

# Characterizing the plasmidome of *Enterococcus faecium*

Anita Schürch, Sergio Arredondo Alonso, Jukka Corander, Rob Willems

## Background

Plasmids are important vehicles for rapid evolution and adaptation in *Enterococcus faecium*. To date, less than 100 different plasmids from this species have been described, and this collection is highly biased toward plasmids carrying antibiotic resistance determinants such as vancomycin resistance genes.

## Objectives

Capturing the diversity of plasmids in a set of 1644 *E. faecium* isolates by nanopore long read sequencing of a pre-selected subset of 50 *E. faecium* from different sources

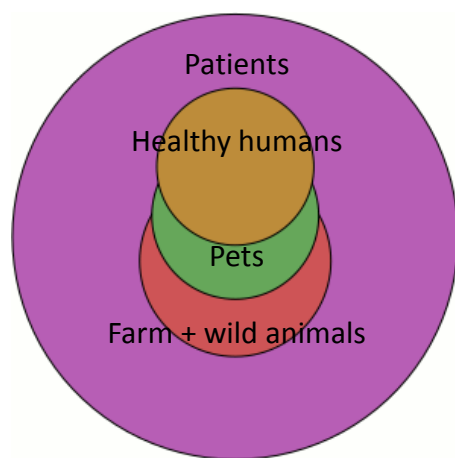


Figure 1 Distribution of sources of 1644 *E. faecium* isolates

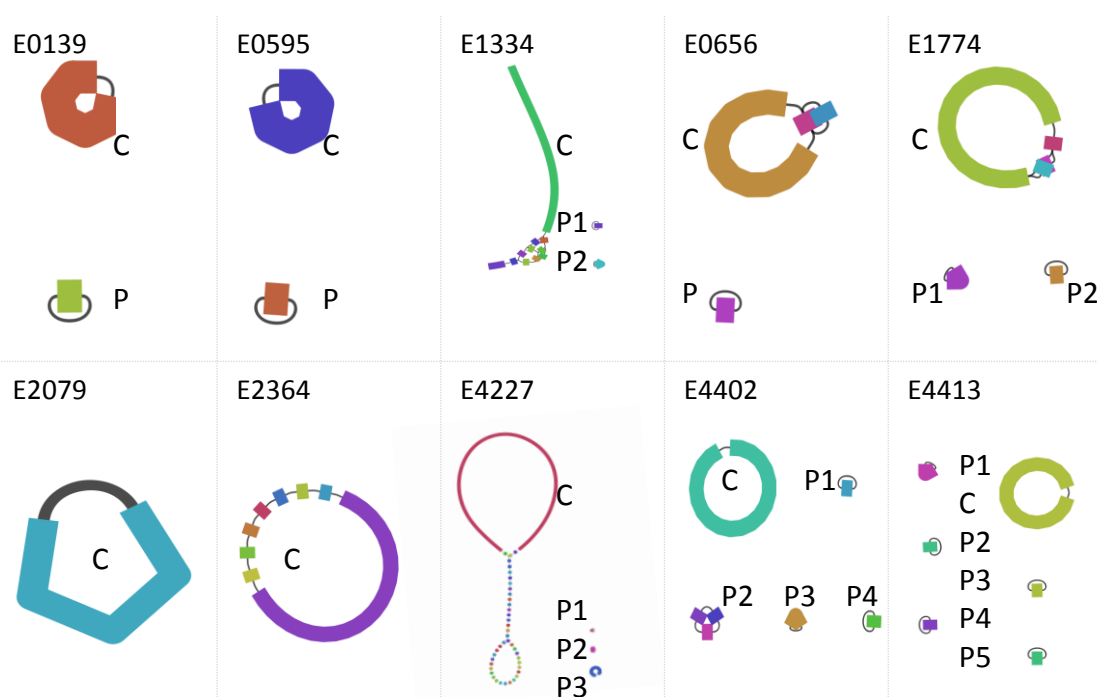


Figure 3. Assembly graphs of 10 fully sequenced isolates. Graphs were visualized with bandage. C = Chromosome. P = Plasmid

## Predicted pan-plasmidome

- ✓ 20,386 orthologous groups in 1644 strains (median 234 genes) with coverage difference from chromosome
- ✓ Most OG occur only once (Figure 2B)
- ✓ 311 OG are present in 15% or more of the strains, ie. vanA (Table 1)

OG	Freq	BBH
soj_1	1119	PrgP (plasmid)
wecD	906	streptothricin N-acetyltransferase
aadK	894	aminoglycosid nucleotidyltransferase
vanH	717	Vancomycin resistance gene cluster
vanX	716	
walR	712	
vanA	698	

Table 1. Selection of 7 of 20 most prevalent OG with frequency and best blast hit (BBH)

## Conclusions

- ✓ The pan-plasmidome predicted from Illumina-sequenced isolates can serve as prior to selection of a set of strains with diverse plasmid content for further long read sequencing
- ✓ Preliminary results from completely assembled plasmids indicate that a median of 21% of the plasmid sequences were predicted by PlasmidSPAdes from short reads
- ✓ Two isolates were missing plasmids, several isolates shared (part of) one plasmid

## Methods

- Plasmid reconstruction
  - ✓ Illumina sequencing of 1644 isolates (Fig 1), followed by plasmid reconstruction with PlasmidSPAdes
  - ✓ Annotation of potential plasmids with Prokka
  - ✓ Pan-plasmidome reconstruction with Roary
  - ✓ Dimensionality reduction with t-sne
  - ✓ K-means clustering into 50 clusters
  - ✓ Selection of strain closest to each centroid (Euclidian distance, Figure 2A)
- Completion of plasmid sequences of selected strains
  - ✓ Multiplexing and sequencing with nanopore
  - ✓ Hybrid assembly with Unicycler

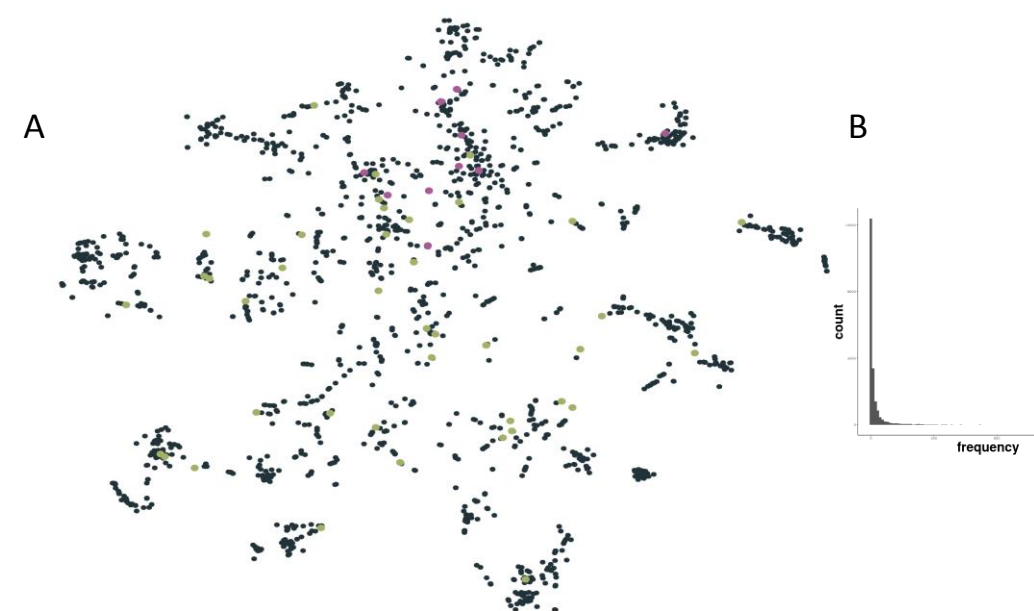


Figure 2. A. t-distributed stochastic neighbor embedding (tSNE) of presence/absence data of orthologous groups (OG) of potential plasmid origin. Green: 50 isolates selected for nanopore sequencing. Pink: 10 completed isolates. B. Distribution of 20,386 potential plasmid OG among 1644 isolates

## Results

### Completed plasmids (based on long read sequencing of 10 of the 50 isolates)

- ✓ 20 plasmids from 10 genomes (range 0-5 plasmids per isolate, Figure 3)
- ✓ 2 isolates without plasmids (both isolates from pigs)
- ✓ Almost identical plasmid found in three isolates (Figure 4A)
- ✓ No entire assembled plasmid was correctly predicted by PlasmidSPAdes
- ✓ A median of 21% of the confirmed plasmid sequences were predicted by PlasmidSPAdes from short reads (Fig 4B)

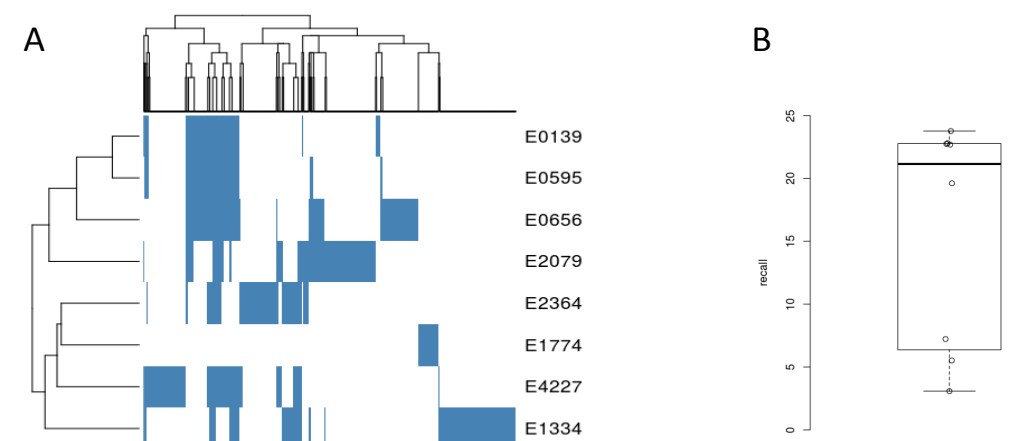


Figure 4. A. Presence/absence matrix of sequenced plasmid OG of 8 completed isolates. B. Comparison between predicted and confirmed plasmid sequence