

# High-Resolution Melting Genotyping of *Enterococcus faecium* Based on Multilocus Sequence Typing Derived Single Nucleotide Polymorphisms

Steven Y. C. Tong<sup>1\*</sup>, Shirley Xie<sup>2</sup>, Leisha J. Richardson<sup>1</sup>, Susan A. Ballard<sup>2</sup>, Farshid Dakh<sup>1</sup>, Elizabeth A. Grabsch<sup>2</sup>, M. Lindsay Grayson<sup>2</sup>, Benjamin P. Howden<sup>2</sup>, Paul D. R. Johnson<sup>2</sup>, Philip M. Giffard<sup>1</sup>

**1** Menzies School of Health Research, Darwin, Northern Territory, Australia, **2** Department of Infectious Diseases and Microbiology, Austin Health, Heidelberg, Victoria, Australia

## Abstract

We have developed a single nucleotide polymorphism (SNP) nucleated high-resolution melting (HRM) technique to genotype *Enterococcus faecium*. Eight SNPs were derived from the *E. faecium* multilocus sequence typing (MLST) database and amplified fragments containing these SNPs were interrogated by HRM. We tested the HRM genotyping scheme on 85 *E. faecium* bloodstream isolates and compared the results with MLST, pulsed-field gel electrophoresis (PFGE) and an allele specific real-time PCR (AS kinetic PCR) SNP typing method. *In silico* analysis based on predicted HRM curves according to the G+C content of each fragment for all 567 sequence types (STs) in the MLST database together with empiric data from the 85 isolates demonstrated that HRM analysis resolves *E. faecium* into 231 “melting types” (MelTs) and provides a Simpson’s Index of Diversity (*D*) of 0.991 with respect to MLST. This is a significant improvement on the AS kinetic PCR SNP typing scheme that resolves 61 SNP types with *D* of 0.95. The MelTs were concordant with the known ST of the isolates. For the 85 isolates, there were 13 PFGE patterns, 17 STs, 14 MelTs and eight SNP types. There was excellent concordance between PFGE, MLST and MelTs with Adjusted Rand Indices of PFGE to MelT 0.936 and ST to MelT 0.973. In conclusion, this HRM based method appears rapid and reproducible. The results are concordant with MLST and the MLST based population structure.

**Citation:** Tong SYC, Xie S, Richardson LJ, Ballard SA, Dakh F, et al. (2011) High-Resolution Melting Genotyping of *Enterococcus faecium* Based on Multilocus Sequence Typing Derived Single Nucleotide Polymorphisms. PLoS ONE 6(12): e29189. doi:10.1371/journal.pone.0029189

**Editor:** Tarek Msadek, Institut Pasteur, France

**Received:** July 28, 2011; **Accepted:** November 22, 2011; **Published:** December 16, 2011

**Copyright:** © 2011 Tong et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This work was supported by an Australian National Health and Medical Research Council Postdoctoral Training Fellowship (508829) to ST and by the Austin Hospital Department of Pathology to SX. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have read the journal’s policy and have the following conflicts: PG is an inventor on a patent application that describes the software “Minimum SNPs” and is eligible for royalty payments if this is commercialized. The patent is “Assessing Data Sets. PCT/AU01/00527”. This does not alter the authors’ adherence to all the PLoS ONE policies on sharing data and materials. The other authors have declared that no conflicts of interest exist.

\* E-mail: Steven.tong@menzies.edu.au

## Introduction

*Enterococcus faecium* emerged as a hospital-associated pathogen in the 1990s in the USA, Europe and Australia. The acquisition of resistance to ampicillin and vancomycin and spread of successful vancomycin-resistant clones has resulted in an increasing burden of vancomycin-resistant *E. faecium* (VREfm) as a difficult to treat nosocomial infection [1].

Pulsed-field gel electrophoresis (PFGE) has been considered the reference method for tracking outbreaks due to its high discriminatory power [2]. However PFGE remains difficult to standardize between laboratories, and the high genomic plasticity of *E. faecium* means that using PFGE data to track strains over long time periods and to infer the large scale population structure can be problematic [1]. In contrast, multilocus sequence typing (MLST) provides portable genotyping data that unambiguously reveal the population structure and long-term patterns of dissemination [3]. However, MLST is laborious and expensive. Furthermore, whole genome sequencing costs have significantly decreased and may in the future make it increasingly difficult to justify MLST.

However, in the meantime, a rapid, simple and low cost *E. faecium* genotyping method that provides results concordant with

MLST would be useful for *E. faecium* monitoring and the initial investigation of suspected VREfm outbreaks. In this study a typing method was developed based upon the identification and high-resolution melting (HRM) interrogation of DNA fragments containing resolution-optimized sets of SNPs derived from the *E. faecium* MLST database. The use of such SNP sets in allele specific real-time PCR (AS kinetic PCR) based bacterial genotyping has been previously described [4–6]. The adaptation of this approach to HRM technology is a recently reported innovation [7,8]. HRM analysis of PCR fragments incorporates information from not only the key SNP of interest but also neighboring SNPs. This additional information can result in a surprisingly improved discriminatory power [7]. We compared and validated the HRM method against PFGE, MLST and an AS kinetic PCR method [6] using a set of 85 well characterized *E. faecium* bloodstream isolates.

## Methods

### *Enterococcus faecium* isolates

A collection of 85 *E. faecium* blood culture isolates previously described and typed by MLST were included in this study [9]. The collection comprised 34 VREfm and 51 vancomycin-susceptible

(VSEfm) isolates, encompassed 17 STs and reflected the epidemiology of *E. faecium* over a 12 year period at a single Australian institution. The isolates were collected as part of an infection control quality improvement project and formal approval by the local ethics committee was not required.

**DNA extraction**

DNA was prepared using the DNeasy Tissue Kit (Qiagen Pty Ltd., Victoria, Australia) and the modified protocol for isolation of genomic DNA from Gram-positive bacteria as described in the manufacturer’s instructions.

**Pulsed-field gel electrophoresis**

PFGE of *SmaI*-digested genomic DNA agarose plugs was performed as previously described [10]. Gels were analyzed with GelCompar II version 3.5 (Applied Maths NV, Sint-Martens-Latem, Belgium) using Dice Coefficient and represented by unweighted pair group method with arithmetic mean (UPGMA) with 0.75% optimization and 1% tolerance. Isolates with similarity of 80% or more were considered a clonal cluster.

**Selection of SNPs**

Resolution optimized SNPs from the concatenated *E. faecium* MLST database at <http://efaecium.mlst.net/> were derived using the program ‘Minimum SNPs’ [11]. SNPs that did not change the %G+C content (C↔G, A↔T) were excluded as these are not usually detected by HRM. Primers were designed using Primer Express 2.0.0 (Applied Biosystems) and obtained from Prologo (Mulgrave, Victoria, Australia). Sequences of primers are given in **Table 1**.

**HRM-based typing**

Reactions were performed on a RotorGene 6000 device (Corbett, Mortlake, NSW, Australia). Reactions volumes comprised 5 µL 2 × Platinum® SYBR® Green qPCR SuperMix-UDG (Invitrogen Life Technologies, Mt Waverley, Australia), 1 µM of each primer and 2 µL of DNA (1:10 dilution) in a final volume of 10 µL. The temperature cycling parameters were: 95°C for 2 min; 30 cycles of 95°C for 5 s, 55°C for 20 s; followed by 72°C for 2 min and 50°C for 30 secs. Amplification was followed by HRM (72–88°C) increasing by 0.1°C at each step. Runs were performed in laboratories in two separate institutions by multiple operators and results were completely concordant and reproducible. Normalization temperatures for analysis of each fragment were standardized (**Table 1**). The results for the real-time PCR experiments were analysed using the Rotorgene 6000 software v 1.7 (Corbett) and data was exported to Microsoft Excel 2007 and GraphPad Prism v5.01 (GraphPad Software, Inc, California, USA). Three known isolates were included in each run for calibration purposes. To determine ‘across run comparisons’ the inter-run difference in the melting temperature (T<sub>m</sub>) for these three isolates was determined, and the mean T<sub>m</sub> difference to the closest 0.1°C was used to calibrate the other samples. Use of the three known control isolates in each run also facilitated the accurate assignment of HRM alleles (see below).

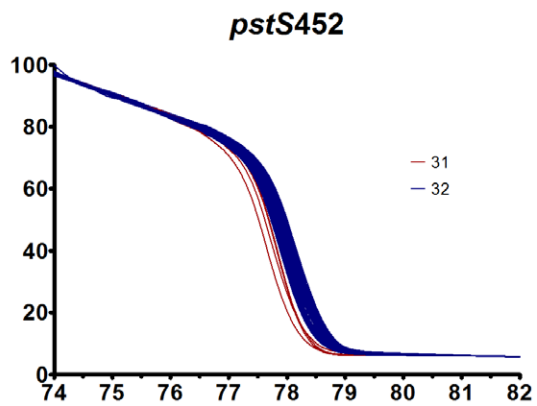
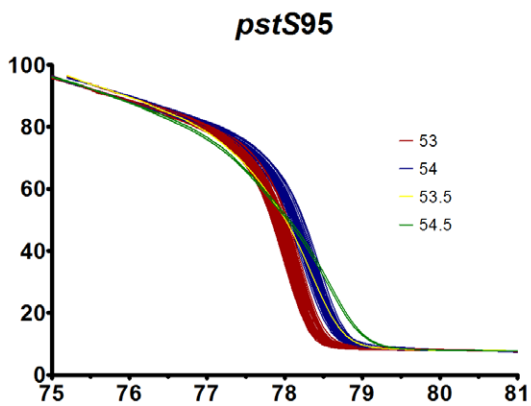
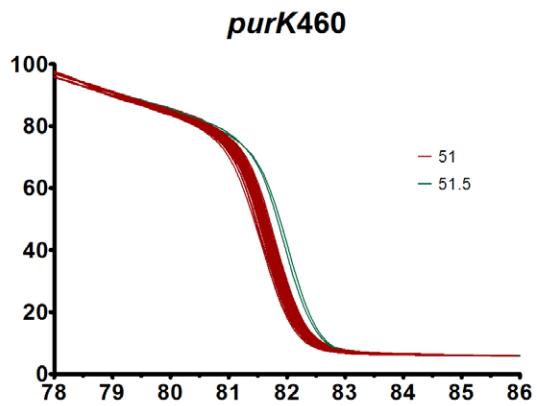
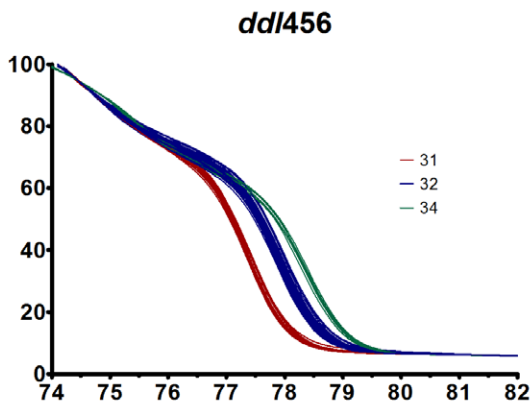
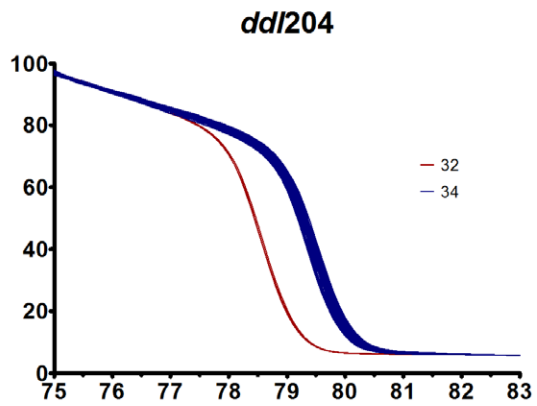
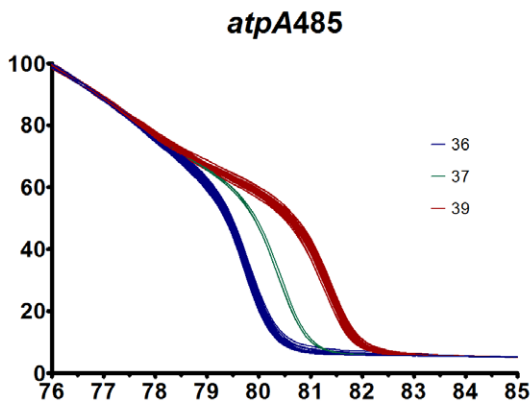
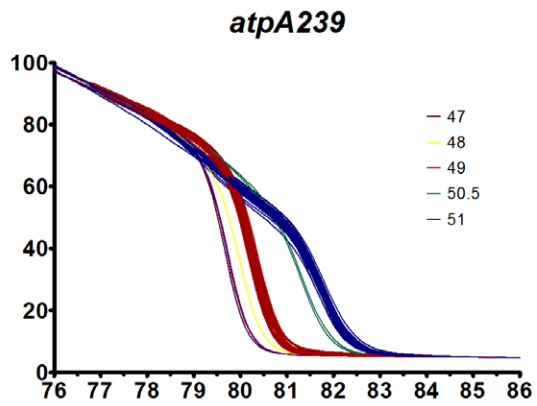
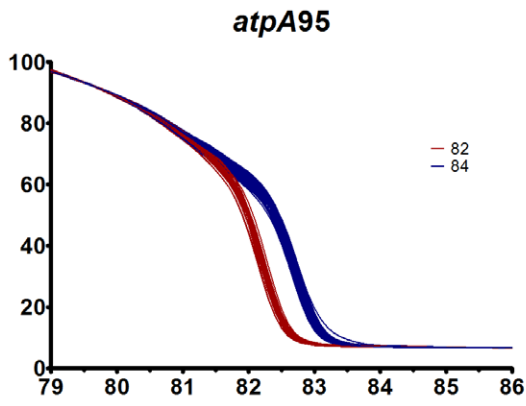
**HRMType® software**

We used a previously described Stata/IC 10.1 (StatCorp, College Station, Texas, USA) “do\_file” named HRMType® to assemble a predicted key for translating between MLST and multilocus HRM data [7]. This procedure is based on the premise

**Table 1.** The eight fragments: Primers, fragment sizes and normalization regions.

Region name	Primer sequence (5’-3’)	Fragment Size (bp)	Normalization regions (°C)
atpA95	F CCTTGGCGATTTCGAGTCC	180	79–80 85–86
	R CCTGTTTGCATTGGTTCGTTAAC		
atpA239	F GTTAACGAACCAATGCAACAGG	108	76–77 85–86
	R ATACAGATCATATCTTGACCTTTTTGG		
atpA485	F TCTACAGTTCGTACACAAGTTGAAAC	87	76–77 84–85
	R CACCCATTGCAGTACCAGCA		
ddl204	F GAAGGTTCTTTGCTTTATCCGATG	79	75–76 82–83
	R CGAGAATCATACTGATAGGCTGTTG		
ddl456	F GTGAAGTCGTAAGAGACGTAGC	97	74–75 81–82
	R TAACATCGTGAAGCTAACTTCGC		
purK460	F GGTGGCAGGAAATGGTCAAG	107	78–79 85–86
	R GCAATCACACGGGAATACG		
pstS95	F TGTATTTGCAGAAGAAAGAGACG	141	75–76 80–81
	R TGATCTTTCCCGCAACTT		
pstS452	F CAAGCACTAAGTATCGACGGTGTAG	79	74–75 81–82
	R TCTAAGAATTTTTGGACTTCAGGG		

doi:10.1371/journal.pone.0029189.t001



**Figure 1. High-resolution melting curves for the eight fragments.** The normalized fluorescence (y-axis) is plotted against temperature in °C (x-axis) and shows the curves for 85 *E. faecium* isolates. Each graph is labelled with the amplified SNP and combines two HRM runs that have been calibrated against each other; different curves are represented by a different color and accordingly labelled. doi:10.1371/journal.pone.0029189.g001

that the  $T_m$  of each amplified fragment is in large part dependent on the absolute G+C content. HRMType<sup>®</sup> converts each MLST database-specified variant of each HRM fragment into a predicted HRM allele on the basis of G+C content. For example, for the *atpA95* fragment, in the current *E. faecium* MLST database, the numbers of G+Cs are 79, 80, 82, 83, 84, 85. HRMType<sup>®</sup> then converts each known ST into a predicted multilocus HRM type that is termed a “Melting Type” (MeIT), and displays each ST with a corresponding MeIT. Our convention is that the HRM alleles are numbered according to the absolute number of G+C residues and the MeITs are numbered sequentially commencing with ST1 corresponding to MeIT 1. If during experimental work a fragment (e.g., *atpA239* for ST178) does not produce a curve as predicted according to G+C content, the unpredicted curve can be assigned a new HRM allele number. For this fragment (e.g., *atpA239*), this new HRM allele number is assigned for all other STs that have the identical sequence in the fragment (e.g., there are 20 other STs with the same *atpA239* fragment sequence as ST178). HRMType<sup>®</sup> also calculates the cumulative Simpson’s Index of Diversity (*D*) [12] in comparison to MLST as each fragment is added to the set.

**In silico SNP genotyping**

We adapted HRMType<sup>®</sup> to allow the input of the specific eight SNPs used by Rathnayake et al. [6] and produce a translation key between STs and the concatenated sequence of SNPs, which we have termed ‘SNP types’. For each of the 85 isolates, SNP types were determined by *in silico* analysis of the MLST sequence.

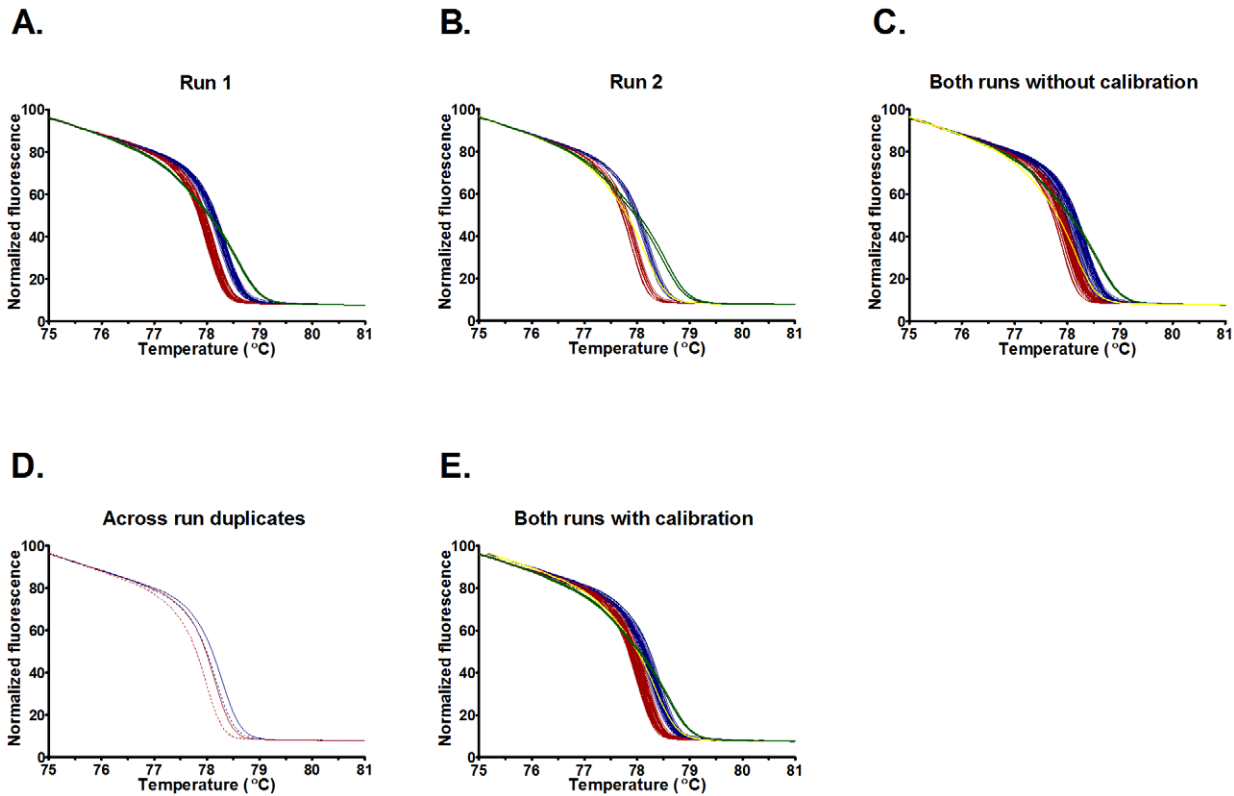
**Determination of the concordance between HRM typing and the MLST defined *E. faecium* population structure**

The Adjusted Rand Indices (ARI) were calculated using Comparing Partitions [13]. To generate the goeBURST diagram we used the algorithms and software as described by Francisco et al. [14] and found at <http://goeburst.phyloviz.net/>. This allows for the overlaying of the MeITs onto a snapshot of the MLST derived population structure of *E. faecium*.

**Results**

**Identification of a resolution optimized set of SNPs**

A set of eight SNPs within the MLST loci was identified on the basis of maximized *D*. The SNPs were (locus name followed by



**Figure 2. High-resolution melting curves for *pstS95*.** The graphs represent (A and B) HRM curves for two separate runs demonstrating discrimination of curves 53 (red) and 54 (blue) when each run is considered separately; (C) raw data from both runs combined without calibration, demonstrating an inability to discriminate curves 53 and 54; (D) raw data from two isolates included in both runs (solid red and dashed red lines for curve 53 isolate and solid blue and dashed blue for curve 54 isolate) demonstrating a shift in the curves when runs are compared without calibration; (E) data from both runs combined with calibration demonstrating discrimination of curves 53 and 54. doi:10.1371/journal.pone.0029189.g002

nucleotide position within that locus): *atpA* 95, *atpA* 239, *atpA* 485, *ddl* 204, *ddl* 456, *purK* 460, *pstS* 95 and *pstS* 452. Amplified fragments incorporating these SNPs ranged in size from 79 to 180 bp (Table 1). *In silico* analysis based on predicted  $T_m$  differences suggested that these eight fragments would partition the 567 sequence types (STs) currently in the MLST database into 230 Melting Types (MeTs) with a  $D$  of 0.991, where the MLST database is regarded as providing  $D=1$ . This indicates a probability of 0.991 that any two randomly selected STs will be discriminated according to their MeT.

Notably, the resolution optimized SNPs were identified in only four of the seven MLST loci. Examination of the MLST loci determined that these are the four most polymorphic loci with the largest number of alleles and individually they also provide the most discriminatory power compared to MLST. Specifically, the number of alleles ( $n$ ) and  $D$  for the seven loci are (in order of discriminatory power): *atpA* ( $n = 58$ ,  $D = 0.911$ ), *ddl* ( $n = 45$ ,  $D = 0.843$ ), *purK* ( $n = 49$ ,  $D = 0.841$ ), *pstS* ( $n = 66$ ,  $D = 0.724$ ), *gdh* ( $n = 42$ ,  $D = 0.643$ ), *adk* ( $n = 25$ ,  $D = 0.557$ ), *gyd* ( $n = 27$ ,  $D = 0.498$ ).

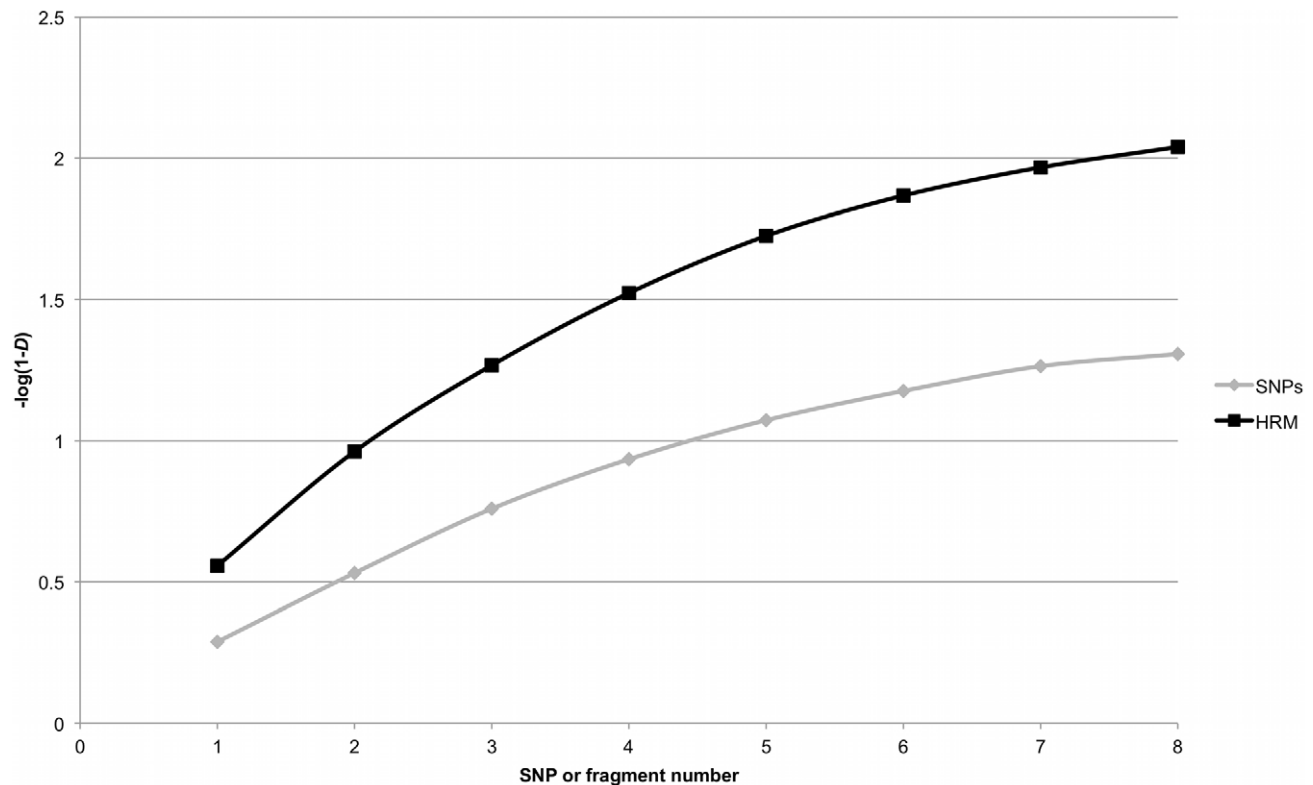
### High-resolution melting curve analysis

We performed HRM genotyping for these eight fragments on 85 *E. faecium* isolates [9]. The HRM curves from each of the eight fragments for all 85 isolates are displayed in Figure 1. The displayed curves have been calibrated for inter-run variability, however we found that, with the exception of fragment *pstS95*, the raw data from the curves could have been exported without calibration and the curves correctly assigned. For fragment *pstS95* (Figure 2), curves 53 and 54 were discriminated in each individual

run, but when raw data from the two runs were combined, these curves could not be discriminated. An inter-run shift of 0.2°C for the calibrating isolates was evident; following calibration, curves 53 and 54 were discriminated in the combined dataset. The HRM curves produced were in accordance with G+C content, with minor exceptions for three isolates. For two ST178 isolates unpredicted curves were seen at *atpA239*, *purK460*, and *pstS95*; and for an ST495 isolate there was an unpredicted curve at *pstS95* and the predicted curve 50 at *purK460* could not be reliably discriminated from curve 51. The key generated by HRMType® was adjusted to take into account the unpredicted curves and converted into an Excel based key to facilitate ongoing interpretation (Results S1). Following this adjustment, there were 231 MeTs for the 567 STs in the MLST database and a  $D$  of 0.991. In comparison, *in silico* SNP genotyping with the previously described eight SNPs provides 61 SNP types and a  $D$  of 0.951. The cumulative resolving power of HRM typing compared to SNP genotyping was found to be clearly superior; in fact, using four HRM fragments provides significantly higher resolution than all eight SNPs (Figure 3).

### Comparison of HRM typing with MLST, PFGE and SNP genotyping

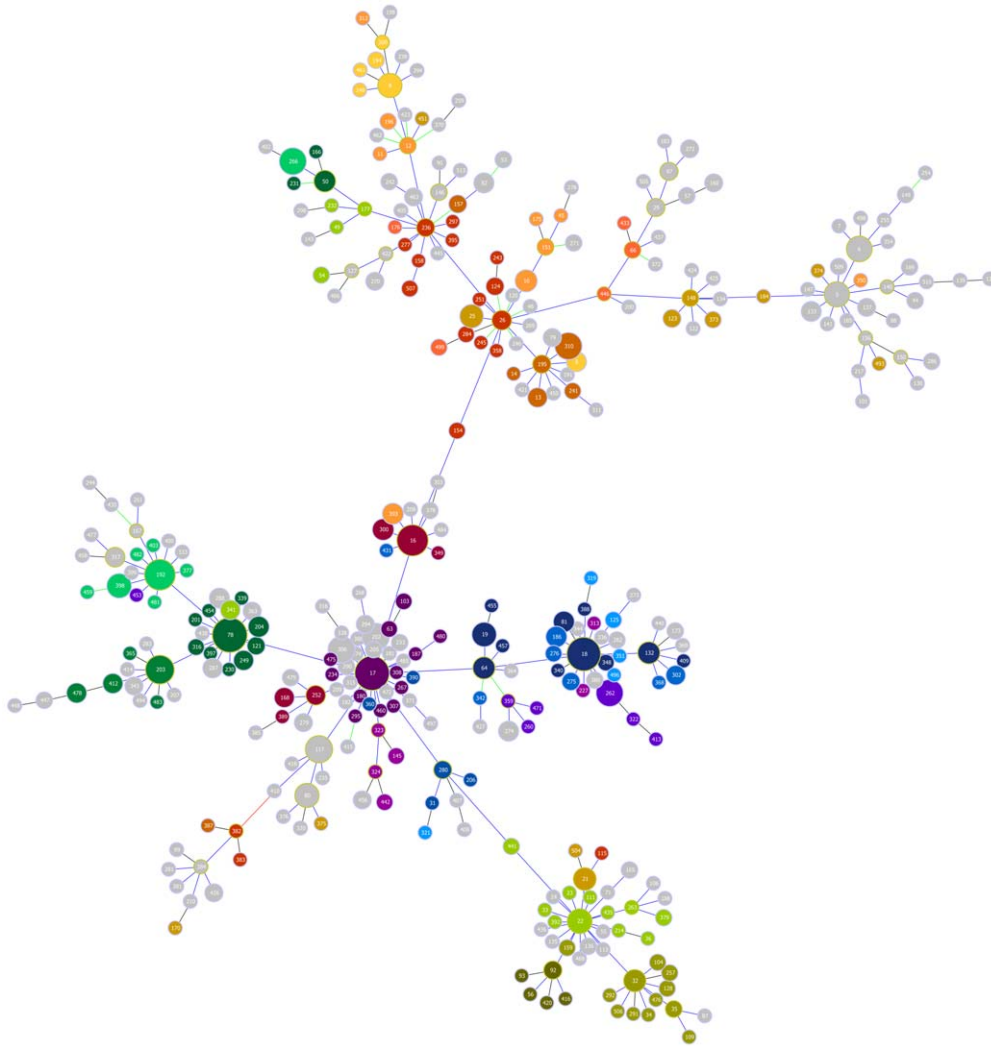
We compared the discriminatory power and concordance of genotyping results between HRM typing, MLST, PFGE and SNP genotyping for the 85 isolates (Figure 4; Results S2). The 85 isolates were partitioned into 13 PFGE patterns, 17 STs, 14 MeTs and 8 SNP types with  $D$  values of 0.77, 0.76, 0.75 and 0.64 respectively. The ARI for PFGE to ST was 0.908, PFGE to MeT 0.936, and ST to MeT 0.973, indicating excellent concordance between these three methods.



**Figure 3. Comparison of HRM and SNP methods.** Cumulative Simpson's Index of Diversity on a log scale for high-resolution melting and single nucleotide polymorphism methods against all 567 s multilocus sequence types. doi:10.1371/journal.pone.0029189.g003







**Figure 5. goeBURST diagram of the *Enterococcus faecium* clonal complex 17 population.** goeBURST diagram of the *Enterococcus faecium* clonal complex 17 population from the multilocus sequence type (ST) database with superimposed melting types (MeITs). The top 26 MeITs (each with  $\geq 5$  associated STs) are represented by different colors. Other STs in grey are represented by smaller MeITs that each have  $< 5$  associated STs. STs are indicated by numbers in the circles, the size of each circle represents the relative abundance of that ST, and lines indicate single locus variants. doi:10.1371/journal.pone.0029189.g005

of the eight fragments being interrogated. Third, as the MLST database is updated with new STs being deposited, the HRMType<sup>®</sup> translation key will need to be updated. For this purpose, frequently updated keys can be found at <http://menzies.edu.au/research/tropical-and-emerging-infectious-disease/bacterial-genotyping>. Fourth, we have validated this method on a collection of local isolates that may not be representative of the global collection of STs, although the collection does contain major clones such as ST17, 18, 22, 32, and 203.

The HRM typing method shows a high level of concordance with the *E. faecium* population structure, as inferred by goeBURST analysis. Instances of non-concordance are very likely due to the limitations of algorithms such as eBURST [15] and goeBURST [14]. These limitations are particularly acute in recombinogenic species such as *E. faecium* [16] and care must be taken in the interpretation of such diagrams. Specifically, a large interlinked complex such as CC17 may simply reflect intensive sampling of part of the population by MLST, and it is possible for quite unrelated STs to be part of such complexes, provided they are linked by several intermediate STs. Conversely, a large interlinked

complex in a recombinogenic species may define so many essentially equivalent topologies that there will be instances of closely related STs being separated by anomalously large numbers of intermediate STs. Consistent with this, we determined that STs that are not resolved by HRM typing, and yet are distant on the goeBURST diagram, are actually much more closely related than the goeBURST diagram indicates. Additionally, the large interlinked CC17 likely incorporates several instances of *E. faecium* clonal radiation of diversification, with the identified founders being ST17, ST18, ST22, ST26, ST78 and ST192 [1]. The HRM method resolves all of these STs. It was therefore concluded that the concordance of population structure and HRM typing is close to as good as it can be, given the finite resolving power of HRM typing, and the considerable inherent difficulties in inferring and depicting the population structures of recombinogenic bacterial species. Given that the HRM method is based on SNPs derived from MLST, any limitations of MLST in depicting the population structure of *E. faecium* will also be present with HRM typing.

With recombinogenic bacterial species, the maximum resolving power that may be obtained from a given number of SNPs is high

because linkage between the allelic states of the SNPs is reduced or obliterated by horizontal transfer of chromosome fragments, thus making it possible for resolving power to increase logarithmically rather than linearly as SNPs are added to a typing set [8]. Even if only the “nucleating” SNPs are taken into account, the SNPs defined many more than eight genotypes, thus indicating that *E. faecium* recombination has greatly increased the resolving power conferred by a given number of SNPs. Interestingly, these SNPs may be clustered in certain loci, and for *E. faecium* the unbiased algorithm used to identify SNPs to maximize resolving power determined these SNPs to be in only four of the seven MLST loci. Not surprisingly, these are the four most polymorphic and discriminatory MLST loci. Use of SNPs from four, rather than seven, MLST loci is further justified in that results from HRM typing are concordant with MLST. Our methods for identifying regions optimized for genotyping and interpreting the data with reference to databases of sequence variation are not restricted to the HRM platform or MLST databases. The methods could be adapted to identify fragments containing resolution optimized SNPs across a genome and to generate a key for other platforms such as with base composition analysis using PCR/electron-spray mass spectrometry [17].

Depending on the monetary, time and capillary sequencing resources available, MLST can be resource intensive for many laboratories. HRM typing utilizes eight PCR reactions with little downstream processing and is thus significantly cheaper than MLST, which requires an initial seven PCR reactions followed by capillary sequencing and analysis. Although less discriminatory than MLST, HRM typing can be used as an initial genotyping

scheme to rapidly screen isolates at low cost while providing results that are concordant with MLST. For example, it can be used to identify globally recognized and endemic strains of *E. faecium* and only where indicated, such as during a local outbreak, would PFGE, MLST or whole genome sequencing be required. Even then HRM is a useful tool to focus and reduce the requirement for these more highly discriminating methods. Although MLST continues to be widely used and whole genome analysis using next generation sequencing technologies is becoming the gold standard genotyping methodology (and will probably replace MLST), we feel that there continues to be an important role for rapid, low cost methodologies that make use of existing instrumentation present in many less resourced laboratories.

## Supporting Information

### Results S1 Melting type and multilocus sequence type conversion key.

(XLS)

### Results S2 Results for HRM typing, MLST, PFGE and SNP genotyping for 85 isolates.

(XLS)

## Author Contributions

Conceived and designed the experiments: ST MG PJ PG. Performed the experiments: ST SX LR SB FD EG. Analyzed the data: ST SB EG BH. Contributed reagents/materials/analysis tools: ST. Wrote the paper: ST PG. Critical review of the manuscript: SB EG BH.

## References

- Willems RJ, van Schaik W (2009) Transition of *Enterococcus faecium* from commensal organism to nosocomial pathogen. *Future Microbiol* 4: 1125–1135.
- Chuang YC, Wang JT, Chen ML, Chen YC (2010) Comparison of an automated repetitive-sequence-based PCR microbial typing system with pulsed-field gel electrophoresis for molecular typing of vancomycin-resistant *Enterococcus faecium*. *J Clin Microbiol* 48: 2897–2901.
- Homan WL, Tribe D, Poznanski S, Li M, Hogg G, et al. (2002) Multilocus sequence typing scheme for *Enterococcus faecium*. *J Clin Microbiol* 40: 1963–1971.
- Huygens F, Inman-Bamber J, Nimmo GR, Munckhof W, Schooneveldt J, et al. (2006) *Staphylococcus aureus* genotyping using novel real-time PCR formats. *J Clin Microbiol* 44: 3712–3719.
- Price EP, Thiruvankataswamy V, Mickan L, Unicomb L, Rios RE, et al. (2006) Genotyping of *Campylobacter jejuni* using seven single-nucleotide polymorphisms in combination with *flaA* short variable region sequencing. *J Med Microbiol* 55: 1061–1070.
- Rathnayake IU, Hargreaves M, Huygens F (2011) Genotyping of *Enterococcus faecalis* and *Enterococcus faecium* Isolates by Use of a Set of Eight Single Nucleotide Polymorphisms. *J Clin Microbiol* 49: 367–372.
- Richardson LJ, Tong SY, Towers RJ, Huygens F, McGregor K, et al. (2011) Preliminary validation of a novel high resolution melt-based typing method based on the multilocus sequence typing scheme of *Streptococcus pyogenes*. *Clin Microbiol Infect* 17: 1426–1434.
- Lilliebridge RA, Tong SYC, Giffard PM, Holt DC (2011) The utility of high-resolution melting analysis of SNP nucleated PCR amplicons - a MLST based *Staphylococcus aureus* typing scheme. *PLoS ONE* 6: e19749. doi:10.1371/journal.pone.0019749.
- Johnson PD, Ballard SA, Grabsch EA, Stinear TP, Seemann T, et al. (2010) A sustained hospital outbreak of vancomycin-resistant *Enterococcus faecium* bacteremia due to emergence of vanB *E. faecium* sequence type 203. *J Infect Dis* 202: 1278–1286.
- Ballard SA, Pertile KK, Lim M, Johnson PD, Grayson ML (2005) Molecular characterization of vanB elements in naturally occurring gut anaerobes. *Antimicrob Agents Chemother* 49: 1688–1694.
- Robertson GA, Thiruvankataswamy V, Shilling H, Price EP, Huygens F, et al. (2004) Identification and interrogation of highly informative single nucleotide polymorphism sets defined by bacterial multilocus sequence typing databases. *J Med Microbiol* 53: 35–45.
- Hunter PR, Gaston MA (1988) Numerical index of the discriminatory ability of typing systems: an application of Simpson's index of diversity. *J Clin Microbiol* 26: 2465–2466.
- Carrico JA, Silva-Costa C, Melo-Cristino J, Pinto FR, de Lencastre H, et al. (2006) Illustration of a common framework for relating multiple typing methods by application to macrolide-resistant *Streptococcus pyogenes*. *J Clin Microbiol* 44: 2524–2532.
- Francisco AP, Bugalho M, Ramirez M, Carrico JA (2009) Global optimal eBURST analysis of multilocus typing data using a graphic matroid approach. *BMC Bioinformatics* 10: 152.
- Feil EJ, Li BC, Aanensen DM, Hanage WP, Spratt BG (2004) eBURST: Inferring Patterns of Evolutionary Descent among Clusters of Related Bacterial Genotypes from Multilocus Sequence Typing Data. *J Bacteriol* 186: 1518–1530.
- Turner KM, Hanage WP, Fraser C, Connor TR, Spratt BG (2007) Assessing the reliability of eBURST using simulated populations with known ancestry. *BMC Microbiol* 7: 30.
- Hall TA, Sampath R, Blyn LB, Ranken R, Ivy C, et al. (2009) Rapid molecular genotyping and clonal complex assignment of *Staphylococcus aureus* isolates by PCR coupled to electrospray ionization-mass spectrometry. *J Clin Microbiol* 47: 1733–1741.