Functional characterization of a gene cluster responsible for inositol catabolism associated with hospital-adapted isolates of Enterococcus faecium

Top, Janetta; Baan, Jery; Bisschop, Adinda; Arredondo-Alonso, Sergio; van Schaik, Willem; Willems, Rob J L

DOI:
10.1099/mic.0.001085

Functional characterization of a gene cluster responsible for inositol catabolism associated with hospital-adapted isolates of *Enterococcus faecium*

Running title: inositol metabolism in *E. faecium*

Janetta Top[1*], Jery Baan[1], Adinda Bisschop[1], Sergio Arredondo-Alonso[1], Willem van Schaik[1,2], and Rob J.L. Willems[1].

[1] Department of Medical Microbiology, University Medical Center Utrecht, Utrecht, The Netherlands

[2] Institute of Microbiology and Infection, College of Medical and Dental Sciences, University of Birmingham, Birmingham, UK.

[*]Corresponding author. Mailing address: Department of Medical Microbiology, Room G04.614, University Medical Center Utrecht, Heidelberglaan 100, 3584CX Utrecht, The Netherlands. Phone: +31-887551318. E-mail: j.top@umcutrecht.nl

Keywords: *Enterococcus faecium,* hospital-associated, inositol catabolism, ICE*Efm1*

Repository: Illumina MiSeq reads of two *E. faecium* strains have been deposited in the European Nucleotide Archive (ENA) with accession number PRJEB43191.

ABSTRACT

Enterococcus faecium is a nosocomial, multidrug-resistant pathogen. Whole genome sequence studies revealed that hospital-associated *E. faecium* isolates are clustered in a separate clade A1. Here, we investigated the distribution, integration site and function of a putative *iol* gene cluster that encodes for *myo*-inositol (MI) catabolism. This *iol* gene cluster was found as part of an ~20 kbp genetic element (*iol* element), integrated in ICE*Efm1* close to its integrase gene in *E. faecium* isolate E1679.

Among 1,644 *E. faecium* isolates, ICE*Efm1* was found in 789/1,227 (64.3%) clade A1 and 3/417 (0.7%) non-clade A1 isolates. The *iol* element was present at a similar integration site in 180/792 (22.7%) ICE*Efm1*-containing isolates. Examination of the phylogenetic tree revealed genetically closely related isolates that differed in presence/absence of ICE*Efm1* and/or *iol* element, suggesting either independent acquisition or loss of both elements.

E. faecium iol gene cluster containing isolates E1679 and E1504 were able to grow in minimal medium with only *myo*-inositol as carbon source , while the *iolD*-deficient mutant in E1504 (E1504Δ*iolD*) lost this ability and an *iol* gene cluster negative recipient strain gained this ability after acquisition of ICE*Efm1* by conjugation from donor strain E1679. Gene expression profiling revealed that the *iol* gene cluster is only expressed in the absence of other carbon sources. In an intestinal colonization mouse model the colonization ability of E1504Δ*iolD* mutant was not affected relative to the wild-type E1504 strain.

In conclusion, we describe and functionally characterise a gene cluster involved in MI catabolism that is associated with the ICE*Efm1* island in hospital-associated *E. faecium* isolates. We were unable to show that this gene cluster provides a competitive advantage during gut colonisation in a mouse model. Therefore, to what extent this gene cluster contributes to the spread and ecological specialisation of ICE*Efm1*-carrying hospital-associated isolates remains to be investigated.

47 Abbreviations: ICE*Efm1,* integrative conjugative element *Efm1*; MI, *myo*-inositol; M1-MI, minimal

48 medium containing *myo*-inositol

49

INTRODUCTION

Enterococcus faecium is a commensal of the gastrointestinal tract, but also an important cause of nosocomial infections, especially in immunocompromised patients (1). An important contributing factor is that these *E. faecium* isolates have acquired resistance to almost all available antibiotics, including ampicillin, gentamicin and vancomycin and less frequently against the more recently introduced antibiotics linezolid, daptomycin and tigecycline (2). Previous whole genome sequencing (WGS)-based studies revealed a split in the *E. faecium* population in a hospital-associated clade (clade A) and community associated clade (clade B) (3,4). Clade A was further subdivided in clade A1, mainly representing hospital-associated isolates and clade A2, mainly representing animal isolates (5). Recently it became clear that the clade A2 animal isolates do not form a monophyletic subclade and no longer support the split of clade A isolates into two single subclades (2,6–8). In a recent study analyzing 1,644 clade A isolates, 98% of hospital associated isolates clustered in clade A1, representing the most frequent source in this clade (89%) (9). Isolates clustering outside clade A1 are now considered non-clade A1 (9).

In addition to antibiotic resistance, several virulence factors have been identified to be enriched among the hospital-associated *E. faecium* isolates, including cell-wall associated proteins involved in biofilm formation like the Enterococcal surface protein (Esp), a distinct subset of Secreted antigen A (SagA) containing a specific number of serine-threonine repeats, the biofilm and endocarditis-associated permease A (BepA) (10–12), and microbial surface components recognizing adhesive matrix molecules (MSCRAMMs), see review (13). Furthermore, hospital-associated *E. faecium* isolates are enriched for the presence of two genomic islands (putatively) involved in carbohydrate metabolism (14,15). For one of these islands encoding for a phosphotransferase system (PTS), deletion of *ptsD*, predicted to encode the enzyme IID subunit of this PTS, significantly impaired the ability of *E. faecium* to colonize the murine intestinal tract during antibiotic treatment (15).

The aforementioned *esp* gene is encoded on ICE*Efm1* (16,17). In 2010, we published the first comparative genomic study of *E. faecium,* which included in total 7 *E. faecium* genomes (18). ICE*Efm1* was present in three of these genomes, including the hospital-associated isolates E1162, U317 and E1679. Comparison of ICE*Efm1* in these isolates revealed the presence of a genetic element of ~20 kbp in strain E1679, predicted to include a cluster of genes encoding for inositol catabolism, designated as *iol* element (18). Inositol is widely found in natural environments like soil, plants and aquatic environments (19) and exists in various isomeric forms such as *myo*-, D-*chiro*-, *scyllo*- and *neo*-inositol. Of these, *myo*-inositol (MI) is by far the most prevalent form in nature (19). In this study, we determined the presence of ICE*Efm1* and the *iol* element in a collection of 1,644 clade A isolates described previously (9). Furthermore, we functionally characterized the *iol* gene cluster and investigated whether the capability of *E. faecium* to catabolize MI could provide those strains a competitive advantage in a mouse colonization model.

MATERIALS AND METHODS

Bacterial strains, plasmids and growth conditions

E. faecium and *Escherichia coli* strains and plasmids used in this study are listed in Table 1. The *E. coli* strains DH5α (Invitrogen) and EC1000 (20) were grown in Luria-Bertani (LB) medium. *E. faecium* was grown on blood agar (BA; tryptic soy agar supplemented with 5% sheep red blood cells); BD, Alphen aan den Rijn, The Netherlands), in brain heart infusion (BHI) medium or supplemented with 1% *myo*-inositol (BHI-MI) and in a previously described M1 medium that minimizes growth of *E. faecium* when no carbon source is added and consists of 10 g of tryptone and 0.5 g of yeast extract in 1 l of phosphate-buffered saline (PBS; 138 mM NaCl, 2.7 mM KCl, 140 mM Na2HPO4, 1.8 mM KH2PO4, adjusted to pH 7.4 with HCl) (21), supplemented with different carbon sources, including 1% *myo*-inositol (M1-MI) and 1% D-*chiro*-inositol (M1-DCI) at 37°C. For *E. faecium*, the antibiotics gentamicin and spectinomycin were used in concentrations of 300 µg/ml. For *E. coli*, gentamicin and

spectinomycin were used in concentrations of 30 µg/ml and 100 µg/ml, respectively. All antibiotics were obtained from Sigma-Aldrich (Saint Louis, MO).

Bioinformatic analysis

In order to determine the presence of ICE*Efm1* and the *iol* element in the 1,644 clade A isolates described in a previous study (9), the complete genome sequence of strain AUS0085 (accession number NC_021994), which contains ICE*Efm1* including the *iol* element, was used as reference to generate an ad-hoc whole genome MLST (wgMLST) scheme using Seqsphere+ version 5.0.0 (Ridom GmbH, Münster, Germany, https://www.ridom.de/seqsphere/). In AUS0085, ICE*Efm1* encompass locus tags EFAU085-02788 (encoding an LPXTG protein) to EFAU085-02871 (encoding the integrase). Genome assemblies were imported into SeqSphere+ as Fasta files. Using this ad-hoc wgMLST schema, the presence/absence of ICE*Efm1* and the *iol* element was determined in the set of 1,644 isolates and added to the metadata (Suppl. Table 1) of the 1,644 clade A isolates of the previously described microreact project, thereby generating the updated project, accessible at https://microreact.org/project/pmCxZKBhMrTAL85aNKXbAh. Patristic distances were extracted from the core-genome based tree using the cophenetic function available in the R package ape (version 5.4-1) (22,23).

Next, we determined the integration site of the *iol* element in ICE*Efm1* as identified in strain *E. faecium* E1679 (18) using the recently described Panaroo pipeline (version 1.2.3) (24). Panaroo was run on 'strict' mode and the function 'panaroo-gene-neighbourhood' (https://gtonkinhill.github.io/panaroo/#/post/gene_neighbourhood) was considered to explore the genome graph neighborhood for the first and last gene of the *iol* element, respectively. The genome sequence of *E. faecium* strain E7356 was used as reference since the entire ICE*Efm1* including the *iol* element was located on a single contig. In this strain the first and last gene of the *iol* element is represented by 'E7356_00286' and 'E7356_00303', respectively. Presence of putative transcription terminators were predicted using RNAfold (http://rna.tbi.univie.ac.at/).

Determination of growth curves

A BioScreen C instrument (Oy Growth Curves AB, Helsinki, Finland) was used to monitor growth of *iol* gene cluster-containing strains, the effects of *iolD* deletion on bacterial growth and acquisition of the *iol* gene cluster in different media as indicated. All strains were grown overnight in BHI. Cells were inoculated at an initial OD_{660} of 0.0025 into 300 µl M1-MI and/or M1-DCI incubated in the Bioscreen C system at 37°C with continuous shaking and absorbance of 600 nm (A600) was recorded every 15 min for 18 hours. Each experiment was performed in triplicate.

Mobilization of ICE*Efm1* using *myo*-inositol as selection marker

The mobilization of ICE*Efm1* was studied by performing filter-mating experiments as previously described (18). The ICE*Efm1*-containing strain *E. faecium* E1679 was used as donor and a rifampicin- and fusidic acid-resistant derivative of *E. faecium* O12 (25) designated *E. faecium* E4658 and a previously used strain *E. faecium* BM4105RF (18) were used as recipient. After filter-mating, aliquots of 200 µl were spread on M1-MI plates containing rifampicin and fusidic acid, both at 25 µg/ml, to select for transconjugants. For quantification of the number of cells of the donor and recipient, 10 µl of a 10-fold dilution series was plated on respectively M1-MI plates and BHI plates containing rifampicin and fusidic acid. Transconjugation efficiency was calculated as the viable counts on the M1-MI plates with rifampicin and fusidic acid divided by the viable counts on BHI plates with rifampicin and fusidic acid. PFGE on SmaI-digested total DNA was performed as described previously (26) to confirm that the transconjugant had the same genetic background as E4658. Lambda Ladder PFGE Marker (New England Biolabs, Ipswich MA) was used to estimate the sizes of the PFGE fragments. Southern blotting and probe hybridization was performed as described previously (27). The probes used in the hybridizations were generated by PCR with the primer-pairs iolA-F and adh-RT-R (Suppl. Table 2). In addition, WGS was performed to determine the integration site of ICE*Efm1* of the recipient strain E4658 and its transconjugant E7855 as previously described (28). In brief, Genomic DNA was isolated from overnight cultures in BHI broth using the Wizard Genomic DNA Purification Kit (Promega), according to the manufacturer's instructions. Library preparation for

sequencing was performed using the Nextera XT Kit. Sequencing runs were generated using 250 nucleotide paired-end sequencing on the MiSeq platform (Illumina). Raw reads were trimmed, assembled into contigs using SPades (vs. 3.6.2) and annotated using PROKKA (vs. 1.11 (29)). Illumina MiSeq reads of these two *E. faecium* strains have been deposited in the European Nucleotide Archive (ENA) with accession number PRJEB43191.

Generation of a *iolD* markerless deletion mutant

The previously described Cre-*lox* system was used to generate a markerless deletion mutant in the *iolD* gene (30). We were unable to generate this mutant in strain E1679 in which the *iol* gene cluster was first identified, since this strain was resistant to both gentamicin and spectinomycin, which are selection markers for the generation of targeted mutants. As these antibiotics are required in the generation of deletion mutants, we screened for other *iol* gene cluster-containing strains and found that strain *E. faecium* E1504 was susceptible for both antibiotics and was therefore used to generate an *iolD* mutant. For the amplification of the 5`-flanking region, we used primers iolDUp-F-XhoI and iolDUp-R-EcoRI and for the 3`-flanking regions primers iolDDn-F-EcoRI and iolDDn-R-SmaI (Suppl. Table 2). Generation of a marked deletion mutant was performed as described (30) and was confirmed by PCR using the *iolD* check-up and check-down primers (Suppl. Table 2). Removal of the gentamicin resistance marker was obtained by electroporation of pWS3-Cre into the marked deletion mutants as described (30). Loss of the marker was confirmed by PCR using the *iolD* check-up and check-down primers.

RNA isolation, reverse transcription and quantitative real-time RT-PCR (qRT-PCR)

In order to investigate whether the *iol* gene cluster is organized as an operon, *E. faecium* E1504 was grown in M1-MI to an OD_{600} of 0.3 prior to RNA isolation. To investigate growth condition-dependent expression of the *iol* gene cluster, E1504 was grown in BHI, BHI-MI, and M1-MI. RNA isolation, cDNA synthesis and quantitative real time PCR (qRT-PCR) was performed as previously described (10). In brief, RNA was isolated using TRI Reagent (Ambion) according to the manufacturer's protocol. For

the operon structure, first strand synthesis using Maxima reverse transcriptase (Thermo Scientific, St. Leon-Rot, Germany) was performed using gene specific primers on the 5'-end of each gene, indicated with RT-R (Suppl. Table 2). The presence of intergenic cDNA was subsequently determined by PCR using the same gene specific primer in combination with a 3'-end located primer of its upstream located gene (Suppl. Table 2). As negative control, the same procedure for cDNA synthesis was followed but without adding reverse transcriptase. As positive control for the PCR, purified genomic E1504 DNA was included.

To determine growth condition-dependent expression of *lacI, iolD, iolA* and the *sss*-like gene, cDNA was synthesized from RNA using Maxima First strand cDNA synthesis kit for RT-qPCR (Thermo Scientific, St. Leon-Rot, Germany). In addition, quantitative PCRs using primers indicated with "q" (Suppl. Table 2) on the synthesized total cDNAs were performed using Maxima® SYBR Green/ROX qPCR Master Mix (Thermo Scientific) using a StepOneTM Realtime PCR system (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 the *tufA* gene was used as a reference for the determination of relative expression levels (Suppl. Table 2)(31). This experiment was performed with three biological replicates in a single technical experiment.

Promoter mapping using 5' RACE

E. faecium E1504 was grown in M1-MI to an OD_{600} of 0.3. Total RNA was isolated as previously described. We used the 5' RACE kit (Rapid amplification of cDNA ends, Invitrogen, The Netherlands) to map the promoter of the inositol gene cluster according to the manufacturers' protocol. After first strand synthesis using gene specific primers 1 (GSP1) (Suppl. Table 2), a nested PCR with GSP2 primers was performed to amplify the product and cloned in pGEM-T Easy TA cloning vector (Promega, Madison, WI). Inserts were sequenced to determine the cDNA end.

***In vivo* mouse colonization model**

Specific pathogen-free 10-wk-old female wildtype BALB/c mice were purchased from Charles River

(Maastricht, The Netherlands). The animals were housed in rooms with a controlled temperature and a 12-h light-dark cycle. They were acclimatized for 1 week prior to the experiment and received standard rodent chow (www.sdsdiets.com) and water ad libitum.

Intestinal colonization by wild-type E1504 and E1504Δ*iolD* (inoculum of $1x10^4$ CFU/300 µl Todd Hewitt Broth) was tested as previously described (10).In brief, two days before inoculation of bacteria, mice were administered subcutaneous injections of ceftriaxone (Roche, Woerden, The Netherlands; 100 µl per injection, 12 mg/ml) two times daily and one time at the day of inoculation. For the remaining duration of the experiment, cefoxitin (0.125 g/l) was added to sterile drinking water . Collection of samples and determination of bacterial outgrowth was performed as previously described (10). For statistical analysis the unpaired two-tailed Student's t-test was applied.

Ethics Statement

This study was approved by the Animal Ethics Committee Utrecht and the Animal Welfare Body Utrecht as part of a project, which was licensed by the Central Authority for Scientific Procedures on Animals (CCD) (license number: AVD115002016568).

RESULTS

Identification and distribution of ICE*Efm1* and the *iol* element in 1,644 clade A isolates

ICE*Efm1* was identified in 789/1,227 (64.3%) clade A1 isolates, including 786 hospital-associated, one human non-hospital associated and two pet isolates (Fig. 1A), but in only 3/417 (0.7%) non-clade A1 isolates, all of which originated from hospitalized patients (Fig. 1A). The *iol* element, encoding 5 genes with unknown function and the *iol* gene cluster (Suppl. Fig. 1), was identified in 180/792 (22.7%) of the ICE*Efm1*-containing isolates (Fig. 1A). The integration site of the *iol* element in ICE*Efm1* was determined by a genome graph neighborhood analysis considering the first and last genes of this element (E7356_00303, E7356_00286), respectively (Suppl. Fig. 2 A and B). For the first gene (E7356_00303 in the reference), in 168/180 of the *iol* element carrying isolates, the same neighboring genes were identified up- and downstream (Suppl. Fig 2A). Although in 57 of these

isolates some variation was observed in either: i) presence of additional hypothetical genes not present in E7356 and/or ii) absence of E7356 genes (e.g. E7359_00299). In 8/180 isolates, the first gene of the *iol* element was found on a small contig and therefore only a small number of neighboring genes could be identified (Suppl. Fig 2A). For 4/180 isolates, also due to very fragmented genome assemblies, the analysis to identify the neighboring genes for the first gene of the *iol* element failed. However, for all 180 isolates the gene neighboring analysis revealed the same genes up- and downstream of the last gene of the *iol* element (E7356_00286) (Suppl. Fig 2B). Based on these results, we assume that in all strains the *iol* element is integrated at the same site.

The previously generated phylogenetic tree based on the core gene alignment for 1,644 *E. faecium* isolates (9) was used to determine the distribution of ICE*Efm1* with and without the *iol* element among previously defined sequence clusters (SC)(9) (Fig. 1A and Suppl. Table 1). ICE*Efm1* was identified among 18 SCs and most prevalent in SC1 (291/792, 36.7%) and SC13 (124/792, 15.7%) (Fig. 1A). The *iol* element was variably present in 13 of these 18 SCs, e.g. in ICE*Efm1* containing SC1 isolates the *iol* element was detected in 82/291 (28.1%) isolates, while it was absent in ICE*Efm1* containing SC13 isolates (Fig. 1A). Examination of the phylogenetic tree using microreact revealed several pairs of genetically closely related isolates with and without ICE*Efm1*, e.g. *E. faecium* E1469 (2000, Madrid, Spain) and E1738 (2002, Madrid, Spain) *(patristic distance = 1.6e-5)* (Fig. 1B) and genetically closely related ICE*Efm1* positive isolates with and without the *iol* element, e.g. *E. faecium* E2359 (2004, Singapore) and E6958 (2010, Portugal) *(patristic distance = 1.11e-4)* (Fig. 1C). These findings suggest acquisition or loss of ICE*Efm1* with and without the *iol* element among hospital-associated *E. faecium* isolates.

In silico analysis of the inositol catabolic pathway as compared to _Bacillus subtilis_

In order to infer the potential role of the *E. faecium iol* gene cluster, we compared the organization and presence/absence of *iol* genes with the previously functionally characterized *iol* gene cluster of *B. subtilis*, which encodes a functional catabolic pathway for MI and DCI (32) (Fig. 2 and Suppl. Fig. 3).

This comparison revealed a difference in gene synteny and the absence of four genes in the *E. faecium iol* gene cluster encoding for IolF (minor MI transporter), IolI (2-keto-myo-inositol (2KMI) isomerase), IolJ (6-phospho-5-dehydro-2-deoxy-D-gluconate aldolase) and IolH (protein with unknown function) (Fig. 2). Based on the MI and DCI catabolic pathway as determined for *B. subtilis,* the absence of these genes suggested that the *E. faecium iol* gene cluster might not be functional (Supp. Fig. 3). However, the *E. faecium iol* gene cluster putatively encodes MI and/or DCI transporters represented by the last two genes annotated as sodium/*myo*-inositol cotransporter (SSS) and sugar phosphotransferase system (PTS), respectively, potentially acting as alternative for IolF or IolT (the major transporter of MI in *B. subtilis* encoded outside the *B. subtilis iol* gene cluster) (Fig. 2 and Suppl. Fig. 3). BLASTp search of *B. subtilis* IolT and IolJ against the complete genome of *E. faecium* AUS0085 revealed the presence of (putative) alternative proteins with shared conserved domains with locus-tags EFAU085_02406 (37% identity to IolT) and EFAU085_00643 (45% identity to IolJ), respectively. In contrast, BLASTp revealed no proteins with significant identity or shared domains for IolI.

In *B. subtilis,* expression of the *iol* gene cluster is regulated by a transcriptional repressor IolR encoded upstream of the *iol* gene cluster (Fig. 2) (33). In *E. faecium* AUS0085, a putative transcriptional regulator, annotated as LacI, is encoded directly upstream the *iol* gene cluster (Fig. 2), but with no similarity with IolR of *B. subtilis.* To determine putative domains of LacI, a protein BLAST was performed, which revealed that LacI contains an N-terminal helix-turn-helix motive with similarity to the LacI family of transcriptional repressors (Suppl. Fig. 4). In addition, the protein contains a large domain with similarity to a periplasmic sugar-binding domain, predicted to be involved in the transport of sugar-containing molecules across cellular membrane. These domains suggest that LacI is the regulator of the *iol* gene cluster, but could also be involved in transport of MI and/or other sugars (Suppl. Fig. 4).

Finally, the *E. faecium iol* gene cluster contains one gene which is absent in the *B. subtilis* gene

cluster, i.e. *adh*. BLAST-P analysis revealed that Adh contains domains that belong to the family of iron-containing alcohol dehydrogenase, most of which have not been characterized.

Growth of *E. faecium* on *myo*-inositol and *D-chiro*-inositol

In order to investigate whether the *iol* gene cluster encodes for a functional MI and/or DCI catabolic pathway, we generated growth curves of *E. faecium* E1679 (Fig. 3A) and another ICE*Efm1*- and *iol* gene cluster-containing strain E1504 in minimal medium M1 containing 1% MI (M1-MI) (Fig. 3B). In addition, E1504 was also grown in M1 with 1% DCI (M1-DCI) (Fig. 3B). Both strains were able to grow in M1-MI, indicating that they are able to use *myo*-inositol for their metabolism, although E1679 seemed to grow better as this strain reached a higher final OD. In contrast, no growth was observed in the presence of DCI, which is likely due to the absence of IolI in the *E. faecium iol* gene cluster. As a control, E1504 was grown in M1 in the absence of a carbon source, but as expected no growth was observed comparable to growth in the presence of DCI.

Generation of an ICE*Efm1* transconjugant and a markerless *iolD* mutant

Two different approaches were used to further confirm that the *iol* gene cluster encodes for MI catabolism. For the first approach, we investigated whether the ability of MI catabolism encoded by ICE*Efm1* is transferable. To this aim, we used MI for the selection of transconjugants using *E. faecium* E1679 as donor and *E. faecium* E4658 and the previously used *E. faecium* BM4105RF as recipient strains (Table 1). No transconjugants were obtained using the BM4105RF as recipient. In contrast, when using E4658 as recipient transconjugants were obtained at a frequency of 1×10^{-8} per recipient cell. The transfer of ICE*Efm1* to E4658 was confirmed by PFGE analysis (Fig. 4A) with subsequent Southern blotting using an internal fragment of the *iol* gene cluster as probe (Fig. 4B). This revealed that a single fragment in the transconjugant, which we designated E7855, had increased in size and hybridized with the internal fragment of the *iol* gene cluster as probe. WGS of E4658 and E7855 confirmed integration of ICE*Efm1* in the 3'-end of the E4658 *rpsI* gene. For the phenotypic characterization, we compared the growth capabilities of the donor, recipient and transconjugant in

M1-MI medium (Fig. 3A). No growth was observed for the recipient E4658. However, the transconjugant E7855 was able to grow in M1-MI medium similar to the donor strain E1679, indicating that acquisition of the *iol* gene cluster resulted in the capability to use MI as carbon source for growth (Fig. 3A).

For the second approach, we constructed a markerless *iolD* mutant strain in *E. faecium* strain E1504, designated E1504Δ*iolD* (Table 1). For the characterization of the E1504Δ*iolD* mutant, we compared growth capabilities of E1504 wildtype and E1504Δ*iolD* mutant in M1-MI medium, which revealed that the E1504Δ*iolD* mutant strain was not able grow (Fig. 3B).

From these two approaches we can conclude that acquisition of the *iol* gene cluster results in the ability to utilize MI, while deletion of *iolD* results in loss of this ability.

The *iol* gene cluster is only expressed in the absence of other carbon sources

Next, we determined growth medium dependent RNA expression levels on a selection of genes encoded on the *iol* gene cluster, including its putative transcriptional regulator *lacI* and three genes dispersed over the *iol* gene cluster, i.e. *iolD, iolA* and *sss* (Fig. 2). First, the expression of the genes was determined in *E. faecium* E1504 grown in BHI used as control and in BHI and M1 supplemented with 1% MI, BHI-MI and M1-MI, respectively (Fig. 5). For none of the four genes, difference in expression levels were observed when E1504 was grown in BHI-MI compared to the BHI control. However, compared to BHI, *iolD, iolA* and *sss* were highly expressed in M1-MI, but no difference was observed for *lacI* (Fig. 5). These findings suggest that the *iol* genes, but not *lacI,* are only expressed in the absence of other carbon sources.

Transcriptional organization of the *iol* gene cluster

The transcription start site of the *iol* gene cluster was identified at 32-bp upstream the *iolC* startcodon using 5'- RACE analysis (Fig. 6). A putative promoter region, including -35 and -10 boxes were identified 44-bp and 69-bp upstream of the *iolC* start codon (Fig. 6). To investigate the transcriptional organization of the *iol* gene cluster, we determined whether the *iol* genes are

transcribed as a single RNA molecule. In total, 12 PCRs were performed on synthesized cDNA from strain *E. faecium* E1504 (Suppl. Fig. 5A). The expected sizes for PCR products (indicated in Suppl. Table 2) were obtained between all genes encoded on the *iol* gene cluster, except for PCR-1 amplifying the intergenic region of *lacI* and *iolC* and PCR-12 amplifying the intergenic region between the *pts* gene and its downstream region (Suppl. Fig. 5B). A predicted transcriptional terminator is located downstream the *lacI* and *pts* genes with a ΔG of -53.20 kcal/mol and ΔG of -121.40 kcal/mol, respectively (Suppl. Fig. 5A, C and D). Downstream *lacI,* we identified two pairs of inverted repeats, which are part of the predicted transcriptional terminator, but the inverted repeat identified upstream from the -35 box is the predicted DNA binding site for the LacI repressor (Fig. 6 and Suppl. Fig. 5C). Taken together, these findings indicate that *lacI* and the *iol* gene cluster are not part of the same operon, but that the *iol* genes, *sss* and *pts* genes are indeed organized as a single operon.

The *iol* gene cluster has no role in intestinal colonization in a mouse

In order to investigate whether *iol* gene cluster containing strains would have a selective advantage in their ability to colonize the gut, we orally inoculated two different groups of mice with either wild-type E1504 or its E1504Δ*iolD* mutant strain to compare intestinal colonization rates. Compared to previous studies (10,15), we used a lower inoculum, i.e. 1 x 10^4 cfu in 300 µl instead of 1 x 10^8 cfu in 300 µl. We hypothesized that using a lower inoculum might result in clearer differences in colonization rates between the two groups very early after inoculation. However, the data showed that already at day 1 after inoculation similar amounts of 1 x 10^{10} *E. faecium* cfu/gr feces were identified in both mice groups inoculated with either the E1504 wildtype or the E1504Δ*iolD* mutant (Fig. 7A). These high colonization rates remained until the end of the experiment at day 10, when mice were sacrificed. Similarly, in the colon and cecum, no difference in colonization levels was observed between wild-type and mutant strains. A small difference was observed in the ileum, where a small reduction in colonization of the wild-type strain was observed, but this difference was not statistically significant (Fig. 7B).

DISCUSSION

In this work, we describe the distribution and functional characterization of the *E. faecium iol* gene cluster, which was identified as part of an ~20 kbp genetic element, designated *iol* element, integrated in ICE*Efm1* of *E. faecium* E1679 (18).

The *iol* element was identified among 23% of the ICE*Efm1*-containing *E. faecium* isolates and these ICE*Efm1*-containing *E. faecium* isolates with or without the *iol* element belonged to multiple sequence clusters within clade A1. Furthermore, the *iol* element was always found integrated at the same site in ICE*Efm1* as in *E. faecium* E1679. This suggests that ICE*Efm1* containing the *iol* element was acquired in the early stage of the clade A1 evolution and subsequently lost in some branches during evolution. Examples for such events were found by close examination of branches containing pairs of genetically closely related isolates not only with and without ICE*Efm1,* but also pairs of ICE*Efm1*-containing isolates with and without the *iol* element. However, we cannot exclude that during the evolution of clade A1 both elements were acquired independently. Gene neighborhood analysis revealed variation in gene content in the region between the integrase (first gene of ICE*Efm1*) and the first gene of the *iol* element, indicating that these elements are not 100% identical between isolates. In summary, it remains unclear whether both elements were acquired or lost, but these findings are in line with previous findings that the genomes of *E. faecium* are highly dynamic (34).

For the functional analysis of the *iol* gene cluster, we first performed an *in silico* comparison with the *iol* gene cluster of *B. subtilis,* which has been studied in detail (32,33). This comparison revealed a different gene synteny and the absence of several genes in the *E. faecium iol* gene cluster. Variation in the organization of *iol* gene clusters has also been described for other gram-positive bacteria, e.g. *Enterococcus faecalis* OG1RF (35)*, Lactobacillus casei* BL23 (36) and *Corynebacterium glutamicum* (37) and in gram-negative bacteria, e.g. *Salmonella enterica* serovar Typhimurium*, Klebsiella pneumonia* and *Yersinia enterocolitica* (38). We further showed that despite the lack of specific *iol* genes relative to *B. subtilis*, both *E. faecium* E1679 and E1504 were able to grow in minimal medium

with MI as sole carbon source and that the *iol* gene cluster is responsible for this phenotype. The observation that the tested *E. faecium* strains were not able to grow in minimal medium with DCI can be explained by the absence of the *iolI* gene. The only species that contained an identical gene order compared to *E. faecium* was *E. faecalis* (35). Bourgogne *et al.* did not perform a detailed characterization of the *E. faecalis iol* gene cluster, but only mentioned that transposon insertion mutants in the *iolB* and *iolG2* genes failed to grow on MI (35). Their annotation of the last gene of the *E. faecalis* gene cluster as IolT, which has 92% amino acid identity with the sodium/*myo*-inositol cotransporter (SSS) protein of *E. faecium,* is likely based on the domain predictions from this protein, as there is only low overall similarity with the major MI transporter IolT (15%) and minor MI transporter IolF (18%) of *B. subtilis*. In addition, in *E. faecalis* a similar *lacI* gene encoded upstream the *iol* gene cluster was considered as putative regulator (35). In contrast to *E. faecium,* the *E. faecalis iol* gene cluster is not encoded on a large mobile genetic element like ICE*Efm1* but instead is located on a hot spot for integration, as other types of insertion elements were identified at the same position in other *E. faecalis* isolates (35,39).

The *iol* gene cluster in *E. faecium* is organized as an operon and transcribed as a polycistronic mRNA molecule, comparable to *B. subtilis* (33). Upstream of *iolC*, we identified the transcriptional start site and the putative binding sites for the assumed transcriptional repressor LacI. However, we were not able to confirm this assumption as the generation of a *lacI* mutant strain was unsuccessful and no differential expression was observed in M1-MI compared to BHI. The latter could also be the result of constitutively expression of *lacI.*

Hospital associated *E. faecium* isolates belonging to clade A1 are able to colonize the dysbiotic gut of hospitalized patients at high densities, which contributes significantly to subsequent clinical infections and hospital transmission (40). In order to investigate whether the *iol* gene cluster contributes to high density gut colonization, we compared the colonization capacity of *E. faecium* strain E1504 and its E1504Δ*iolD* mutant in a mouse intestinal colonization model, mimicking colonization of dysbiotic gut. In this model, we did not observe a difference in the capacity to

colonize the gut between wild-type and the *iolD* mutant strain. Only for the small intestines, a small, but not significant decreased colonization was observed for E1504 wild-type strain. Absence of MI in the gut mice as an explanation for this lack of difference in colonization rate between wild-type and *iol* mutant is not likely as inositol is present in the food pellets for mice according to the product sheet of the food supplier (www.sdsdiets.com). Furthermore, increased inositol metabolites have been found in mice treated with clindamycin (41). In our mouse colonization study, mice were treated with cephalosporins, therefore we cannot directly translate the published findings to our study. The most probable explanation for the lack of difference in colonization rates between wild-type and *iolD* mutant is that the *iol* gene cluster is only expressed in the absence of other carbon sources. In our colonization model, mice were fed with rodent chow that contained other carbon sources, hence it is likely that under these conditions the *iol* gene cluster is not expressed. Regulation of expression of the *iol* gene cluster by other carbon sources has also been observed for other species, e.g. *S. enterica* serovar Typhimurium, *C. perfringens* and *L. casei* (36,38,42).

For *Legionella pneumonia* it has been shown that utilization of inositol provides this species a selective advantage for intracellular survival in amoebae and macrophages (43). When inositol was added to *L. pneumophila*-infected amoebae or macrophages, intracellular growth of a parental strain was promoted, but not of the *iolT* or *iolG* mutant. Growth stimulation by inositol was restored by complementation of the mutant strains (43). Macrophage survival has also been experimentally shown for *E. faecium* (31,44). Future research should reveal whether *E. faecium* isolates containing the *iol* gene cluster have a selective advantage and can persist longer in macrophages.

Author statements

Authors and contributors

List of Orcid IDs:

Janetta Top: 0000-0002-4620-8128; Jery Baan: 0000-0002-7913-0873; Sergio Arredondo-Alonso: 0000-0002-4445-2442; Willem van Schaik: 0000-0001-5832-0988; Rob Willems: 0000-0002-7370-0812

Conceptualization, JT and RJLW; validation, JT, JB and AB; formal analysis JT, JB and SA; investigation, JT, JB and AB; writing of original draft, JT, JB, WvS and RJLW; visualization, JT and JB; funding acquisition, WvS and RJLW.

Conflicts of interest

The authors declare no conflicts of interest.

Funding information

SA and RJLW: This study was supported by the Joint Programming Initiative in Antimicrobial Resistance (JPIAMR Third call, STARCS, JPIAMR2016-AC16/00039).

WvS was supported by a VIDI grant (917.13.357) from the Netherlands Organization for Scientific Research (VIDI: 917.13.357) and a Royal Society Wolfson Research Merit Award (WM160092).

References

1. Weiner LM, Webb AK, Limbago B, Dudeck MA, Patel J, Kallen AJ, et al. Antimicrobial-Resistant Pathogens Associated with Healthcare-Associated Infections: Summary of Data Reported to the National Healthcare Safety Network at the Centers for Disease Control and Prevention, 2011-2014. Infect Control Hosp Epidemiol. 2016;37(11):1288–301.
2. Guzman Prieto AM, van Schaik W, Rogers MRC, Coque TM, Baquero F, Corander J, et al. Global emergence and dissemination of enterococci as nosocomial pathogens: Attack of the clones? Vol. 7, Frontiers in Microbiology. 2016.
3. Galloway-Peña 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).
4. Palmer KL, Godfrey P, Griggs A, Kos VN, Zucker J, Desjardins C, 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):1–11.
5. Lebreton F, van Schaik W, McGuire AM, Godfrey P, Griggs A, Mazumdar V, et al. Emergence of epidemic multidrug-resistant *Enterococcus faecium* from animal and commensal strains. MBio. 2013;4(4).
6. Raven KE, Reuter S, Reynolds R, Brodrick HJ, Russell JE, Török ME, et al. A decade of genomic history for healthcare-associated *Enterococcus faecium* in the United Kingdom and Ireland. Genome Res. 2016;26(10):1388–96.
7. Gouliouris T, Raven KE, Ludden C, Blane B, Corander J, Horner CS, et al. Genomic surveillance of *Enterococcus faecium* reveals limited sharing of strains and resistance genes between livestock and humans in the United Kingdom. MBio. 2018;9(6):1–15.
8. Gouliouris T, Raven KE, Moradigaravand D, Ludden C, Coll F, Blane B, et al. Detection of vancomycin-resistant *Enterococcus faecium* hospital-adapted lineages in municipal

wastewater treatment plants indicates widespread distribution and release into the environment. Genome Res. 2019;29(4):626–34.

9. Arredondo-Alonso S, Top J, McNally A, Puranen S, Pesonen M, Pensar J, et al. Plasmids shaped the recent emergence of the major nosocomial pathogen *Enterococcus faecium*. MBio. 2020;11(1):1–17.

10. Top J, Paganelli FL, Zhang X, van Schaik W, Leavis HL, van Luit-Asbroek M, et al. The *Enterococcus faecium* Enterococcal Biofilm Regulator, EbrB, Regulates the *esp* Operon and Is Implicated in Biofilm Formation and Intestinal Colonization. PLoS One. 2013;8(5).

11. Paganelli FL, de Been M, Braat JC, Hoogenboezem T, Vink C, Bayjanov J, et al. Distinct SagA from hospital-associated clade A1 *Enterococcus faecium* strains contributes to biofilm formation. Appl Environ Microbiol. 2015;81(19):6873–82.

12. Paganelli FL, Huebner J, Singh K V., Zhang X, Van Schaik W, Wobser D, et al. Genome-wide Screening Identifies Phosphotransferase System Permease BepA to Be Involved in *Enterococcus faecium* Endocarditis and Biofilm Formation. J Infect Dis. 2016;214:189–95.

13. Hendrickx APA, Van Schaik W, Willems RJL. The cell wall architecture of *Enterococcus faecium*: from resistance to pathogenesis. Future Microbiology. 2013. p. 993–1010.

14. Heikens E, Van Schaik W, Leavis HL, Bonten MJM, Willems RJL. Identification of a novel genomic island specific to hospital-acquired clonal complex 17 *Enterococcus faecium* isolates. Appl Environ Microbiol. 2008;74(22):7094–7.

15. Zhang X, Top J, De Been M, Bierschenk D, Rogers M, Leendertse M, et al. Identification of a genetic determinant in clinical *Enterococcus faecium* strains that contributes to intestinal colonization during antibiotic treatment. J Infect Dis. 2013;207(11):1780–6.

16. Top J, Sinnige JC, Majoor EAM, Bonten MJM, Willems RJL, Van Schaik W. The recombinase IntA is required for excision of *esp*-containing ICE*Efm1* in *Enterococcus faecium*. J Bacteriol. 2011;193(4):1003–6.

17. Leavis H, Top J, Shankar N, Borgen K, Bonten M, Van Embden J, et al. A Novel Putative

Enterococcal Pathogenicity Island Linked to the *esp* Virulence Gene of *Enterococcus faecium* and Associated with Epidemicity. J Bacteriol. 2004;186(3):672–82.

18. van Schaik W, Top J, Riley DR, Boekhorst J, Vrijenhoek JEP, Schapendonk CME, 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).

19. Turner BL, Papházy MJ, Haygarth PM, McKelvie ID. Inositol phosphates in the environment. Philos Trans R Soc B Biol Sci. 2002;357(1420):449–69.

20. Leenhouts K, Buist G, Bolhuis A, Ten Berge A, Kiel J, Mierau I, et al. A general system for generating unlabelled gene replacements in bacterial chromosomes. Mol Gen Genet. 1996;253:217–24.

21. Zhang X, Vrijenhoek JEP, Bonten MJM, Willems RJL, Van Schaik W. A genetic element present on megaplasmids allows *Enterococcus faecium* to use raffinose as carbon source. Environ Microbiol. 2011;13(2):518–28.

22. Paradis E, Claude J, Strimmer K. APE: Analyses of phylogenetics and evolution in R language. Bioinformatics. 2004;20(2):289–90.

23. Fourment M, Gibbs MJ. PATRISTIC: A program for calculating patristic distances and graphically comparing the components of genetic change. BMC Evol Biol. 2006;6:1–5.

24. Tonkin-Hill G, MacAlasdair N, Ruis C, Weimann A, Horesh G, Lees JA, et al. Producing polished prokaryotic pangenomes with the Panaroo pipeline. Genome Biol. 2020;21(180).

25. Willems RJL, Top J, Van Den Braak N, Van Belkum A, Mevius DJ, Hendriks G, et al. Molecular diversity and evolutionary relationships of Tn*1546*-like elements in enterococci from humans and animals. Antimicrob Agents Chemother. 1999;43(3):483–91.

26. Murray BE, Singh K V., 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):2059–63.

27. Heikens E, Bonten MJM, Willems RJL. Enterococcal surface protein *esp* is important for biofilm formation of *Enterococcus faecium* E1162. J Bacteriol. 2007;189(22):8233–40.
28. Ahmed MO, Elramalli AK, Baptiste KE, Daw MA, Zorgani A, Brouwer E, et al. Whole Genome Sequence Analysis of the First Vancomycin-Resistant *Enterococcus faecium* Isolates from a Libyan Hospital in Tripoli. Microb Drug Resist. 2020;26(11):1390–8.
29. Seemann T. Prokka: Rapid prokaryotic genome annotation. Bioinformatics. 2014;30(14):2068–9.
30. Zhang X, Paganelli FL, Bierschenk D, Kuipers A, Bonten MJM, Willems RJL, et al. Genome-wide identification of ampicillin resistance determinants in *Enterococcus faecium*. PLoS Genet. 2012;8(6).
31. Lebreton F, van Schaik W, Sanguinetti M, Posteraro B, Torelli R, Le Bras F, et al. AsrR Is an Oxidative Stress Sensing Regulator Modulating *Enterococcus faecium* Opportunistic Traits, Antimicrobial Resistance, and Pathogenicity. PLoS Pathog. 2012;8(8).
32. Yoshida KI, Yamaguchi M, Morinaga T, Kinehara M, Ikeuchi M, Ashida H, et al. *myo*-inositol catabolism in *Bacillus subtilis*. J Biol Chem. 2008;283(16):10415–24.
33. Yoshida KI, Aoyama D, Ishio I, Shibayama T, Fujita Y. Organization and transcription of the *myo*-inositol operon, *iol*, of *Bacillus subtilis*. J Bacteriol. 1997;179(14):4591–8.
34. Bayjanov JR, Baan J, Rogers MRC, Troelstra A, Willems RJL, van Schaik W. *Enterococcus faecium* genome dynamics during long-term asymptomatic patient gut colonization. Microb Genomics. 2019;5(7).
35. Bourgogne A, Garsin DA, Qin X, Singh K V., Sillanpaa J, Yerrapragada S, et al. Large scale variation in *Enterococcus faecalis* illustrated by the genome analysis of strain OG1RF. Genome Biol. 2008;
36. Yebra MJ, Zúñiga M, Beaufils S, Pérez-Martínez G, Deutscher J, Monedero V. Identification of a gene cluster enabling *Lactobacillus casei* BL23 to utilize *myo*-inositol. Appl Environ Microbiol. 2007;73(12):3850–8.

547 37. Krings E, Krumbach K, Bathe B, Kelle R, Wendisch VF, Sahm H, et al. Characterization of *myo*-
548 inositol utilization by *Corynebacterium glutamicum*: The stimulon, identification of
549 transporters, and influence on L-lysine formation. J Bacteriol. 2006;188(23):8054–61.
550 38. Kröger C, Fuchs TM. Characterization of the *myo*-inositol utilization island of *Salmonella*
551 *enterica* serovar Typhimurium. J Bacteriol. 2009;191(2):545–54.
552 39. Pöntinen AK, Top J, Arredondo-Alonso S, Tonkin-Hill G, Freitas AR, Novais C, et al. Apparent
553 nosocomial adaptation of *Enterococcus faecalis* predates the modern hospital era. Nat
554 Commun [Internet]. 2021;12(1). Available from: http://dx.doi.org/10.1038/s41467-021-
555 21749-5
556 40. Ruiz-Garbajosa P, De Regt M, Bonten M, Baquero F, Coque TM, Cantón R, et al. High-density
557 fecal *Enterococcus faecium* colonization in hospitalized patients is associated with the
558 presence of the polyclonal subcluster CC17. Eur J Clin Microbiol Infect Dis. 2012;31(4):519–22.
559 41. Jump RLP, Polinkovsky A, Hurless K, Sitzlar B, Eckart K, Tomas M, et al. Metabolomics analysis
560 identifies intestinal microbiota-derived biomarkers of colonization resistance in clindamycin-
561 treated mice. PLoS One. 2014;9(7).
562 42. Kawsar HI, Ohtani K, Okumura K, Hayashi H, Shimizu T. Organization and transcriptional
563 regulation of myo-inositol operon in *Clostridium perfringens*. FEMS Microbiol Lett.
564 2004;235(2):289–95.
565 43. Manske C, Schell U, Hilbi H. Metabolism of *myo*-inositol by *Legionella pneumophila* promotes
566 infection of amoebae and macrophages. Appl Environ Microbiol. 2016;82(16):5000–14.
567 44. Gröbner S, Fritz E, Schoch F, Schaller M, Berger AC, Bitzer M, et al. Lysozyme activates
568 *Enterococcus faecium* to induce necrotic cell death in macrophages. Cell Mol Life Sci.
569 2010;67(19):3331–44.
570
571

Table 1. Bacterial strains and plasmids

Strain or plasmid	Relevant characteristics [a]	Reference or source
Strains		
E. faecium		
E1679	Clinical blood isolate; Brazil; AmpR, VanR, GenR, SpcS; ICE*Efm1*+	18
E1504	Clinical blood isolate; Spain; AmpR, VanS, GenS, SpcS; ICE*Efm1*+	9
E1504Δ*iolD*	Markerless deletion mutant of *iolD* of E1504; GenS; ICE*Efm1+*	This study
E4658	Pig isolate; Netherlands; RifR, FusR, VanR; ICE*Efm1*-	25
E7855	Transconjugant; E1679 ICE*Efm1* donor, O12 acceptor	This study
BM4105RF	Recipient strain; France; RifR, FusR; ICE*Efm1*-	18
E. coli		
DH5α	*E. coli* host strain for routine cloning	Invitrogen
EC1000	MC1000 *glgB*::*repA*	20
Plasmids		
pWS3	Shuttle plasmid; ts in gram-positive hosts; SpcR	30
pEF39	pWS3: *ebrB* fusion with gentamicin resistance casssette cloned in the EcoRI site of the *ebrB*gene fusion fragment plasmid for generating an *ebrB* marked mutation; SpcR, GenR	This study
pEF40	pWS3: *esp* fusion with gentamicin resistance casssette cloned in the EcoRI site of the *esp*gene fusion fragment plasmid for generating an *esp* marked mutation; SpcR, GenR	This study
pWS3-Cre	pWS3 derivative expressing Cre in *E. faecium*	30

[a]Amp, ampicillin; Van, vancomycin; Chl, chloramphenicol; Gen, gentamicin; Spc, spectinomycin;
ICE*Efm1+*, *E. faecium ebrB* containing pathogenicity island

572

573

LEGENDS TO FIGURES

Fig. 1. RAxML tree based on 955 *E. faecium* core genes in 1,644 clade A isolates (adapted from Arredondo et al, 2020 (9)). A: Nodes are colored to sources. Metadata blocks indicate the previously defined sequence clusters (SC), the presence/absence of ICE*Efm1* and presence/absence of the *iol* element. The arrow indicates the previous defined split of clade A1. B and C indicate the position of the subtrees from panel B and C. B: Subtree of genetically closely related isolates where ICE*Efm1* is present/absent. Nodes are colored to presence/absence of ICE*Efm1*. ESP: Spain; NLD: the Netherlands; PRT: Portugal. C: Subtree of genetically closely related isolates where the *iol* element is present/absent. Nodes are colored to presence/absence of ICE*Efm1*. Absence is indicated with green and presence is indicated with yellow. NLD: the Netherlands; PRT: Portugal; SGP: Singapore.

Fig. 2. Structural organization of the *iol* gene clusters of *Bacillus subtilis* and strain *E. faecium* E1504. Blue arrows, *iol* genes that are present in both gene clusters; red arrows, *B. subtilis iol* genes that are absent in *E. faecium;* orange arrow, the *E. faecium* putative transcriptional repressor; light purple arrow, a transposase and green arrows, two additional genes absent in *B. subtilis*. The protein functions encoded by the *iol* genes are provided in the table. The *B. subtilis iol* gene cluster has been described by Yoshida *et al.* (32).

Fig. 3. Growth curves of *E. faecium* to determine the ability to grow on inositol. A: Overnight cultures of E1679 (donor), E4658 (recipient), E7855 (transconjugant) were inoculated at an initial cell density of OD_{600} 0.0025 in M1 medium with 1% *myo*-inositol (M1-MI). B: Overnight cultures of wild-type E1504 and E1504Δ*iolD* were inoculated at an initial cell density of OD_{600} 0.0025 in M1 medium with 1% *myo*-inositol (M1-MI) and for wild-type E1504 also in M1 medium in absence (M1) and presence of1% *D-chiro*-inositol (M1-DCI). Growth was measured every 15 minutes for 18 hours.

Fig. 4. Transfer of ICE*Efm1* from *E. faecium* E1679 to *E. faecium* E4658. A: representative SYBR safe stained PFGE gel of SmaI-digested chromosomal DNA of the donor strain (E1679; lane 1), the recipient strain (E4658; lane 2) and the transconjugant (E7855; lane 3). The gel band that has shifted in the recipient strain due to the insertion of ICE*Efm1* is indicated by the white arrow. B: Southern blot of the PFGE gel hybridized using an internal fragment of the *iol* gene cluster as probe.

Fig. 5. qRT-PCR analysis of *lacI, iolD, iolA* and *sss* expression ratios. *E. faecium* E1504 at OD_{660} 0.3 grown in BHI, BHI-MI and M1-MI. The data from the qRT-PCR were normalized using *tufA* as an internal standard. The differences in gene expression (log2- transformed data) relative to growth in BHI are shown (n=1).

Fig. 6. *iolC* promoter mapping. In red stop codon *lacI* and start codon *iolC* and transcription start (+1). Putative -35 and -10 sequences are underlined, in blue putative transcriptional terminator of *lacI*, in green putative binding site for *lacI* repressor.

Fig. 7. Intestinal colonization. Mice were orally inoculated with E1504 (blue dots) and E1504Δ*iolD* mutant (red dots). A: Numbers of *E. faecium* E1504 and E1504Δi*olD* were determined in fecal pellets of mice at day 1, 3 and 10 after inoculation. B: After 10 days of colonization numbers of E1504 and E1504Δ*iolD* were determined in the colon, cecum and ileum.

Suppl. Fig. 1: Organization of the *iol* element encoded on ICE*Efm1* in strain *E. faecium* E1679 (adapted from van Schaik et al. (18)). The *iol* element is encoding for five genes of unknown function (grey arrows) and the *iol* gene cluster, including the *iol* genes indicated in blue, a putative transcriptional repressor (red), a transposase (yellow) and two other genes (green).

Suppl. Fig. 2: Panaroo gene neighborhood analysis to determine the integration site of the *iol* element using strain *E. faecium* E7356 as reference. A: Neighboring genes for E7356_0303 (light blue), the first gene of the *iol* element as determined for 176/180 isolates. B: Neighboring genes for E7356_0286 (red), the last gene of the *iol* element. Green box: presence of the gene; grey box: absence of gene; blue box: presence of hypothetical gene different from E7356; orange box: presence of a transposase not present in E7356; yellow: gene was not identified as neighboring gene due to contig breaks in these isolates.

Suppl. Fig 3: Genetic organization, MI catabolic pathway and functional activities of the *B. subtilis iol* genes (adapted from Yoshida et al., 2008 (32)). A: *B. subtilis iol* genes proven to encode the enzymes involved in the various reaction steps of the MI catabolic pathway are shown. Blue boxes indicate *iol* genes also identified in the *E. faecium iol* gene cluster. B: Overview protein functions *iol* genes. C: Overview of the compounds that are converted by the Iol enzymes

Suppl. Fig. 4: Graphic summary of LacI protein BLAST result.

Suppl. Fig. 5. Structural organization of the *iol* gene cluster in strain *E. faecium* E1504. A: Genomic organization of the *iol* gene cluster, including the *iol* genes indicated in blue, a putative transcriptional repressor (red), a transposase (yellow) and two other genes (green). Also indicated are the predicted transcription terminators (RNAfold), PCR and primer sites for the determination of transcriptional organization. Blue arrows indicate the primers that were used to synthesize cDNA and that were used in combination with the upstream located primer in red for the amplification of the intergenic region. B: PCR results for the 12 PCRs as indicated in panel A. As negative control PCR-2 was also performed on RT- cDNA. As positive control PCR-2 was performed on E1504 genomic DNA. C: downstream *lacI* predicted terminator (RNAfold). D: downstream *pts* predicted terminator (RNAfold).

651