


METHOD

Open Access



# Comprehensive antibiotic-linked mutation assessment by resistance mutation sequencing (RM-seq)

Romain Guérillot<sup>1</sup>, Lucy Li<sup>1</sup>, Sarah Baines<sup>1</sup>, Brian Howden<sup>1</sup> , Mark B. Schultz<sup>1,2,3</sup>, Torsten Seemann<sup>4,2</sup>, Ian Monk<sup>1</sup>, Sacha J. Pidot<sup>1</sup>, Wei Gao<sup>1</sup>, Stefano Giulieri<sup>1</sup>, Anders Gonçalves da Silva<sup>1,2,3</sup>, Anthony D'Agata<sup>3</sup>, Takehiro Tomita<sup>3</sup>, Anton Y. Peleg<sup>5,6</sup>, Timothy P. Stinear<sup>1,2†</sup> and Benjamin P. Howden<sup>1,2,3,7\*†</sup>

## Abstract

Mutation acquisition is a major mechanism of bacterial antibiotic resistance that remains insufficiently characterised. Here we present RM-seq, a new amplicon-based deep sequencing workflow based on a molecular barcoding technique adapted from Low Error Amplicon sequencing (LEA-seq). RM-seq allows detection and functional assessment of mutational resistance at high throughput from mixed bacterial populations. The sensitive detection of very low-frequency resistant sub-populations permits characterisation of antibiotic-linked mutational repertoires in vitro and detection of rare resistant populations during infections. Accurate quantification of resistance mutations enables phenotypic screening of mutations conferring pleiotropic phenotypes such as in vivo persistence, collateral sensitivity or cross-resistance. RM-seq will facilitate comprehensive detection, characterisation and surveillance of resistant bacterial populations (<https://github.com/rguerillot/RM-seq>).

**Keywords:** Antibiotic resistance, Resistance mutations, Deep sequencing, *Staphylococcus aureus*, *Mycobacterium tuberculosis*, Rifampicin, Daptomycin

## Background

Antimicrobial resistance is on the rise and is responsible for millions of deaths every year [1]. Bacterial populations consistently and rapidly overcome the challenge imposed by the use of a new antibiotic. Their remarkable ability to quickly develop resistance is due to their capacity to exchange genes and to their high mutation supply rate. Multidrug-resistant bacteria are therefore becoming increasingly prevalent and drug susceptibility testing (DST) is now central to avoid antibiotic misuse and minimise the risk of inducing the emergence of new resistant clones. Over recent years, genomics has become a powerful tool to understand, combat and control the rise of resistance

[2, 3]. Nevertheless, a precise definition of resistance at the genomic level is crucial to enable fast, culture-independent DST by high-throughput sequencing in the clinical context and to track and fight the spread and persistence of resistant clones globally [3, 4].

The genomic basis of resistance is relatively straightforward to establish for resistance conferred by acquisition of a specific gene. The repertoire of resistance genes (resistome) is now well defined and there are several curated databases and software prediction tools for resistance genes detection [5–7]. In contrast, comprehensive lists of mutations that confer antibiotic resistance are lacking, despite equivalent clinical relevance. Resistance to major classes of antimicrobials including quinolones, beta-lactams, rifamycins, aminoglycosides, macrolides, sulphonamides, polymyxins, glycopeptides and lipopeptides can all occur via mutations. In some species, such as *Mycobacterium tuberculosis*, resistance to all therapeutic agents is mediated by mutations [8].

Resistance mutations can be effectively selected in vitro, and so genome sequence comparisons of resistant clones

\* Correspondence: [bhowden@unimelb.edu.au](mailto:bhowden@unimelb.edu.au)

†Timothy P. Stinear and Benjamin P. Howden contributed equally to this work.

<sup>1</sup>Department of Microbiology and Immunology, The University of Melbourne at the Doherty Institute for Infection & Immunity, Melbourne, Victoria, Australia

<sup>2</sup>Doherty Applied Microbial Genomics, The University of Melbourne at the Peter Doherty Institute for Infection & Immunity, Melbourne, Victoria, Australia



derived from sensitive ancestral clones after antibiotic exposure have permitted the identification of numerous resistance-associated mutations [9–12]. From these studies, it is apparent that the mutational landscape for a single antibiotic combination within a specific bacterium can be broad [13–16]. Therefore, standard approaches relying on sequence comparisons of single pairs of isogenic mutants are not practical to extensively define the mutational resistome.

Resistance mutations commonly arise in genes encoding the primary drug target or central regulatory genes, such as *gyrA*, *parC*, *rpsL*, *gidB*, *rpoB*, *23S rRNA*, *rplC*, *rplD* and *walKR* (for quinolone, aminoglycoside, rifampicin, linezolid and glycopeptide resistance) [16, 17]. Because of their implications in central cell processes, such as DNA replication, translation, transcription and cell-wall metabolism regulation, mutations arising in these genes have been associated with a broad range of pleiotropic effects in addition to the antibiotic resistance that they cause [17]. An increasing body of literature shows that antibiotic resistance mutations can lead to broader negative therapeutic consequences through cross-resistance to other antimicrobials [18–20], increased biofilm formation [21], increased virulence [22–25] and enhanced immune evasion [25–28]. However, there is currently no efficient method to identify pleiotropic mutations. Comprehensively identifying mutations associated with antibiotic cross-resistance and increased risk of therapeutic failure will provide crucial information for future personalised medicine and will help to improve therapeutics guidelines through a greater understanding of the drivers and consequences of mutational resistance. At an epidemiological and evolutionary level, understanding why specific resistance mutations are preferentially selected might provide a rational basis for development of effective measures to combat the rise of resistance.

In this study, we developed an innovative workflow called resistance mutation sequencing (RM-seq) that enables the unbiased quantification of resistance alleles from complex in vitro-derived resistant clone libraries, selectable under any experimental condition, allowing identification and characterisation of mutational resistance and its consequences. Here we investigated mutational resistance in *Staphylococcus aureus* and *Mycobacterium tuberculosis* and demonstrated that complex resistant subpopulations can be effectively characterised in vitro or detected in vivo using RM-seq.

## Methods

### In vitro selection of rifampicin-resistant clones

All experiments were conducted with *S. aureus* USA300 strain NRS384, acquired from BEI resources. Rifampicin-resistant colonies were selected from 20 independent overnight heart infusion (HI) 10 mL broth cultures ( $5 \times 10^9$

CFU/mL) inoculated from single colonies. Cultures were pelleted at 10 min at 3000g and re-suspended in 200  $\mu$ L of HI broth ( $2.5 \times 10^{11}$  CFU/mL). These concentrated overnight cultures were then pooled and plated on HI plates supplemented with rifampicin at 0.006, 0.5, 1, and 4 mg/L. Given that the spontaneous resistance rate for rifampicin in *S. aureus* is  $\sim 2 \times 10^{-8}$  [29], 20 to 30 plates inoculated with 75  $\mu$ L were necessary to recover  $\sim 10,000$  resistant clones after 48-h incubation at 37 °C. All resistant colonies were recovered by scraping the plate flooded with 2 mL of phosphate buffered saline (PBS). After washing the pooled clone libraries in PBS, aliquots were used for genomic DNA extraction and RM-seq library preparation and stocked in 25% glycerol at  $-80$  °C.

### Amplicon library preparation and deep sequencing

Genomic DNA was extracted from 1 mL aliquots adjusted to an OD<sub>600</sub> of 5 in HI broth. Cells were pelleted and washed twice in PBS and genomic DNA was extracted using the DNeasy Blood & Tissue Kit (QIAGEN). Random 16 bp barcodes were introduced by performing 8 cycles of linear PCR with the primer *x\_rmseq\_F* (Additional file 1: Table S1) using the following PCR mix: 2  $\mu$ L of *x\_rmseq\_F* (5 nM), 1  $\mu$ L of genomic DNA (6 ng/ $\mu$ L), 12.5  $\mu$ L Phusion® High-Fidelity PCR Master Mix (2X, New England BioLabs Inc.) and 6  $\mu$ L H<sub>2</sub>O. The following PCR cycle conditions were used: 30 s at 98 °C, then 8 cycles of 10 s at 98 °C, 30 s at 50 °C, 30 s 72 °C, and a 2-min elongation step at 72 °C. Following the final cycle of the linear PCR, samples were cooled to 25 °C and the nested exponential PCR were performed by immediately adding 3.5  $\mu$ L of a primer mix containing 2  $\mu$ L of primer *x\_rmseq\_R* (100 nM), 0.6  $\mu$ L forward and 0.6  $\mu$ L reverse Nextera XT Index Kit primers (10  $\mu$ M), 0.3  $\mu$ L H<sub>2</sub>O. The PCR conditions above were then used for a further 25 cycles. The resulting amplicons comprising Illumina adaptor and indices was purified with Agencourt® AMPure® XP magnetic beads (Beckman Coulter) using beads/sample volume ratio of 0.8. Purified amplicons were then normalised at 4 nM according to expected size and measured DNA concentrations (Qubit™ dsDNA HS Assay Kit). Amplicons with different indices were pooled and the sequencing library was diluted to 15 pM with 10% *phiX* control spike and sequenced on Illumina Miseq or Nextseq using Reagent Kit v3 to produce 300 bp or 150 bp paired-end reads. Sequencing reads of RM-seq experiments are available from NCBI/ENA/DDJ under BioProject number PRJNA399605.

### Bioinformatics analysis pipeline

The RM-seq pipeline processes raw reads after demultiplexing by the Illumina sequencing instrument. The pipeline uses *bwa mem* read aligner (0.7.15-r1140) [30] to map

reads to a reference locus (*rpoB*) and *samtools* (v1.3) [31] to remove unmapped and low-quality reads from the read sets. Then *pear* (v0.9.10) [32] is used to merge paired reads. Merged reads sharing identical barcodes are aligned using *Clustal Omega* (v1.2.1) [33]. Cons from the EMBOSS suite (v6.6.0.0) [34] is used to collapse the alignments into single error-corrected consensus reads. To speed-up processing, read alignment and consensus sequence generation tasks are executed in parallel using *GNU parallel* [35]. Unique consensus DNA sequences are identified via clustering using the *cd-hit-est* module of the *CD-HIT* (v4.7) software [36]. Resultant unique representative consensus sequences are translated to amino acids using *getorf* and annotated at the protein and nucleotide level using *diffseq*, both modules of the EMBOSS suite. The annotated effect of mutation is then re-associated to each barcode in the final output table. The RM-seq bioinformatics pipeline is available from Github (<https://github.com/rguerillot/RM-seq>).

#### Construction of *rpoB* mutants by allelic exchange

Allelic exchange experiments were performed using shuttle vector pIMAY-Z [37] with some modifications. Full-length *rpoB* sequences corresponding to the 19 different *rpoB* alleles reconstructed by allelic exchange in the *S. aureus* NRS384 strain were obtained by performing PCR overlap extension with Phusion High-Fidelity DNA Polymerase (New England Biolabs) and introducing *rpoB* codon mutations to the primer tails (Additional file 1: Table S1). Gel-purified *rpoB* amplicons were then joined with pIMAY-Z using Seamless Ligation Cloning Extract (SLiCE) cloning [38] and transformed into *Escherichia coli* strain IM08B [37] to allow CC8-like methylation of the plasmid and bypass the *S. aureus* restriction barrier. The presence of a cloned *rpoB* insert in pIMAY-Z plasmid was then confirmed by colony PCR using primers pIMAY-Z-MCSF and pIMAY-Z-MCSR. Purified plasmid was then electroporated into *S. aureus* and plated on HI supplemented with chloramphenicol at 10 mg/L and X-gal (5-bromo-4-chloro-3-indolyl- $\beta$ -d-galactopyranoside; Melford) at 100 mg/L and grown 48 h at 30 °C. Blue colonies were picked and grown in HI broth at 37 °C without Cm selection pressure overnight to allow loss of the pIMAY-Z thermosensitive plasmid. Double cross-overs leading to allelic replacement of the wild type with the desired rifampicin-resistant *rpoB* alleles were directly selected by plating cultures on HI plates supplemented with 0.06 mg/L of rifampicin. Rifampicin-resistant and chloramphenicol-sensitive colonies arising at a frequency higher than  $10^{-3}$  were considered as potentially positive clones for allelic exchange as spontaneous rifampicin resistance arises at a much lower frequency of  $\sim 2 \times 10^{-8}$  (resistant clones per culture) [29] in the wild-type strain. Clones were then colony purified on HI plates before glycerol storage and extraction of

genomic DNA. To validate the allelic exchange procedure, the whole genome sequence of all reconstructed strains was determined with the Illumina Miseq or Nextseq 500 platforms, using Nextera XT paired-end libraries ( $2 \times 300$  bp or  $2 \times 150$  bp respectively). To ensure that no additional mutations were introduced during the allelic exchange procedure, reads of all mutant strains were mapped to the reference NRS384 genome [37] using Snippy (v 2.9) (<https://github.com/tseemann/snippy>). The results of the SNP/indel calling of the reconstructed mutants were then compared with our NRS384 WT reference isolate. The SNP/indel profile for each mutant is presented in Additional file 1: Table S2 and sequence reads have been deposited under BioProject numbers PRJNA360176 and PRJNA399605.

#### Antibiotic susceptibility testing and time kill assays

Rifampicin and daptomycin MIC were measured using E-tests (BioMérieux) on Mueller-Hinton plates supplemented with 50 mg/L  $\text{Ca}^{2+}$  following the manufacturer's instructions. For daptomycin time kill assays, 10 mL of brain heart infusion (BHI) broth supplemented with 8 mg/L daptomycin and 50 mg/L  $\text{Ca}^{2+}$  was inoculated with  $10^6$  CFU/mL of an overnight culture. Cultures were incubated at 37 °C with constant shaking, and samples were collected at 3, 6 and 24 h time points. Cell survival after daptomycin exposure was assessed by calculating the ratio of the CFU at 3, 6 and 24 h on the CFU of the initial inoculum ( $10^6$  CFU/mL) and taking the average colony counts of duplicate BHI agar plates. All daptomycin time kill assays were performed in biological triplicate.

#### Mutant differential abundance analysis after daptomycin selection

Three replicates of daptomycin selection were performed on a rifampicin selected population (in vitro 1 population selected with rifampicin at 0.06 mg/L). A high initial inoculum of  $5 \times 10^8$  CFU/mL was used to recover a sufficient amount of bacterial DNA from surviving cells after daptomycin exposure. After 3 h or 24 h of exposure to daptomycin at 8 mg/L, surviving bacterial populations were pelleted and washed. To remove extracellular DNA resulting from daptomycin-induced cell death, cell pellets were incubated 45 min at 37 °C with 1  $\mu$ L of Amplification Grade DNase I (1 U/ $\mu$ L Invitrogen) in 5  $\mu$ L of 10X DNase I reaction buffer and 44  $\mu$ L laboratory grade  $\text{H}_2\text{O}$ . Then, DNase I was inactivated with 5  $\mu$ L of 25 mM EDTA (pH 8.0) and 10 min incubation at 65 °C. Genomic DNA was extracted, and *rpoB* mutant abundance was assessed by RM-seq as described above. Differential abundance analysis of the mutant before and after daptomycin exposure was performed with the R *DESeq2* (1.10.1) package [39], using the count of mutation calculated from table output of the RM-seq data

processing pipeline. DESeq2 analysis was performed with all mutations count superior to 1 using default parameters and Cooks cut-off set to false. The Wald statistical test performed by *DESeq2* to estimate the significance of the changes in mutation abundance after exposure to daptomycin was used to screen *rpoB* mutations that were associated with increased or decreased tolerance to daptomycin. The detailed explanation of this test is described in [39]. Wald test *P* values were adjusted for multiple testing using the procedure of Benjamini and Hochberg [40].

### Mouse infection model

Wild-type 6-week-old female BALB/c mice were injected via the tail vein with approximately  $2 \times 10^6$  colony-forming units (CFU) in a volume of 100  $\mu$ L PBS. The mice were monitored every 8 h until completion of the experiment and were euthanized after 1 day or 7 days post-infection. Bacteria from the liver, kidney, and spleen were recovered by mechanical homogenisation in 1 mL of phosphate buffered saline (PBS), serially diluted and plated on BHI plates. Colonies forming after overnight incubation at 37 °C were pooled and assessed by RM-seq.

### Detection of resistant sub-populations of *M. tuberculosis* from sputum samples

DNA was extracted from isolates cultured from sputum specimens as previously described [41]. RM-seq libraries were prepared as described above using the *rmseq* primers specific to *pncA*, *ethA*, and *rpoB* resistance determining regions (Additional file 1: Table S1). Deep sequencing was performed on a Nextseq 500 platform using Reagent Kit v3 to produce 150 bp overlapping paired-end reads and analyses were performed using the RM-seq bioinformatics analysis pipeline. Primary *M. tuberculosis* culture and phenotypic susceptibility testing was performed using the radiometric BACTEC 460TB system (Becton Dickinson).

## Results

### The RM-seq workflow

RM-seq is an amplicon-based, deep sequencing technique founded on the single molecule barcoding method [42]. Here we have adapted the LEA-seq barcoding method first described for unbiased detection of 16S rRNA gene alleles, in order to identify and quantify at high-throughput, mutations that confer resistance to a given antibiotic [42, 43]. RM-seq can take advantage of the ability of bacteria to quickly develop resistance in vitro to identify and functionally characterise resistance-associated mutations at high throughput. A large and genetically diverse population of resistant clones that encompass the mutational landscape of resistance is selected (Fig. 1a). In order to maximise the genetic diversity, a large number of resistant

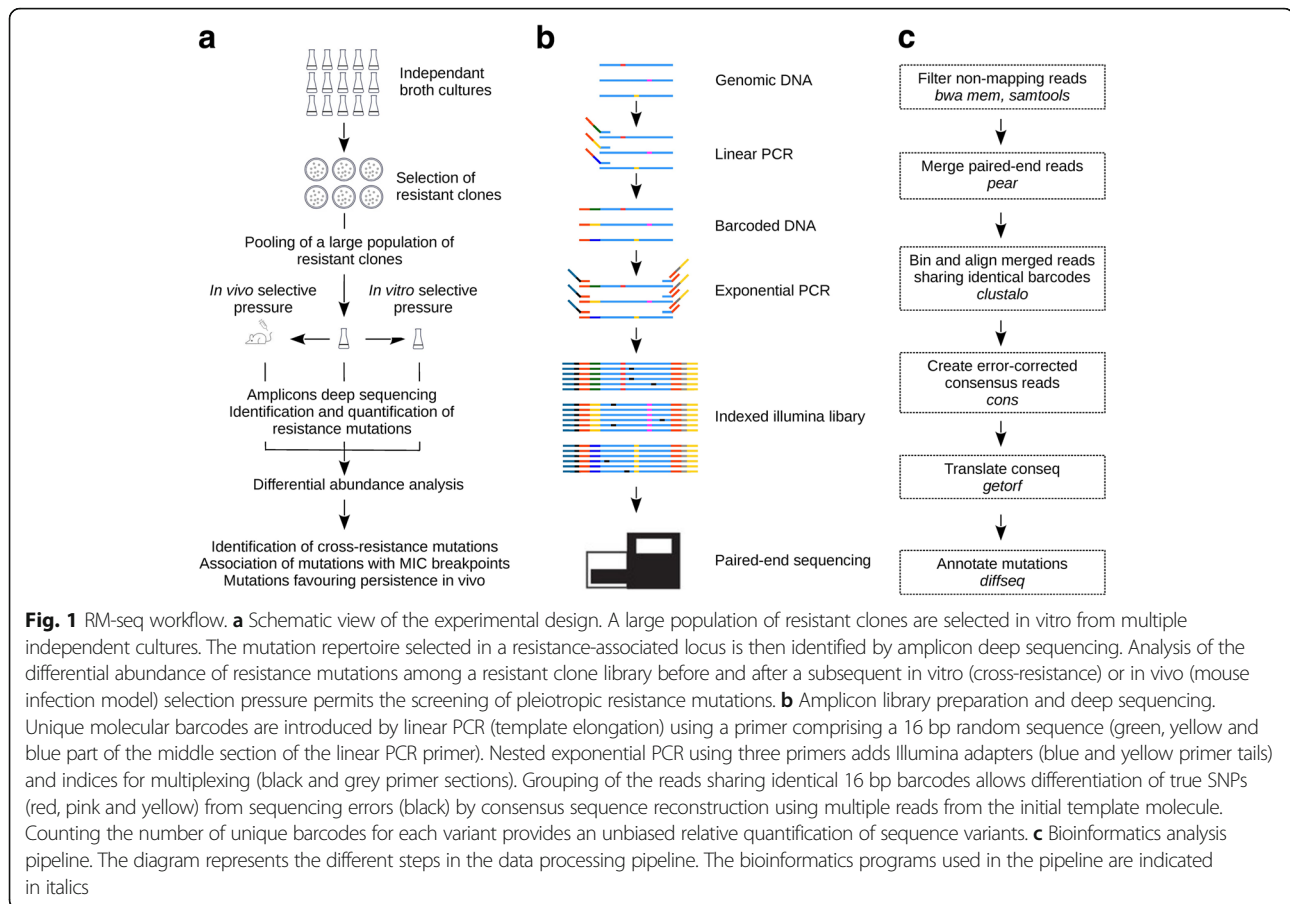
clones (~10,000) are pooled from multiple independent culture and genomic DNA of the mixed resistant population is extracted and the mutational repertoire interrogated by amplicon deep sequencing.

The high sensitivity and the accurate quantification of the frequency of all the selected mutations in a given genetic locus, enabled screening of complex, mixed libraries of resistant clones. In theory, genetic interactions can be tracked and associated with any selectable pleiotropic phenotype of interest (e.g. cross-resistance to other antimicrobials, immune evasion) by measuring the relative abundance of resistant clones before and after selection. Specific mutations that favour the growth or survival under in vitro or in vivo test condition will increase in frequency within the population and be readily detected by RM-seq.

Unbiased allele quantification and a low error rate are enabled by single-molecule barcoding during the PCR amplicon library preparation (Fig. 1b). Sequencing reads sharing identical barcodes are grouped to create consensus sequences of the genetic variants initially present in the population. The single-molecule barcoding step has two major advantages. Firstly, it allows error correction of the sequenced DNA and thus high confidence in calling of resistance-associated mutations that might occur at a frequency well below the inherent error rate (~1% [44]) of the sequencer. Secondly, it permits accurate quantification of allele frequencies by correcting for the amplification bias introduced during the exponential PCR step. The RM-seq bioinformatics pipeline takes as input the raw reads and outputs a table of all annotated substitutions, insertions and deletions identified in the selected population given the original sequence (the target locus sequence before selection). A diagram of the steps in the data analysis pipeline is presented in Fig. 1c (RM-seq analysis tool is available from <https://github.com/rguerillot/RM-seq>).

### Sensitive and quantitative detection of single-nucleotide variants in complex bacterial populations

To assess the capability of the RM-seq protocol to detect and quantify rare genetic variants from mixed populations of resistant bacteria, we first evaluated its error correction efficiency. We sequenced at high depth a 270 bp region comprising the rifampicin resistance-determining region (RRDR) of a *S. aureus* rifampicin susceptible isolate (wild-type strain NRS384). By counting incorrect nucleotide calls at each position after aligning raw reads to the WT sequence, we found an average error rate per position of  $2.8 \times 10^{-2} \pm 1.7 \times 10^{-2}$  (standard deviation [SD]), which is commonly observed for the Miseq instrument [44]. Merging forward and reverse reads replaces the lower quality score bases from one read by the higher quality score of the paired read and



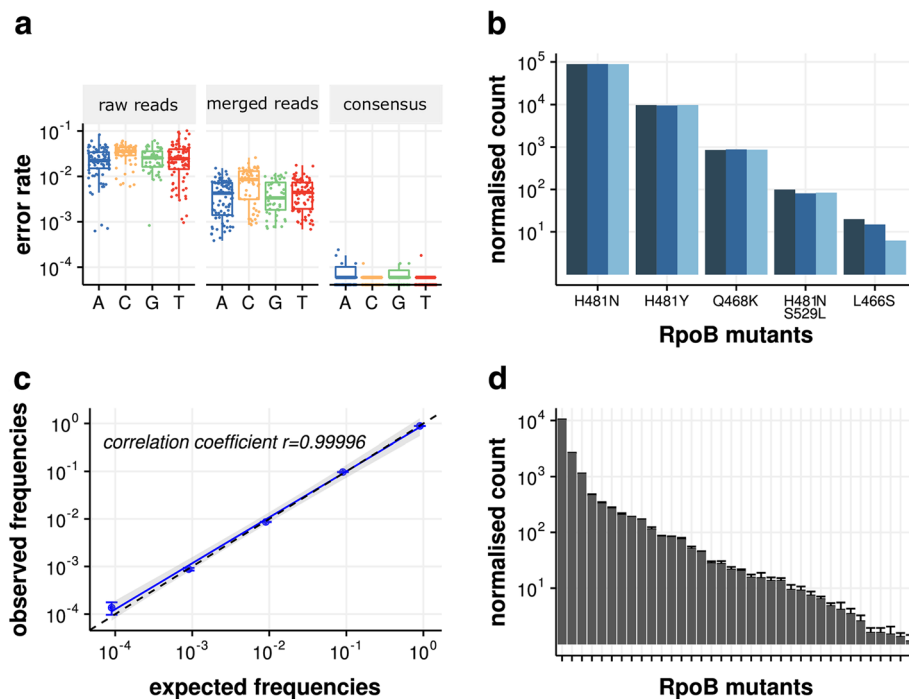
reduced the error rate by an order of magnitude to  $5.6 \times 10^{-3} \pm 4.4 \times 10^{-3}$  (SD). By reconstructing consensus reads supported by at least 10 reads, the RM-seq further reduced the error rate by three orders of magnitude to  $1.16 \times 10^{-5} \pm 3.1 \times 10^{-5}$  (SD) (Fig. 2a). At the protein level no further mutations were observed among the 16,516 consensus reads generated (error rate  $< 6 \times 10^{-5}$ ).

We then tested the performance of RM-seq genetic variant quantification on a defined population of genetically reconstructed rifampicin-resistant clones. Six different double or single nucleotide variants (SNV) representing different rifampicin-resistant *rpoB* mutants were mixed at a relative CFU frequency of 0.9, 0.09, 0.009, up to 0.000009. We applied RM-seq protocol three times independently from three different genomic DNA extractions obtained from this mock community. After library preparation and sequencing on the Illumina MiSeq platform, we obtained 1.8–2.2 million raw reads per library, which yielded between 32,433 and 35,496 error-corrected consensus reads, supported by 10 reads or more. At this sequencing depth, the mutants ranging from a relative frequency of  $\sim 1$  to  $10^{-4}$  were readily identified in all three replicates. The normalised count of the different mutants showed little variation between the replicate experiments

(Fig. 2b) and we observed a very good correlation between the expected mutant frequencies and the observed frequencies after RM-seq (Fig. 2c). We also assessed the technical variability of the detection and quantification of RM-seq by independently processing three times the same complex population of in vitro selected resistant clones ( $\sim 10,000$  colonies). The relative standard error (RSE) of variant quantification ranged from 0.3% for the most frequent to 38% for rarest variants and the median RSE was 11% (Fig. 2d).

#### High-throughput identification of rifampicin resistance mutations

In order to comprehensively characterise the mutational repertoire associated with rifampicin resistance, we applied RM-seq on the RRDR of three independent pools of  $\sim 10,000$  colonies capable of growing on agar supplemented with 0.06 mg/L of rifampicin (European Committee on Antimicrobial Susceptibility Testing [EUCAST] non-susceptibility clinical breakpoint). In total, we identified 72 different predicted protein variants; among these, 34 were identified in the three independent resistant populations, 17 variants were identified among two resistant populations, and 21 were identified in a single selection



**Fig. 2** Assessments of the RM-seq protocol. **a** Error correction evaluation. RM-seq error correction combining merging of paired-end reads with consensus sequence determination from grouped reads sharing identical barcode allows a three order of magnitude reduction in false SNP calling when compared with raw reads calling for the different base. **b** Quantification of populations of *S. aureus* *rpoB* mutants. Three independent assessments of *rpoB* mutants from three independent genomic DNA preparations originating from a defined population are presented by the different blue bars (technical replicates). **c** Correlation of observed versus expected SNV frequencies. Blue points represent means and error bars represent SEM of three technical replicates. The blue line represents the linear regression of the frequencies measured by RM-seq and the dashed line represent the perfect correlation between expected and observed frequencies. **d** Quantification of *S. aureus* *rpoB* mutants from a complex population of in vitro selected rifampicin-resistant mutants. Columns represent mean normalised counts of the different *rpoB* mutations that were observed among all triplicates, and error bars represent SEM

experiment (Fig. 3). According to our recent extensive literature review of the alleles previously associated with rifampicin resistance [45], 30 mutations were previously associated with rifampicin non-susceptibility and 42 alleles identified by RM-seq represent new associations.

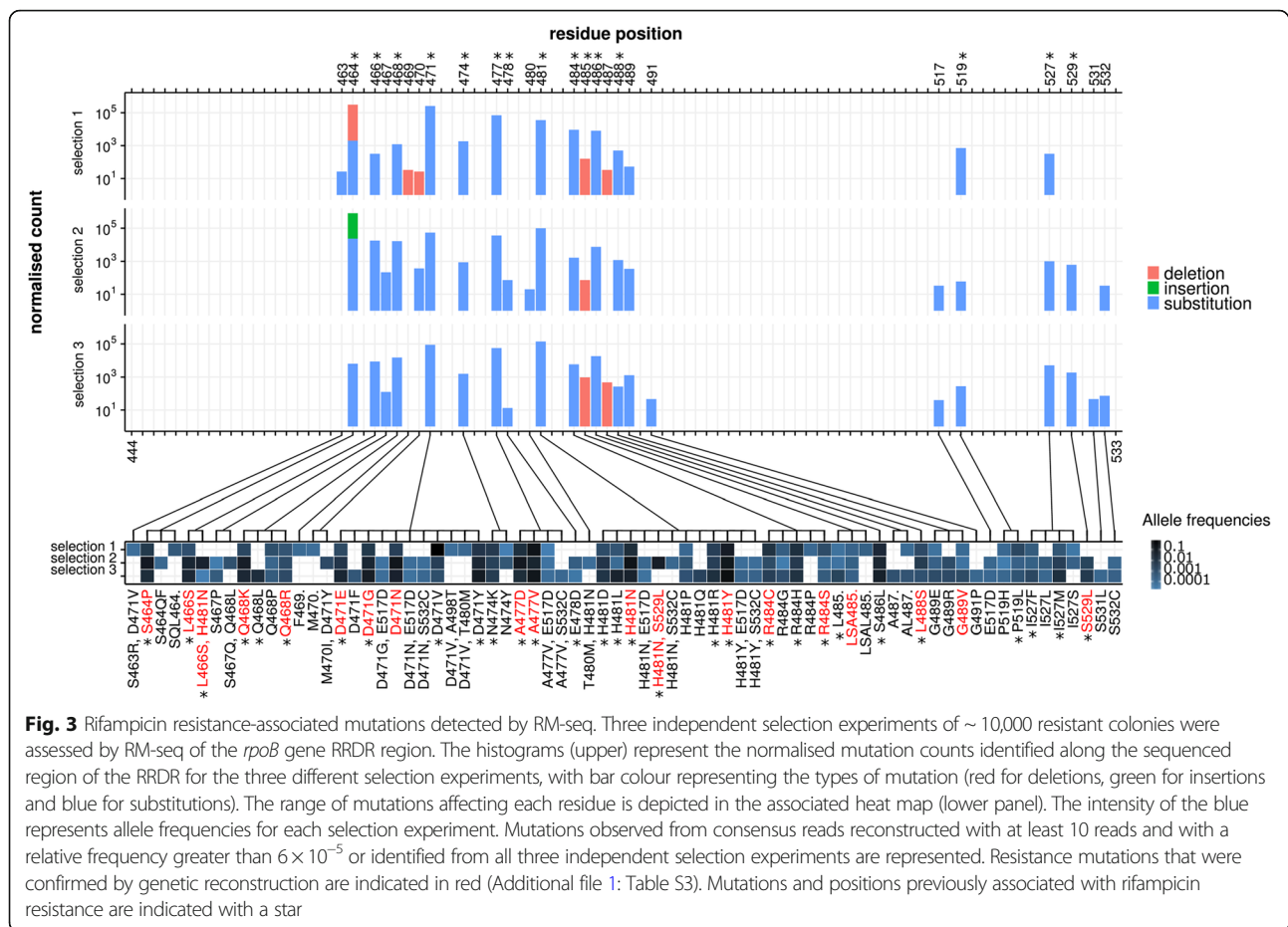
By looking at the different mutated positions, 21 amino acid positions were repeatedly affected along the RRDR with a similar pattern of mutation frequency at these positions. We observed that 11 different amino acid positions have never previously been associated with rifampicin resistance. The 3D structure modelling of *S. aureus* RpoB protein from *Escherichia coli* RpoB-rifampicin structure showed that the mutated positions were all in close proximity ( $\leq 10$  Å) to the rifampicin binding pocket of the beta-subunit of the RNA polymerase (Additional file 2: Figure S1). Therefore, amino acid sequence alteration at these positions are likely to reduce the rifampicin-RNA polymerase affinity and thus to promote resistance. Interestingly, several residues in close proximity to the rifampicin binding pocket were never affected, suggesting that amino acid substitution at these locations do not impair rifampicin binding or that functional constraints make changes to these positions lethal for *S. aureus*. The vast

majority of the variants led to amino acid substitutions and several positions, such as 471 and 481, were found to be affected by a high number of different substitutions (11 and 12, respectively). We also observed one complex insertion (S464QF) and eight different deletions. Positions 485 and 487 represented deletion hotspots, as they were affected by single, triple and quadruple residue deletions (L485., LSA485., LSAL485.) and single and double deletions (A487. and AL487.), respectively.

We used allelic exchange and site-directed mutagenesis in the WT susceptible background (rifampicin MIC 0.012 mg/L) to reconstruct 19 different *rpoB* alleles that were identified by RM-seq. After whole genome sequencing was used to ensure no secondary non-synonymous mutations or insertion/deletion were introduced (Additional file 1: Table S2), we confirmed that all these mutations resulted in rifampicin non-susceptibility or resistance with rifampicin MICs above 0.095 mg/L (Additional file 1: Table S3).

#### High-throughput genotype to phenotype associations of resistance mutations with clinical breakpoints

To test if RM-seq could be applied to link a repertoire of resistance mutations to a particular resistance threshold, we



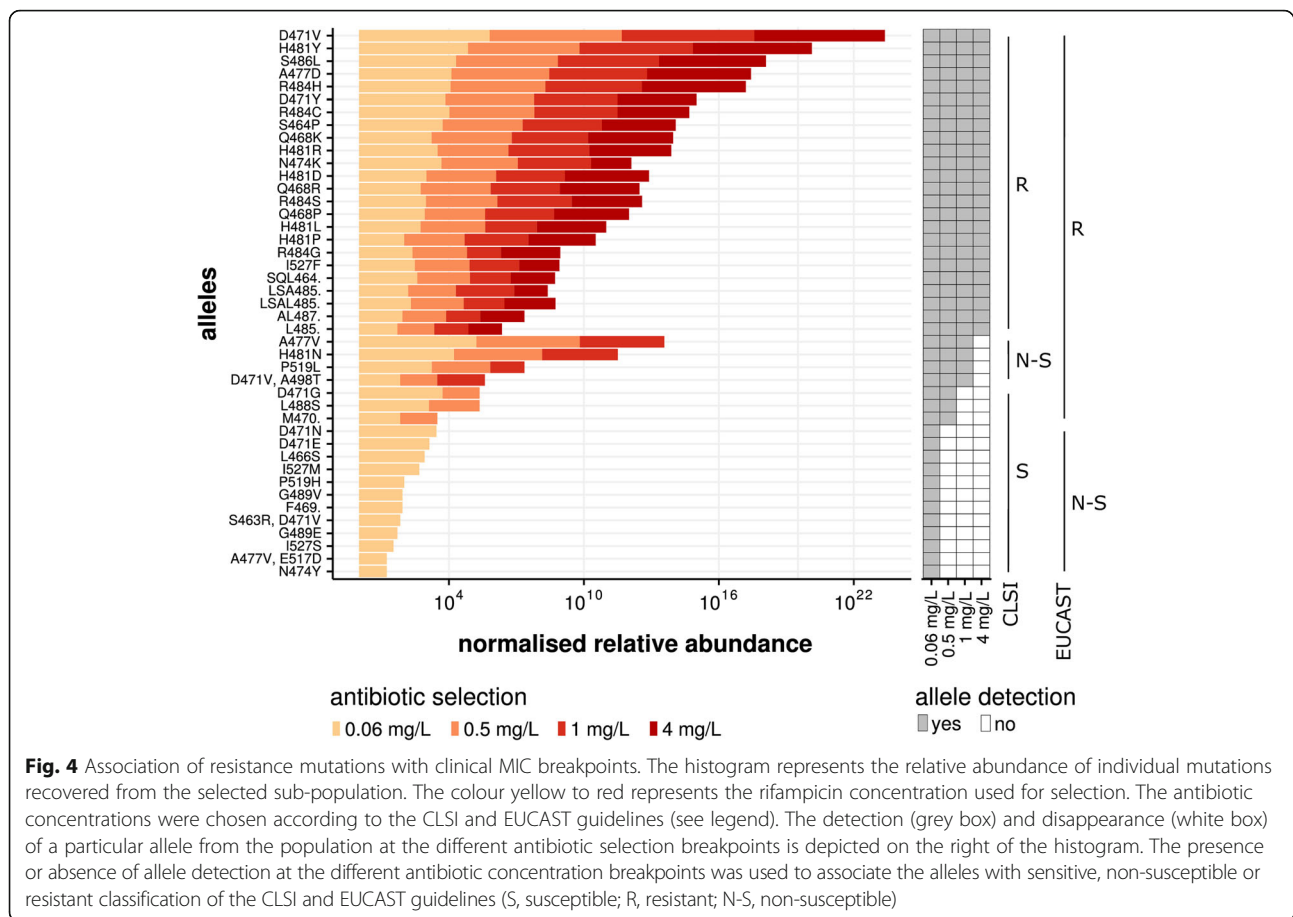
selected rifampicin-resistant clones, grown on plates supplemented with different concentrations of antibiotic (in this case, rifampicin). To select resistant sub-populations, we used the most widely used clinical resistance breakpoints from the guidelines of the European Committee on Antimicrobial Susceptibility Testing (EUCAST) and the Clinical & Laboratory Standards Institute (CLSI) [46, 47]. Therefore, we selected sub-populations growing on plates supplemented with rifampicin at concentrations of 0.06 mg/L (EUCAST non-susceptibility), 0.5 mg/L (EUCAST resistance), 1 mg/L (CLSI non-susceptibility) and 4 mg/L (CLSI resistance). The result of resistant sub-population detection and quantification by RM-seq associated with the different antibiotic concentration thresholds is presented in Fig. 4. Among 43 mutations, 24 mutations were detected at all antibiotic concentration thresholds and therefore would be classified as resistance-conferring mutations by both guidelines. Among the 19 other mutations detected, four were associated with resistance levels ranging from 1 to 4 mg/L, three with resistance ranging from 0.5 to 1 mg/L and the remaining 12 with resistance ranging from 0.006 to 0.5 mg/L. Interestingly, *S. aureus* with any of these last 12 alleles, selected only at low antibiotic concentrations, would be classified as non-susceptible by EUCAST and susceptible by CLSI.

Similarly, *S. aureus* with three mutations associated with resistance by EUCAST would be classified as susceptible by CLSI (Fig. 4).

We used mutants reconstructed by allelic-exchange to verify that the resistance level predicted by RM-seq matched the MIC conferred by a particular allele. Among 17 reconstructed mutants tested, 16 showed MICs in complete accord with the RM-seq prediction (Additional file 1: Table S3). One mutant (D471G) with a borderline measured MIC of 0.5 mg/L was predicted to have an MIC superior to 0.5 and inferior or equal to 1 despite showing clear reduction in abundance on 0.5 mg/L plate by RM-seq (Fig. 4).

#### High-throughput screening of resistance mutations associated with antimicrobial cross-resistance or collateral sensitivity

In order to evaluate if RM-seq can be used to characterise pleiotropic resistance mutations that confer an increased or decreased susceptibility to a second antibiotic (cross-resistance or collateral sensitivity respectively), we followed the differential abundance of resistance mutations of a complex rifampicin-resistant population after selection with a second antibiotic, daptomycin. We



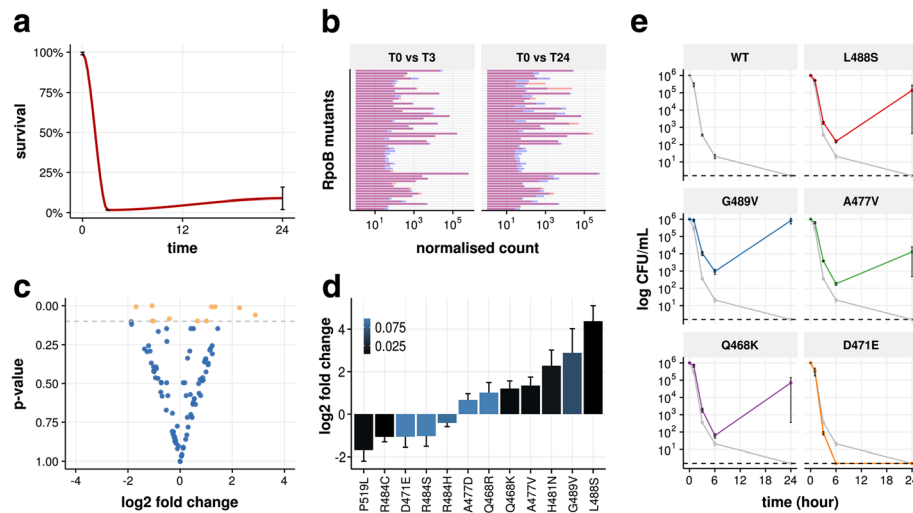
chose daptomycin because it is a last-line antibiotic used against multidrug-resistant *S. aureus*, commonly deployed in combination therapy with rifampicin to treat complicated infections [48–50]. Furthermore, some *rpoB* mutations have been previously associated with subtle changes in daptomycin MIC [51, 52]. We screened for pleiotropic effects on daptomycin resistance by performing three independent time killing experiments using a large in vitro-derived population of rifampicin-resistant clones. Daptomycin concentrations of 8 mg/L corresponding to the minimal plasma concentration commonly reached during standard antibiotic therapy were used [53]. Survival of the rifampicin-resistant population at 3 h represented 1.6% ( $\pm 0.1$  SEM) of the initial inoculum and bacterial regrowth was observed to 8.8% ( $\pm 6.9$  SEM) at 24 h (Fig. 5a). The abundance of all rifampicin resistance mutations were then quantified by RM-seq for the initial bacterial population (the inoculum) and the surviving population at 3 h and 24 h after daptomycin exposure for the three independent killing experiments (Fig. 5b).

We tested for significant differential abundance of all the different mutations detected (Fig. 5c). After 3 h of daptomycin treatment, one mutation appeared to

increase in frequency (Q468K) and another decreased (P519L) but the null hypothesis (no change) could not be rejected ( $p > 0.05$  after correction for multiple testing [Wald test]). At 24 h of daptomycin selection, differential abundance of these two mutations increased, together with 10 other rifampicin resistance mutations when compared with the mutant abundance in the initial population (Fig. 5d). All the rifampicin resistance mutations that were previously identified as conferring decreased susceptibility to daptomycin ( $n = 6$ ) were found enriched after daptomycin selection, and four mutations had significant fold changes at the 24-h time point ( $p < 0.05$ , Wald test). These experiments show that changes in relative allele abundance as measured by RM-seq are concordant with changes in daptomycin susceptibility [52, 54].

In order to validate the use of RM-seq as a screening method to identify new mutations that confer cross-resistance or collateral sensitivity, we introduced in the wild-type strain by allelic exchange seven rifampicin-resistant mutations that were significantly enriched and three mutations that were significantly rarefied after daptomycin selection. Among these mutations, MIC testing validated six of the seven rifampicin





**Fig. 5** Screening of resistance mutations associated with cross-resistance or collateral sensitivity. **a** Daptomycin selection (8 mg/L) of a pooled population of in vitro selected rifampicin-resistant clones. Survival was quantified by CFU counting on BHI agar plates at 3 h and 24 h of exposure. Error bars represent  $\pm$  SEM of three independent exposures to daptomycin. **b** Rifampicin-resistant mutant quantification of rifampicin mutant before and after 3 h or 24 h of daptomycin exposure. Each bar of the histogram represents the averaged normalised count of the different *rpoB* mutants in the population. Average quantification of the three replicates at  $T=0$  and after daptomycin exposure are indicated by blue and red bars respectively. Bars are superimposed for each mutant and overlap of the bars are coloured in purple. Increases and decreases in allele frequencies after daptomycin exposure are indicated by red and blue bars respectively on the top of purple bars. **c** Volcano plot showing fold change in *rpoB* alleles frequency after 24 h of daptomycin exposure. Each dot represents a different *rpoB* mutant. Orange dots represent mutants with  $p$  value  $< 0.1$  by Wald test. **d** Rifampicin resistance mutations associated with significant fold change after 24 h of daptomycin treatment. Mutations with positive and negative  $\log_2$  fold change are predicted to be associated with cross-resistance and collateral sensitivity to daptomycin, respectively. The intensity of the blue coloration of the bars represents adjusted  $p$  values (Wald test). **e** Daptomycin time kill assays. Rifampicin-resistant mutants were assessed in triplicates (biological replicates), points represent the mean survival at each time point and error bars SD. Dashed lines represent detection limit

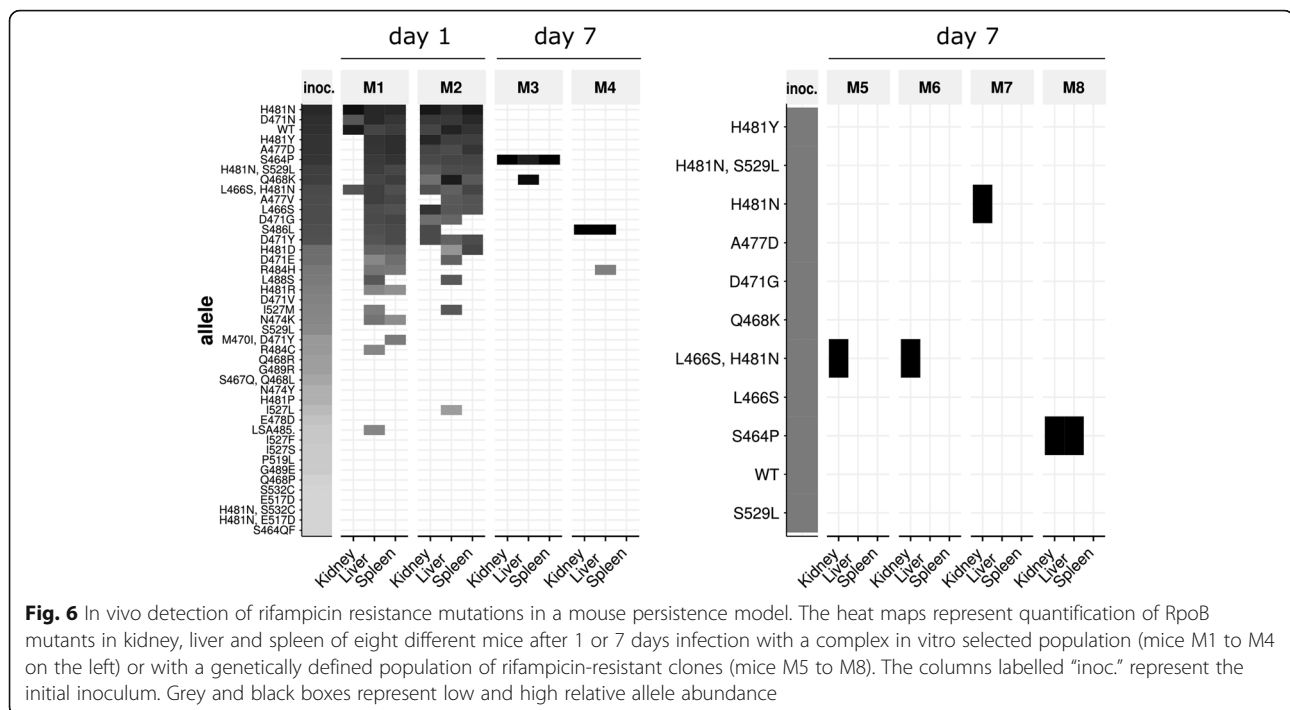
resistance mutations as decreasing susceptibility to daptomycin and one mutation as increasing the daptomycin susceptibility (Additional file 1: Table S3). We then performed daptomycin time kill assays and found that even though the D471E mutation did not show a decreased MIC to daptomycin (Additional file 1: Table S3), this mutant was less tolerant to daptomycin (Fig. 5e), concordant with the RM-seq prediction which demonstrated reduced abundance of this mutation after daptomycin exposure (Fig. 5d). Similarly, rifampicin resistance mutations L488S, G489V, A477V and Q468K were clearly associated with increased tolerance to daptomycin killing (Fig. 5e).

Taken together, our data demonstrate that RM-seq can identify pleiotropic resistance mutations conferring changes in susceptibility to a secondary antibiotic from large pool of resistant clone selected in vitro after exposure to a primary antibiotic.

#### Tracking-resistant clones in vivo in a mouse infection model

The relationship between resistance selection, in vivo fitness cost and pathogenicity has been a long standing research topic [24, 27, 55–57]. In a proof-of-principle experiment to investigate the dynamics and fitness of

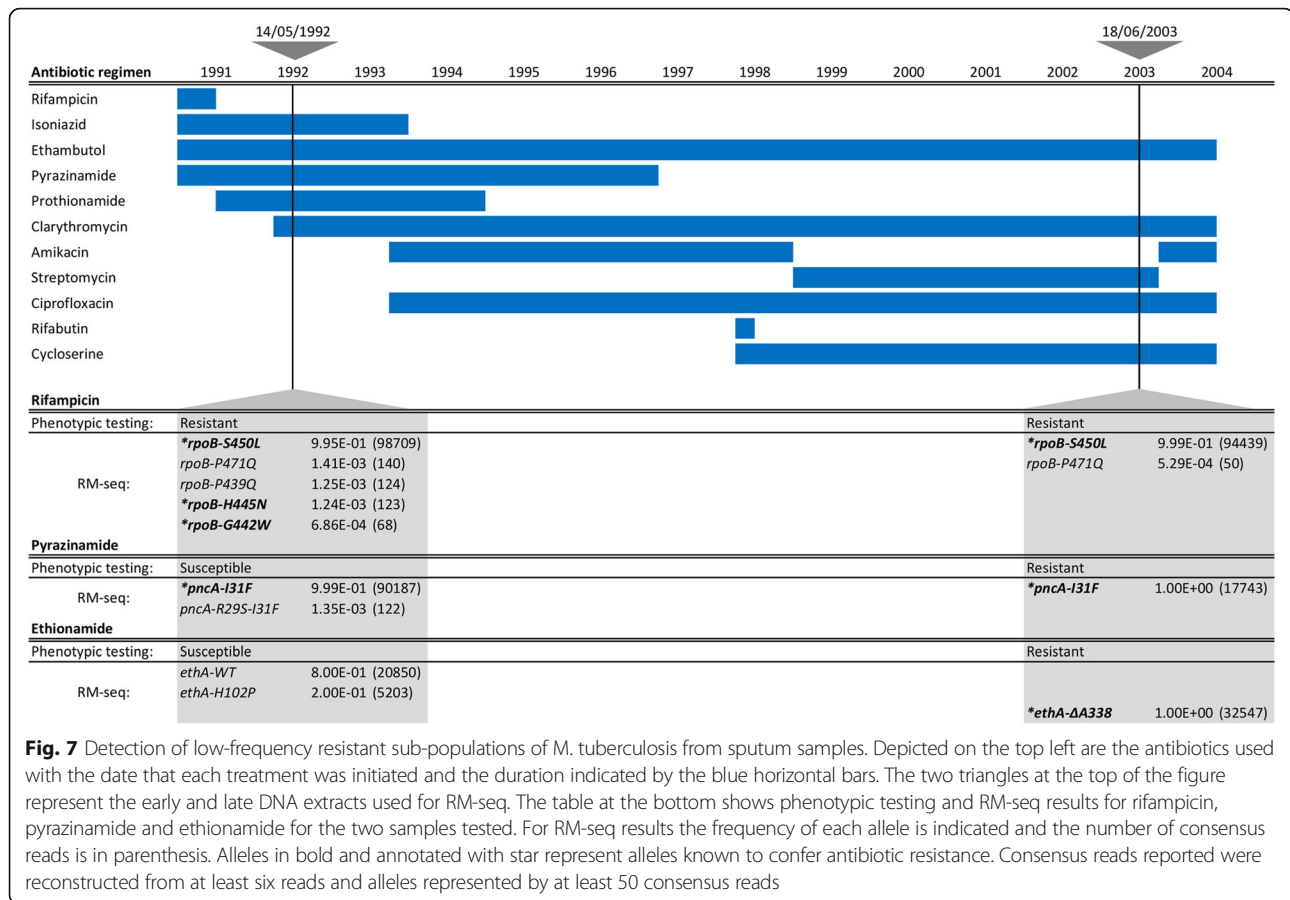
resistance mutations in vivo, we followed the abundance of rifampicin resistance mutation by RM-seq in a mouse model of persistent infection. Six-week-old BALB/c mice were injected via the tail vein with a complex, in vitro-derived population of rifampicin-resistant mutants that also included susceptible WT clones. We then quantified the abundance of RpoB mutants in the inoculum and at 1 and 7 days post-infection in the kidney, liver and spleen of the mice (Fig. 6). At 24-h post-infection, we recovered a diverse set of RpoB mutants with different relative abundances in the two mice tested. The diversity of mutants appeared to be reduced when compared with the inoculum in the different organs and several initially abundant mutants were not recovered showing a rapid clearance of several inoculated clones. Interestingly, at 7 days post-infection, we observed a drastic reduction in resistant clone diversity with only a small number of clones dominating. This result supports the concept that the establishment of *S. aureus* infection in the mouse is highly clonal, following a “bottleneck” in which very few bacterial cells establish infectious foci or abscesses in invaded organs [58]. Despite the intravenous inoculum containing a diversity of resistant clone, we observed that within a given mouse different organs were infected with the same clones.



We then infected mice with an inoculum comprising an equal amount of 10 reconstructed RpoB mutants together with the wild-type-susceptible strain. After 7 days of infection, four mice were analysed for RpoB mutant abundance by RM-seq. As observed in the previous experiment, the wild-type allele did not persist after 7 days, showing that resistant clones are not outcompeted by the wild-type clone for persistence in the mouse model; even without antibiotic selective pressure (Fig. 6). Three out of four mice were infected with clones encoding the H481N mutation, which has been found to be the most frequent mutation among sequenced *S. aureus* human isolates [45], two had the L466S, H481N double mutation and one had H481N only. Intriguingly, because the mice were infected simultaneously with 11 different clones, the probability is low that at least two mice would become infected with the L466S, H481N by chance ( $p = 0.043$ ). The probability is also low that at least three mice would become randomly infected by a clone encoding the H481N mutation ( $p = 0.057$ ). Given the relatively small number of mice investigated here, no conclusions can be drawn on the potential competitive advantage of specific resistance mutations in vivo. Nevertheless, we show here that RM-seq can be used to follow the dynamics of complex populations of clones in a mouse infection model and that the design of complex multi-clone competition assays in vivo is achievable with RM-seq.

#### Detection of low-frequency resistant sub-populations of *M. tuberculosis* from sputum samples

A primary motivation for developing RM-seq is to reduce inappropriate antimicrobial therapy by allowing the early detection of low-frequency drug-resistant sub-populations that can arise during antimicrobial therapy. Treatment of tuberculosis, caused by infection with *M. tuberculosis*, could be significantly improved by an accurate and sensitive amplicon sequencing method. This is because phenotypic testing of resistance takes weeks to obtain a result and current rapid molecular diagnostic methods only detect a handful of commonly occurring mutations and have low sensitivity for the detection of resistant sub-populations [59]. To assess the potential applicability of RM-seq for clinical detection of resistant sub-population, we retrospectively applied RM-seq on genomic DNA extracted from cultured sputum sample of a previously reported example of chronic pulmonary multidrug-resistant tuberculosis that had been investigated by whole genome sequencing [60]. Here we investigated the emergence of resistance mutations from two samples (sampling interval of 11 years) of three different loci in the genes *rpoB*, *pncA* and *ethA* associated with resistance to rifampicin, pyrazinamide and ethionamide. The multiple changes that were made to the treatment regimen are summarised in Fig. 7. Using RM-seq, we found four other low frequencies of *rpoB* mutants in addition to the dominant *rpoB*-S450L



alleles previously associated with rifampicin resistance in this case (Fig. 7) [61]. Among those, *rpoB*-H445N (frequency of  $1.24 \times 10^{-3}$ ) and *rpoB*-G442W (frequency of  $6.86 \times 10^{-4}$ ) represent known rifampicin resistance-conferring alleles [62, 63]. In the later sputum samples collected 12 years after the end of rifampicin treatment, these low frequency sub-populations of rifampicin-resistant clones were not detected but the dominant rifampicin-resistant population harbouring mutation *rpoB*-S450L persisted together with a low-frequency population harbouring the *rpoB*-P471Q allele. In association with pyrazinamide resistance the resistant allele *pncA*-I31F dominated the population after a year and half of treatment together with a low frequency of double mutant sub-population represented by the allele *pncA*-R29S-I31F. The resistant mutant *pncA*-I31F was also detected on the later isolate. Surprisingly, early samples were susceptible to pyrazinamide as established by phenotypic testing despite a high prevalence of the *pncA*-I31F-resistant allele. For ethionamide resistance, the wild-type version of the gene *ethA* was initially dominant in the population (frequency of  $8 \times 10^{-1}$ ) together with a low-frequency allele not associated with resistance *ethA*-H102P (frequency of  $2 \times 10^{-1}$ ). The ethionamide

resistance mutation *ethA*-ΔA338 causing a frameshift in the gene was readily detected in accordance with phenotypic testing in the later sputum sample. Thus, RM-seq was able to identify low-frequency sub-populations of antibiotic-resistant *M. tuberculosis*.

## Discussion

In this study, we designed and validated a new high-throughput workflow call RM-seq that enables fast and comprehensive characterisation of antibiotic resistance mutations. We show that a straightforward molecular PCR-based barcoding step coupled with high-throughput sequencing significantly reduces background sequencing noise and permits accurate identification and quantification of rare resistance mutations in complex bacterial populations. By applying RM-seq on large pools of in vitro selected rifampicin-resistant *S. aureus* clones, we demonstrate that the mutational resistome of a resistance locus can be defined. We found that the range of rifampicin resistance mutations in *S. aureus* is broader than previously understood, highlighting the inadequacy of our understanding of the genetic basis of resistance. Here, 72 mutations were associated with rifampicin resistance in *S. aureus*. In comparison, the Comprehensive

Antibiotic Resistance Database (CARD) only contains six rifampicin resistance mutations [5]. As the RM-seq protocol can be applied on any combination of microorganisms and resistance, its use has the potential to greatly enhance current knowledge on microbial adaptation to antibiotic exposure.

One limitation of RM-seq is the size limit of the sequenced region that can be interrogated by a single amplicon (270 bp with fully overlapping reads). This limitation is imposed by the maximum read length of Illumina® paired-end sequencing technology. Nevertheless, because RM-seq is compatible with standard Nextera® indexing primers, up to 384 resistance targets can be multiplexed in a single sequencing run. Furthermore, when performing read sub-sampling simulation, we found that the number of high-quality consensus reads increase almost linearly with the number of reads when performing low-depth sequencing (Additional file 2: Figure S2). As little as 140,000 reads would be sufficient to obtain 10,000 consensus reads supported by 10 reads. Theoretically, all resistance variants arising among more than 350 different targeted regions of 270 bp would be accurately identified from a mixed population 1000 bacterial clones using a single MiSeq run (94,500 bp, ~ 86 different genes). This kind of experimental design would be valuable to characterise the genetic basis of poorly defined resistance mechanisms or to determine all the resistance mutation arising in a particular gene. During the preparation of this manuscript, we scanned the full genes *mprF* (2523 bp) and *cls2* (1482 bp) for mutations conferring daptomycin resistance in *S. aureus* by sequencing 10 and six amplicons respectively (manuscript in preparation). For the development of diagnostic tool, multiple resistance hotspots could be assessed by RM-seq using a similar design.

RM-seq trades sequencing depth with accuracy of detection and require a relatively high-sequencing depth (Additional file 1: Table S4) to be effective. Compared with widely used PCR deep sequencing approaches, the method repeatedly sequences the same template DNA molecule to allow consensus-based sequence error correction. This permits accurate detection and quantification of sequence variants that occur at a frequency below 10%. When mutations are expected to occur at a frequency above 10%, then barcoded amplicon deep sequencing is not necessary (see error correction comparison of Fig. 2a).

The application of RM-seq is not restricted to the high-throughput identification of resistance mutations and can also be used to characterise the phenotypic impact of specific resistance mutations. We demonstrate here that differential mutation abundance analysis can be performed to link subsets of mutations with clinical resistance breakpoints and to identify resistance mutations that favour survival or multiplication in particular conditions.

Comparisons of allele frequencies in mixed populations before and after exposure to a second antibiotic permitted the identification of specific resistance mutations that confer cross-resistance and collateral sensitivity. The demonstration that several specific rifampicin resistance mutations can prevent bacterial clearance by daptomycin in vitro can have potential clinical implication regarding the usage rifampicin and daptomycin in combination therapy. A deeper understanding of how evolution of microbial resistance towards a given antibiotic influences susceptibility or resistance to other drugs would have profound impact as it could be exploited to fight resistance rise through combination therapy or by the temporal cycling of different antibiotics [18, 64, 65].

We also showed that the persistence of resistance alleles can be followed during experimental infection (murine blood stream infection model). As it is known that specific resistance mutations can favour pathogenesis and immune evasion [24–26], RM-seq can be used to screen for resistance mutations that increase or decrease survival against ex vivo selective pressures (e.g. whole blood killing, phagocytosis, antimicrobial peptide killing, complement killing) or that favour colonisation or tissue invasion (e.g. biofilm formation, cell attachment, intracellular persistence). A better characterisation of critical resistance mutations that confer cross-resistance or that impact pathogenesis would permit both improving antibiotic resistance surveillance and drug management if a higher therapeutic risk is confirmed.

Fast and culture-independent molecular diagnostic tools have revolutionised pathogen identification and resistance typing in clinical settings. We show here that RM-seq can be used to detect very low frequency sub-population of resistant clones from patients infected by *M. tuberculosis*. The development of diagnostic tools based on the combination of PCR-based barcoding and massively parallel sequencing represents a promising approach for the next generation of genetic-based diagnostics. RM-seq has potential advantages over standard quantitative and molecular probe-based diagnostic tests. For instance, RM-seq would be more sensitive than the current best practice platform for rifampicin resistance detection in *M. tuberculosis*, GeneXpert, as this platform fails to identify sub-populations of rifampicin-resistant strains representing less than 10% of the population [59] and digital PCR and qPCR assays that have been validated for rare mutations with frequencies-of-occurrence not lower than 0.1% [66]. This property of RM-seq may have important clinical implications as similarly to molecular test, most phenotypic tests fail to detect heterogeneous resistance with resistance allele frequency below 1%, and lower frequency of resistance have been frequently described [2, 13, 67]. RM-seq detection is not conditional on the affinity of short DNA probes;

therefore, all sensitive and resistant variant can be detected and differentiated at the sequence level. Nevertheless, like all PCR-based sequence assays, RM-seq variant detection and quantification assumes that the primers bind equally efficiently in all alleles in the population and consequently primers should be designed outside potentially variable regions. Taken together, diagnostic tools based on molecular barcoding and deep sequencing have the potential to perform better than current state of the art diagnostic tests by accurately detecting pre-existing rare resistant sub-population as well as uncommon resistance mutations.

## Conclusions

We expect that RM-seq will be a valuable tool for the comprehensive characterisation of the mutational resistance repertoire. A deeper understanding of resistance at the DNA level will be the basis for improved genomic surveillance of antibiotic-resistant pathogens, optimised antibiotic treatment regimens, and can ultimately lead to precision medicine approaches for treating microbial infections.

## Additional files

**Additional file 1: Table S1.** Primer sequences used for mutation reconstruction and RM-seq. **Table S2.** Whole genome sequencing and SNP/indel calling of reconstructed *rpoB* mutants. **Table S3.** Rifampicin and daptomycin MIC of reconstructed mutants. **Table S4.** Characteristics of RM-seq sequencing libraries. (XLSX 34 kb)

**Additional file 2: Figure S1.** 3D model of *S. aureus* RpoB protein in complex with rifampicin. **Figure S2.** Prediction of the number of consensus reads at different sequencing depths. (PDF 648 kb)

## Abbreviations

BHI: Brain heart infusion; CFU: Colony-forming unit; CLSI: Clinical & Laboratory Standards Institute; *E. coli*: *Escherichia coli*; EUCAST: European committee on antimicrobial susceptibility testing; *M. tuberculosis*: *Mycobacterium tuberculosis*; PBS: Phosphate buffered saline; RM-seq: Resistance mutation sequencing; *S. aureus*: *Staphylococcus aureus*; SLICE: Seamless ligation cloning extract; WT: Wild type

## Funding

This work was supported by the National Health and Medical Research Council (NHMRC), Australia project grant (GNT1066791) and Research Fellowship to TPS (GNT1008549) and Practitioner Fellowship to BPH (GNT1105905). Doherty Applied Microbial Genomics is funded by the Department of Microbiology and Immunology at The University of Melbourne.

## Availability of data and materials

Sequencing reads are available from NCBI/ENA/DBJ under BioProject numbers PRJNA360176 and PRJNA399605. The RM-seq analysis pipeline code is available from Github (<https://github.com/rguerillot/RM-seq>).

## Authors' contributions

RG, TPS and BPH designed and planned the project. RG, LL, SB, BH, WG, AD and TT designed and performed the laboratory experiments. RG, TT, IM, SP, SG, AG, MBS, TS, AYP, TPS and BPH provided intellectual input and analysed the data. The manuscript was drafted by RG, TPS and BPH. All authors reviewed and contributed to the final manuscript.

## Ethics approval and consent to participate

Samples from a TB patient were collected as part of the Victorian Department of Health and Human Services TB Control Programme, under the Public Health and Wellbeing Act 2008 (<https://www2.health.vic.gov.au/about/legislation/public-health-and-wellbeing-act>). Informed consent was obtained from the patient. This research was performed in accordance with the Declaration of Helsinki. Mouse experiments were performed in accordance with protocols approved by the Biochemistry & Molecular Biology, Dental Science, Medicine, Microbiology & Immunology, and Surgery Animal Ethics Committee of the University of Melbourne (approval number 1212591).

## Competing interests

The authors declare that they have no competing interests.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Author details

<sup>1</sup>Department of Microbiology and Immunology, The University of Melbourne at the Doherty Institute for Infection & Immunity, Melbourne, Victoria, Australia. <sup>2</sup>Doherty Applied Microbial Genomics, The University of Melbourne at the Peter Doherty Institute for Infection & Immunity, Melbourne, Victoria, Australia. <sup>3</sup>Microbiological Diagnostic Unit Public Health Laboratory, The University of Melbourne at the Peter Doherty Institute for Infection & Immunity, Melbourne, Victoria, Australia. <sup>4</sup>Melbourne Bioinformatics, The University of Melbourne, Melbourne, Victoria, Australia. <sup>5</sup>Department of Infectious Diseases, The Alfred Hospital and Central Clinical School, Monash University, Melbourne, Victoria, Australia. <sup>6</sup>Infection and Immunity Theme, Monash Biomedicine Discovery Institute, Department of Microbiology, Monash University, Clayton, Victoria, Australia. <sup>7</sup>Infectious Diseases Department, Austin Health, Heidelberg, Victoria, Australia.

Received: 14 April 2018 Accepted: 24 July 2018

Published online: 31 August 2018

## References

- World Health Organization. Antimicrobial resistance: global report on surveillance 2014: World Health Organization; 2014. <http://www.who.int/drugresistance/documents/surveillance-report/en/>. Accessed 27 July 2018.
- Köser CU, Ellington MJ, Peacock SJ. Whole-genome sequencing to control antimicrobial resistance. *Trends in Genetics*. 2014;30(9):401–7.
- Schürch AC, van Schaik W. Challenges and opportunities for whole-genome sequencing based surveillance of antibiotic resistance. *Ann N Y Acad Sci*. 2017;1388:108–20.
- Van Belkum A, Dunne WM. Next generation antimicrobial susceptibility testing. *Journal of clinical microbiology*. 2013;JCM-00313. <https://doi.org/10.1128/JCM.00313-13>.
- McArthur AG, Waglechner N, Nizam F, Yan A, Azad MA, Baylay AJ, et al. The comprehensive antibiotic resistance database. *Antimicrob Agents Chemother*. 2013;57:3348–57.
- de Man, Tom JB, and Brandi M. Limbago. SSTAR, a stand-alone easy-to-use antimicrobial resistance gene predictor. *mSphere* 2016;1.1:e00050–15.
- Liu B, Pop M. ARDB - antibiotic resistance genes database. *Nucleic Acids Res*. 2009;37(Database issue):D443–7.
- Smith T, Wolff KA, Nguyen L. Molecular biology of drug resistance in *Mycobacterium tuberculosis*. *Curr Top Microbiol Immunol*. 2013;374:53–80.
- Feng J, Lupien A, Gingras H, Wasserscheid J, Dewar K, Légaré D, et al. Genome sequencing of linezolid-resistant *Streptococcus pneumoniae* mutants reveals novel mechanisms of resistance. *Genome Res*. 2009;19:1214–23.
- Livemore DM, Warner M, Jamroz D, Mushtaq S, Nichols WW, Mustafa N, et al. In vitro selection of ceftazidime-avibactam resistance in enterobacteriaceae with KPC-3 carbapenemase. *Antimicrob Agents Chemother*. 2015;59:5324–30.
- Chen CJ, Lin MH, Shu JC, Lu JJ. Reduced susceptibility to vancomycin in isogenic *Staphylococcus aureus* strains of sequence type 59: tracking evolution and identifying mutations by whole-genome sequencing. *J Antimicrob Chemother*. 2014;69:349–54.

12. Mwangi MM, Wu SW, Zhou Y, Sieradzki K, de Lencastre H, Richardson P, et al. Tracking the in vivo evolution of multidrug resistance in *Staphylococcus aureus* by whole-genome sequencing. *Proc Natl Acad Sci U S A*. 2007;104:9451–6.
13. Howden BP, Peleg AY, Stinear TP. The evolution of vancomycin intermediate *Staphylococcus aureus* (VISA) and heterogenous-VISA. *Infect Genet Evol*. 2014;21:575–82.
14. Barbosa C, Trebosch V, Kemmer C, Rosenstiel P, Beardmore R, Schulenburg H, Jansen G. Alternative evolutionary paths to bacterial antibiotic resistance cause distinct collateral effects. *Molecular biology and evolution*. 2017;34(9): 2229–44.
15. Howden BP, Davies JK, Johnson PDR, Stinear TP, Grayson ML. Reduced vancomycin susceptibility in *Staphylococcus aureus*, including vancomycin-intermediate and heterogeneous vancomycin-intermediate strains: Resistance mechanisms, laboratory detection, and clinical implications. *Clin Microbiol Rev*. 2010;23:99–139.
16. Händel N, Schuurmans JM, Feng Y, Brul S, Ter Kuile BH. Interaction between mutations and regulation of gene expression during development of de novo antibiotic resistance. *Antimicrob Agents Chemother*. 2014;58:4371–9.
17. Hershberg R. Antibiotic-independent adaptive effects of antibiotic resistance mutations. *Trends Genet*. 2017;33(8):521–8.
18. Rodriguez De Evgrafov M, Gumpert H, Munck C, Thomsen TT, Sommer MO. Collateral resistance and sensitivity modulate evolution of high-level resistance to drug combination treatment in *Staphylococcus aureus*. *Mol Biol Evol*. 2015;32:1175–85.
19. Jugheli L, Bzekalava N, de Rijk P, Fissette K, Portaels F, Rigouts L. High level of cross-resistance between kanamycin, amikacin, and capreomycin among *Mycobacterium tuberculosis* isolates from Georgia and a close relation with mutations in the *rrs* gene. *Antimicrob Agents Chemother*. 2009;53:5064–8.
20. Sacco E, Cortes M, Josseume N, Bouchier C, Dubée V, Hugonnet J-E, et al. Mutation landscape of acquired cross-resistance to glycopeptide and  $\beta$ -lactam antibiotics in *Enterococcus faecium*. *Antimicrob Agents Chemother*. 2015;59:5306–15.
21. Yu J, Wu J, Francis KP, Purchio TF, Kadurugamuwa JL. Monitoring in vivo fitness of rifampicin-resistant *Staphylococcus aureus* mutants in a mouse biofilm infection model. *J Antimicrob Chemother*. 2005;55:528–34.
22. Helms M, Simonsen J, Mølbak K. Quinolone resistance is associated with increased risk of invasive illness or death during infection with *Salmonella* serotype typhimurium. *J Infect Dis*. 2004;190:1652–4.
23. Smani Y, López-Rojas R, Domínguez-Herrera J, Docobo-Pérez F, Martí S, Vila J, et al. In vitro and in vivo reduced fitness and virulence in ciprofloxacin-resistant *Acinetobacter baumannii*. *Clin Microbiol Infect*. 2012;18(1):E1–4.
24. Beceiro A, Tomás M, Bou G. Antimicrobial resistance and virulence: a successful or deleterious association in the bacterial world? *Clin Microbiol Rev*. 2013;26:185–230.
25. Gao W, Cameron DR, Davies JK, Kostoulis X, Stepnell J, Tuck KL, et al. The RpoB H481Y rifampicin resistance mutation and an active stringent response reduce virulence and increase resistance to innate immune responses in *Staphylococcus aureus*. *J Infect Dis*. 2013;207:929–39.
26. Bæk KT, Thøgersen L, Mogensen RG, Møllergaard M, Thomsen LE, Petersen A, et al. Stepwise decrease in daptomycin susceptibility in clinical *Staphylococcus aureus* isolates associated with an initial mutation in *rpoB* and a compensatory inactivation of the *clpX* gene. *Antimicrob Agents Chemother*. 2015;59:6983–91.
27. Cameron DR, Howden BP, Peleg AY. The interface between antibiotic resistance and virulence in *Staphylococcus aureus* and its impact upon clinical outcomes. *Clin Infect Dis*. 2011;53:576–82.
28. Miskinyte M, Gordo I. Increased survival of antibiotic-resistant *Escherichia coli* inside macrophages. *Antimicrob Agents Chemother*. 2013;57:189–95.
29. O'Neill AJ, Cove JH, Chopra I. Mutation frequencies for resistance to fusidic acid and rifampicin in *Staphylococcus aureus*. *J Antimicrob Chemother*. 2001; 47:647–50.
30. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv*. 2013;0:3.
31. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence alignment/map format and SAMtools. *Bioinformatics*. 2009;25:2078–9.
32. Zhang J, Kobert K, Flouri T, Stamatakis A. PEAR: a fast and accurate Illumina paired-end reAd mergeR. *Bioinformatics*. 2014;30:614–20.
33. Sievers F, Higgins DG. Clustal Omega. *Curr Protoc Bioinforma*. 2014;48:3.13. 1–16.
34. Rice P, Longden I, Bleasby A. EMBOSS: the European molecular biology open software suite. *Trends Genet*. 2000;16:276–7.
35. Tange O. GNU parallel: the command-line power tool. *USENIX Mag*. 2011;36: 42–7.
36. Fu L, Niu B, Zhu Z, Wu S, Li W. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics*. 2012;28:3150–2.
37. Monk IR, Tree JJ, Howden BP, Stinear TP, Foster TJ. Complete bypass of restriction systems for major *Staphylococcus aureus* lineages. *MBio*. 2015;6: e00308–15.
38. Zhang Y, Werling U, Edelman W. SLICE: a novel bacterial cell extract-based DNA cloning method. *Nucleic Acids Res*. 2012;40(8):e55.
39. Love MI, Huber W, Anders S, Lönnerstedt I, Speed T, Robinson M, et al. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15:550.
40. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the royal statistical society. Series B (Methodological)*. 1995;1:289–300.
41. Ross BC, Raios K, Jackson K, Sievers A, Dwyer B. Differentiation of *Mycobacterium tuberculosis* strains by use of a nonradioactive southern blot hybridization method. *J Infect Dis*. 1991;163:904–7.
42. Faith JJ, Guruge JL, Charbonneau M, Subramanian S, Seedorf H, Goodman AL, et al. The long-term stability of the human gut microbiota. *Science*. 2013;341:1237439.
43. Kivioja T, Vähärautio A, Karlsson K, Bonke M, Enge M, Linnarsson S, et al. Counting absolute numbers of molecules using unique molecular identifiers. *Nat Methods*. 2011;9:72–4.
44. Schirmer M, Ijaz UZ, D'Amore R, Hall N, Sloan WT, Quince C. Insight into biases and sequencing errors for antimicrobial sequencing with the Illumina MiSeq platform. *Nucleic Acids Res*. 2015;43:e37.
45. Guérillot R, da Silva AG, Monk I, Giulieri S, Tomita T, Alison E, Porter J, Pidot S, Gao W, Peleg AY, Seemann T. Convergent Evolution Driven by Rifampin Exacerbates the Global Burden of Drug-Resistant *Staphylococcus aureus*. *mSphere*. 2018;3(1):e00550–17.
46. EUCAST. Antimicrobial susceptibility testing for bacteria: EUCAST; 2015. [http://www.eucast.org/ast\\_of\\_bacteria/](http://www.eucast.org/ast_of_bacteria/). Accessed 27 July 2018.
47. CLSI. Performance standards for antimicrobial susceptibility testing. CLSI supplement M100S. CLSI Suppl. M100S. Wayne: Clin. Lab. Stand. Inst; 2016.
48. Forrest GN, Tamura K. Rifampin combination therapy for nonmycobacterial infections. *Clin. Microbiol. Rev*. 2010;23(1):14–34.
49. Saleh-Mghir A, Muller-Serieys C, Dinh A, Massias L, Crémieux AC. Adjunctive rifampin is crucial to optimizing daptomycin efficacy against rabbit prosthetic joint infection due to methicillin-resistant *Staphylococcus aureus*. *Antimicrob Agents Chemother*. 2011;55:4589–93.
50. Garrigós C, Murillo O, Euba G, Verdaguer R, Tubau F, Cabellos C, et al. Efficacy of usual and high doses of daptomycin in combination with rifampin versus alternative therapies in experimental foreign-body infection by methicillin-resistant *Staphylococcus aureus*. *Antimicrob Agents Chemother*. 2010;54:5251–6.
51. Cui L, Isii T, Fukuda M, Ochiai T, Neoh HM, Da Cunha Camargo ILB, et al. An RpoB mutation confers dual heteroresistance to daptomycin and vancomycin in *Staphylococcus aureus*. *Antimicrob Agents Chemother*. 2010; 54:5222–33.
52. Aiba Y, Katayama Y, Hishinuma T, Murakami-Kuroda H, Cui L, Hiramatsu K. Mutation of RNA polymerase  $\beta$ -subunit gene promotes heterogeneous-to-homogeneous conversion of  $\beta$ -lactam resistance in methicillin-resistant *Staphylococcus aureus*. *Antimicrob Agents Chemother*. 2013;57:4861–71.
53. Reiber C, Senn O, Muller D, Kullak-Ublick G, Corti N. Therapeutic drug monitoring of daptomycin: a retrospective monocentric analysis. *Ther Drug Monit*. 2015;37:634–40.
54. Berti AD, Baines SL, Howden BP, Sakoulas G, Nizet V, Proctor RA, et al. Heterogeneity of genetic pathways toward daptomycin nonsusceptibility in *Staphylococcus aureus* determined by adjunctive antibiotics. *Antimicrob Agents Chemother*. 2015;59:2799–806.
55. Holmes NE, Turnidge JD, Munchhof WJ, Robinson JO, Korman TM, O'Sullivan MVN, et al. Antibiotic choice may not explain poorer outcomes in patients with *Staphylococcus aureus* bacteremia and high vancomycin minimum inhibitory concentrations. *J Infect Dis*. 2011;204:340–7.
56. Gao W, Chua K, Davies JK, Newton HJ, Seemann T, Harrison PF, et al. Two novel point mutations in clinical *Staphylococcus aureus* reduce linezolid susceptibility and switch on the stringent response to promote persistent infection. *PLoS Pathog*. 2010;6:e1000944.

57. Beceiro A, Tomás M, Bou G. Antimicrobial resistance and virulence: a beneficial relationship for the microbial world? *Enferm Infecc Microbiol Clin*. 2012;30:492–9.
58. McVicker G, Prajsnar TK, Williams A, Wagner NL, Boots M, Renshaw SA, et al. Clonal expansion during *Staphylococcus aureus* infection dynamics reveals the effect of antibiotic intervention. *PLoS Pathog*. 2014;10:e1003959.
59. Zetola NM, Shin SS, Tumedí KA, Moeti K, Ncube R, Nicol M, et al. Mixed *Mycobacterium tuberculosis* complex infections and false-negative results for rifampin resistance by genexpert MTB/RIF are associated with poor clinical outcomes. *J Clin Microbiol*. 2014;52:2422–9.
60. Meumann EM, Globan M, Fyfe JAM, Leslie D, Porter JL, Seemann T, et al. Genome sequence comparisons of serial multi-drug-resistant *Mycobacterium tuberculosis* isolates over 21 years of infection in a single patient. *Microb genomics*. 2015;1:e000037.
61. Donnabella V, Martiniuk F, Kinney D, Bacerdo M, Bonk S, Hanna B, et al. Isolation of the gene for the beta subunit of RNA polymerase from rifampicin-resistant *Mycobacterium tuberculosis* and identification of new mutations. *Am J Respir Cell Mol Biol*. 1994;11:639–43.
62. Ramaswamy SV, Dou S, Rendon A, Yang Z, Cave MD, Graviss EA. Genotypic analysis of multidrug-resistant *Mycobacterium tuberculosis* isolates from Monterrey, Mexico. *J Med Microbiol*. 2004;53:107–13.
63. Pozzi G, Meloni M, Iona E, Orrù G, Thoresen OF, Ricci ML, et al. *rpoB* Mutations in multidrug-resistant strains of *Mycobacterium tuberculosis* isolated in Italy. *J Clin Microbiol*. 1999;37:1197–9.
64. Pál C, Papp B, Lázár V. Collateral sensitivity of antibiotic-resistant microbes. *Trends Microbiol*. 2015;23:401–7.
65. Imamovic L, Sommer MOA. Use of collateral sensitivity networks to design drug cycling protocols that avoid resistance development. *Sci Transl Med*. 2013;5:204ra132.
66. Whale AS, Bushell CA, Grant PR, Cowen S, Gutierrez-Aguirre I, O'Sullivan DM, et al. Detection of rare drug resistance mutations by digital PCR in a human influenza A virus model system and clinical samples. *J Clin Microbiol*. 2016; 54:392–400.
67. Eilertson B, Maruri F, Blackman A, Herrera M, Samuels DC, Sterling TR. High proportion of heteroresistance in *gyrA* and *gyrB* in fluoroquinolone-resistant *Mycobacterium tuberculosis* clinical isolates. *Antimicrob Agents Chemother*. 2014;58:3270–5.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

