The rise and the fall of a *Pseudomonas aeruginosa* endemic lineage in a hospital

Marie Petitjean[1,2], Paulo Juarez[2], Alexandre Meunier[1], Etienne Daguindau[3], Hélène Puja[2], Xavier Bertrand[1,2], Benoit Valot[2,4]† and Didier Hocquet[1,2,4,5,*],†

Abstract

The biological features that allow a pathogen to survive in the hospital environment are mostly unknown. The extinction of bacterial epidemics in hospitals is mostly attributed to changes in medical practice, including infection control, but the role of bacterial adaptation has never been documented. We analysed a collection of *Pseudomonas aeruginosa* isolates belonging to the Besançon Epidemic Strain (BES), responsible for a 12year nosocomial outbreak, using a genotype-to-phenotype approach. Bayesian analysis estimated the emergence of the clone in the hospital 5 years before its opening, during the creation of its water distribution network made of copper. BES survived better than the reference strains PAO1 and PA14 in a copper solution due to a genomic island containing 13 metal-resistance genes and was specifically able to proliferate in the ubiquitous amoeba *Vermamoeba vermiformis*. Mutations affecting amino-acid metabolism, antibiotic resistance, lipopolysaccharide biosynthesis, and regulation were enriched during the spread of BES. Seven distinct regulatory mutations attenuated the overexpression of the genes encoding the efflux pump MexAB-OprM over time. The fitness of BES decreased over time in correlation with its genome size. Overall, the resistance to inhibitors and predators presumably aided the proliferation and propagation of BES in the plumbing system of the hospital. The pathogen further spread among patients via multiple routes of contamination. The decreased prevalence of patients infected by BES mirrored the parallel and convergent genomic evolution and reduction that affected bacterial fitness. Along with infection control measures, this may have participated in the extinction of BES in the hospital setting.

DATA SUMMARY

We confirm that all supporting data, code and protocols have been provided within the article or through Supplementary Material.

INTRODUCTION

The epidemic curves of bacterial pathogens in hospitals are mostly wave-shaped, with a rapid increase in cases followed by a slower decline [1]. During pathogen outbreaks in hospital settings, patients are infected directly or indirectly, either from the environment or from other patients. However, the exact source of the outbreak (e.g. a patient, a healthcare worker, the environment) is almost never documented and often limited to the identification of an index case due to the lag between contamination of the healthcare setting and detection of the pathogen in the index case [2, 3]. Aside from antibiotic resistance, the biological features that allow a pathogen to enter and survive in the aggressive hospital environment are mostly unknown. However, genomic analyses are able to estimate the origin of an outbreak by calculating the time of emergence of the most recent common ancestor and the genetic features that could have favoured its expansion [4, 5].

Whole-genome sequencing analysis has already been used to identify genetic elements gained, lost, or mutated during outbreaks of bacterial pathogens, but the phenotypic consequences of such mutations are rarely determined [6–8]. Other studies have illustrated the difficulty of phenotypically confirming *in silico* predictions [9]. Although the extinction of epidemics in hospitals is mostly attributed to infection control or more widely to changes in medical practice [10], the role of bacterial adaptation has never been considered.

We addressed this question by exploring the genotypic and phenotypic variations of an epidemic strain of *P. aeruginosa* that infected more than 250 patients over 12 years in a single hospital [2]. *P. aeruginosa* is a ubiquitous Gram-negative species that exhibits such genomic plasticity that it allows its adaptation to a variety of niches, from the wild environment to that of disinfected hospitals and patients treated with antibiotics [11]. This opportunistic pathogen is a leading cause of nosocomial infections, as well as chronic lung infections in cystic fibrosis (CF) patients [12]. In addition to its high intrinsic resistance to a wide variety of antibiotics, *P. aeruginosa* can readily acquire resistance through mutations or horizontal transfer of resistance determinants. Thus, resistance to all currently available anti-pseudomonal antibiotics is frequent [13], justifying its inclusion in the list of serious threats by the CDC [14]. *P. aeruginosa* is frequently responsible for outbreaks in the hospital setting, in which the source of contamination is mostly unknown or speculated, at best [3]. Water points-of-use have often been incriminated in *P. aeruginosa* hospital outbreaks. However, sinks are themselves seeded by clones of human origin, blurring the tracks back to the source of hospital contamination [15].

We used WGS analysis and a genotype-to-phenotype approach of such a long-term outbreak of *P. aeruginosa* as a model to (*i*) date the true origin of the contamination of the hospital, (*ii*) identify the biological features that favoured its installation, (*iii*) identify parallel and convergent evolution selected during the outbreak, and (*iv*) correlate the fitness of the strain with its genomic evolution.

METHODS

Hospital setting and isolate collection

A prolonged outbreak of *P. aeruginosa* in the University Hospital of Besançon, a 1300-bed teaching hospital in France, was investigated. The construction of its facilities began in 1978 and the main building was commissioned in 1983, with all pipes of the drinking-water installation being made of copper. All *P. aeruginosa* isolates retrieved from patients hospitalized between May 1997 and April 2008 were typed by pulsed field gel electrophoresis (PFGE) to assess the number of infected patients. Patients with *P. aeruginosa* isolates for which the pulsotype shared the same band profile with the index case (*n*=276) were included in the cohort [2]. None of them suffered from cystic fibrosis. Then, 54 isolates retrieved from 54 patients (BES-1 to BES-54) were selected for sequencing. They were evenly distributed over time (one per quarter) and considered to be representative of the outbreak (Table S1). The first 27 isolates were recovered in

Impact Statement

Bacterial pathogens are responsible for nosocomial outbreaks, but the sources of hospital contamination are mostly unclear and the role of bacterial evolution in the extinction of outbreaks has never been considered. Here, we found that an epidemic strain of the pathogen *Pseudomonas aeruginosa* contaminated the drinking water network of a hospital due to its tolerance to copper and predatory amoeba, both present in the water pipes. The extinction of the outbreak occurred concomitantly with parallel and convergent genome evolution and a reduction in the size of the bacterial genome that correlated with the fitness of the pathogen. Our data suggest that pathogen evolution participated in the extinction of an outbreak in a hospital setting.

the early stages of the outbreak (before January 2001) and the other 27 isolates were considered to be late isolates (Table S1).

DNA sequencing and analysis

Bacterial DNA was isolated from overnight cultures on Mueller-Hinton agar using the Genomic-tip kit (Qiagen) and paired-end sequenced with Illumina NextSeq, at 2×150 bp. Reads were subsampled before assembly to limit the coverage to 80×. The contigs were built with SPAdes [16]. Local and external outgroup isolates of *P. aeruginosa* of the same sequence type (ST) were used to root the phylogenetic tree (Table S1) [3]. We previously sequenced BES-1, the first clinical isolate of BES identified in May 1997, using PacBio technology and used it as a reference [17]. STs were determined *in silico* using the PubMLST database (https://pubmlst.org/) and the homemade pipeline pyMLST (https://github.com/bvalot/pyMLST). The BES pangenome was built by combining the BES-1 circular genome with non-redundant contigs from the other BES isolates using the Ragout tool [18]. For this purpose, the assemblies of each BES isolate were aligned against the BES-1 genome using Mummer [19]. Genes of the BES pangenome were detected using Prodigal and then annotated with BLAST [20, 21]. Resistance and virulence genes were sought using ResFinder and the Virulence Factor DataBase, respectively [22, 23]. The reads from all isolates were then aligned against the BES pangenome using BWA and the alignments used for further analysis (https://github.com/lh3/bwa). The presence or absence of each gene was tested using FeatureCounts [24]. SNPs were identified using freebayes with a minimum coverage of ten reads per isolate and a minimum quality of 30 (https://github.com/ekg/freebayes). Non-synonymous SNPs and SNPs in promoter regions (RBS and promoter sequences at −10 and −35) were further sorted. SNPs and gained or lost gene blocks were clustered in relation to the phylogenetic tree with MixOmics [25]. IS insertion events were detected using panISa software [26]. The boundaries of each potential IS were sought in

the BES-1 genome to exclude chromosomal rearrangement before searching for them in the BES genomes to assess the length of the potential ISs. The presence of the potential IS was confirmed by PCR and sequencing with specific IS1 and IS2 primers (Table S2).

Phylogeny and most recent common ancestor (MRCA) calculation

The 265 core genes of the BES isolates that harboured at least one SNP were concatenated for phylogenetic analysis, producing a 397062 bp alignment with 285 SNPs. MrBayes software was used to build the maximum clade credibility tree [27]. We evaluated the molecular clock signal present on data with tempEST [28] and found a correlation of 0.72 between time and root-to-tip divergence. Therefore, the isolation dates of the BES and outgroup isolates were used as calibration nodes to calculate the age of the MRCA.

RT-qPCR experiments

mexB expression was assessed in all BES isolates by RT-qPCR as previously described [29] with specific primers for the housekeeping gene *rpsL*, and *mexB* (Table S2). After each assay, a dissociation curve confirmed the specificity of all PCR amplicons. The mRNA levels of *mexB* were normalized to that of the reference gene *rpsL* and expressed as a ratio to the levels in the reference strain *P. aeruginosa* PAO1. Based on previous results, all BES isolates with at least two-fold higher *mexB* expression than that in PAO1 were considered to be MexAB-OprM overproducers [29].

Phenotype determination

Susceptibility to antipseudomonal compounds was tested using agar dilution, as recommended by the CLSI [30]. We determined the serotype of all isolates with specific antiserum (Bio-Rad). Metabolic changes in BES isolates were assessed using procedures detailed in Fig. S2.

Survival of the BES-4 isolate and its ΔGI-7 mutant in a copper solution

For all tested *P. aeruginosa* isolates, an overnight culture in Luria-Bertani broth (LBB, Bio-Rad) was diluted 1/200 into fresh LBB and further incubated at 37 °C with gentle shaking until mid-log phase. The bacterial suspension was centrifuged and the pellet was washed once and resuspended in sterile deionized water. A 150 μg l^{-1} copper (II) sulphate solution (Sigma-Aldrich) was inoculated with approximately 10^8 bacteria before incubation for 24 h at 22 °C in the dark. The concentration of 150 μg l^{-1} was chosen as it was the highest value retrieved in the drinking water of our hospital [31]. Initial and final bacterial concentrations were determined by plating appropriate dilutions on LB agar (Bio-Rad) and incubating overnight at 37 °C. The ratios of the initial and final bacterial concentrations determined the survival rates. We repeated each experiment six times independently for all isolates. The reference strains of *P. aeruginosa* PAO1 (ST549), PA14 (ST235), and a clinical isolate of ST235 (strain PA1646) [31] were used as controls.

Survival of *P. aeruginosa* in free-living amoebae

The survival and growth of the BES strain was evaluated in the free-living amoeba *Vermamoeba vermiformis* 172A, isolated from the drinking water installation of the Lausanne Hospital (Switzerland) [32]. The protozoans were grown in PYNFH medium (LGC standards, Molsheim, France) as monolayers at 25 °C before their harvest. *P. aeruginosa* strains were cultured on Mueller–Hinton agar (MHA) for 24 h at 37 °C, harvested, and then suspended in PAS (Page's amoeba salines). Briefly, 1.5×10^5 amoebae were incubated with *P. aeruginosa* at a MOI of 5 (five times more bacteria) for 4 h at 25 °C. Non-internalized bacteria were then killed by a 2 h incubation with 0.3 mg ml^{-1} amikacin. The co-cultures were washed three times with PAS to remove the antibiotic and further incubated in PYNFH for 24 h at 25 °C. We then lysed the amoebae with 0.4% Triton X100 for 30 min at 25 °C. Appropriate dilutions of the initial and final bacterial inoculum were plated on MHA plates to calculate the proportion of surviving bacterial cells.

Chromosomal deletion of the genomic island GI-7

The GI-7 region was deleted from isolate BES-4 using overlapping PCR and homologous recombination detailed elsewhere [31]. The deletion of GI-7 in BES-4 isolate was confirmed by PCR and sequencing.

BES coculture

Mid-log phase BES isolates were co-cultured in LBB at a 1:1 ratio against the mutant BES-4ΔGI-7 as a reference. We extracted and purified the DNA (QIAamp DNA Mini Kit, Qiagen) from 48 h co-cultures. DNA concentrations were determined using QUBIT 2.0 and adjusted to 0.1 ng μl^{-1}. The *mutL* (present in all isolates) and *copB* (specifically absent from BES-4ΔGI-7) genes were quantified using specific primers and Taqman probes against standard curves (Table S2). qPCR was performed on a 7500 Fast Real Time PCR system and the results analysed using 7500 software v 2.3 (Applied Biosystems). The competition index (CI, proportion of the tested BES isolate in the co-culture after 48 h of incubation) was calculated as the ratio of the concentrations of *copB* and *mutL*. In other words, a CI of 0.5 was neutral. The CI of BES-4 against BES-4ΔGI-7 was 0.53±0.01, indicating that the deletion of GI-7 did not alter bacterial fitness.

Data availability

All genomic data were deposited in NCBI under the project accession number PRJNA399056. Biosample details are given in the Table S1.

RESULTS AND DISCUSSION

Population structure and most recent common ancestor of *P. aeruginosa* BES

Between May 1997 and April 2008, 276 patients were infected with the *P. aeruginosa* BES in the University Hospital of
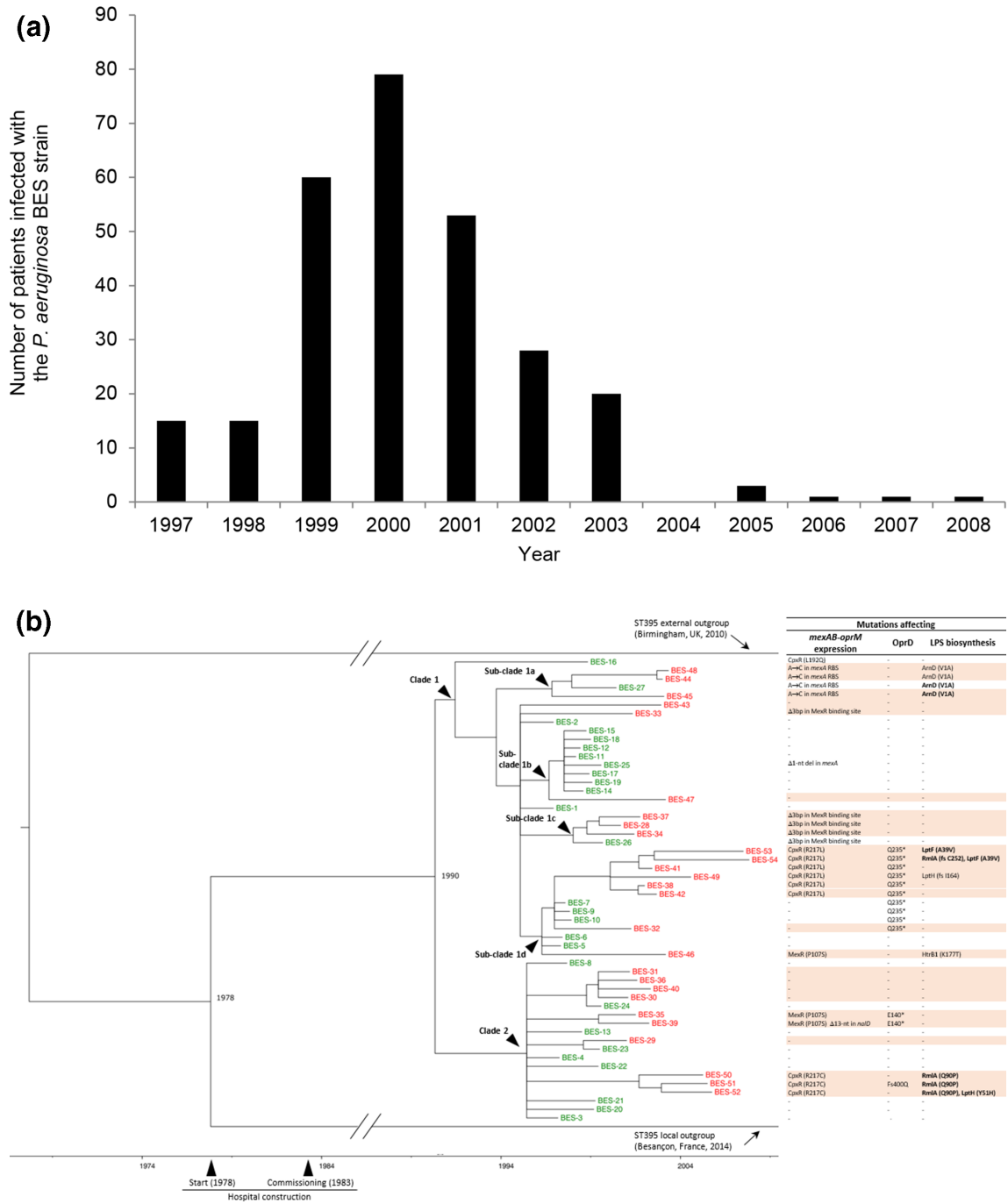
Fig. 1. Temporal data and phylogenetic tree of the *P. aeruginosa* Besançon Epidemic Strain (BES) in a University Hospital. (a) Number of new patients contaminated by or infected with *P. aeruginosa* BES (belonging to the widespread clone ST395) by year during the 12 years of the outbreak in the University Hospital of Besançon (France). (b) The left panel shows the tree built with MrBayes using the following parameters: GTR substitution model and 1000000 cycles with the MCMC algorithm, with sampling of the chain every 1000 cycles. We defined two major clades, 1 and 2, with clade 1 divided into four sub-clades, 1a to 1d. The isolates in green and red are early (before January 2001) and late (after January 2001) isolates, respectively. Isolates in black were outgroups. The right panel indicates significant mutations that affected the expression of *mexAB-oprM*, OprD, or lipopolysaccharide (LPS) biosynthesis. Fs, frameshift. Mutations shown in bold indicate non-O:6 isolates.

Besançon (Fig. 1a). We sequenced the full genomes of 54 isolates (retrieved from 54 patients) evenly distributed over time and considered to be representative of the outbreak (Table S1, available in the online version of this article). *In silico* MLST showed the sequenced isolates to belong to the widespread clone ST395 [33]. Plasmids were absent from the all genomes. Phylogenetic analysis using 285 SNPs in the core genome showed that BES isolates fell into the two major clades, 1 and 2, with clade 1 being divided into four sub-clades, 1a to 1d (Fig. 1b). There was no clear link between the date of isolation or hospital ward and position in the phylogenetic tree (Table S1). Although the pulsed-field gel electrophoresis pattern of BES isolates was the same [2], WGS analysis showed two hidden parallel outbreaks simultaneously occurring in the same hospital (Fig. 1b), as already observed for other Gram-negative pathogens [34]. The dating of emergence of the clone ST395 at the hospital required the calculation of the clock rate of our ST395 population, that was estimated to be 3.96×10^{-6} mutations per site per year by Bayesian analysis (95% highest posterior density, 3.08×10^{-6}–4.96×10^{-6}), which is higher than that estimated for the *P. aeruginosa* species [35]. With the addition of an ST395 outgroup isolate retrieved in our hospital, but not belonging to BES, to the database (Table S1), Bayesian analysis dated the putative common ancestor that linked the BES ST395 isolates from our hospital to 1990 (95% confidence interval [CI]=1987–1993), 7 years before its first isolation from a patient in May 1997. Furthermore, using an external ST395 outgroup isolated from the hospital of Birmingham (United Kingdom), we estimated the emergence of clone ST395 at the hospital during its construction in 1978 (95% CI=1970–1985) (Fig. 1b) during the creation of the water distribution network. The introduction of ST395 during the construction of a hospital was already suspected in the United Kingdom [3]. This is consistent with the frequent isolation of the ST395 clone in water [36, 37] and the probable role of drinking water as a source of *P. aeruginosa* infection [38, 39]. Overall, the ST395 clone has presumably contaminated the water network of the hospital before its opening but the pathogen further spread among patients by other routes of transmission (e.g. contact with the environment, other patients, or healthcare workers).

BES *P. aeruginosa* is tolerant to copper

The date of emergence of the common ancestor of ST395 (1978) led us to suspect the water distribution system of our hospital as the source of the outbreak. The water distribution and plumbing systems of our hospital are made of copper and are hostile environments for bacterial development, with nutrient-poor conditions, inhibitors (e.g. chlorine and copper-ions), and predators (free-living amoeba). We therefore postulated that BES could cope with those harsh conditions. Analysis of the BES genome sequences identified a 37-kb genome island (GI) that harboured an array of thirteen genes encoding copper transporters (Fig. S1) in all isolates of BES [17]. This GI shared 99.7% identity with GI-7 of a *P. aeruginosa* ST308 copper-tolerant

strain [31], also retrieved from water networks of hospitals. Our previous work showed that GI-7 likely emerged in environmental *Pseudomonas* species before its transfer to the *P. aeruginosa*. This GI belongs to the core genome of ST395, ST308, and ST309, and has been found sporadically in ST111 and ST253 isolates [31]. We phenotypically confirmed that the representative isolate BES-4 survived better than the *P. aeruginosa* reference strains PAO1 and PA14 (Fig. 2a) in a 150 µg l⁻¹ copper (II) sulphate solution, correponding to the highest copper concentration retrieved in the drinking water of our hospital [31]. We then tested the mutant strain BES-4ΔGI-7, deleted of the entire sequence of GI-7 by chromosomal recombination, under the same conditions to confirm the role of GI-7 in the tolerance of BES-4 to copper in solution. The deletion of GI-7 decreased the tolerance to copper of the BES strain relative to its wild-type parent, reducing it to that of the comparators PAO1 and PA14 (Fig. 2a). However, the tolerance to copper-ions under these conditions was not fully specific to GI-7, as the representative of the high-risk clone ST235 (isolate PA1646, without GI-7) showed copper tolerance similar to that of BES-4 (Fig. 2a). However, no ST235-specific determinants for tolerance to copper-ions have been identified [40]. Interestingly, a sampling campaign in 2017 found that the water distribution network of the University Hospital of Besançon was contaminated by a GI-7 containing ST308 clone [31]. Overall, these results confirm that GI-7 can help waterborne *P. aeruginosa* (ST308, ST395) to tolerate copper in solution.

BES *P. aeruginosa* proliferates in free-living amoebae

Protozoans are predators for bacteria in water [41]. *Vermamoeba (Hartmannella) vermiformis* is ubiquitous in drinking-water systems and resistant to commonly used disinfection methods [42]. It has already been shown that a strain of *P. aeruginosa* isolated from the water of a hospital can resist *V. vermiformis* grazing [43]. We thus tested the survival of BES-4 in the reference strain of *V. vermiformis* 172A versus that of the same controls used to test copper tolerance to further investigate whether this mechanism may contribute to the resistance of *P. aeruginosa* BES in the drinking-water network. BES better resisted the grazing of this protozoan than the comparators and was even able to proliferate inside the amoeba (Fig. 2b). Copper resistance has been implicated in the tolerance to amoeba, but the deletion of GI-7 had no effect on grazing resistance [44] (Fig. 2b). The resistance mechanism of BES to protozoan grazing remains to be identified. Hence, protozoans may have been responsible for the proliferation and propagation of BES in the water system and, as already demonstrated for *Legionella pneumophila*, the intracellular environment protected the bacterial pathogen from chlorine disinfection [45]. However, such resistance to protozoan grazing could have broader consequences. Pathogens use related strategies for intracellular survival in both amoeba and human macrophages, which in turn share similar
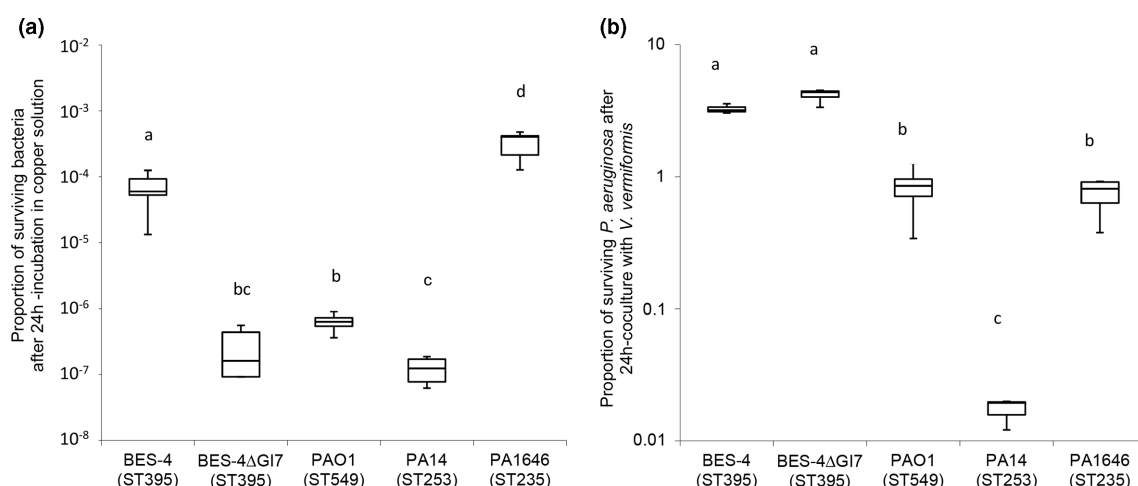
Fig. 2. *P. aeruginosa* BES is tolerant to copper-ions and proliferates in the free-living amoebae *Vermamoeba vermiformis*. (a) Survival of *P. aeruginosa* BES-4, with or without GI-7, in a 150 µg l⁻¹ copper (II) sulphate solution after 24 h of incubation shown on a decimal logarithm scale. Values are shown as the mean ratio ±SD (≥ 5 independent experiments). PCR experiments confirmed that the genomes of PAO1, PA14, and PA1646 (ST235) did not harbour GI-7. (b) The *P. aeruginosa* BES (ST395) proliferates in the free-living amoebae *Vermamoeba vermiformis*. The reference strain *V. vermiformis* 172A was inoculated with *P. aeruginosa* strains at a multiplicity of infection (MOI) of 5 and incubated for 24 h at 25 °C. Values are shown as the mean ratio ±SD (three technical replicates and ≥3 biological replicates). Statistical analysis was performed by ANOVA. Letters depict statistical dissimilarity using Tukey HSD (honestly significant difference) post hoc test at *P* < 0.01. For the two panels, y-axis represent the ratio of the quantification on LB agar plates of the living cells at the end of the experiments and that of the initial inoculum.

phagocytic mechanisms. Consequently, the coincidental evolution hypothesis suggests that virulence factors arose as a response to other selective pressures, such as predation, rather than for virulence *per se* [46].

Overall, the resistance to an inhibitor (copper) and a predator (free-living amoeba) presumably aided the proliferation and propagation of *P. aeruginosa* BES in the plumbing system during the building of the hospital, 5 years before its opening. This finding stresses the need for strict adherence to the procedures of commissioning water systems in healthcare premises [47].

Biological functions altered during the 12-year spread of *P. aeruginosa* BES

After having identified the likely source of contamination of the hospital building and the associated biological features, we explored the evolution of BES during its spread. During their interaction with humans, bacterial pathogens can accumulate mutations, allowing them to better adapt to the host, evade the immune response, or increase their resistance to antibiotics [48]. Our collection covered the 12-year epidemic and enabled the identification of evolution of the same strain in multiple individuals. We then searched for biological features selected or lost during the outbreak. Variant calling retrieved 458 SNPs in the pangenome of the BES population, amongst which 220 were non-synonymous or present in the promoting regions of genes (Table S3). Among them, 68 mutations affected 55 genes encoding proteins for which the function was known. We compared the proportion of altered functions in early

and late isolates and identified four major functions that were affected (χ^2 test, *P* < 0.05): amino-acid metabolism, antibiotic resistance, lipopolysaccharide (LPS) biosynthesis, and regulation.

Although late isolates had mutations in genes involved in amino-acid metabolism, their growth in minimal media was unaffected (Fig. S2).

The biosynthetic pathway of LPS is encoded by several genes responsible for the biosynthesis of lipid A, O-antigen, and transport towards the outer membrane [49]. Eleven isolates, of which ten were late isolates, displayed parallel evolution affecting genes known to be involved in LPS biosynthesis, either in O-antigen assembly (*rmlA*), LPS export (*lptF*, *lptH*), or lipid A biosynthesis (*htrB1*, *arnD*) (Fig. 1b). Serotyping confirmed that most (7 of 11) of these isolates were polyagglutinable, confirming an altered LPS. This contrasted with all the other isolates of BES, for which serotype O:6 was found (Fig. 1b). LPS structures can be modified in response to environmental pressures (e.g. divalent cations, antimicrobial peptides) and surface interactions. Modification of the LPS of *P. aeruginosa* is an important factor in the adaptation of this pathogen to chronic infection. Isolates BES-27,–44, -45,–46, and −48 harboured mutations that could have altered the lipid A biosynthesis (Fig. 1b). Reduced LPS immunostimulatory potential, due to lipid A modification, may contribute to evasion of the immune system and thus the survival of the pathogen over the course of a chronic infection [50]. However, LPS extracted from all the isolates of BES
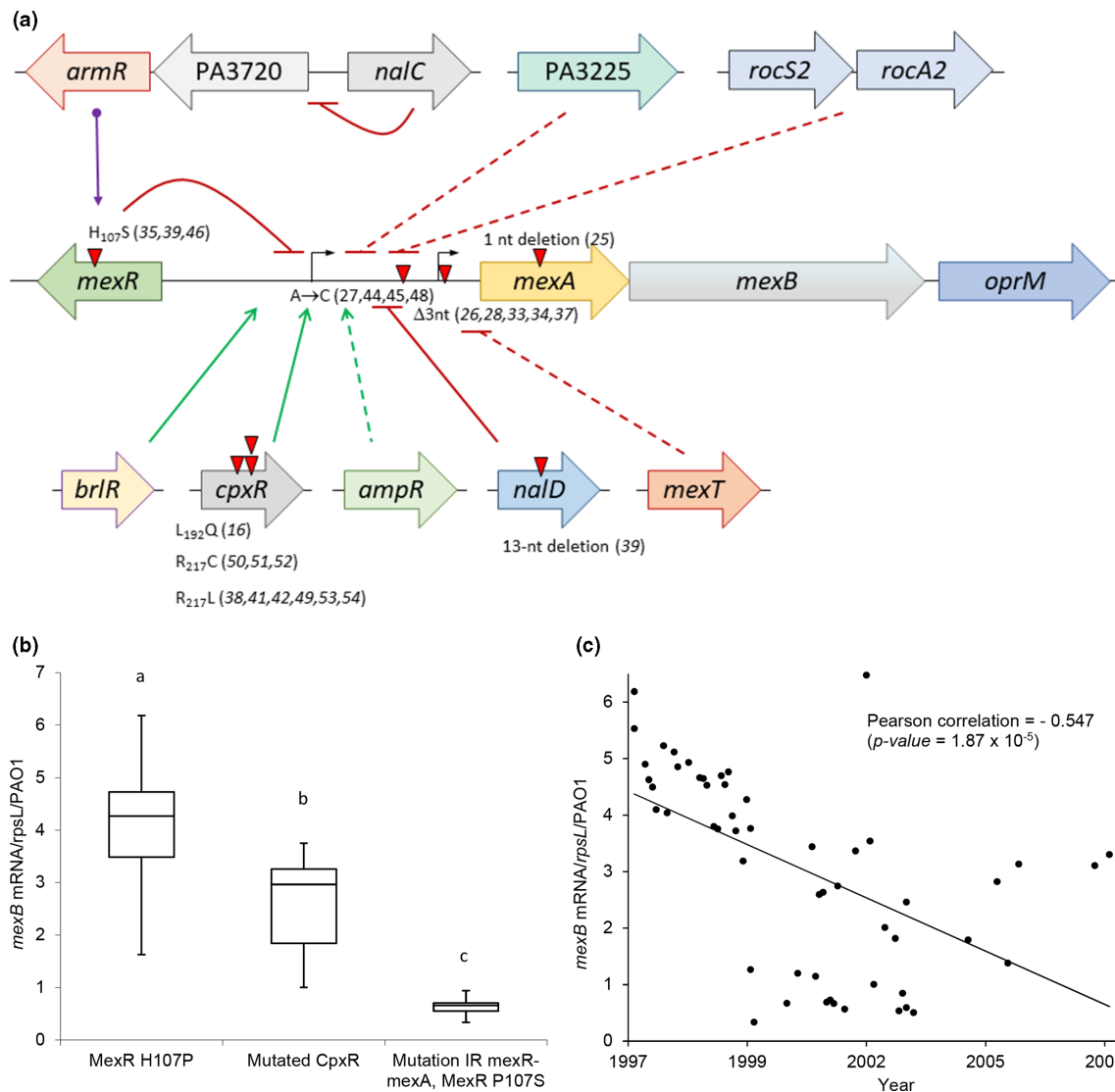
Fig. 3. *P. aeruginosa* BES attenuates the overexpression of the RND efflux pump MexAB-OprM during its spread via parallel and convergent evolution. (a) Mutations affecting *mexAB-oprM* expression in *P. aeruginosa* BES in the regulatory network of *mexAB-oprM*. Local repressor MexR ensures basal expression of the *mexAB-oprM* operon. The repressor NalD directly binds to the proximal promoter of the operon. NalC indirectly modulates *mexAB-oprM* expression by repressing the expression of the *armR* gene, the product of which acts as an anti-repressor of MexR. Mutations in these three repressors have been reported in *nalB*, *nalD*, and *nalC* mutants, respectively. Two direct activators of *mexAB-oprM* are CpxR and BrlR. Finally, the products of four other genes have been reported to have an indirect negative (*mexT*, *rocS1/2-A2*, *PA3225*) or positive (*ampR*) impact on *mexAB-oprM* expression (64). Red triangles indicate mutations occurring in regions involved in the regulation of *mexAB-OprM* expression, with the number of the affected BES isolates in parenthesis. (b). Expression of *mexB* in isolates of BES sorted by the type of mutation in the regulatory pathway of *mexAB-oprM*. ANOVA with a Tukey HSD test, with a/b, a/c, and b/c being significantly different; $P < 0.001$. (c) Attenuation of *mexAB-oprM* overexpression in *P. aeruginosa* BES over time.

stimulated human peripheral blood mononuclear cells with the same intensity (Fig. S3).

Antibiotic resistance of *P. aeruginosa* BES

Early isolates of BES were classified as multidrug-resistant [51], with decreased susceptibility to aminoglycosides (gentamicin, tobramycin), fluoroquinolones (ciprofloxacin), and β-lactam compounds [52]. Such decreased susceptibility may be explained by chromosomal mutations in resistance-associated genes. Indeed, β-lactam resistance can be explained by observed mutations in the regulating genes *ampR* (M1L in AmpR) and *ampD* (A134V and T139M in AmpD), responsible for increased production of the cephalosporinase AmpC and alterations in *mexR* (H107P in MexR), a regulator of the multi-drug efflux pump MexAB–OprM. All isolates also displayed canonical mutations in the quinolone-resistance determining regions
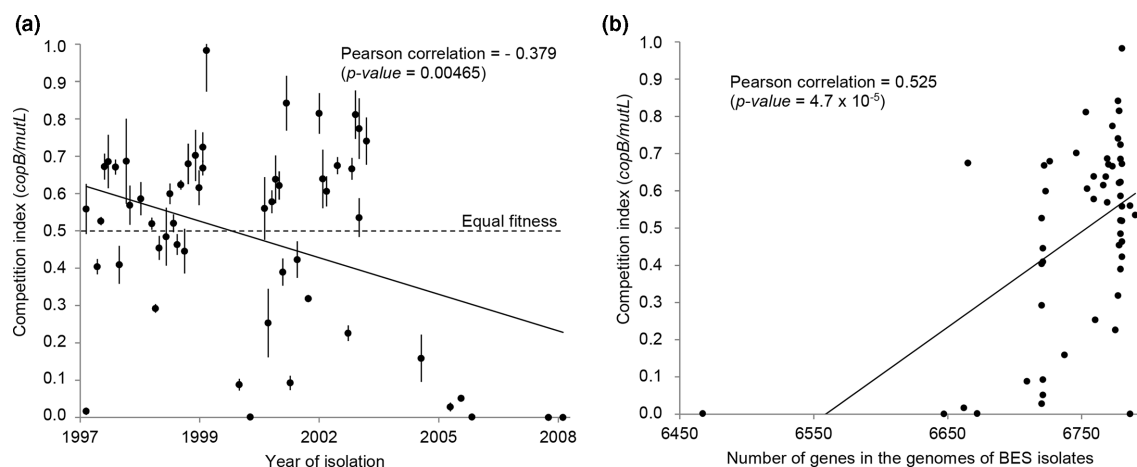
Fig. 4. The fitness of *P. aeruginosa* BES decreases in relation to its genome size during a hospital-associated outbreak. (a) *P. aeruginosa* isolates of BES were cocultured in Luria-Bertani broth with the reference isolate *P. aeruginosa* BES-4ΔGI-7. After 48 h of incubation, we determined the *copB/mutL* ratio, which represents the proportion of the tested isolate of BES (harbouring *copB* and *mutL*) in the co-culture relative to that of the isolate BES-4ΔGI-7 (harbouring only *mutL*) and used it to calculate the competition index (CI), as described in Methods. Data are shown as the mean±SD of three biological replicates. The CI of each BES isolate is plotted against its time of isolation. The dotted line (CI=0.5) represents an equal fitness. (b) The fitness of *P. aeruginosa* BES correlates with the size of the genome.

of GyrA (T83I) and ParC (S87L), explaining the increased resistance to fluoroquinolones. In addition, all BES isolates harboured *ant(2′)-Ia*, coding for a 2′-aminoglycoside nucleotidyl-transferase that is involved in aminoglycoside resistance [13, 53]. BES did not acquire foreign resistance determinants during the outbreak. Sequence analysis of three early and ten late isolates showed three different mutations in *oprD* (Fig. 1b) that correlated with higher resistance to imipenem (Welch Two Sample Test, $P=2.5 \times 10^{-5}$). Resistance to carbapenem via OprD mutation, certainly favoured by antibiotic treatment, has been shown to enhance fitness *in vivo* and to facilitate mucosal colonization and dissemination to the spleen [54].

Modulation of *mexAB-OprM* overexpression

The resistance-nodulation-division efflux pump MexAB–OprM is expressed at basal levels by wild-type cells and regulated by a complex regulatory network of at least ten direct or indirect regulators [55] (Fig. 3a). Early isolates displayed an alteration of the DNA-binding domain of the repressor protein MexR (H107P, substituting the polar histidine with the non-polar proline). This resulted in the overexpression of *mexB* relative to that of the PAO1 reference strain (Fig. 3b) and contributed to enhanced resistance to the MexAB-OprM substrates aztreonam and ciprofloxacin (Fig. S4) [29]. Other mutations in MexAB-OprM-regulating regions arose independently during the spread of BES and were found in 19 of the 27 late isolates. Hence, 12 isolates of BES carried non-synonymous mutations in *mexR*, *cpxR*, or *nalD*, and seven had mutations in the promoter region of the *mexAB-oprM* operon (Fig. 3a). We verified the impact of these mutations on the expression of the efflux pump by measuring *mexB* expression in all

isolates of BES by RT-qPCR. Almost all the isolates (29 of 31, 94%) that had the same mutations in the *mexAB-oprM* regulation pathway as BES-1 overexpressed *mexB* (median value, 4.3; Fig. 3b). Ten late isolates with a H107P MexR background had undergone three types of mutations in the regulator CpxR (Fig. 3a), leading to an intermediate level of *mexB* expression (median value, 3.0; Fig. 3b). Moreover, 11 isolates harboured mutations in the *mexA-mexR* intergenic region (MexR binding site, and *mexA* RBS region) or had a P107S substitution in MexR (returning a polar amino acid – serine – to position 107), expressed *mexB* at a level similar to that of the reference strain PAO1 (median value, 0.7; Fig. 3b). Efflux expression, as measured by *mexB* expression, also decreased over time (Fig. 3c).

Overall, seven distinct regulatory mutations attenuated the initial *mexAB-oprM* overexpression of BES (Fig. 3a). MexAB-OprM plays a role in antibiotic resistance, but can also export virulence determinants that allow *P. aeruginosa* to be invasive and cause infection [53]. The attenuation of *mexAB-oprM* overexpression has been shown to reduce the invasiveness of the bacteria in a cell model [56]. It is thus possible that convergent evolution decreasing the overproduction of this pump contributed to the persistence of BES in hospitalized patients. Moreover, the overexpression of *mexAB-oprM* reduces the survival of *P. aeruginosa* in the environment [57]. As expected, attenuation of the overexpression of *mexAB-oprM* reduced the level of resistance to aztreonam and ciprofloxacin (Fig. S4). The regulatory mutations in BES contrasted with the alterations in *mexAB-oprM* genes that are selected in *P. aeruginosa* populations in CF airways, leading to hypersensitivity to antibiotics [58]. A functional MexAB-OprM pump is necessary for the full virulence of *P. aeruginosa* in cell models [59]. However,

its overproduction has no effect on the invasiveness of the pathogen, advocating against the enhanced survival of MexAB-OprM-overproducing mutants in protozoans [56]. This fine-tuning of *mexAB-oprM* expression possibly allows a trade-off between antibiotic resistance, virulence, and survival in the environment. Such efflux modulation may then be important for survival and persistence of the clone in hospital setting.

Insertion or deletion of gene blocks during the spread of BES

P. aeruginosa can evolve through the acquisition or loss of genome fragments [6]. Accessory gene profiling showed clear differences in accessory gene content among isolates, with gene blocks specific to certain clusters (Table S4). For example, isolates BES-1 and BES-28, which were retrieved 4 years apart, independently underwent a ca. 130-kb deletion (as BES-26 did not), including genes encoding quorum-sensing regulators (LasI, LasR, and RsaL) and the flagellar operon region *fli* [17]. The late isolate BES-52 lost a 278 kb fragment that encompassed *hmgA*, of which deletion leads to an easily recognisable brown colony phenotype due to homogentisate accumulation. These melanogenic mutants are probably adapted to life in diverse bacterial communities in chronically-infected patients [60]. Despite the relatively high plasticity observed throughout the outbreak, the decrease over time in the number of genes per genome was not statistically significant (Spearman correlation test, $P=0.064$; Fig. S5). The epidemiology of *P. aeruginosa* BES at our hospital (Fig. 1a) supports the idea of emergence and extinction over time. As genome reduction can alter the bacterial growth rate [61], we evaluated the fitness of each BES isolate by coculturing them with an early representant of BES and measuring by qPCR the proportion of each competitor after 48 h of incubation. Coculture was preferred since it permits detection of fitness differences that are ≤ 1% per generation, undetectable through the assessment of growth rates in single culture [62]. Indeed, the fitness of BES decreased over time (Fig. 4a) and correlated with the size of the genome (Fig. 4b). More specifically, non-O:6 isolates ($n=7$) showed lower fitness in coculture than O:6 isolates (Welch two-sample test, $P=0.012$).

Limitations and strengths

This study had several limitations. Although we selected randomly distributed isolates throughout the period of the outbreak, the selection and inclusion criteria were not systematic and dependent on the availability of the isolates. Thus, the data presented here may be affected by a sampling bias that could have obscured epidemiological and evolutionary signals. This retrospective study did not allow water sampling at the time of emergence of the clone in the hospital (ca. 1978) that could have showed water contamination with BES and *V. vermiformis*. However, others have showed that *V. vermiformis* is endemic in hospital plumbing systems [63].

Here, we made use of a rare and valuable resource (i.e. longitudinally collected isolates from a variety of infection types in a single hospital from a period of more than a decade). While adaptation of human pathogens within a single niche has been described for evolution *in vitro*, very little research has focused on bacterial adaptive processes within the course of an outbreak. We attempted to validate *in silico* predictions by wet lab experiments. Hence, our genotype-to-phenotype approach correlated tolerance to copper, resistance to antibiotics, modulation of the expression of efflux pump, and bacterial fitness with genome content. However, sequence analysis predicted auxotrophy for tryptophan and impaired growth in histidine and asparagine, which were not confirmed *in vitro*. This highlights the need to control the functional effects of the mutations observed in WGS data.

Conclusions

This whole-genome phylogenetic analysis of a *P. aeruginosa* outbreak suggests contamination of the water distribution system before the opening of the hospital with the widespread clone of *P. aeruginosa* ST395. Its resistance to copper and protozoans presumably helped the strain to survive in the water distribution system. Although BES has presumably contaminated the water network of the hospital before its opening, it might have further spread among patients by routes of transmission independent of water. The prevalence of patients infected by the epidemic clone ST395 started to decrease in 2001, concomitantly with parallel and convergent evolution. Indeed, independent mutations affected the amino-acid metabolism pathways, led to the mutation of porin OprD, and altered the LPS synthesis pathway. It also modulated the overexpression of the efflux pump MexAB-OprM, of which its fine-tuning may be important for survival and persistence of the clone. We found that the fitness of BES decreased over time, in correlation with genome length. The parallel adaptation of the ST395 isolates to patients or environment may have decreased its persistence in these niches and thereby its patient-to-patient transmissibility. This, along with the implemented infection control measures (patient and staff cohorting, increased hand hygiene, antibiotic stewardship) may have contributed to the extinction of this epidemic pathogen in a hospital setting.

D.H. designed the research; M.P., P.J., A.M., H.P., and B.V. performed the research; B.V. contributed new reagents/analytical tools; M.P., P.J., E.D., B.V., and D.H. analysed the data; and M.P., P.J., X.B., B.V., and D.H. wrote the paper.

References

1. Otter JA, Doumith M, Davies F, Mookerjee S, Dyakova E, *et al*. Emergence and clonal spread of colistin resistance due to multiple mutational mechanisms in carbapenemase-producing *Klebsiella pneumoniae* in London. *Sci Rep* 2017;7:12711.

2. Bertrand X, Bailly P, Blasco G, Balvay P, Boillot A, *et al*. Large outbreak in a surgical intensive care unit of colonization or infection with *Pseudomonas aeruginosa* that overexpressed an active efflux pump. *Clin Infect Dis* 2000;31:E9–E14.

3. Quick J, Cumley N, Wearn CM, Niebel M, Constantinidou C, *et al*. Seeking the source of *Pseudomonas aeruginosa* infections in a recently opened hospital: an observational study using whole-genome sequencing. *BMJ Open* 2014;4:e006278.

4. Eppinger M, Pearson T, Koenig SS, Pearson O, Hicks N, *et al*. Genomic epidemiology of the Haitian cholera outbreak: a single introduction followed by rapid, extensive, and continued spread characterized the onset of the epidemic. *mBio* 2014;5:e01721.

5. Merker M, Blin C, Mona S, Duforet-Frebourg N, Lecher S, *et al*. Evolutionary history and global spread of the *Mycobacterium tuberculosis* Beijing lineage. *Nat Genet* 2015;47:242–249.

6. Mathee K, Narasimhan G, Valdes C, Qiu X, Matewish JM, *et al*. Dynamics of *Pseudomonas aeruginosa* genome evolution. *Proc Natl Acad Sci USA* 2008;105:3100–3105.

7. Koser CU, Holden MT, Ellington MJ, Cartwright EJ, Brown NM, *et al*. Rapid whole-genome sequencing for investigation of a neonatal MRSA outbreak. *N Engl J Med* 2012;366:2267–2275.

8. Buhl M, Kastle C, Geyer A, Autenrieth IB, Peter S, *et al*. Molecular evolution of extensively drug-resistant (XDR) *Pseudomonas aeruginosa* strains from patients and hospital environment in a prolonged outbreak. *Front Microbiol* 2019;10:1742.

9. Marsh JW, Mustapha MM, Griffith MP, Evans DR, Ezeonwuka C, *et al*. Evolution of outbreak-causing carbapenem-resistant *Klebsiella pneumoniae* ST258 at a tertiary care hospital over 8 years. *mBio* 2019;10.

10. French CE, Coope C, Conway L, Higgins JP, McCulloch J, *et al*. Control of carbapenemase-producing *Enterobacteriaceae* outbreaks in acute settings: an evidence review. *J Hosp Infect* 2017;95:3–45.

11. Aujoulat F, Roger F, Bourdier A, Lotthe A, Lamy B, *et al*. From environment to man: genome evolution and adaptation of human opportunistic bacterial pathogens. *Genes (Basel)* 2012;3:191–232.

12. Lyczak JB, Cannon CL, Pier GB. Establishment of *Pseudomonas aeruginosa* infection: lessons from a versatile opportunist. *Microbes Infect* 2000;2:1051–1060.

13. Breidenstein EBM, de la Fuente-Núñez C, Hancock REW. *Pseudomonas aeruginosa*: all roads lead to resistance. *Trends Microbiol* 2011;19:419–426.

14. CDC. 2019. Antibiotic Resistance Threats in the United States. In: *Department of Health and Human Services, CDC*. U.S: Atlanta, GA, 2019.

15. Hota S, Hirji Z, Stockton K, Lemieux C, Dedier H, *et al*. Outbreak of multidrug-resistant *Pseudomonas aeruginosa* colonization and infection secondary to imperfect intensive care unit room design. *Infect Control Hosp Epidemiol* 2009;30:25–33.

16. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, *et al*. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 2012;19:455–477.

17. Petitjean M, Martak D, Silvant A, Bertrand X, Valot B, *et al*. Genomic characterization of a local epidemic *Pseudomonas aeruginosa* reveals specific features of the widespread clone ST395. *Microb Genom* 2017;3:e000129.

18. Kolmogorov M, Raney B, Paten B, Pham S. Ragout - a reference-assisted assembly tool for bacterial genomes. *Bioinformatics* 2014;30:i302-9.

19. Kurtz S, Phillippy A, Delcher AL, Smoot M, Shumway M, *et al*. Versatile and open software for comparing large genomes. *Genome Biol* 2004;5:R12.

20. Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, *et al*. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 2010;11:119.

21. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol* 1990;215:403–410.

22. Zankari E, Hasman H, Cosentino S, Vestergaard M, Rasmussen S, *et al*. Identification of acquired antimicrobial resistance genes. *J Antimicrob Chemother* 2012;67:2640–2644.

23. Chen L, Xiong Z, Sun L, Yang J, Jin Q. VFDB 2012 update: toward the genetic diversity and molecular evolution of bacterial virulence factors. *Nucleic Acids Res* 2012;40:641–645.

24. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 2014;30:923–930.

25. Rohart F, Gautier B, Singh A, Lê Cao K-A. mixOmics: An R package for omics feature selection and multiple data integration. *PLoS Comput Biol* 2017;13:e1005752.

26. Treepong P, Guyeux C, Meunier A, Couchoud C, Hocquet D, *et al*. panISa: ab initio detection of insertion sequences in bacterial genomes from short read sequence data. *Bioinformatics* 2018;34:3795–3800.

27. Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, *et al*. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol* 2012;61:539–542.

28. Rambaut A, Lam TT, Max Carvalho L, Pybus OG. Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). *Virus Evol* 2016;2:vew007.

29. Hocquet D, Nordmann P, El Garch F, Cabanne L, Plésiat P. Involvement of the MexXY-OprM efflux system in emergence of cefepime resistance in clinical strains of *Pseudomonas aeruginosa*. *Antimicrob Agents Chemother* 2006;50:1347–1351.

30. CLSI. Method for dilution antimicrobial susceptibility tests for bacteria that grow aerobically. In: *CLSI document M7-A10 Approved standard*, 10th edn. Wayne, Pa: National Committee for Clinical Laboratory Standards, 2015.

31. Jeanvoine A, Meunier A, Puja H, Bertrand X, Valot B, *et al*. Contamination of a hospital plumbing system by persister cells of a copper-tolerant high-risk clone of *Pseudomonas aeruginosa*. *Water Res* 2019;157:579–586.

32. Thomas V, Herrera-Rimann K, Blanc DS, Greub G. Biodiversity of amoebae and amoeba-resisting bacteria in a hospital water network. *Appl Environ Microbiol* 2006;72:2428–2438.

33. Horcajada JP, Montero M, Oliver A, Sorli L, Luque S, *et al*. Epidemiology and treatment of multidrug-resistant and extensively drug-resistant *Pseudomonas aeruginosa* infections. *Clin Microbiol Rev* 2019;32.

34. Deleo FR, Chen L, Porcella SF, Martens CA, Kobayashi SD, *et al*. Molecular dissection of the evolution of carbapenem-resistant multilocus sequence type 258 *Klebsiella pneumoniae*. *Proc Natl Acad Sci USA* 2014;111:4988–4993.

35. Marvig RL, Johansen HK, Molin S, Jelsbak L. Genome analysis of a transmissible lineage of *Pseudomonas aeruginosa* reveals pathoadaptive mutations and distinct evolutionary paths of hypermutators. *PLoS Genet* 2013;9:e1003741.

36. Martin K, Baddal B, Mustafa N, Perry C, Underwood A, *et al*. Clusters of genetically similar isolates of *Pseudomonas aeruginosa* from multiple hospitals in the UK. *J Med Microbiol* 2013;62:988–1000.

37. Teixeira P, Tacao M, Alves A, Henriques I. Antibiotic and metal resistance in a ST395 *Pseudomonas aeruginosa* environmental isolate: A genomics approach. *Mar Pollut Bull* 2016;110:75–81.

38. Schelstraete P, Van Daele S, De Boeck K, Proesmans M, Lebecque P, *et al*. *Pseudomonas aeruginosa* in the home environment of newly infected cystic fibrosis patients. *Eur Respir J* 2008;31:822–829.

39. Heirali A, McKeon S, Purighalla S, Storey DG, Rossi L, *et al*. Assessment of the microbial constituents of the home environment of individuals with cystic fibrosis (CF) and their association with lower airways infections. *PLoS One* 2016;11:e0148534.

40. Treepong P, Kos VN, Guyeux C, Blanc DS, Bertrand X, *et al*. Global emergence of the widespread *Pseudomonas aeruginosa* ST235 clone. *Clin Microbiol Infect* 2018;24:258–266.

41. Matz C, Kjelleberg S. Off the hook - how bacteria survive protozoan grazing. *Trends Microbiol* 2005;13:302–307.

42. Thomas JM, Ashbolt NJ. Do free-living amoebae in treated drinking water systems present an emerging health risk? *Environ Sci Technol* 2011;45:860–869.

43. Cateau E, Imbert C, Rodier MH. *Hartmanella vermiformis* can be permissive for *Pseudomonas aeruginosa*. *Lett Appl Microbiol* 2008;47:475–477.

44. Hao X, Luthje F, Ronn R, German NA, Li X, *et al*. A role for copper in protozoan grazing - two billion years selecting for bacterial copper resistance. *Mol Microbiol* 2016;102:628–641.

45. Donlan RM, Forster T, Murga R, Brown E, Lucas C, *et al*. Legionella pneumophila associated with the protozoan *Hartmannella vermiformis* in a model multi-species biofilm has reduced susceptibility to disinfectants. *Biofouling* 2005;21:1–7.

46. Sun S, Noorian P, McDougald D. Dual role of mechanisms involved in resistance to predation by protozoa and virulence to humans. *Front Microbiol* 2018;9:1017.

47. Department of Health UK. *Health Technical Memorandum 04-01 Part A: Design, Installation and Commissioning*. 2016.

48. Smith EE, Buckley DG, Wu Z, Saenphimmachak C, Hoffman LR, *et al*. Genetic adaptation by *Pseudomonas aeruginosa* to the airways of cystic fibrosis patients. *Proc Natl Acad Sci USA* 2006;103:8487–8492.

49. King JD, Kocincova D, Westman EL, Lam JS. Review: Lipopolysaccharide biosynthesis in *Pseudomonas aeruginosa*. *Innate Immun* 2009;15:261–312.

50. Maldonado RF, Sa-Correia I, Valvano MA. Lipopolysaccharide modification in Gram-negative bacteria during chronic infection. *FEMS Microbiol Rev* 2016;40:480–493.

51. Magiorakos AP, Srinivasan A, Carey RB, Carmeli Y, Falagas ME, *et al*. Multidrug-resistant, extensively drug-resistant and pandrug-resistant bacteria: an international expert proposal for interim standard definitions for acquired resistance. *Clin Microbiol Infect* 2012;18:268–281.

52. Hocquet D, Bertrand X, Köhler T, Talon D, Plésiat P. Genetic and phenotypic variations of a resistant *Pseudomonas aeruginosa* epidemic clone. *Antimicrob Agents Chemother* 2003;47:1887–1894.

53. Piddock LJ. Multidrug-resistance efflux pumps - not just for resistance. *Nat Rev Microbiol* 2006;4:629–636.

54. Skurnik D, Roux D, Cattoir V, Danilchanka O, Lu X, *et al*. Enhanced in vivo fitness of carbapenem-resistant oprD mutants of *Pseudomonas aeruginosa* revealed through high-throughput sequencing. *Proc Natl Acad Sci USA* 2013;110:20747–20752.

55. Tian ZX, Yi X-X, Cho A, O'Gara F, Wang YP. CpxR activates MexAB-OprM efflux pump expression and enhances antibiotic resistance in both laboratory and clinical nalB-type isolates of Pseudomonas aeruginosa. *PLoS Pathog* 2016;12:e1005932.

56. Hirakata Y, Kondo A, Hoshino K, Yano H, Arai K, *et al*. Efflux pump inhibitors reduce the invasiveness of *Pseudomonas aeruginosa*. *Int J Antimicrob Agents* 2009;34:343–346.

57. Sanchez P, Linares JF, Ruiz-Diez B, Campanario E, Navas A, *et al*. Fitness of in vitro selected *Pseudomonas aeruginosa nalB* and *nfxB* multidrug resistant mutants. *J Antimicrob Chemother* 2002;50:657–664.

58. Vettoretti L, Plesiat P, Muller C, El Garch F, Phan G, *et al*. Efflux unbalance in *Pseudomonas aeruginosa* isolates from cystic fibrosis patients. *Antimicrob Agents Chemother* 2009;53:1987–1997.

59. Hirakata Y, Srikumar R, Poole K, Gotoh N, Suematsu T, *et al*. Multidrug efflux systems play an important role in the invasiveness of *Pseudomonas aeruginosa*. *J Exp Med* 2002;196:109–118.

60. Hocquet D, Petitjean M, Rohmer L, Valot B, Kulasekara HD, *et al*. Pyomelanin-producing *Pseudomonas aeruginosa* selected during chronic infections have a large chromosomal deletion which confers resistance to pyocins. *Environ Microbiol* 2016;18:3482–3493.

61. Tsuchiya K, Cao YY, Kurokawa M, Ashino K, Yomo T, *et al*. A decay effect of the growth rate associated with genome reduction in *Escherichia coli*. *BMC Microbiol* 2018;18:101.

62. Binet R, Maurelli AT. Fitness cost due to mutations in the 16S rRNA associated with spectinomycin resistance in *Chlamydia psittaci* 6BC. *Antimicrob Agents Chemother* 2005;49:4455–4464.

63. Rohr U, Weber S, Michel R, Selenka F, Wilhelm M. Comparison of free-living amoebae in hot water systems of hospitals with isolates from moist sanitary areas by identifying genera and determining temperature tolerance. *Appl Environ Microbiol* 1998;64:1822–1824.