communications biology

A Nature Portfolio journal

Article

https://doi.org/10.1038/s42003-025-08156-y

Genetic landscape and evolution of *Acinetobacter pittii*, an underestimated emerging nosocomial pathogen

Shengke Wang[1,5], Yan Zhou[2,5], Yuezhuo Wang[3], Keshu Tang[1], Danqi Wang[1], Jiawen Hong[1,4], Pengcheng Wang[1], Sheng Ye[1], Jie Yan[2] ✉, Shengkai Li[3] ✉, Zhemin Zhou[3] ✉ & Jimei Du[1] ✉

As a member of *Acinetobacter calcoaceticus-baumannii* complex, *Acinetobacter pittii* has been an emerging concern in nosocomial infection due to its increasing prevalence and multidrug resistance (MDR). However, its population structure remains broadly unknown, hampering efficient tracing of its transmission and evolution. In this study, we developed a distributed core genome multilocus sequence typing (dcgMLST) for *A. pittii* based on 750 genomes and employed it to map the genetic landscape and evolution of *A. pittii*. The results demonstrated that two hierarchical clustering (HC) levels effectively correspond to genetic diversity from species (HC1100) to natural populations (HC450), as well as that a predominant lineage, HC1100_4, accounts for 33.9% of *A. pittii* strains. Subsequent analysis revealed that specific gene gain and loss events within HC1100_4 are linked to adaptations to environmental stress. Moreover, we identified a cluster of multidrug-resistant plasmids PT_712 responsible for the dissemination of bla_{NDM-1} genes within the genus of *Acinetobacter*. This study provides a framework for characterizing genetic diversity, evolutionary dynamics, molecular population distribution, and tracing of *A. pittii*, which has the potential to improve infection control strategies and public health policy.

Pathogenic bacteria in genus *Acinetobacter*, particularly those of *Acinetobacter calcoaceticus-baumannii* (ACB) complex, pose severe threats to human health worldwide due to their wide adaption capacity to different environments and multidrug resistance (MDR)[1]. *Acinetobacter baumannii* (*A. baumannii*) is a leading cause of nosocomial infections, and carbapenem-resistant *A. baumannii* has been regarded a critical threat, calling for efficient and prompt controlling measures[2]. Other ACB bacteria, while less renowned, were also frequently associated with 11%–37% of nosocomial infections caused by *Acinetobacter* worldwide, demanding further investigations[3–5].

Initially recognized as genospecies 3, *Acinetobacter pittii* (*A. pittii*) was proposed as a formal species in the ACB complex since 2011[6]. In recent decades, the potential of *A. pittii* as a nosocomial pathogen has likely been underestimated due to the lack of proper techniques to distinguish it from *A. baumannii*. With advancements in diagnostic technology, *A. pittii* has been increasingly identified in nosocomial infections[7,8], with its detection rates in blood samples occasionally surpassing those of *A. baumannii*[9,10]. This highlights *A. pittii* as an emerging nosocomial pathogen on a global scale[11]. Although its clinically drug resistance is not as severe as that of *A. baumannii*[12], there has been a rise in the isolation multi-drug resistant strains[13], carrying genes responsible for carbapenemases[14], and gene mutations associated with colistin resistance[15]. Given that *A. pittii* inhabitats a wide range of environments, including water[16], soil[17], wild animal[18], and food[19], it may have the capacity to disseminate drug resistant factors, such as antimicrobial resistances gene bla_{NDM}[20]. Considering its widespread environmental presence and increasing prevalence in both nosocomial and community infection across various regions[11,21], it is imperative to investigate the population structure and genetic characteristics of *A. pittii*, especially under the framework of the One Health strategy.

The core genomic multi-locus sequence typing (cgMLST) schemes facilitated the rapid population assignments and genetic analysis of many bacterial pathogens, and the distributed cgMLST (dcgMLST) schemes

[1]Wenzhou Key Laboratory of Sanitary Microbiology, Department of Microbiology and Immunology, School of Laboratory Medicine, Institute of One Health, Wenzhou Medical University, Wenzhou, China. [2]School of Medicine, Zhejiang University, Hangzhou, China. [3]Key Laboratory of Alkene-carbon Fibres-based Technology & Application for Detection of Major Infectious Diseases, MOE Key Laboratory of Geriatric Diseases and Immunology, Cancer Institute, Suzhou Medical College, Soochow University, Suzhou, China. [4]Taizhou Hospital of Zhejiang Province Affiliated to Wenzhou Medical University, Linhai, Zhejiang, China. [5]These authors contributed equally: Shengke Wang, Yan Zhou. ✉e-mail: med_bp@zju.edu.cn; Zxzzzhu@Gmail.com; zmzhou@suda.edu.cn; djm@wmu.edu.cn

further enabled decentralized use and sharing of cgMLST data[22]. In this study, we complied the largest collection to date of 750 *A. pittii* genomes, sourced from public databases (726 genomes) and our own sequencing results (24 genomes). Utilizing this dataset, we established a dcgMLST scheme specific tailored for *A. pittii* and conducted practical application. Based on this scheme, we hierarchically separated *A. pittii* genomes into three distinct clustering levels reflecting vary levels of genetic diversity, and revealed the global predominance of a major lineage, HC1100_4. Moreover, we identified a cluster of multidrug-resistant IncX plasmids responsible for the dissemination of bla_{NDM-1} genes across *Acinetobacter* species.

Results

Distributed cgMLST scheme and the species tree based on 726 *A. pittii* genomes

We retrieved all 726 *A. pittii* genomes from NCBI GenBank (as of April 25th, 2024), to build a comprehensive dataset (Supplementary Data 1) encompassing strains isolated from 30 countries/regions across six continents (Fig. 1a). All strains were isolated in the past 34 years, with over 70% isolated after 2010. Over half (53.5%) of the strains were from East Asia, followed by North America (25.2%) and Europe (13.4%). Some *A. pittii* strains may be sourced from same hospital as they are genetically almost identical. To reduce the over-representation of such strains, we estimated the average nucleotide identities (ANIs) of all genomes and built a subset of 217 representative sequences by selecting one random sequence from each of the 217 single-linkage clusters (SLCs) with a threshold of > 98%. We estimated a pan-genome of 18,882 genes for the representative genomes using PEPPAN[23], including 2840 soft-core genes that were shared by ≥ 95% of the representative genomes. These soft-core genes were also shared by 94–100% of the genomes in the whole dataset, and were used to establish the dcgMLST[22], which uses the MD5 hash function to assign each gene a 128-bit integer allelic ID based on its sequence. Using the dcgMLST scheme, we compiled all *A. pittii* genomes into allelic profiles of no more than 2840 integers. Pairwise comparisons of genetically closely related *A. pittii* genomes revealed a linear correlation between the dcgMLST alleles and the core genomic SNPs, with approximately 1.33 SNPs for each allelic difference (Supplementary Fig. 1a).

We estimated the genetic relationships of all 726 *A. pittii* genomes using the SLC algorithm in pHierCC[24]. The results included multiple hierarchical clustering (HC) configurations designated as HC[threshold] (e.g., HC450, HC1100), the number indicates the specific SLC threshold used for clustering. Among all 2840 possible levels, HC450 and HC1100 generated the most separating clusters (Fig. 1b and Supplementary Fig. 1b) according to their Silhouette scores. We plotted the frequencies of genome pairs with different dcgMLST allelic differences and found that these two aforementioned levels all fell into valleys with substantially fewer genome pairs (Fig. 1b), indicating their representation of genetically separated populations in *A. pittii* (Supplementary Fig. 1c).

The maximum-likelihood phylogenetic tree of all *A. pittii* genomes displayed eleven major clusters with the top four of HC1100_4 ($n = 246$), HC1100_15 ($n = 59$), HC1100_11 ($n = 46$), and HC1100_2 ($n = 36$). Compared to the HC1100, MLST has some unassigned genomes and some misassignments, such as ST63 and ST64 cluster (Fig. 1c). To further evaluate the performance of dcgMLST on population assignments, we compared the phylogenetic tree of HC1100, HC450 with Pasteur's ST, PopPUNK and HierBaps by plotting a Sankey diagram (Supplementary Fig. 2a, b) and calculated their pairwise AMI values (Table 1). As shown in Tabel 1, HC1100 and HC450 performed well on the phylogenetic tree, with almost all HC1100s highly corresponding to the HierBaps clusters, HC450s to the PopPUNK clusters.

Geographical distribution and resistome among *A. pittii* HC1100 populations

All *A. pittii* genomes were separated into 148 HC1100 clusters and 215 HC450 clusters. Particularly, we identified a predominant cluster, HC1100_4, that accounted for over 1/3 of the available *A. pittii* genomes. HC1100_4 has been the most abundant lineage worldwide over the past two decades in all major geographic regions (Fig. 2a). Furthermore, HC1100_4 was subdivided into 23 HC450 clusters, many of which exhibited geographic specificities. For example, over 80% of HC450_4 and HC450_76 strains were from East Asia, all HC450_33 strains were from North America. Notably, HC1100_4 was less commonly found in human-associated sample (32%) than others from environments (40%) (Fig. 2a). Meanwhile, 122 of the 148 HC1100 clusters contained only 1 and 2 genomes, represented by long branches diverging from the root of the phylogeny. The majority of these rare HC1100 clusters were found in East Asia (48%), North America (18%), and Europe (9%), coincident with the isolation frequencies of the strains.

We identified fourteen classes of antibiotic resistance genes (ARGs) in *A. pittii* genomes (Fig. 2b). The majority of the strains carried genes associated with resistance to cephalosporin (98.1%), followed by those to aminoglycoside (67.1%), carbapenem (22.7%), macrolide (16.3%), sulfonamide (15.6%) and others. Notably, bla_{OXA-58} (7.5%) and bla_{NDM-1} (6.8%) were the predominant carbapenemase gene and both widely found in 10 and 14 HC1100 clusters, respectively. Additionally, a few strains also carried *qnrD*/*qnrVC6* (0.4%) and *mcr-1.1* (0.1%) genes, associated with resistance to either quinolone, and colistin, respectively. There were considerable (22.9%) MDR strains, predominantly isolated from China and Japan (Supplementary Data 1). However, no significant association was found between the number of ARG and geographic distribution ($P = 0.2$) (Supplementary Fig. 3). From the perspect of sources, resistance pattern was observed distinct among human-associated, wild animal, and environmental samples (Fig. 2c, d). Analysis of all 57 complete genomes demonstrated that most ARGs on the chromosome were scattered as mobile genetic elements (MGEs) across the genomes (Supplementary Fig. 4).

HC1100_4 adaptation related core-genome evolution

To investigate the genetic context of the predominant HC1100_4 lineage, we compared its accessory genes with those in its neighboring lineages of HC1100_2 and HC1100_128. To this end, a Random Forest model based on all 18,882 pan-genes was trained to find genes differentiating these HC1100 clusters and those differentiating them with other HC1100s. We found that, compared to other HC1100 clusters, the most recent common ancestor of these three lineages experienced a genome expansion, with enrichment of 328 accessory genes while only depleting 178 genes. Subsequently, after the divergence of HC1100_4 from HC1100_2 and HC1100_128, it experienced a genome reduction, with over twice of more gene depletions (200 genes) than enrichments (78 genes). We categorized the most enriched or depleted genes into Clusters of Orthologous Genes (COGs) and found substantial over-representation of genes associated with stress responses, including biofilm formation, metal response, and oxidative stress. Furthermore, we found that 64% of the genes enriched in HC1100_4 fell into six genomic islands (GEIs) in the chromosome (Fig. 3a–c). These GEIs were unlikely associated with MGEs because we did not find integrases or transposons around them. Instead, genes from the same island were functionally associated, likely representing results of long-term natural adaptations.

Specifically, GEI 1 includes three hypothetical genes along with *pepN* and *aldA*, which are involved in alanine hydrolysis metabolism and defense against hypochlorite, respectively[25,26]. GEI 2 consists of five genes related to capsule and LPS synthesis, potentially involved in antigen synthesis[27]. GEI 3 comprises six genes associated with environmental resistance and metal ion tolerance. *djlA* may work synergistically with *dnaJ* to promote bacterial growth[28]. *cusR* and *czcO* help bacteria survive in high-metal environments[29,30]. *mdtA* plays a key role in MDR as efflux pump[31]. *arsC* and *arsH* help bacteria tolerate arsenic compounds and may contribute to oxidative stress response[32]. GEI 4 includes a set of genes that collectively encode subunits of L-tartrate dehydratase[33]. GEI 5 contains the *gcvA* and *fabG* genes which ensure bacterial survival and adaptability under different environmental conditions[34,35]. GEI 6 encompasses six genes. *acrE* involved in bacterial MDR[36], *srrA* senses environmental signals and regulates metabolism[37], *argS* encodes arginyl-tRNA synthetase[38], *ribX* and *ribY*

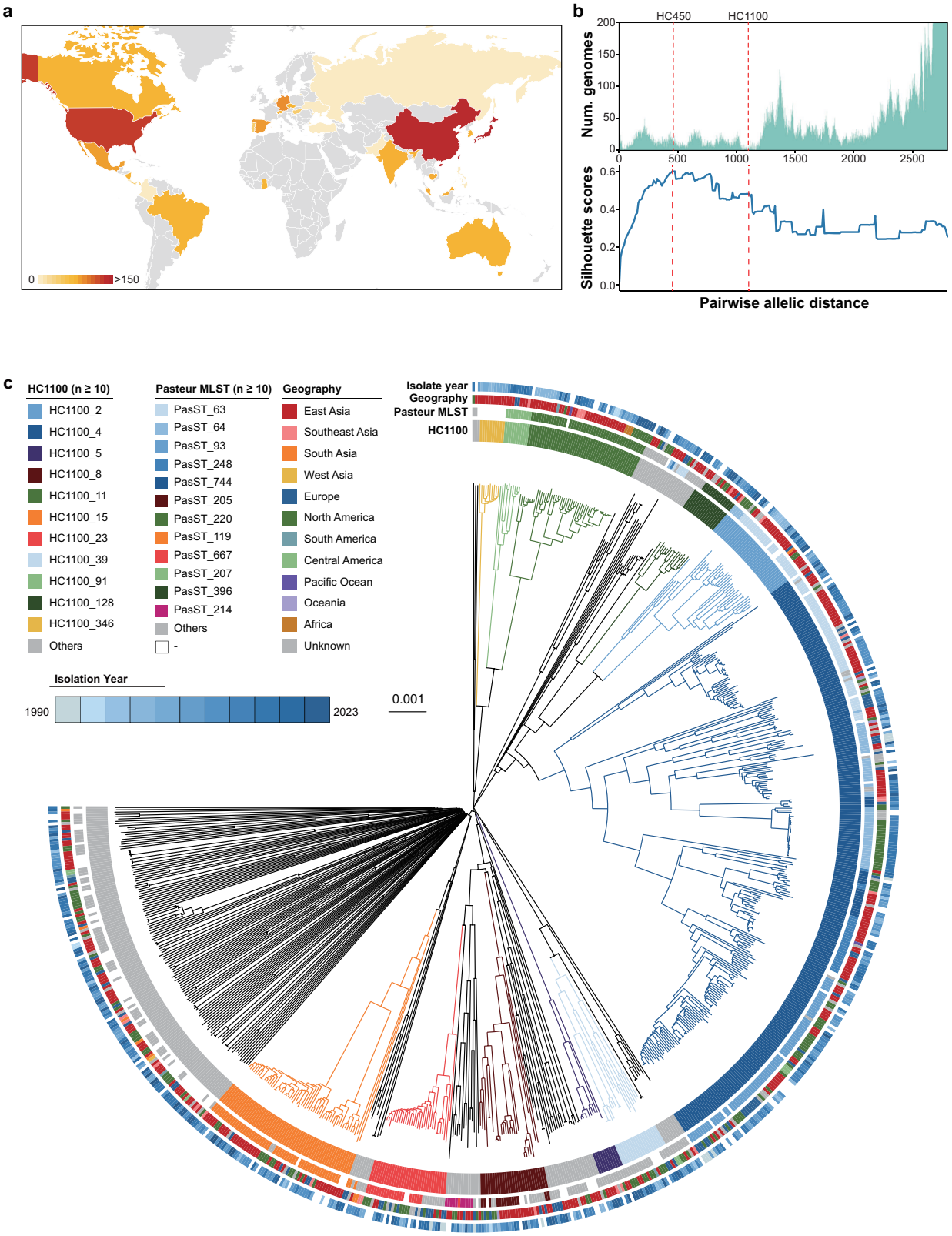

Fig. 1 | Population structure of *A. pittii* genomes. a Geographic distribution of *A. pittii* genomes. The block color represents the number of strains. **b** Frequencies of genome pairs with different dcgMLST allelic differences and Silhouette scores of different pairwise allelic distances. Two levels, HC450 and HC1100, fell into valleys with substantially fewer genome pairs, which genetically separated natural populations in *A. pittii*. HC450 exhibited the farthest separation distances between lineages within its cluster, with a Silhouette score of 0.60. HC1100 represented the optimal hierarchical clustering scheme (interval peak) for pairwise allelic distances greater than 1000 core alleles, achieving a Silhouette score of 0.48. **c** The evolutionary tree of *A. pittii* based on 2840 core genes in the 217 representative genomes. HC1100, Pasteur MLST, geography, and isolate year were shown in the outer rings. The map was modified from open-source data in [https://github.com/antvis/L7]. Source Data is provided as Source Data.

Table 1 | Statistical analysis results of four clustering methods

AMI value vs.	HC450	HC1100	HC1100 (HC1100_4 excluded)	HierBaps	PopPUNK	MLST
HC450	–	–	–	0.7045	0.9631	0.8039
HC1100	–	–	–	0.9352	0.7328	0.7568
HC1100 (HC1100_4 excluded)	–	–	–	0.9093	0.8485	0.8888
HierBaps	0.7045	0.9352	0.9093	–	0.6927	0.7404
PopPUNK	0.9631	0.7328	0.8485	0.6927	–	0.7907
MLST	0.8039	0.7568	0.8888	0.7404	0.7907	–

involved in riboflavin (vitamin B2) biosynthesis[39], and *hpxO* functions in carbon and nitrogen acquirement during nucleotide metabolism[40]. Overall, HC1100_4 experienced substantial gene loss while acquiring six GEIs. Such gene content variations might be associated with the predominance of the lineage, giving it survival advantages in various environments, including MDR, heavy metal tolerance, and diversified metabolic capabilities.

dcgMLST typing platform and application

We constructed an online platform for *A. pittii*, available on the database PathoBase hosted by Soochow University. Population assignment, genome annotation, and plasmid prediction results will be yield in 5 to 20 min after uploading genome or sequencing data (Fig. 4a).

We applied the PathoBase to analysis 24 *A. pittii* collected in this study (Table 2 and Fig. 4b). The cohort was divided into 16 distinct HC1100s and 17 HC450s. The HC1100_4 was the largest cluster, accounting for five isolates from three cities, followed by four HC1100 clusters that each contained two isolates and 11 singletons. Notably, three isolates identified carrying $bla_{\text{NDM-1}}$, one in chromosome (HC1100_65) and two on PT_712 plasmids (HC1100_65, HC1100_4). We used the PT_712 in *A. pittii* WZ-35 (NZ_CP168074.1) (https://www.ncbi.nlm.nih.gov/nuccore/NZ_CP168074.1/) as the reference sequence, which indicated the PT_712 plasmids in *A. pittii* were approximately 270 KB in length, sharing the majority of backbone genes and carrying the same set of ARGs of $bla_{\text{NDM-1}}$, *aph(3')-VIa*, *aac(3)-IId*, $bla_{\text{OXA-58}}$, *msr(E)*, *mph(E)*, and *floR* (Fig. 4c).

Plasmid driven $bla_{\text{NDM-1}}$ transmission in *Acinetobacter*

For further understanding of PT_712, we assigned plasmid types (PTs) for all plasmids in 57 completely sequenced genomes. Twenty-three PTs were identified as resistant PT (Fig. 5a). PT_712 and PT_3057 were the two most frequent (6/21) plasmids carrying the most ARGs, with an average of 9.3 ARGs in PT_712 and 6.7 ARGs in PT_3057, while other PTs carried an average of 0.1 to 1.5 ARGs. Apart from common resistant gene *msr(E)* (6/6) and *mph(E)* (6/6), *aac(3)-IId* (5/6), *floR* (4/6), and *sul2* (4/6), it is particularly noteworthy that PT_712 carring $bla_{\text{NDM-1}}$ (3/6) and bla_{IMP} (3/6) that confer resistant to carbapenem[41,42]. A study in Thailand showed that PT_712 enables the *A. pittii* strain A436 resistance to meropenem[43], another study from China reported the *A. pittii* strain AP2044 carrying PT_712 exhibited a distinct elevated MIC value (≥ 1024 mg/L) for third-generation cephalosporins[44]. Both projects confirmed that the plasmid PT_712 possess and enhance the multidrug-resistant capability of *A. pittii*.

PT_712 was identified in 5 HC1100s of *A. pittii*. We suppose it may transfer among different populations, and even across *Acinetobacter* species. We scanned 6949 publicly available *Acinetobacter* genomes with clear source by PathoBase, and identified 20,752 plasmids from 654 PTs. PT_712 was the most frequently detected plasmid among the genus (Supplementary Data 2), with the highest prevalence of 27.4% in *A. baumannii*, followed by *A. pittii* (11.3%) and *A. ursingii* (8.5%). Notably, 38.7% PT_712 plasmids in *Acinetobacter* carried $bla_{\text{NDM-1}}$ genes (Supplementary Data 3). Phylogenetic reconstruction of PT_712 plasmids showed that two PT_712 plasmids identified in this study fell into a China-enriched deep clade in the tree. There were other clades of PT_712 primarily found in the USA and South America countries, exhibiting geographic specificity (Fig. 5b and Supplementary Data 3).

We then built a plasmid-species interaction network of *Acinetobacter* by connecting each species with the PTs predicted in it, and visualized it in Gephi (Supplementary Data 4). We also measured the importance of each node in the network using two centrality metrics of hub and average weighted degree (Fig. 5c). Remarkably, PT_712 ranked the fourth in node hub score (0.11) and the third in average weighted degree (7.3), underscoring its significant role in transferring horizontal gene transfers (Fig. 5d, e).

Discussion

A. pittii is a member of the ACB complex that associated with nosocomial infections[6]. It is often overlooked, mainly due to its relatively low drug resistance and detection rate in comparison to *A. baumannii*. However, emerging evidence indicates that MDR *A. pittii* strains are increasingly implicated in hospital-acquired infections[11,45,46]. The interpretation of clinically relevant genomic information from *A. pittii* are challenging, primarily due to the limited understanding of its genetic framework. In this study, we developed a distributed core genome multilocus sequence typing (dcgMLST) scheme for *A. pittii*, based on an analysis of 750 genomes. This approach has unveiled a global lineage distribution of *A. pittii*. Furthermore, we applied the data analysis pipeline of the dcgMLST scheme to *A. pittii* strains from four cities in East China, identifying a plasmid carrying the $bla_{\text{NDM-1}}$ gene, which is capable of horizontal transfer within the genus of *Acinetobacter*.

Currently, there is no accurate and specific typing scheme for *A. pitti*, scientists always implement the MLST frameworks developed for *A. baumannii*[47] for clinical and epidemic application. As clinical and public health interests in *A. pittii* rise, the necessity for advanced genomic tools becomes evident. The cgMLST scheme is widely recognized for monitoring disease outbreaks and tracing transmission pathways[22,48]. By utilizing stable genetic markers, cgMLST facilitates robust and reliable comparisons of genetic relatedness across strains, thereby minimizing the impact of indels and enhancing its application in epidemiological and evolutionary research[49]. However, cgMLST schemes necessitate the reconstruction of a central database for the nomenclature of added alleles and sequence types. The species-specific dcgMLST scheme we developed for *A. pittii* can nomenclature added alleles and sequence types directly, and update the central database automatically[22]. HierBaps[50,51] and PopPUNK[52] are widely used tools for species clustering analysis. The AMI values between HC1100 and HierBaps, as well as HC450 and PopPUNK, were significantly closer than those observed with MLST (Table 1), with almost all HC1100s and HC450s showing highly correspondence to the HierBaps and PopPUNK clusters, respectively (Supplementary Fig. 2). Moreover, HierBaps has limited scalability and is unable to assign types to newly introduced genomes, while PopPUNK clusters are unstable and may merge when strains are added[50–52]. All these suggested that hierarchical clustering based dcgMLST offers satisfying phylogenetic compatibility, outperform the other schemes.

An important outcome from the population structure analysis is the extensive global dissemination and persistent detection rates of HC1100_4, which constitutes 33.9% of *A. pittii* genomes. This phenomenon is

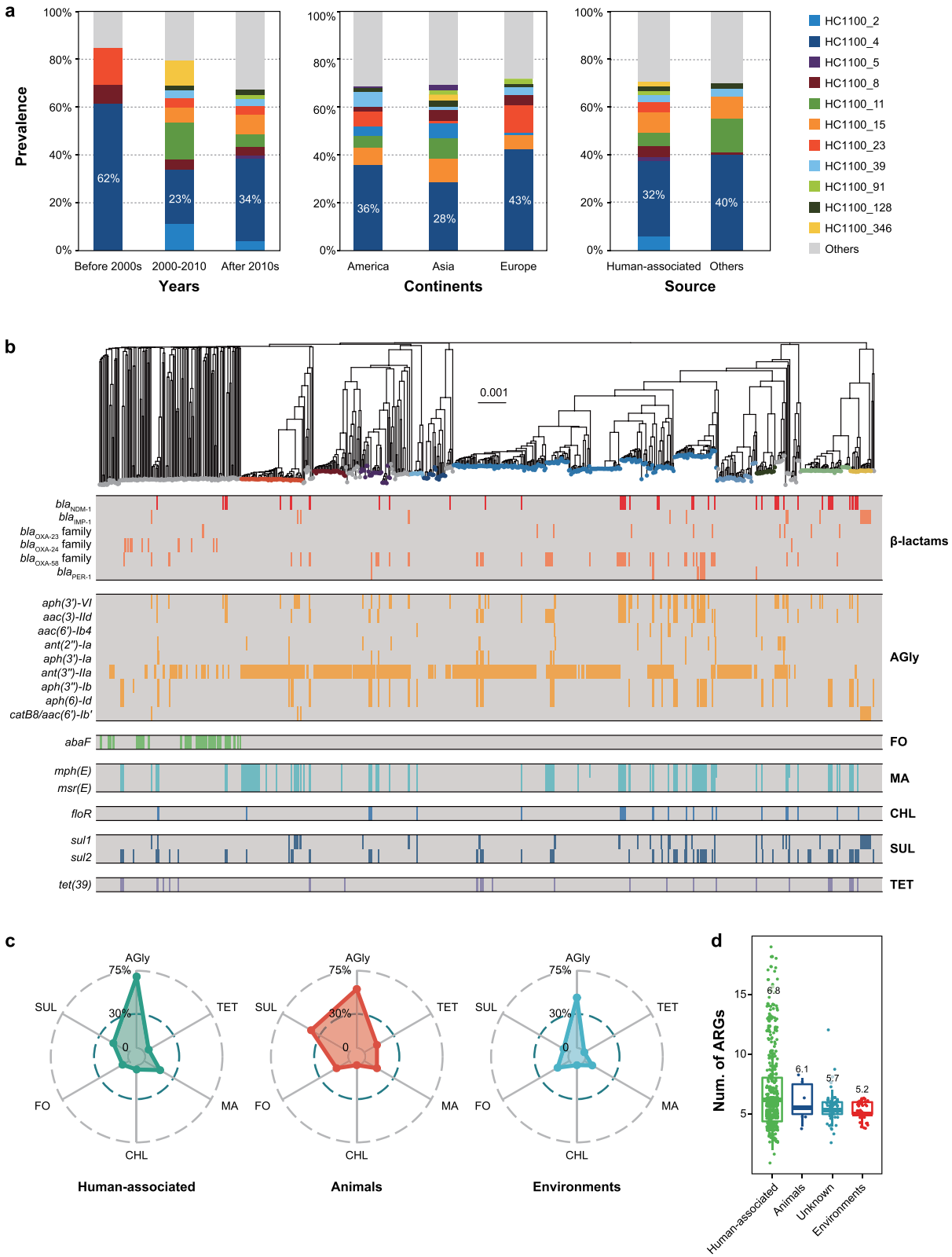

Fig. 2 | Distribution and antibiotic resistance gene (ARG) profile of *A. pittii*. **a** Prevalence of HC1100s in different periods ($n_{\text{Before 2000s}} = 13$, $n_{\text{2000-2010}} = 97$, $n_{\text{After 2010s}} = 511$), geography ($n_{\text{America}} = 179$, $n_{\text{Asia}} = 386$, $n_{\text{Europe}} = 89$) and host ($n_{\text{Human-associated}} = 580$, $n_{\text{Others}} = 87$). **b** Prevalence of ARGs in HC1100s. The filled squares by different colors indicate the resistance genes were carried, and blank mean not. AGly aminoglycoside, FO fosfomycin, MA macrolide, CHL chloramphenicol, SUL sulfonamide, TET tetracycline. **c** The prevalence of ARGs in *A. pittii* across different hosts ($n_{\text{Human-associated}} = 580$, $n_{\text{Animals}} = 21$, $n_{\text{Environments}} = 66$). **d** The number of ARGs in *A. pittii* across different hosts ($n_{\text{Human-associated}} = 580$, $n_{\text{Animals}} = 21$, $n_{\text{Unknown}} = 59$, $n_{\text{Environments}} = 66$). Source Data is provided as Source Data.

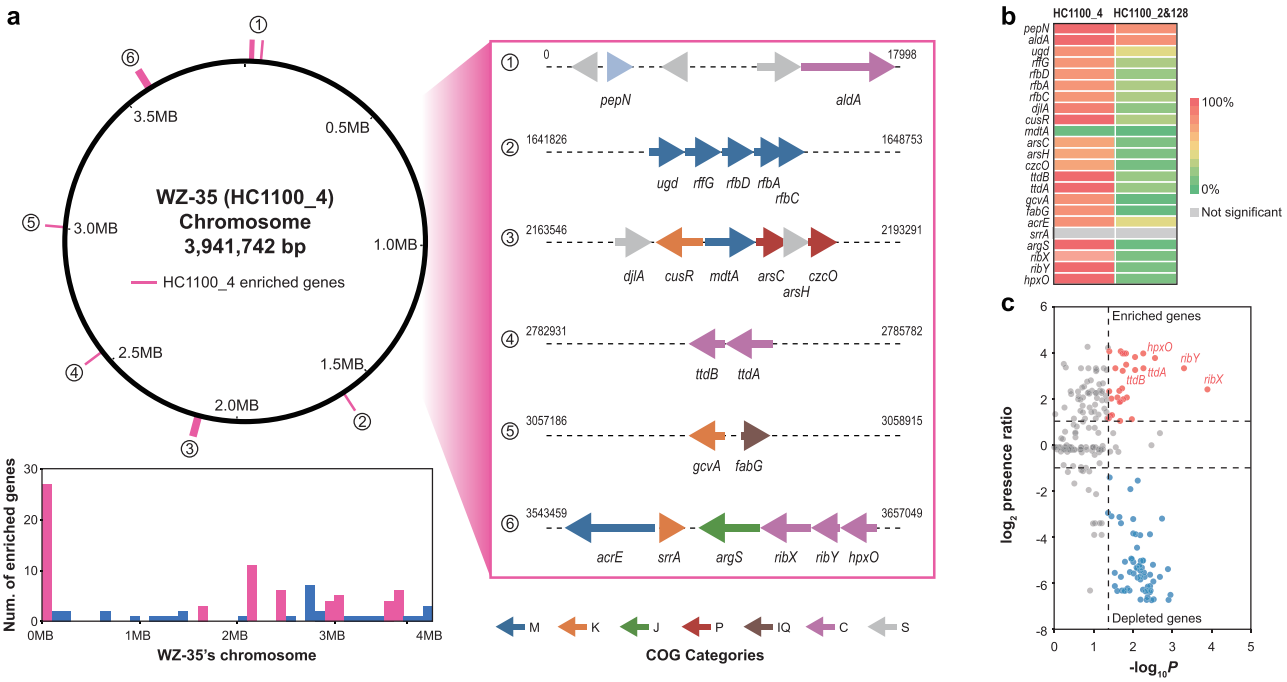

Fig. 3 | Enriched genes of *A. pittii* HC1100_4 lineage. a Genes enriched in HC1100_4 fell into six genomic islands in the chromosome. The genome circle map of WZ-35 (HC1100_4) chromosome with annotated locations of enriched genes. The number of enriched genes at different positions on the chromosome of WZ-35 was shown in the bar graph. Top six genomic islands with well-defined functions were marked in pink and specifically shown to contain genes. Using COGs to clarify the function of each gene by different colors. **b** The prevalence of enriched genes on six genomic islands in lineage HC1100_4 compared to its pre-lineages HC1100_2 and HC1100_128. **c** Enriched genes and depleted genes in the HC1100_4 lineage compared to its pre-lineages HC1100_2 and HC1100_128.

reminiscent of the International Clone 2 observed in *A. baumannii*[53]. By analyzing the pan-genome of *A. pittii*, it was revealed that the HC1100_4 lineage exhibits frequent variation in gene composition, primarily involving acquired genes related to environmental adaptability. Given its widespread distribution and environmental adaptability, an increase in infection and drug-resistant strains is likely inevitable, necessitating heightened attention and research on *A. pittii*. To facilitate practical application, we have integrated the dcgMLST scheme and genome analysis workflow for *A. pittii* into PathoBase. Application of this scheme to 24 *A. pittii* genomes demonstrated that the dcgMLST on PathoBase is effective for genome typing and population analysis, including the prediction of plasmids, ARGs, and virulence factors. Furthermore, our analysis revealed that PT_712 was exclusively predicted within the genus *Acinetobacter* without specificity to any particular species in the phylogenetic tree (Fig. 5b). This suggests the potential for cross-species transmission of the plasmid. Notably, a significant finding was that there are already some strains of *A. pittii* harbored $bla_{\text{NDM-1}}$, bla_{IMP} carrying PT_712 plasmids. As $bla_{\text{NDM-1}}$ and bla_{IMP} confer resistance to carbapenems[41,42], the last resort antibiotics for severe gram-negative infections, the spread of $bla_{\text{NDM-1}}$-carrying PT_712 is particularly concerning.

We acknowledge the limitations inherent in this study, which include the relatively small number of strains in the database and the geographically biased distribution of strain origins. These limitations may contribute to potential instability in strain typing and introduce biases in the geographical representation of strain distribution.

In conclusion, we have dressed the gap in knowledge regarding the population structure of *A. pittii* by developing a dcgMLST scheme. From a global perspective, our findings indicated that *A. pittii* exhibits considerable genetic diversity and possesses relatively fewer ARGs, compared to *A. baumannii*. It had a dominant epidemic lineage HC1100_4, whose success may be partially attributed to the acquisition of certain key genes. Furthermore, we identified the presence of plasmid PT_712 across multiple *Acinetobacter* species, suggesting its potential for inter-species transmission within the genus. *A. pittii* warrants increased attention from both clinical and environmental perspectives, particularly within the framework of the One Health strategy.

Methods

Antimicrobial susceptibility testing

The antimicrobial susceptibility analysis was carried out using the Kirby-Bauer disk diffusion method. The minimum inhibitory concentration (MIC) breakpoints were interpreted according to CLSI-2023[54]. Fifteen antibiotics belonging to 5 categories were tested: Ceftazidime (CAZ), Cefepime (FEP), Ceftriaxone (CRO), Imipenem (IPM), Meropenem (MEM), Piperacillin/Tazobactam (TZP), Ticarcillin/Clavulanic acid (TCC), Tobramycin (TOB), Gentamicin (GEN), Amikacin (AMK), Levofloxacin (LVX), Ciprofloxacin (CIP), Trimethoprim/ Sulfamethoxazole (SXT), Minocycline (MNO), and Doxycycline (DOX).

Whole genome sequencing and data collection

Twenty-four strains of *A. pittii* isolated in 2015, 2021, and 2023 were retrieved from the strain repositories of four local hospitals located in different cities across East China. Ethical approval was reviewed and given by the Ethics Committee of Taizhou Hospital (KL20240118). With Ethics Committee's waiver of informed consent, anonymous host information of these strains was recorded in Supplementary Data 5. Total DNA of each isolate was purified and recovered by a silica gel column (D3146, HiPure Bacterial DNA kit) after incubation. Paired-end libraries with insert sizes of ~300 bp were prepared following Illumina standard genomic DNA library preparation procedure (VAHTS Universal DNA Library Prep kit for Illumina V3) and sequenced on an Illumina NovaSeq 6000 using the S4 reagent kits (v1.5) according to the manufacturer's instructions. The sequencing reads of each isolate were quality-trimmed and assembled into contigs using EToKi[49]. In addition, 799 assembled genomes were downloaded from the NCBI GenBank database on April 25, 2024. To ensure the quality of the genomes, we kept only assemblies with N50 values more than 10 KB, and total sizes between 3.69 and 4.32 MB were retained, after which using

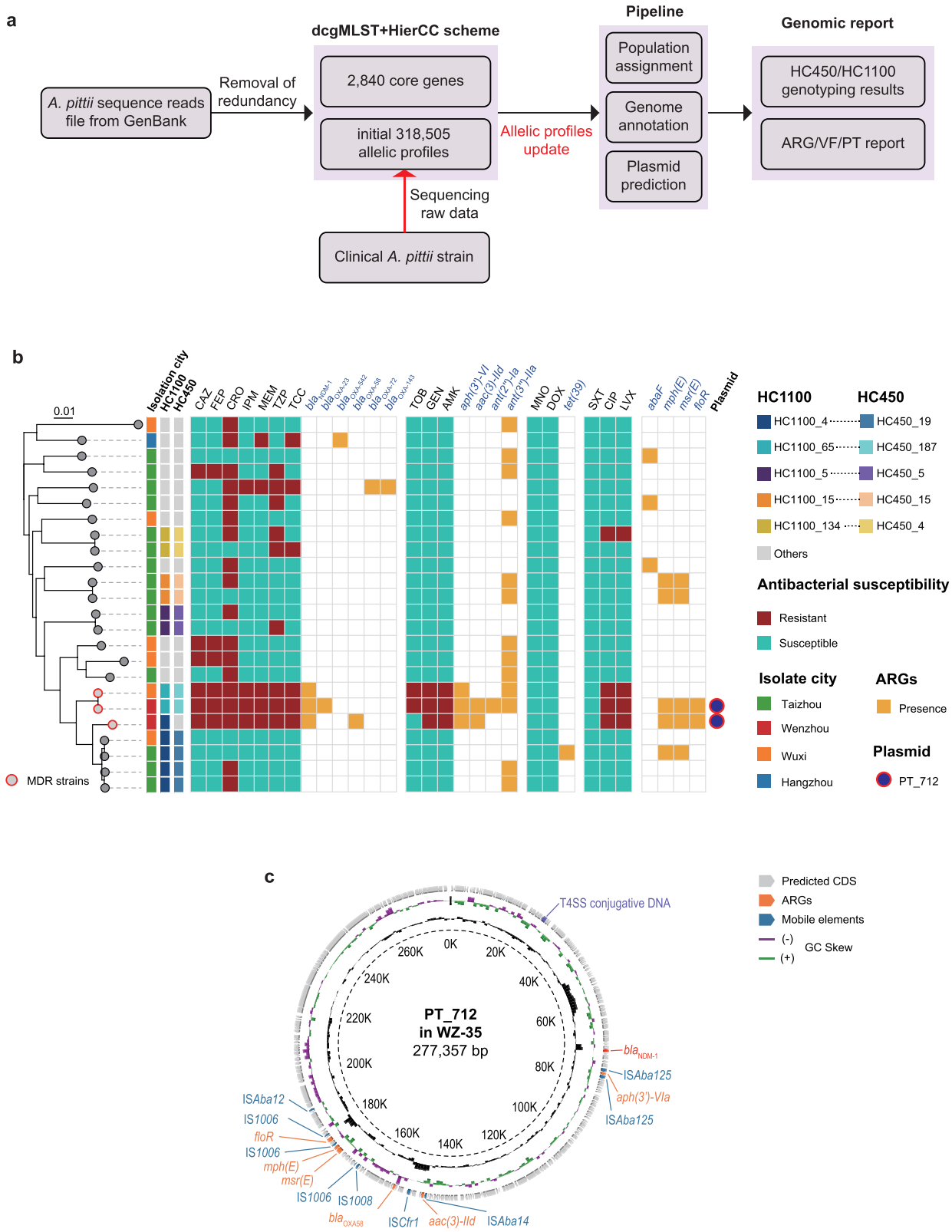

Fig. 4 | Application of dcgMLST to 24 clinical *A. pittii*. **a** Flowchart of the dcgMLST scheme applied to the analysis of collected strains. **b** Twenty-four *A. pittii* isolates collected from East China. The antimicrobial susceptibility test (AST) was performed to test the susceptibility of 24 isolates. Three isolates were MDR. ARGs and plasmids were annotated by differently colored squares. Three MDR isolates carried bla_{NDM-1}, two of which were from PT_712. Phylogenetic trees and typing results of these strains were shown by using dcgMLST. CAZ ceftazidime, FEP cefepime, CRO ceftriaxone, IPM imipenem, MEM meropenem, TZP piperacillin/tazobactam, TCC ticarcillin/clavulanic acid, TOB tobramycin, GEN gentamicin, AMK amikacin, LVX levofloxacin, CIP ciprofloxacin, SXT trimethoprim/sulfamethoxazole, MNO minocycline, DOX doxycycline. **c** The structure and functional genes of plasmid PT_712 in WZ-35. Source Data is provided as Source Data.

Table 2 | Genome summary of 24 clinical *A. pittii* from East China

Strain	NCBI assembly	Location	Isolate year	MLST (Pasteur)	HC450	HC1100
TZ-01	GCA_039887985.1_ASM3988798v1_genomic.fna	Taizhou, Zhejiang	2021	–	493	493
TZ-02	GCA_039887875.1_ASM3988787v1_genomic.fna	Taizhou, Zhejiang	2021	93	19	4
TZ-04	GCA_039887665.1_ASM3988766v1_genomic.fna	Taizhou, Zhejiang	2022	–	668	668
TZ-05	GCA_039887655.1_ASM3988765v1_genomic.fna	Taizhou, Zhejiang	2022	–	705	705
TZ-06	GCA_039887775.1_ASM3988777v1_genomic.fna	Taizhou, Zhejiang	2022	795	134	134
TZ-07	GCA_039887295.1_ASM3988729v1_genomic.fna	Taizhou, Zhejiang	2022	795	134	134
TZ-08	GCA_039887475.1_ASM3988747v1_genomic.fna	Taizhou, Zhejiang	2022	–	662	662
TZ-10	GCA_039887695.1_ASM3988769v1_genomic.fna	Taizhou, Zhejiang	2023	93	19	4
TZ-11	GCA_039887425.1_ASM3988742v1_genomic.fna	Taizhou, Zhejiang	2023	205	157	8
TZ-13	GCA_039887635.1_ASM3988763v1_genomic.fna	Taizhou, Zhejiang	2023	119	15	15
TZ-14	GCA_039887235.1_ASM3988723v1_genomic.fna	Taizhou, Zhejiang	2023	119	15	15
TZ-15	GCA_039887595.1_ASM3988759v1_genomic.fna	Taizhou, Zhejiang	2023	93	19	4
TZ-17	GCA_039887375.1_ASM3988737v1_genomic.fna	Taizhou, Zhejiang	2023	629	5	5
TZ-18	GCA_039887285.1_ASM3988728v1_genomic.fna	Taizhou, Zhejiang	2023	–	711	711
TZ-19	GCA_039887245.1_ASM3988724v1_genomic.fna	Taizhou, Zhejiang	2023	629	5	5
WX-20	GCA_039887275.1_ASM3988727v1_genomic.fna	Wuxi, Jiangsu	2023	–	607	607
WX-21	GCA_039887975.1_ASM3988797v1_genomic.fna	Wuxi, Jiangsu	2022	–	611	611
WX-22	GCA_039888005.1_ASM3988800v1_genomic.fna	Wuxi, Jiangsu	2022	93	19	4
WX-23	GCA_039887525.1_ASM3988752v1_genomic.fna	Wuxi, Jiangsu	2022	249	66	66
WX-25	GCA_039887995.1_ASM3988799v1_genomic.fna	Wuxi, Jiangsu	2022	214	181	139
WX-26	GCA_039887485.1_ASM3988748v1_genomic.fna	Wuxi, Jiangsu	2022	1611	187	65
HZ-33	GCA_039887735.1_ASM3988773v1_genomic.fna	Hangzhou, Zhejiang	2023	–	714	714
WZ-35	GCA_039888065.1_ASM3988806v1_genomic.fna	Wenzhou, Zhejiang	2015	64	17	4
WZ-38	GCA_039888245.1_ASM3988824v1_genomic.fna	Wenzhou, Zhejiang	2015	1611	187	65

fastANI v1.33[55], and only 726 genomes that exhibited ≥95% ANIs to the reference GCF_000191145.1 [PHEA-2, CP002177] were kept in downstream analysis.

Phylogenetic analysis and genomic annotation

We used EToKi align module to align all the genomes with GCF_000191145.1 [PHEA-2, CP002177]. Phylogeny of *A. pittii* was estimated using EToKi phylo module based on 460,000 SNPs after removal of recombinant regions using RecHMM. The phylogenetic trees in Figs. 1 and 2 were visualized using the iTOL v5 platform for clear and comprehensive representation[56]. The 7-gene MLST ST of each genome was in silico predicted after comparing the genomic sequences using BLASTn v2.11.0[57] onto the allelic sequences for Pasteur's schemes, both hosted on PubMLST[58]. MLST profile was then clustered into different CCs with eBurst v1.0.5[59]. For sequence types (STs) that were not documented in the PubMLST database during our study period, including those resulting from previously unreported combinations of the seven MLST gene loci or the detection of alleles, we classified them as unassignable. ARGs in each genome were predicted using AMRFinderPlus v3.11.26[60], and the virulence genes were predicted based on BLASTp searches against the VFDB 2022 release[61], with ≥ 90% identity and ≥ 70% coverage, respectively. Furthermore, the types of OCL and cps were predicted using Kaptive v2.0.3[62]. Finally, plasmids were predicted by KleTy PlasT module using its default setting[63].

Characterization of core genome and dcgMLST scheme

To identify core genes suitable for the dcgMLST scheme, we first built a set of representative genomes that retain most of the genetic diversities while removing redundancy caused by genetically virtually similar genomes. To do that, we assessed the pair-wise genetic distances of all genomes using Kssd v1.1[64], and grouped genomes into SLCs that shared ≥ 99.5% identity. One sequence with the greatest N50 value was chosen for each cluster, resulting in the final set of 217 representative genomes annotated using Prokka v1.14.6[65].

DTy pipeline was employed to build the dcgMLST based on pan-genome estimated by PEPPAN v1.0.5[22,23]. Reference sequences for each pan gene were selected by choosing one allele for each cluster of ≥ 90% identities using EToKi MLSTdb. The MLSTdb module also identifies potential paralogous genes with ≥ 90% identities, which were removed from the scheme. Based on the DTy results, we extracted a subset of 2840 core genes from them by selecting each gene that was (1) present in ≥ 95% of genomes, and (2) maintained intact open reading frames in > 94% of its alleles. These core genes were used as the basis of the dcgMLST scheme.

All genomes were then genotyped using the dcgMLST scheme by DTy again. Each allele in the dcgMLST was designated based on a 128-digit MD5 hash value of its sequence, rather than an arbitrary sequential integer from a central database in the traditional cgMLST schemes. This allowed each genome to be characterized as a collection of up to 318,505 MD5 hash values, which each represented a unique core gene allelic sequence. This method was described by Zhong et al. [22]. We then hierarchically grouped the allelic profiles of all genomes into multi-level clusters using pHierCC in its development mode and statistically evaluated the consistencies and cohesiveness of each cluster using the pHCCeval module in the pHierCC package[24]. Hierarchical levels assigned by pHierCC pipeline showed natural population structures inside *A. pittii*, which could be recognized by Silhouette's score.

Based on the species tree described above, we compared dcgMLST clusters to other methods. The MLST CCs were estimated as clusters of single-locus variants of the 7-gene MLST profiles by the eBURST algorithm implemented in the goeBURST software[59]. Furthermore, we used

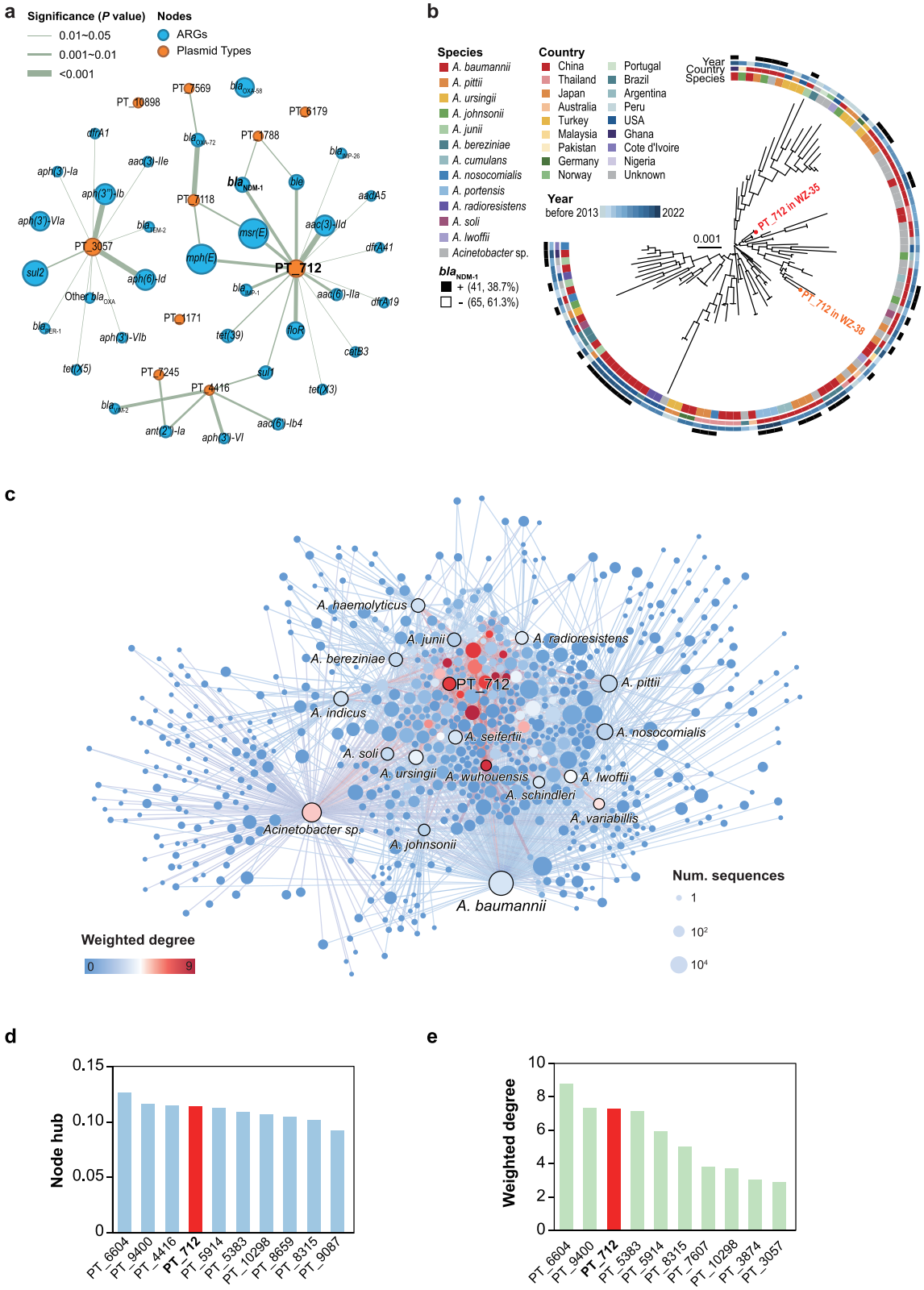

Fig. 5 | The significance of PT_712 in *Acinetobacter*. **a** The network diagram illustrating the relationship between plasmids and their associated ARGs in the complete genomes of *A. pittii*. **b** The phylogenetic tree of PT_712. The 106 PT_712 and the two collected PT_712 were used for cluster analysis. The species, geography, isolate year, and bla_{NDM-1} were annotated in the outer circle of the phylogenetic tree. The plasmid PT_712 has the characteristic of specific transmission in the genus *Acinetobacter*. And they were sporadic throughout the world and randomly carries bla_{NDM-1}. **c–e** Importance of PT_712 in the genus *Acinetobacter*. **c** Network plots of PTs and all *Acinetobacter sp*. **d** The hub values of the top 10 plasmids were shown with the bar graph, with PT_712 as the red column. **e** The weighted degree values of the top 10 plasmids were shown with the bar graph, with PT_712 as the red column. Source Data is provided as Source Data.

PopPUNK to create a database and used the bgmm (Gaussian Mixture Modeling algorithms) model to separate them into 212 clusters[52]. Finally, we also explored HierBaps to obtain the 1-level clusters, which eventually resulting in 144 results[50,51]. All clustering schemes were pairwise compared using AMI value, which can be reproduced following formula below[66]:

$$\mathrm{AMI}(U,V)=\frac{\mathrm{MI}(U,V)-E[\mathrm{MI}(U,V)]}{\max(H(U),H(V))-E[\mathrm{MI}(U,V)]}$$

$\mathrm{MI}(U, V)$ is the Mutual Information between clustering results U and V. $E[\mathrm{MI}(U, V)]$ is the expected value of Mutual Information, accounting for the influence of random factors. $H(U)$ and $H(V)$ are the entropies of clustering results U and V, respectively.

Pan-genome diversification analysis by Random Forest

Pan-genome was concluded by PEPPAN as described above, resulting in a presence/absence matrix for 18,882 ortholog gene within *A. pittii*. To discriminate the HC1100 specific genes, we passed the gene presence/absence matrix to a python module 'RandomForestClassifer' from sklearn ensemble package[67] to train a model with parameters 'n_estimators = 1000, random_state = 42'. The importance of each feature was exported by Random Forest model's attribution. The matrix and the Python code were available at https://github.com/Naclist/Wang-et-al-repo-for-A.-pittii.git and can be reproduced following the documentation.

We employed a random forest model to screen for feature genes. During the construction of the random forest model, we optimized the parameters through grid search to ensure optimal model performance. Specifically, we adjusted hyperparameters such as the depth of decision trees, the minimum number of samples required for node splitting, and the number of bootstrap iterations. Based on the trained random forest model, we extracted the feature importance scores for each gene, which reflect the degree of influence each gene has on predicting the target phenotype. To ensure the robustness of the selection, we repeated this process on an independent dataset. Subsequently, we assessed the statistical significance of the feature genes using Pearson's correlation analysis. The significance of the association score (P-value) for each gene was determined by comparing the observed association scores with a null distribution generated through Monte Carlo simulations, which modeled the random gain and loss of HC1100 unrelated traits along the clonal phylogeny. For functional inference, significantly associated genes ($P < 0.001$) were mapped to the aforementioned coding sequence pan-genome, and the alleles of each gene locus were annotated using COG (Clusters of Orthologous Groups of proteins) classification[68]. This approach ensures robust identification of genes linked to specific traits and their functional characterization.

We integrated an analysis pipeline for *A. pittii*, which includes tools of dcgMLST/HierCC-based population assignment, antibiotic resistance gene annotation, and plasmid prediction. This integrated pipeline has been deployed on PathoBase, an open platform operated by Soochow University. Registered users can upload, assemble, analyze, and download genomic data online. The pipeline is efficient, can complete analysis of a 4 MB *A. pittii* assembled genome within 5 min and raw data for 20 min.

Statistics and reproducibility

Statistical analyses were conducted using established bioinformatics pipelines and custom algorithms to ensure robustness. Hierarchical clustering (HC) of *A. pittii* genomes was performed using pHierCC with thresholds (e.g., HC450, HC1100) optimized via Silhouette scores to reflect natural population structures. Adjusted Mutual Information (AMI) was calculated to compare clustering consistency across methods (HierBaps, PopPUNK, MLST). A Random Forest model (n_estimators = 1000, grid search-optimized hyperparameters) was trained on the pan-genome presence/absence matrix (18,882 genes) to identify lineage-specific gene enrichment/depletion. Statistical significance of feature genes ($P < 0.001$) was assessed using Pearson's correlation and Monte Carlo simulations. Antimicrobial resistance prevalence across hosts and regions was analyzed with descriptive statistics.

The study utilized 750 *A. pittii* genomes (726 public, 24 newly sequenced), spanning 30 countries and 34 years, to minimize sampling bias. Redundant genomes were filtered by average nucleotide identity (ANI ≥ 98%), generating a representative subset of 217 strains. Sequencing data for new isolates were processed with EToKi and assembled using Illumina protocols (paired-end libraries, NovaSeq 6000). Quality control included N50 (> 10 KB) and size thresholds (3.69 to 4.32 MB). Phylogenetic trees were reconstructed from 460,000 SNPs, with recombination regions removed using RecHMM. All analyses, including dcgMLST allele assignment (MD5 hash-based), plasmid typing, and ARG annotation, were automated via PathoBase.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The assembled genome and raw reads for each *A. pittii* isolates have been deposited in the NCBI BioProject database under accession code PRJNA1114632. PT_712 plasmids' sequences are both available on GitHub at https://github.com/Naclist/Wang-et-al-repo-for-A.-pittii.git (DOI: 10.5281/zenodo.15301292). PT_712 in WZ-35 genome can be found under Nucleotide accession number NZ_CP168074.1. Source data of each Figure and Table can be found in Supplementary Data, Supplementary Table, and are available from the corresponding author on reasonable request. Source data are provided as Supplementary Data .

Code availability

PathoBase link: https://pathobase.iotabiome.com/. Documentation for PathoBase and code for PlasT are both available on GitHub at https://github.com/Naclist/Wang-et-al-repo-for-A.-pittii.git (https://doi.org/10.5281/zenodo.15301292).

Received: 21 October 2024; Accepted: 1 May 2025;
Published online: 13 May 2025

References

1. Castanheira, M., Mendes, R. E. & Gales, A. C. Global epidemiology and mechanisms of resistance of *Acinetobacter baumannii-calcoaceticus* complex. *Clin. Infect. Dis.* **76**, S166–S178 (2023).
2. Müller, C. et al. A global view on carbapenem-resistant *Acinetobacter baumannii*. *mBio* **14**, e0226023 (2023).
3. Wisplinghoff, H. et al. Nosocomial bloodstream infections due to *Acinetobacter baumannii*, *Acinetobacter pittii* and *Acinetobacter nosocomialis* in the United States. *J. Infect.* **64**, 282–290 (2012).
4. Chusri, S. et al. Clinical outcomes of hospital-acquired infection with *Acinetobacter nosocomialis* and *Acinetobacter pittii*. *Antimicrob. Agents Chemother.* **58**, 4172–4179 (2014).
5. Wang, X., Chen, T., Yu, R., Lü, X. & Zong, Z. *Acinetobacter pittii* and *Acinetobacter nosocomialis* among clinical isolates of the *Acinetobacter calcoaceticus-baumannii* complex in Sichuan, China. *Diagn. Microbiol. Infect. Dis.* **76**, 392–395 (2013).
6. Nemec, A. et al. Genotypic and phenotypic characterization of the *Acinetobacter calcoaceticus–Acinetobacter baumannii* complex with the proposal of *Acinetobacter pittii* sp. nov. (formerly *Acinetobacter* genomic species 3) and *Acinetobacter nosocomialis* sp. nov. (formerly *Acinetobacter* genomic species 13TU). *Res. Microbiol.* **162**, 393–404 (2011).
7. Jain, A. L. et al. Characteristics of invasive *Acinetobacter* species isolates recovered in a pediatric academic center. *BMC Infect. Dis.* **16**, 346 (2016).

8. Jeong, S. et al. Identification of *Acinetobacter* species using matrix-assisted laser desorption ionization-time of flight mass spectrometry. *Ann. Lab. Med.* **36**, 325–334 (2016).
9. Vasconcelos, A. T. R. et al. The changing epidemiology of *Acinetobacter* spp. producing OXA carbapenemases causing bloodstream infections in Brazil: a BrasNet report. *Diagn. Microbiol. Infect. Dis.* **83**, 382–385 (2015).
10. Pailhoriès, H., Tiry, C., Eveillard, M. & Kempf, M. *Acinetobacter pittii* isolated more frequently than *Acinetobacter baumannii* in blood cultures: the experience of a French hospital. *J. Hosp. Infect.* **99**, 360–363 (2018).
11. He, X. et al. Genetic characterization of plasmid-borne bla_{OXA-58} and bla_{OXA-72} in *Acinetobacter pittii* in Shaanxi, China. *J. Glob. Antimicrob. Resist.* **38**, 167–172 (2024).
12. Worku, S. et al. Molecular characterization of carbapenemase and extended spectrum beta-lactamase producing *Acinetobacter baumannii* isolates causing surgical site infections in Ethiopia. *BMC Infect. Dis.* **24**, 459 (2024).
13. Roy, S., Morita, D., Bhattacharya, S., Dutta, S. & Basu, S. Novel sequence type of carbapenem-resistant *Acinetobacter pittii* ST1451 with enhanced virulence isolated from septicaemic neonates in India. *J. Antimicrob. Chemother.* **79**, 779–783 (2024).
14. Bello-López, E. et al. *Acinetobacter pittii*: the emergence of a hospital-acquired pathogen analyzed from the genomic perspective. *Front. Microbiol.* **15**, 1412775 (2024).
15. Chapartegui-González, I., Lázaro-Díez, M. & Ramos-Vivas, J. Genetic resistance determinants in clinical *Acinetobacter pittii* genomes. *Antibiotics (Basel)* **11**, 676 (2022).
16. Yang, Y. et al. Metagenomics of high-altitude groundwater reveal different health risks associated with antibiotic-resistant pathogens and bacterial resistome in the latitudinal gradient. *Water Res.* **262**, 122032 (2024).
17. Jia, W. et al. Antibiotic resistome in landfill leachate and impact on groundwater. *Sci. Total Environ.* **927**, 171991 (2024).
18. Li, J. et al. *Acinetobacter pittii*, an emerging new multi-drug resistant fish pathogen isolated from diseased blunt snout bream (Megalobrama amblycephala Yih) in China. *Appl. Microbiol. Biotechnol.* **101**, 6459–6471 (2017).
19. Carvalheira, A., Casquete, R., Silva, J. & Teixeira, P. Prevalence and antimicrobial susceptibility of *Acinetobacter* spp. isolated from meat. *Int. J. Food Microbiol.* **243**, 58–63 (2017).
20. Pasteran, F. et al. Emergence of genetically unrelated NDM-1-producing *Acinetobacter pittii* strains in Paraguay. *J. Antimicrob. Chemother.* **69**, 2575–2578 (2014).
21. Souhail, B. et al. First report of *Acinetobacter pittii* acute community-acquired pneumonia in an immunocompetent patient in France following a heat wave. *BMC Infect. Dis.* **24**, 35 (2024).
22. Zhong, L. et al. Distributed genotyping and clustering of *Neisseria* strains reveal continual emergence of epidemic meningococcus over a century. *Nat. Commun.* **14**, 7706 (2023).
23. Zhou, Z., Charlesworth, J. & Achtman, M. Accurate reconstruction of bacterial pan- and core genomes with PEPPAN. *Genome Res.* **30**, 1667–1679 (2020).
24. Zhou, Z., Charlesworth, J. & Achtman, M. HierCC: a multi-level clustering scheme for population assignments based on core genome MLST. *Bioinformatics* **37**, 3645–3646 (2021).
25. Chandu, D., Kumar, A. & Nandi, D. PepN, the major Suc-LLVY-AMC-hydrolyzing enzyme in *Escherichia coli*, displays functional similarity with downstream processing enzymes in archaea and eukarya. *J. Biol. Chem.* **278**, 5548–5556 (2003).
26. Imber, M. et al. The aldehyde dehydrogenase AldA contributes to the hypochlorite defense and is redox-controlled by protein S-bacillithiolation in *Staphylococcus aureus*. *Redox Biol.* **15**, 557–568 (2018).
27. Amor, P. A. & Whitfield, C. Molecular and functional analysis of genes required for expression of group IB K antigens in *Escherichia coli*: characterization of the *his*- region containing gene clusters for multiple cell-surface polysaccharides. *Mol. Microbiol.* **26**, 145–161 (1997).
28. Genevaux, P., Schwager, F., Georgopoulos, C. & Kelley, W. L. The *djlA* gene acts synergistically with *dnaJ* in promoting *Escherichia coli* growth. *J. Bacteriol.* **183**, 5747–5750 (2001).
29. Guffanti, A. A., Wei, Y., Rood, S. V. & Krulwich, T. A. An antiport mechanism for a member of the cation diffusion facilitator family: divalent cations efflux in exchange for K^+ and H^+. *Mol. Microbiol.* **45**, 145–153 (2002).
30. Fu, B. et al. Metal-induced sensor mobilization turns on affinity to activate regulator for metal detoxification in live bacteria. *Proc. Natl. Acad. Sci. USA***117**, 13248–13255 (2020).
31. Guerrero, P. et al. Characterization of the BaeSR two-component system from *Salmonella typhimurium* and its role in ciprofloxacin-induced *mdtA* expression. *Arch. Microbiol.* **194**, 453–460 (2012).
32. Rong, Q. et al. Sb(III) resistance mechanism and oxidation characteristics of *Klebsiella aerogenes* X. *Chemosphere* **293**, 133453 (2022).
33. Oshima, T. & Biville, F. Functional identification of *ygiP* as a positive regulator of the *ttdA-ttdB-ygjE* operon. *Microbiology* **152**, 2129–2135 (2006).
34. Wilson, R. L. & Stauffer, G. V. DNA sequence and characterization of GcvA, a LysR family regulatory protein for the *Escherichia coli* glycine cleavage enzyme system. *J. Bacteriol.* **176**, 2862–2868 (1994).
35. Suo, B. et al. Comparative transcriptomic analysis of *Staphylococcus aureus* reveals the genes involved in survival at low temperature. *Foods* **11**, 996 (2022).
36. Horiyama, T., Nikaido, E., Yamaguchi, A. & Nishino, K. Roles of *Salmonella* multidrug efflux pumps in tigecycline resistance. *J. Antimicrob. Chemother.* **66**, 105–110 (2011).
37. Pragman, A. A., Yarwood, J. M., Tripp, T. J. & Schlievert, P. M. Characterization of virulence factor regulation by SrrAB, a two-component system in *Staphylococcus aureus*. *J. Bacteriol.* **186**, 2430–2438 (2004).
38. Stephen, P. et al. Structure of *Escherichia coli* arginyl-tRNA synthetase in complex with $tRNA^{Arg}$: pivotal role of the D-loop. *J. Mol. Biol.* **430**, 1590–1606 (2018).
39. Grill, S. et al. Identification and characterization of two *Streptomyces davawensis* riboflavin biosynthesis gene clusters. *Arch. Microbiol.* **188**, 377–387 (2007).
40. Hicks, K. A., O'Leary, S. E., Begley, T. P. & Ealick, S. E. Structural and mechanistic studies of HpxO, a novel flavin adenine dinucleotide-dependent urate oxidase from *Klebsiella pneumoniae*. *Biochemistry* **52**, 477–487 (2013).
41. Chatterjee, S., Mondal, A., Mitra, S. & Basu, S. *Acinetobacter baumannii* transfers the bla_{NDM-1} gene via outer membrane vesicles. *J. Antimicrob. Chemother.* **72**, 2201–2207 (2017).
42. Azizi, O. et al. Class 1 integrons in non-clonal multidrug-resistant *Acinetobacter baumannii* from Iran, description of the new bla_{IMP-55} allele in In1243. *J. Med. Microbiol.* **65**, 928–936 (2016).
43. Chopjitt, P. et al. Genomic characterization of clinical extensively drug-resistant *Acinetobacter pittii* isolates. *Microorganisms* **9**, 242 (2021).
44. Ding, Z. et al. Phenotypic and genotypic characteristics of a tigecycline-resistant *Acinetobacter pittii* isolate carrying bla_{NDM-1} and the novel bla_{OXA} allelic variant $bla_{OXA-1045}$. *Front. Microbiol.* **13**, 868152 (2022).
45. Wunderlich, A. et al. Carbapenem resistance in *Acinetobacter pittii* isolates mediated by metallo-β-lactamases. *J. Antimicrob. Chemother.* **78**, 488–496 (2023).

46. Zhang, L. et al. Phenotypic variation and carbapenem resistance potential in OXA-499-producing *Acinetobacter pittii*. *Front. Microbiol.* **11**, 1134 (2020).
47. Pulami, D., Kämpfer, P. & Glaeser, S. P. High diversity of the emerging pathogen *Acinetobacter baumannii* and other *Acinetobacter* spp. in raw manure, biogas plants digestates, and rural and urban wastewater treatment plants with system specific antimicrobial resistance profiles. *Sci. Total Environ.* **859**, 160182 (2023).
48. Li, S., He, Y., Mann, D. A. & Deng, X. Global spread of *Salmonella enteritidis* via centralized sourcing and international trade of poultry breeding stocks. *Nat. Commun.* **12**, 5109 (2021).
49. Zhou, Z. et al. The EnteroBase user's guide, with case studies on *Salmonella* transmissions, *Yersinia pestis* phylogeny, and *Escherichia* core genomic diversity. *Genome Res.* **30**, 138–152 (2020).
50. Tonkin-Hill, G., Lees, J. A., Bentley, S. D., Frost, S. D. W. & Corander, J. RhierBAPS: an R implementation of the population clustering algorithm hierBAPS. *Wellcome Open Res.* **3**, 93 (2018).
51. Cheng, L. et al. Hierarchical and spatially explicit clustering of DNA sequences with BAPS software. *Mol. Biol. Evol.* **30**, 1224–1228 (2013).
52. Lees, J. A. et al. Fast and flexible bacterial genomic epidemiology with PopPUNK. *Genome Res.* **29**, 304–316 (2019).
53. Shelenkov, A., Petrova, L., Zamyatin, M., Mikhaylova, Y. & Akimkin, V. Diversity of international high-risk clones of *Acinetobacter baumannii* revealed in a russian multidisciplinary medical center during 2017-2019. *Antibiotics (Basel)* **10**, 1009 (2021).
54. CLSI. *Performance Standards for Antimicrobial Susceptibility Testing, M100*, 33rd ed. (Clinical and Laboratory Standards Institute, 2023).
55. Jain, C., Rodriguez-R, L. M., Phillippy, A. M., Konstantinidis, K. T. & Aluru, S. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat. Commun.* **9**, 5114 (2018).
56. Letunic, I. & Bork, P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* **49**, W293–W296 (2021).
57. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).
58. Jolley, K. A., Bray, J. E. & Maiden, M. C. J. Open-access bacterial population genomics: BIGSdb software, the PubMLST.org website and their applications. *Wellcome Open Res.* **3**, 124 (2018).
59. Francisco, A. P., Bugalho, M., Ramirez, M. & Carriço, J. A. Global optimal eBURST analysis of multilocus typing data using a graphic matroid approach. *BMC Bioinforma.* **10**, 152 (2009).
60. Feldgarden, M. et al. AMRFinderPlus and the Reference Gene Catalog facilitate examination of the genomic links among antimicrobial resistance, stress response, and virulence. *Sci. Rep.* **11**, 12728 (2021).
61. Liu, B., Zheng, D., Jin, Q., Chen, L. & Yang, J. VFDB 2019: a comparative pathogenomic platform with an interactive web interface. *Nucleic Acids Res.* **47**, D687–D692 (2019).
62. Lam, M. M. C., Wick, R. R., Judd, L. M., Holt, K. E. & Wyres, K. L. Kaptive 2.0: updated capsule and lipopolysaccharide locus typing for the *Klebsiella pneumoniae* species complex. *Microb. Genom.* **8**, 000800 (2022).
63. Li, H. et al. KleTy: integrated typing scheme for core genome and plasmids reveals repeated emergence of multi-drug resistant epidemic lineages in *Klebsiella* worldwide. *Genome Med.* **16**, 130 (2024).
64. Yi, H., Lin, Y., Lin, C. & Jin, W. Kssd: sequence dimensionality reduction by k-mer substring space sampling enables real-time large-scale datasets analysis. *Genome Biol.* **22**, 84 (2021).
65. Seemann, T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* **30**, 2068–2069 (2014).
66. Romano, S., Vinh, N. X., Bailey, J. & Verspoor, K. Adjusting for Chance Clustering Comparison Measures. *J. Mach. Learn Res*. **17**, 1–32 (2016).
67. Breiman, L. Random forests. *Mach. Learn.* **45**, 5–32 (2001).
68. Tatusov, R. L., Galperin, M. Y., Natale, D. A. & Koonin, E. V. The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res.* **28**, 33–36 (2000).

Acknowledgements

We thank Mr. Shenghai Wu, Miss Renjing Hu for providing *A. pittii* strains. We also extend our appreciation to Iotabiome Biotechnology Inc., Suzhou, China for sequencing and assistance, and Key Discipline of Zhejiang Province in Medical Technology (First Class, Category A). This research was supported by the Fundamental Public Welfare Research Program of Zhejiang Province (No. LTGY24H160042), and Fundamental Public Welfare Research Program of Wenzhou Science and Technology Bureau (No. Y20240142).

Author contributions

J.D., Z.Z., S.L., and J.Y. designed the study. S.W., Y.Z., Y.W., K.T., D.W., J.H., P.W., and S.Y. performed experiments, analyses, and wrote the initial manuscript. S.W. and Y.Z. designed and wrote the pipeline. S.W., Y.Z., and S.L. discussed the results. J.D., Z.Z., S.L., and J.Y. directed the project and edited the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s42003-025-08156-y.

Correspondence and requests for materials should be addressed to Jie Yan, Shengkai Li, Zhemin Zhou or Jimei Du.

Peer review information *Communications Biology* thanks Nabil-Fareed Alikhan, Alexandr Nemec, and the other, anonymous, reviewer for their contribution to the peer review of this work. Primary Handling Editors: Robert Kingsley and David Favero. A peer review file is available.