qin X, Galloway-Pena JR, Sillanpaa J, Roh JH, Nallapareddy SR, Chowdhury S, Bourgogne A, Choudhury T, Muzny DM, Buhay CJ *et al*: Complete genome sequence of *Enterococcus faecium* strain TX16 and comparative genomic analysis of *Enterococcus faecium* genomes. *BMC Microbiol* 2012, 12:135.
10. Palmer KL, Godfrey P, Griggs A, Kos VN, Zucker J, Desjardins C, Cerqueira G, Gevers D, Walker S, Wortman J *et al*: Comparative genomics of enterococci: variation in *Enterococcus faecalis*, clade structure in *E. faecium*, and defining characteristics of *E. gallinarum* and *E. casseliflavus*. *MBio* 2012, 3(1):e00318-00311.
11. Leavis HL, Willems RJ, van Wamel WJ, Schuren FH, Caspers MP, Bonten MJ: Insertion sequence-driven diversification creates a globally dispersed emerging multiresistant subspecies of *E. faecium*. *PLoS Pathog* 2007, 3(1):e7.
12. Courvalin P: Vancomycin resistance in gram-positive cocci. *Clin Infect Dis* 2006, 42 Suppl 1:S25-34.
13. Werner G, Coque TM, Hammerum AM, Hope R, Hryniewicz W, Johnson A, Klare I, Kristinsson KG, Leclercq R, Lester CH *et al*: Emergence and spread of vancomycin resistance among enterococci in Europe. *Euro Surveill* 2008, 13(47).
14. Xu X, Lin D, Yan G, Ye X, Wu S, Guo Y, Zhu D, Hu F, Zhang Y, Wang F *et al*: *vanM*, a new glycopeptide resistance gene cluster found in *Enterococcus faecium*. *Antimicrob Agents Chemother* 2010, 54(11):4643-4647.
15. Lebreton F, Depardieu F, Bourdon N, Fines-Guyon M, Berger P, Camiade S, Leclercq R, Courvalin P, Cattoir V: D-Ala-d-Ser VanN-type transferable vancomycin resistance in *Enterococcus faecium*. *Antimicrob Agents Chemother* 2011, 55(10):4606-4612.
16. Fontana R, Aldegheri M, Ligozzi M, Lopez H, Sucari A, Satta G: Overproduction of a low-affinity penicillin-binding protein and high-level ampicillin resistance in *Enterococcus faecium*. *Antimicrob Agents Chemother* 1994, 38(9):1980-1983.

17. Fontana R, Grossato A, Rossi L, Cheng YR, Satta G: Transition from resistance to hypersusceptibility to beta-lactam antibiotics associated with loss of a low-affinity penicillin-binding protein in a *Streptococcus faecium* mutant highly resistant to penicillin. *Antimicrob Agents Chemother* 1985, 28(5):678-683.
18. Williamson R, le Bouguenec C, Gutmann L, Horaud T: One or two low affinity penicillin-binding proteins may be responsible for the range of susceptibility of *Enterococcus faecium* to benzylpenicillin. *J Gen Microbiol* 1985, 131(8):1933-1940.
19. Mainardi JL, Morel V, Fourgeaud M, Cremniter J, Blanot D, Legrand R, Frehel C, Arthur M, Van Heijenoort J, Gutmann L: Balance between two transpeptidation mechanisms determines the expression of beta-lactam resistance in *Enterococcus faecium*. *J Biol Chem* 2002, 277(39):35801-35807.
20. Mainardi JL, Legrand R, Arthur M, Schoot B, van Heijenoort J, Gutmann L: Novel mechanism of beta-lactam resistance due to bypass of DD-transpeptidation in *Enterococcus faecium*. *J Biol Chem* 2000, 275(22):16490-16496.
21. Heikens E, Bonten MJ, Willems RJ: Enterococcal surface protein Esp is important for biofilm formation of *Enterococcus faecium* E1162. *J Bacteriol* 2007, 189(22):8233-8240.
22. Leendertse M, Heikens E, Wijnands LM, van Luit-Asbroek M, Teske GJ, Roelofs JJ, Bonten MJ, van der Poll T, Willems RJ: Enterococcal surface protein transiently aggravates *Enterococcus faecium*-induced urinary tract infection in mice. *J Infect Dis* 2009, 200(7):1162-1165.
23. Sillanpaa J, Nallapareddy SR, Singh KV, Prakash VP, Fothergill T, Ton-That H, Murray BE: Characterization of the *ebp*$_{fm}$ pilus-encoding operon of *Enterococcus faecium* and its role in biofilm formation and virulence in a murine model of urinary tract infection. *Virulence* 2010, 1(4):236-246.
24. Nallapareddy SR, Singh KV, Murray BE: Contribution of the collagen adhesin Acm to pathogenesis of *Enterococcus faecium* in experimental endocarditis. *Infect Immun* 2008, 76(9):4120-4128.
25. Nallapareddy SR, Weinstock GM, Murray BE: Clinical isolates of *Enterococcus faecium* exhibit strain-specific collagen binding mediated by Acm, a new member of the MSCRAMM family. *Mol Microbiol* 2003, 47(6):1733-1747.
26. Choudhury T, Singh KV, Sillanpaa J, Nallapareddy SR, Murray BE: Importance of two *Enterococcus faecium* loci encoding Gls-like proteins for in vitro bile salts stress response and virulence. *J Infect Dis* 2011, 203(8):1147-1154.
27. van Schaik W, Willems RJ: Genome-based insights into the evolution of enterococci. *Clin Microbiol Infect* 2010, 16(6):527-532.
28. van Schaik W, Top J, Riley DR, Boekhorst J, Vrijenhoek JE, Schapendonk CM, Hendrickx AP, Nijman IJ, Bonten MJ, Tettelin H *et al*: Pyrosequencing-based comparative genome analysis of the nosocomial pathogen *Enterococcus faecium* and identification of a large transferable pathogenicity island. *BMC Genomics* 2010, 11:239.
29. Rao CV, Kirby JR, Arkin AP: Design and diversity in bacterial chemotaxis: a comparative study in *Escherichia coli* and *Bacillus subtilis*. *PLoS Biol* 2004, 2(2):E49.
30. Jacob F: Evolution and tinkering. *Science* 1977, 196(4295):1161-1166.
31. Zhang X, Paganelli FL, Bierschenk D, Kuipers A, Bonten MJ, Willems RJ, van Schaik W: Genome-wide identification of ampicillin resistance determinants in *Enterococcus faecium*. *PLoS Genet* 2012, 8(6):e1002804.
32. Lebreton F, van Schaik W, Sanguinetti M, Posteraro B, Torelli R, Le Bras F, Verneuil N, Zhang X, Giard JC, Dhalluin A *et al*: AsrR is an oxidative stress sensing regulator modulating *Enterococcus faecium* opportunistic traits, antimicrobial resistance, and pathogenicity. *PLoS Pathog* 2012, 8(8):e1002834.
33. Goodman AL, Wu M, Gordon JI: Identifying microbial fitness determinants by insertion sequencing using genome-wide transposon mutant libraries. *Nat Protoc* 2011, 6(12):1969-1980.

34. van Opijnen T, Bodi KL, Camilli A: Tn-seq: high-throughput parallel sequencing for fitness and genetic interaction studies in microorganisms. *Nat Methods* 2009, 6(10):767-772.
35. Langridge GC, Phan MD, Turner DJ, Perkins TT, Parts L, Haase J, Charles I, Maskell DJ, Peters SE, Dougan G *et al*: Simultaneous assay of every *Salmonella Typhi* gene using one million transposon mutants. *Genome Res* 2009, 19(12):2308-2316.
36. Hollenbeck BL, Rice LB: Intrinsic and acquired resistance mechanisms in *enterococcus*. *Virulence* 2012, 3(5):421-433.
37. Hood MI, Becker K, Roux CM, Dunman PM, Skaar EP: Genetic determinants of intrinsic colistin tolerance in *Acinetobacter baumannii*. *Infect Immun* 2012.
38. Poirel L, Bonnin RA, Nordmann P: Genetic basis of antibiotic resistance in pathogenic *Acinetobacter* species. *IUBMB Life* 2011, 63(12):1061-1067.
39. Alvarez-Ortega C, Wiegand I, Olivares J, Hancock RE, Martinez JL: Genetic determinants involved in the susceptibility of *Pseudomonas aeruginosa* to beta-lactam antibiotics. *Antimicrob Agents Chemother* 2010, 54(10):4159-4167.
40. Breidenstein EB, Khaira BK, Wiegand I, Overhage J, Hancock RE: Complex ciprofloxacin resistome revealed by screening a *Pseudomonas aeruginosa* mutant library for altered susceptibility. *Antimicrob Agents Chemother* 2008, 52(12):4486-4491.
41. Fajardo A, Martinez-Martin N, Mercadillo M, Galan JC, Ghysels B, Matthijs S, Cornelis P, Wiehlmann L, Tummler B, Baquero F *et al*: The neglected intrinsic resistome of bacterial pathogens. *PLoS One* 2008, 3(2):e1619.
42. Lee S, Hinz A, Bauerle E, Angermeyer A, Juhaszova K, Kaneko Y, Singh PK, Manoil C: Targeting a bacterial stress response to enhance antibiotic action. *Proc Natl Acad Sci U S A* 2009, 106(34):14570-14575.
43. Schurek KN, Marr AK, Taylor PK, Wiegand I, Semenec L, Khaira BK, Hancock RE: Novel genetic determinants of low-level aminoglycoside resistance in *Pseudomonas aeruginosa*. *Antimicrob Agents Chemother* 2008, 52(12):4213-4219.
44. Girgis HS, Hottes AK, Tavazoie S: Genetic architecture of intrinsic antibiotic susceptibility. *PLoS One* 2009, 4(5):e5629.
45. Gilmore MS, Ferretti JJ: Microbiology. The thin line between gut commensal and pathogen. *Science* 2003, 299(5615):1999-2002.
46. Ubeda C, Taur Y, Jenq RR, Equinda MJ, Son T, Samstein M, Viale A, Socci ND, van den Brink MR, Kamboj M *et al*: Vancomycin-resistant *Enterococcus* domination of intestinal microbiota is enabled by antibiotic treatment in mice and precedes bloodstream invasion in humans. *J Clin Invest* 2010, 120(12):4332-4341.
47. Donskey CJ, Chowdhry TK, Hecker MT, Hoyen CK, Hanrahan JA, Hujer AM, Hutton-Thomas RA, Whalen CC, Bonomo RA, Rice LB: Effect of antibiotic therapy on the density of vancomycin-resistant enterococci in the stool of colonized patients. *N Engl J Med* 2000, 343(26):1925-1932.
48. Zhang X, Vrijenhoek JE, Bonten MJ, Willems RJ, van Schaik W: A genetic element present on megaplasmids allows *Enterococcus faecium* to use raffinose as carbon source. *Environ Microbiol* 2011, 13(2):518-528.
49. Cantarel BL, Lombard V, Henrissat B: Complex carbohydrate utilization by the healthy human microbiome. *PLoS One* 2012, 7(6):e28742.
50. Abubucker S, Segata N, Goll J, Schubert AM, Izard J, Cantarel BL, Rodriguez-Mueller B, Zucker J, Thiagarajan M, Henrissat B *et al*: Metabolic reconstruction for metagenomic data and its application to the human microbiome. *PLoS Comput Biol* 2012, 8(6):e1002358.
51. Koropatkin NM, Cameron EA, Martens EC: How glycan metabolism shapes the human gut microbiota. *Nat Rev Microbiol* 2012, 10(5):323-335.
52. Chang DE, Smalley DJ, Tucker DL, Leatham MP, Norris WE, Stevenson SJ, Anderson AB, Grissom JE, Laux DC, Cohen PS *et al*: Carbon nutrition of *Escherichia coli* in the mouse intestine. *Proc Natl Acad Sci U S A* 2004, 101(19):7427-7432.

53. Marco ML, Peters TH, Bongers RS, Molenaar D, van Hemert S, Sonnenburg JL, Gordon JI, Kleerebezem M: Lifestyle of *Lactobacillus plantarum* in the mouse caecum. *Environ Microbiol* 2009, 11(10):2747-2757.
54. Thomas M, Wrzosek L, Ben-Yahia L, Noordine ML, Gitton C, Chevret D, Langella P, Mayeur C, Cherbuy C, Rul F: Carbohydrate metabolism is essential for the colonization of *Streptococcus thermophilus* in the digestive tract of gnotobiotic rats. *PLoS One* 2011, 6(12):e28789.
55. Sonnenburg ED, Zheng H, Joglekar P, Higginbottom SK, Firbank SJ, Bolam DN, Sonnenburg JL: Specificity of polysaccharide use in intestinal bacteroides species determines diet-induced microbiota alterations. *Cell* 2010, 141(7):1241-1252.
56. Begley M, Sleator RD, Gahan CG, Hill C: Contribution of three bile-associated loci, *bsh*, *pva*, and *btlB*, to gastrointestinal persistence and bile tolerance of *Listeria monocytogenes*. *Infect Immun* 2005, 73(2):894-904.
57. Kristich CJ, Wells CL, Dunny GM: A eukaryotic-type Ser/Thr kinase in *Enterococcus faecalis* mediates antimicrobial resistance and intestinal persistence. *Proc Natl Acad Sci U S A* 2007, 104(9):3508-3513.
58. Reynolds MM, Bogomolnaya L, Guo J, Aldrich L, Bokhari D, Santiviago CA, McClelland M, Andrews-Polymenis H: Abrogation of the twin arginine transport system in *Salmonella enterica serovar* Typhimurium leads to colonization defects during infection. *PLoS One* 2011, 6(1):e15800.
59. De Paepe M, Gaboriau-Routhiau V, Rainteau D, Rakotobe S, Taddei F, Cerf-Bensussan N: Trade-off between bile resistance and nutritional competence drives *Escherichia coli* diversification in the mouse gut. *PLoS Genet* 2011, 7(6):e1002107.

Samenvatting

De Gram-positieve bacterie *Enterococcus faecium* is lang beschouwd als een onschadelijke microbe in de darmflora van zoogdieren. In de laatste twee decennia heeft deze bacterie zich echter snel ontpopt tot een belangrijke opportunistische pathogeen die infecties veroorzaakt in patiënten in ziekenhuizen over de hele wereld. Het verwerven van eigenschappen zoals resistentie tegen vele antibiotica en de mogelijkheid om de darmen van gehospitaliseerde patiënten tot een hoog niveau te koloniseren hebben bijgedragen aan de recente ecologische overgang van *E. faecium* van darm-commensaal tot ziekenhuis-pathogeen. *E. faecium* is in staat om ziekenhuis-infecties zoals bacteriëmie, endocarditis, urineweg - en postoperatieve wondinfecties te veroorzaken. Deze infecties leiden tot een hogere patiënt mortaliteit en toenemende kosten van de gezondheidszorg.

Er zijn tot op heden talrijke studies gedaan naar de moleculaire epidemiologie en populatiestructuur van *E. faecium*. Deze onderzoeken leidden tot de identificatie van een klinische sub-populatie van *E. faecium*. Analyses door middel van vergelijkende genomica toonden aan dat klinische *E. faecium* isolaten genetisch verschillend zijn van de darm isolaten van gezonde mensen. Hoewel de kennis over *E. faecium* door epidemiologische studies en genoom sequentie analyses in de afgelopen jaren snel is gestegen, staat het wetenschappelijk onderzoek aan de functionele karakterisering van genetische determinanten van *E. faecium* die bijdragen aan het succes van deze bacterie als darm-commensaal en als ziekenhuis-pathogeen nog in de kinderschoenen. Het onderzoek op dit gebied is lang belemmerd door het ontbreken van geschikte methodes voor genetische studies in *E. faecium*.

Als onderdeel van de studies beschreven in dit proefschrift hebben wij moleculaire methodes en technieken ontwikkeld die de efficiënte genetische manipulatie van *E. faecium* mogelijk maken. Dit heeft studies op basis van functionele genomica mogelijk gemaakt en hierdoor konden de genetische determinanten van antibioticum resistentie en niche aanpassing van *E. faecium* worden geïdentificeerd en bestudeerd, hetgeen uiteindelijk heeft geleid tot de karakterisatie van de moleculaire mechanismen die bijdragen aan het ecologische succes van deze ziekenhuis-pathogeen. In hoofdstuk 2 wordt een set van moleculaire technieken ontwikkeld voor functioneel-genomische studies van *E. faecium*, die ons in staat hebben gesteld om zowel ‘high-throughput’ genoom-wijde analyses als gen-gericht onderzoek uit te voeren in dit organisme. Deze ontwikkelingen hebben geleid tot de identificatie van *E. faecium* genen die bijdragen tot resistentie tegen ampicilline. In hoofdstuk 3 beschrijven wij een genetische determinant die verrijkt is in klinische *E. faecium* isolaten en die bijdraagt aan de kolonisatie van de darm in klinische *E. faecium* stammen. In hoofdstuk 4 worden de genetische determinanten die betrokken zijn in de resistentie tegen galzouten en de transcriptionele respons na galzout stress in *E. faecium* beschreven. Hoofdstuk 5 beschrijft een genetisch element dat onderdeel is van horizontaal

overdraagbare megaplasmiden en waarmee *E. faecium* kan groeien op de niet-verteerbare suiker raffinose. Dit genetisch element komt vaker voor in stammen die geïsoleerd zijn uit de ontlasting van gezonde vrijwillgers dan in klinische isolaten. Hoofdstuk 6 beschrijft de identificatie en transcriptionele analyse van genetische determinanten die betrokken zijn bij de groei op maltodextrine door *E. faecium*. Ten slotte worden de bevindingen en toekomstperspectieven van het onderzoek beschreven in dit proefschrift besproken in hoofdstuk 7.

Dit proefschrift beschrijft de ontwikkeling en implementatie van nieuwe moleculaire technieken voor genoom-wijde en gen-gerichte analyses in *E. faecium*. Deze technieken vormen een platform voor functioneel-genomische studies en maken het mogelijk om onderzoek te verrichten naar de biologie van deze steeds belangrijker wordende ziekenhuis-pathogeen. Door de toepassing van functionele genomica in *E. faecium* hebben wij studies kunnen verrichten ter identificatie en functionele karakterisatie van genetische determinanten die betrokken zijn bij resistentie tegen ampicilline, galzout resistentie, koolhydraat metabolisme en darm-kolonisatie. Deze determinanten dragen belangrijk bij aan het ecologische succes van *E. faecium* als darm-commensaal én als ziekenhuis-pathogeen. De kennis die verkregen is in dit onderzoek kan worden toegepast voor de ontwikkeling van therapeutische strategieën gericht tegen darm-kolonisatie en infectie door *E. faecium*.

Acknowledgements/致谢

On April 28 2008, in an excited, curious, anxious and nervous mood, I took a flight from China to the Netherlands. This was my first trip abroad and I came to Utrecht for an interview for a PhD position. Luckily, I got this position, and four months later I started my PhD program at the UMC Utrecht. After four years of work, my thesis is now finished. During this period I have received great support from many people. I would like to acknowledge all their efforts, and especially, I want to thank the following people who have directly or indirectly contributed to this thesis.

Willem, thank you for helping me to adapt so smoothly to this new "niche". I really appreciate that you showed me and Xiaoman how to get a supermarket cart using a coin, you drove us to IKEA, you helped us with arranging our visa, housing, bank account, taxes… (I still own you many cups of beer). In the lab you taught me how to perform Southern blot, microarray hybridization, enzymatic assays… how to analyze results and how to write a manuscript… you inspirited and encouraged my imagination and creativity, correcting my mistakes, and guiding me into the right direction… I also enjoy our wonderful experiences in Burgers' zoo, in Zwolle "Safari", in Giethoorn, visiting your primary school and at conferences… Rob, it is great that I also had you as my co-promoter, because this equipped me with "double guarantee and dual safety" to ensure that I was on the right path to my destination. Every time when I asked you a question, no matter if it was scientifically or personally, you always patiently explained to me. I really appreciate your advice and ideas for my research, your comments on my manuscript, your guidance and encouragement, and also the delicious food produced by the "BBQ factory" in your garden… Marc, when I started working here you were getting busier with the management of the whole department, but you were always interested in and up-to-date with my research. I appreciate your humorous and inspiring words, which encouraged me to "dare to be good" and pursue greater achievements. I would also like to thank you and Helen for hosting us in Maastricht, and I am looking forward to the next time. Finally, "heel erg bedankt" to my promoters, I have "acquired" a lot of valuable "traits" from this four-year PhD training and pleasant experience with you. These "traits" will not only benefit me for my research, but also for my life, just as it is implicated in the PhD degree—Doctor of Philosophy.

Janetta, it was so nice of you to invite me and Xiaoman to visit your home and have a dinner with your family. We enjoyed it very much! Thank you for your care, help, flowers and gifts for me, Xiaoman and Minghe. It is so nice to work with you together and have you living close by. Also thank you for your help in the animal experiments, for which Prof. Tom van der Poll kindly provided lab space. I would like to give my special thanks to Elena and Fernanda, for helping me arrange the promotion. It is so nice to have you stand together with me in this special event. I wish you both and also Jan and Ana, the other PhD students in our group, good luck with your research and thesis! I am really happy that I will still stay

with you after my promotion. So we can still work together and have fun in bars, and at conferences... Damien, I would like to thank you for your great assistance in the last year of my PhD study. Your work has contributed a lot to my research and created more time for me to deal with my thesis. Mark, thank you for your support and your bioinformatics input contributed greatly to my research. Malbert, Joyce, Iacovos, Lisette, Annemarie and Miranda, thank you all for your help in my experiments. I would also like to thank the former PhD students of the Enterokokken group: Janetta, Helen, Antoni (also for your help and tips for my thesis printing), Esther, Marieke and Masja. Your achievements and experiences have provided me a broader platform and saved me from twists and turns in my research.

I would like to thank all my roommates (a special thank to Rutger, as he demanded) in room G04.643 and G04.527. Every year “caterpillars” come into our rooms and “butterflies” fly away. Time flies, and now it is my turn to leave. I will miss you all and our noisy-lively, busy-relaxing, messy-clean rooms.

Thank you to all my colleagues in the MMB department. I like the harmonious working atmosphere of our department and enjoy working with you.

Thanks to all my friends in the Netherlands!

感谢爸爸妈妈。一直以来他们都在为我无私的奉献着。我成长的每一步都离不开他们的关心、引导和支持。从小学到大学再到现在，我也从一个孩子成为了一个父亲，这让我更加体会到他们对我的无私的关爱。
感谢哥嫂以及我可爱的侄女明序带给我温暖的亲情。
感谢岳父岳母来荷兰探望我们，照顾明和，让我们在异国他乡感受到他们带来亲情。
感谢二舅、三舅、大爷和表舅等家里的长辈们以及兄长们对我学业的关心、引导和帮助。
在外求学期间我常常怀念爷爷和姥爷。这本论文也承载着他们对我的勉励和期许。

感谢吴清民教授和刘金华教授对我的培养和教导。
感谢好友秦堃多年来的帮助，交流和分享。
感谢我在荷兰的朋友们。他们让我的留学生活更加丰富多彩。
感谢我大学时的哥们儿，总是给我惊喜。

明和：谢谢你今天晚上给我分享你的饼干，并且按时睡觉……
小曼：你让我这些年的“寒窗苦读”变得轻松愉快。
临沂 — 济南/北京 — Utrecht ……
让我们一起向下一站出发吧!

List of publications

X. Zhang, J. Top, M. de Been, D. Bierschenk, M. Rogers, M. Leendertse, M. J. Bonten, T. van der Poll, R. J. Willems, W. van Schaik. 2013. Identification of a genetic determinant in clinical *Enterococcus faecium* strains which contributes to intestinal colonization during antibiotic treatment. **J Infect Dis** doi: 10.1093/infdis/jit076.

X. Zhang, F. L. Paganelli, D. Bierschenk, A. Kuipers, M. J. Bonten, R. J. Willems, and W. van Schaik. 2012. Genome-wide identification of ampicillin resistance determinants in *Enterococcus faecium*. **PLoS Genet** 8:e1002804.

F. Lebreton, W. van Schaik, M. Sanguinetti, B. Posteraro, R. Torelli, F. L. Bras, N. Verneuil, X. Zhang, A. Dhalluin, R. J. Willems, R. Leclercq and V. Cattoir. 2012. AsrR is an Oxidative stress sensing regulator modulating *Enterococcus faecium* opportunistic traits, antimicrobial resistance and pathogenicity. **PLoS Pathog** 8(8): e1002834.

X. Zhang, J. E. Vrijenhoek, M. J. Bonten, R. J. Willems, and W. van Schaik. 2011. A genetic element present on megaplasmids allows *Enterococcus faecium* to use raffinose as carbon source. **Environ Microbiol** 13:518-28.

X. Zhang, J. Ren, N. Li, W. Liu, and Q. Wu. 2009. Disruption of the BMEI0066 gene attenuates the virulence of *Brucella melitensis* and decreases its stress tolerance. **Int J Biol Sci** 5:570-7.

Y. L. Cong, J. Pu, Q. F. Liu, S. Wang, G. Z. Zhang, X. L. Zhang, W. X. Fan, E. G. Brown, and J. H. Liu. 2007. Antigenic and genetic characterization of H9N2 swine influenza viruses in China. **J Gen Virol** 88:2035-41.

J. Liu, H. Xiao, F. Lei, Q. Zhu, K. Qin, X. W. Zhang, X. L. Zhang, D. Zhao, G. Wang, Y. Feng, J. Ma, W. Liu, J. Wang, and G. F. Gao. 2005. Highly pathogenic H5N1 influenza virus infection in migratory birds. **Science** 309:1206.

X. Zhang, D. Bierschenk, J. Top, I. Anastasiou, M. J. Bonten, R. J. Willems, and W. van Schaik. Functional genomic analysis of bile salt resistance in *Enterococcus faecium.* Submitted for publication.

J. Top, F. L. Paganelli, X. Zhang, W. van Schaik, H. L. Leavis, M. van Luit-Asbroek, T. van der Poll, M. Leendertse, M. J. Bonten, R. J. Willems. The *Enterococcus faecium* enterococcal biofilm regulator, EbrB, regulates the *esp* gene cluster and is implicated in biofilm formation and intestinal colonization. Submitted for publication.

F. L. Paganelli, R. J. Willems, P. Jansen, X. Zhang, M. J. Bonten, H. L. Leavis. Identification and functional characterization of the major autolysin in *Enterococcus faecium.* Submitted for publication.

X. Zhang, M. Rogers, D. Bierschenk, M. J. Bonten, R. J. Willems, W. van Schaik. A LacI-family regulator activates maltodextrin metabolism of *Enterococcus faecium.* Manuscript in preparation.

X. Zhang, L. Visser, M. J. Bonten, R. J. Willems, W. van Schaik. Identification of conjugative mobile elements in *Enterococcus faecium*. Manuscript in preparation.

- Abstracts and presentations in (inter)national conferences

X. Zhang, J. Top, M. de Been, D. Bierschenk, M. Rogers, M. Leendertse, M. J. Bonten, T. van der Poll, R. J. Willems, W. van Schaik. Identification of a genetic determinant in clinical *Enterococcus faecium* strains which contributes to intestinal colonization during antibiotic treatment. **Gut Day.** Leuven, Belgium; 2012

X. Zhang, F. L. Paganelli, D. Bierschenk, A. Kuipers, M. J. Bonten, R. J. Willems, and W. van Schaik. Genome-wide identification of ampicillin resistance determinants in *Enterococcus faecium*. **112th ASM General Conference**. San Francisco, USA; 2012.

X. Zhang, F. L. Paganelli, D. Bierschenk, A. Kuipers, M. J. Bonten, R. J. Willems, and W. van Schaik. Genome-wide identification of ampicillin resistance determinants in *Enterococcus faecium*. **NVMM & NVvM Scientific Spring Meeting**. Arnhem, The Netherlands; 2012.

X. Zhang, I. Anastasiou, M. J. Bonten, R. J. Willems, and W. van Schaik**.** Comprehensive identification of genes required for antibiotic resistance and bile tolerance in the nosocomial pathogen *Enterococcus faecium*. **100th Anniversary NVvM & Scientific Spring Meeting**. Arnhem, The Netherlands; 2011.

X. Zhang, I. Anastasiou, M. J. Bonten, R. J. Willems, and W. van Schaik**.** Development of Microarray-based Transposon Mapping in *Enterococcus faecium* and its application in the identification of genes required for bile resistance. **ALW Platform Molecular Genetics Annual Meeting.** Lunteren, The Netherlands; 2010.

X. Zhang, I. Anastasiou, M. J. Bonten, R. J. Willems, and W. van Schaik**.** Development of Microarray-based Transposon Mapping (M-TraM) in *Enterococcus faecium* and its application in the identification of genes required for bile resistance. **3rd International ASM Conference on Enterococci**. Partland, USA; 2010.

X. Zhang, J. E. Vrijenhoek, M. J. Bonten, R. J. Willems, and W. van Schaik**.** A large conjugative plasmid confers the ability to utilize α-galactoside sugars as carbon source in *E. faecium*. **NVMM & NVvM Scientific Spring Meeting**. Arnhem, The Netherlands; 2010.

X. Zhang, J. E. Vrijenhoek, M. J. Bonten, R. J. Willems, and W. van Schaik**.** A novel conjugative element from *Enterococcus faecium* involved in α-galactoside metabolism. **ESCMID Conference on Enterococci**. Barcelona, Spain; 2009.

X. Zhang, J. E. Vrijenhoek, M. J. Bonten, R. J. Willems, and W. van Schaik**.** A novel conjugative element from *Enterococcus faecium* involved in α-galactoside metabolism. **FEMS**. Gothenburg, Sweden; 2009.

X. Zhang, J. E. Vrijenhoek, M. J. Bonten, R. J. Willems, and W. van Schaik. Metabolism of α-galactosides in *Enterococcus faecium.* **NVMM & NVvM Scientific Spring Meeting**. Arnhem, The Netherlands; 2009.

Curriculum Vitae

Xinglin Zhang was born on March 26, 1982 in Linyi, China. He started his bachelor study in 2000 at China Agricultural University (CAU), Beijing, majoring in Veterinary Medicine. In 2005 he obtained his B.Sc. degree after completion of his internship under supervision of Prof. Dr. Jinhua Liu, and continued his education at the College of Veterinary Medicine, CAU, as a postgraduate student under supervision of Prof. Dr. Qingmin Wu. In 2008 he obtained his M.Sc degree at CAU. In the same year, he started his PhD training at the Department of Medical Microbiology, University Medical Center Utrecht, and the Infection & Immunity Center, Utrecht University, The Netherlands, under supervision of Dr. Willem van Schaik, Dr. Rob Willems and Prof. Dr. Marc Bonten. The results of his PhD study are described in this thesis.

Xingin Zhang
xinglin.zhang@hotmail.com