Tracking of Antibiotic Resistance Transfer and Rapid Plasmid Evolution in a Hospital Setting by Nanopore Sequencing

Silke Peter,a,b Mattia Bosio,c,i Caspar Gross,d Daniela Bezdan,c Javier Gutierrez,* Philipp Oberhettinger,a,b Jan Liese,a,b Wichard Vogel,f Daniela Dörfel,f,g Lennard Berger,d Matthias Marschal,a,b Matthias Willmann,a,b Ivo Gut,h Marta Gut,h,i Ingo Autenrieth,a,b Stephan Ossowskic,d,h

aInstitute of Medical Microbiology and Hygiene, University of Tübingen, Tübingen, Germany
bGerman Center for Infection Research (DZIF), Partner Site Tübingen, Tübingen, Germany
cCentre for Genomic Regulation (CRG), The Barcelona Institute of Science and Technology, Barcelona, Spain
dInstitute of Medical Genetics and Applied Genomics, University of Tübingen, Tübingen, Germany
eCNAG-CRG, Centre for Genomic Regulation (CRG), Barcelona Institute of Science and Technology (BIST), Barcelona, Spain
fMedical Center, Department of Hematology, Oncology, Immunology, Rheumatology & Pulmonology, University of Tübingen, Tübingen, Germany
gClinical Collaboration Unit Translational Immunology, German Cancer Consortium (DKTK) and German Cancer Research Center (DKFZ), Partner Site Tübingen, Tübingen, Germany
hUniversitat Pompeu Fabra (UPF), Barcelona, Spain
iBarcelona Supercomputing Center, BSC, Barcelona, Spain

Mattia Bosio and Caspar Gross contributed equally to this work.

ABSTRACT Infections with multidrug-resistant bacteria often leave limited or no treatment options. The transfer of antimicrobial resistance genes (ARG) carrying plasmids between bacterial species by horizontal gene transfer represents an important mode of expansion of ARGs. Here, we demonstrate the application of Nanopore sequencing in a hospital setting for monitoring transfer and rapid evolution of antibiotic resistance plasmids within and across multiple species. In 2009, we experienced an outbreak with extensively multidrug-resistant Pseudomonas aeruginosa harboring the carbapenemase-encoding blaIMP-8 gene. In 2012, the first Citrobacter freundii and Citrobacter cronae strains harboring the same gene were detected. Using Nanopore and Illumina sequencing, we conducted comparative analysis of all blaIMP-8 bacteria isolated in our hospital over a 6-year period (n = 54). We developed the computational platform plasmIDent for Nanopore-based characterization of clinical isolates and monitoring of ARG transfer, comprising de novo assembly of genomes and plasmids, plasmid circularization, ARG annotation, comparative genome analysis of multiple isolates, and visualization of results. Using plasmIDent, we identified a 40-kb plasmid carrying blaIMP-8 in P. aeruginosa and C. freundii, verifying the plasmid transfer. Within C. freundii, the plasmid underwent further evolution and plasmid fusion, resulting in a 164-kb megaplasmid, which was transferred to C. cronae. Multiple rearrangements of the multidrug resistance gene cassette were detected in P. aeruginosa, including deletions and translocations of complete ARGs. In summary, plasmid transfer, plasmid fusion, and rearrangement of the ARG cassette mediated the rapid evolution of opportunistic pathogens in our hospital. We demonstrated the feasibility of near-real-time monitoring of plasmid evolution and ARG transfer in clinical settings, enabling successful countermeasures to contain plasmid-mediated outbreaks.

IMPORTANCE Infections with multidrug-resistant bacteria represent a major threat to global health. While the spread of multidrug-resistant bacterial clones is frequently studied in the hospital setting, surveillance of the transfer of mobile genetic elements between different bacterial species was difficult until recent advances in sequencing technologies. Nanopore sequencing technology was applied to track an-
timicrobial gene transfer in a long-term outbreak of multidrug-resistant *Pseudomonas aeruginosa*, *Citrobacter freundii*, and *Citrobacter cronae* in a German hospital over 6 years. We developed a novel computational pipeline, *pathoLogic*, which enables *de novo* assembly of genomes and plasmids, antimicrobial resistance gene annotation and visualization, and comparative analysis. Applying this approach, we detected plasmid transfer between different bacterial species as well as plasmid fusion and frequent rearrangements of the antimicrobial resistance gene cassette. This study demonstrated the feasibility of near-real-time tracking of plasmid-based antimicrobial resistance gene transfer in hospitals, enabling countermeasures to contain plasmid-mediated outbreaks.

**KEYWORDS** plasmids, Nanopore, long read, IMP-8, *Pseudomonas aeruginosa*, *pathoLogic*, *plasmIDent*, genome assembly, horizontal gene transfer, Nanopore sequencing, antimicrobial resistance, plasmid-mediated resistance, surveillance studies

The increase in the number of multidrug-resistant (MDR) bacterial strains has led organizations such as the World Health organization (WHO) and the U.S. Centers for Disease Control and Prevention (CDC) to categorize MDR bacteria as representing a major public health problem (1). Infection of patients with MDR bacteria often leaves only very limited or even no treatment options, thus posing a potentially life-threatening risk to individual patients, in particular, those in intensive care units (2, 3). In addition, infection control measures to prevent spreading are required, resulting in increased efforts with respect to patient care and increased costs for health care providers and public health care systems (4, 5). Although action is needed on different national and international levels, understanding colonization, infection, and transmission routes of these MDR-resistant bacteria in the local hospital setting represents a crucial initial step toward implementation of harmonized, successful strategies to combat infections caused by MDR bacteria (1, 4).

Next-generation sequencing (NGS) has become widely available and has been used successfully to resolve outbreaks and determine transmission routes (see, e.g., the review in reference 6). However, both the clonal transmission of MDR bacteria and the spread of multidrug resistance by horizontal gene transfer (HGT) between different bacterial species represent important modes of expansion of antimicrobial resistance (AMR) genes (7). Although multidrug resistance plasmids and plasmid transfer have been studied in hospital settings, their interrogation is not part of routine infection control practice. Moreover, methods of plasmid characterization and comparisons based on short-read sequences are error prone and unreliable, particularly when larger (>50-kb) plasmids are involved (8), while long-read *de novo* assembly-based plasmid analysis is currently limited to large centers with access to Pacific Biosystems (PacBio) Sequel sequencers (see, e.g., references 9 and 10). Recently, the MinION long-read sequencer (Oxford Nanopore Technologies [ONT]) became more widely available, facilitating fast and inexpensive analysis of multidrug resistance plasmids and horizontal gene transfer and evolution of plasmid-born antimicrobial resistance (AMR) (11, 12). Thus, the technology is potentially suitable for application within the hospital setting. In recent publications, Dong et al. examined the microevolution of *bla*KPC-harboring plasmids in three clinical isolates applying Nanopore technology (13), while Lemon et al. optimized the Nanopore sequencing laboratory workflow and analyzed plasmids from three clinical isolates (11). Long-read sequences substantially increase the contiguity of *de novo* assemblies by spanning repeat regions, resulting in finished microbial genome and plasmid assemblies (14). However, due to the high error rates of Nanopore sequencing, hybrid assemblers such as hybridSPAdes (15) and Unicycler (16) combine long and short reads to achieve a high base level of the accuracy needed for the correct identification of AMR-related genes and variants. In the present study, we aimed to evaluate the application of Nanopore sequencing technology in a hospital setting and
to demonstrate the feasibility of monitoring transfer and rapid evolution of antibiotic resistance plasmids within and across multiple species.

Starting in 2009, our hospital experienced an outbreak caused by an extensively multidrug-resistant *Pseudomonas aeruginosa* clone (17). The strain harbored a carbapenemase-encoding gene (*bla*<sub>IMP-8</sub>), which renders most beta-lactams ineffective, including carbapenems, an antibiotics class of last resort (18). Extensive infectious disease interventions and the establishment of a rectal screening program to identify colonized patients led to a reduction of cases. However, in March 2012, we detected the first *Citrobacter freundii* strain harboring the same carbapenemase-encoding *bla*<sub>IMP-8</sub> gene (19), approximately 2.5 years after the first *P. aeruginosa* *bla*<sub>IMP-8</sub> gene had been detected. Shortly after, the carbapenemase was detected in *Citrobacter cronae* (20). Since *bla*<sub>IMP-8</sub> is rarely encountered in Europe and Germany (21, 22) and has not yet been detected in rectal screening swabs from patients submitted to our hospital for the first time, we hypothesized that horizontal gene transfer had occurred within bacterial strains circulating in our hospital. Therefore, we conducted a sequencing study that included all multidrug-resistant bacteria harboring the *bla*<sub>IMP-8</sub> gene isolated in our hospital over a 6-year period, including patient and environmental isolates. We developed and established a bioinformatics pipeline in order (i) to determine the sequence of the *bla*<sub>IMP-8</sub>-harboring plasmids and characterize all of the AMR genes contained, (ii) to identify potential events of transmission of the plasmids between species, and (iii) to characterize the evolutionary dynamics of the plasmids.

**RESULTS**

Comprehensive analysis platform for antibiotic resistance gene-carrying plasmids. We have developed a comprehensive computational platform for the genomic analysis of clinical isolates and the monitoring of antibiotic resistance gene transfer. *pathoLogic* comprises a hybrid de novo assembly pipeline generating finished genomes and plasmids and performing genome polishing, quality control (QC), annotation, and comparative genome analysis of multiple isolates, as well as visualization of results (Fig. 1). Furthermore, *pathoLogic* integrates the *plasmIDent* method, which confirms the circularity of putative plasmids by ring closure using long reads, performs AMR gene annotation, calculates various sequence properties (e.g., GC content and GC skew and coverage depth), and creates a circular visualization of the annotated plasmid. Finally, sequences of plasmids from multiple isolates of the same or different species are compared in order to identify horizontal gene transfers, structural variations (e.g., AMR gene presence/absence), and point mutations, which can further be utilized for phylogenetic or transmission analysis. *pathoLogic*, *plasmIDent* and a graphical user interface (GUI) are freely available on github (*plasmIDent* pipeline, https://github.com/imgag/plasmIDent; *pathoLogic* pipeline, https://github.com/imgag/pathoLogic).

Characterization of study isolates. In our study, we included all *bla*<sub>IMP-8</sub> AMR gene-positive strains isolated in our hospital from patients or patient-related environmental water sources in the hemato-oncology department over a period of 6 years (*n* = 54). This also comprised the previously reported *P. aeruginosa* outbreak clones (*n* = 34) (17) and one *C. freundii* *bla*<sub>IMP-8</sub> isolate (19), for which Illumina short-read data are available (https://www.ebi.ac.uk/ena/browser/view/PRJEB31907). In order to obtain finished genomes and circularized plasmids, long-read Nanopore sequencing was conducted with all *Citrobacter freundii* (*n* = 8) and *Citrobacter cronae* (*n* = 1) isolates and selected *P. aeruginosa* (*n* = 5) isolates representing different time points (see Table S1 in the supplemental material). Applying the *pathoLogic* pipeline described above enabled us to generate high-quality genomes for all samples. We were able to generate a single circular chromosome along with several circular plasmids in 5 of the 14 samples (Table S2). All other assemblies also had a few large contigs, as indicated by a high NG75 value. Samples with a lower depth of coverage of Nanopore reads (e.g., isolate 9_E_CF) also resulted in more-fragmented assemblies.

Plasmid content and phylogeny of the study isolates. For the first 2.5 years, we observed *bla*<sub>IMP-8</sub> in *P. aeruginosa* isolates from only 26 patients before we first detected
C. freundii and C. cronae carrying bla\textsubscript{IMP-8} (Fig. 2A). The plasmids with relevance to the dynamics of the bla\textsubscript{IMP-8} plasmid evolution are displayed in Fig. 2B. The complete plasmid content of all isolates is summarized in Table S3.

In P. aeruginosa isolates, we detected a 40-kb plasmid carrying the bla\textsubscript{IMP-8} gene (plasmid A, blue). In C. freundii isolates, bla\textsubscript{IMP-8} Plasmid A was found in addition to an 88-kb plasmid (plasmid B, green) without a carbapenemase-encoding gene. Surprisingly, in the C. cronae isolate, a large 164-kb plasmid harboring the bla\textsubscript{IMP-8} gene was detected (plasmid C, red) without any evidence of the presence of plasmid A or plasmid...
FIG 2 (A) Timeline of isolation of \( \text{bla}_{\text{IMP-8}} \) Gram-negative bacteria in the hemato-oncology department over 6 years. Bars represent isolates from patients and the length of their stay in the hospital. Patient 21 was seen only in the outpatient department, marked with an “O.” Environmental isolates are marked with an “X” at the date of isolation. The introduction of a rectal screening program is marked with a black arrow. (B) Overview of plasmids with relevance to the evolution of the \( \text{bla}_{\text{IMP-8}} \) plasmid found in \( \text{P. aeruginosa} \), \( \text{C. freundii} \), and \( \text{C. cronae} \). (C) Maximum likelihood phylogeny of \( \text{Citrobacter} \) species included in the study \((n = 9)\). The \( \text{Citrobacter freundii} \) strains formed two clusters, CF1 \((n = 5)\) and CF2 \((n = 3)\). Strains of cluster CF2 harbored a chromosomal transposon region (black triangle) homologous to the regions of plasmid C. \( \text{C. cronae} \) clustered with the closely related \( \text{C. werkmanii} \) NBRC105721 and DSM17579 strains. The scale bar shows the expected number of nucleotide changes per site. PA, \( \text{P. aeruginosa} \); CF, \( \text{Citrobacter freundii} \); CC, \( \text{Citrobacter cronae} \).
The structures and circular nature of the three plasmids were confirmed by remapping the long-read sequences, resulting in continuous read coverage along the plasmids without breakpoints.

Phylogenetic analysis showed that all of the _P. aeruginosa_ strains were closely related and belonged to a single cluster, indicating clonal spread (data not shown). All isolated _P. aeruginosa_ strains were found to belong to sequence type 308 (ST308). In contrast, the maximum likelihood phylogeny of the _Citrobacter_ isolates revealed a phylogenetically more diverse picture (Fig. 2C). The _C. freundii_ isolates formed two clusters, Cf1 (n = 5) and Cf2 (n = 3), which were clearly distinct (Fig. 2C). Both clusters contained plasmids A and B. Isolates of cluster Cf2 contained an additional plasmid G (Table S3) and a region containing parts of the Tn3 family transposons localized on the chromosome absent in cluster Cf1, which is further described below.

**Comparative genomic analysis and annotation of plasmids.** Next, we performed multiple-sequence alignment of the generated reference sequences of plasmids A, B, and C (Fig. 3). To better understand the chronological order of the horizontal gene transfer (HGT) and fusion events, we first performed an in-depth annotation of plasmid features, including antimicrobial resistance genes, transposons, origin of replication, and GC content (Fig. 3). Plasmid A, which contains the _bla_{IMP-8}_ gene, had average GC content of 59% (Fig. 3, green inner circle). The _bla_{IMP-8}_ gene was located on a class 1 integron together with eight additional antimicrobial resistance genes (Fig. 3, bottom). The integron comprised the _intI1_ integrase gene and AMR genes _bla_{OXA-10}, aac(6)-lb, _bla_{IMP-8}, qacH, aph(3’)-IXV, aadA10, blp_DCA, _sul_._ Plasmid B had a size of approximately 88 kb and substantially lower (~50%) GC content than plasmid A and lacked the _bla_{IMP-8}_ integron. The largest plasmid, plasmid C, with a size of 164 kb, was composed of the entirety of plasmid A, including the class I integron harboring the AMR genes, and plasmid B, as well as two large stretches containing the duplicated regions D1 and D2 (Fig. 3). Therefore, plasmid C most likely resulted from a fusion of plasmids A and B. The two duplicated regions between plasmids A and B harbored a duplicated region (marked in Fig. 3 with a black arrow) composed of parts of Tn3 family transposons, three IS6 family elements, and several AMR genes. Two additional regions containing parts of transposons of the Tn3 family interspersed with additional AMR genes extended one of the fusion regions.

Results of a similarity search for all identified plasmids using NCBI Microbial Nucleotide BLAST are shown in Table S3A. Notably, we found plasmid SDENCHOLpb, which is highly similar (97% identity) to 63% of the sequence of plasmid A. However, plasmid SDENCHOLpb, which was isolated from _Sterolibacterium denitrificans_, lacks the resistance gene cassette found in plasmid A (see Fig. S3 in the supplemental material). SDENCHOLpb was sampled in close geographical proximity to our hospital (the distance from Freiburg, Germany, to Tübingen, Germany, is around 120 km).

**Plasmid content of isolates and plasmid fusion.** In order to determine the plasmid content of all studied isolates, we realigned the Illumina short-read sequences using as a reference assembled plasmid C, which comprises the sequences of plasmids A and B and the duplicated regions D1 and D2 (Fig. 3). The coverage for each strain is displayed in Fig. 4. All _P. aeruginosa_ isolates contained only plasmid A and not plasmid B or C. Sequencing reads of _P. aeruginosa_ that mapped to a small section of the transposon-containing region most likely originated from the chromosome. The picture is more complex for the _Citrobacter_ species, which could be divided into three groups. The _C. cronae_ (28_P_CC) strain contained the complete C plasmid, which was homogeneously covered. The _C. freundii_ isolates formed two groups, one group with plasmid A and B and the second group containing plasmid A and B as well as coverage of the transposon-containing regions. The two groups were found to be identical, with clusters Cf1 and Cf2 distinguished by phylogenetic analysis of the chromosomes.

We further investigated the read coverage distribution for the _C. freundii_ isolates to determine if they harbor only copies of plasmids A and B or instead harbor a combination of copies of plasmids A, B, and C. No continuous short or long reads could be
detected spanning the breakpoints between the plasmid A sequence and the plasmid B sequence and duplicated regions D1 and D2 in either of the two C. freundii clusters (Fig. 4; red lines indicate the breakpoints), suggesting that the short reads mapping between A and B originated from a chromosomal integration of the transposon-
containing regions. Annotation of the assembled chromosomes of *C. freundii* and *C. cronae* isolates confirmed that cluster Cf2 contained the transposon sequence within the chromosomal scaffold whereas cluster Cf1 and *C. cronae* did not (Fig. S1). We conclude that both Cf1 and Cf2 harbor copies only of plasmids A and B, but not of plasmid C, and that Cf2 harbors a copy of the transposon-containing region in the chromosome.

Only isolates of cluster Cf2 show a complete “smear” in the coverage plot across the whole transposon-containing region (13_E_CF, 34_P_CF, and 38_P_CF). In isolates of cluster Cf1, however, we observed only partial coverage of the transposon-containing region (for Cf1.1, 3 isolates, including 9_E_CF, 29_P_CF, and 27_P_CF) or almost no coverage (for Cf1.2, 2 isolates, including 32_P_CF and 30_P_CF). Interestingly, the regions distinguishing the Cf1.1 and Cf1.2 subclasses harbor a mercury resistance
operon (23) present in Cf1.1 but absent in Cf1.2. Pairwise alignment to the full genomes using the nucmer aligner confirmed that these genes are located on plasmid J (Table S3) in isolate 29_P_CF and on noncircular contigs in the other two isolates (9_E_CF and 27_P_CF) of group Cf1.1 (Fig. S1).

In summary, our phylogenetic analysis as well as the comprehensive plasmid annotations indicated that the C. freundii isolates in the Cf1 and Cf2 clusters represent different clones with a mean core single nucleotide polymorphism (SNP) distance of 41,825 nucleotides (minimum, 41,819; maximum, 41,836) and should be treated as separate entities in the identification of plasmid-born horizontal gene transfers.

Deletion and transposition of AMR genes in P. aeruginosa. While the P. aeruginosa isolates homogeneously contained only plasmid A, we observed that the resistance gene cassette for some isolates was not continuously covered with short reads in the reference alignment shown in Fig. 4 (see also Fig. S2). Using short-read-based and long-read-based structural variant detection methods, we identified two types of rearrangement events. First, we found various deletions of resistance genes within the resistance gene cassette in 12 strains, indicated by zero coverage (Fig. S2, white areas flanked by red brackets). Analysis of the resistance genes annotated by ResFinder or CARD on the respective plasmid scaffolds confirmed that these deletions correspond to missing AMR genes in the respective strains (Table S4). Furthermore, all deletions were found to span the sequence from exactly the 5’ end to the 3’ end, consisting of one AMR gene plus the flanking IS element.

Moreover, comparing the resistance gene cassettes of P. aeruginosa isolates 37_P_PA and 39_P_PA, we detected breakpoints between AMR genes without a corresponding drop of coverage, indicating translocation events corresponding to single AMR genes. We therefore performed a multiple-sequence alignment of the class I integrons of the 5 P. aeruginosa isolates for which Nanopore sequences were generated, as the long-read data facilitate the highest-confidence assemblies. Indeed, we identified two structurally different versions of the resistance gene cassette, termed RSC1 and RSC2, the latter likely the result of multiple transposition and deletion events (Fig. 4B). Four isolates harbored wild-type cassette RSC1, while one isolate harbored RSC2. Finally, we aligned the short reads of all 49 P. aeruginosa isolates against the breakpoints distinguishing RSC1 and RSC2. We identified 21 isolates most similar to RSC1 and 9 isolates most similar to RSC2, while 10 isolates could not be uniquely assigned to one or the other, pointing to a third cassette configuration (Fig. S2). Our results indicate that AMR genes on plasmids are subject to strong selective pressure and are frequently removed, likely due to the high cost of transcribing multiple resistance genes.

Rapid plasmid-mediated adaptation: acquisition and loss of AMR genes by horizontal gene transfer and structural rearrangement events. Our findings generated multiple lines of evidence indicating that the rapid gain and loss of AMR genes in opportunistic pathogens in our hospital was mediated by plasmid transfer, merging, and rearrangement, which evolved over multiple distinguishable stages (Fig. 5) in possibly the following sequence of events:

(i) Plasmid A (40 kb) harboring blaIMP-8 and multiple other AMR genes was transferred between P. aeruginosa and C. freundii. Although the direction of transfer cannot be determined with certainty, the fact that P. aeruginosa blaIMP-8 was isolated approximately 2.5 years before the first Citrobacter blaIMP-8 strain was detected suggests a transfer from P. aeruginosa to Citrobacter species. Moreover, the higher GC content of plasmid A points toward an origin of the plasmid from a background with a high level of GC content such as P. aeruginosa (average GC content of 66%). However, the possibility that an unknown intermediate host served as a reservoir for plasmid A cannot be ruled out. Following plasmid transfer to C. freundii, clonal expansion was observed; however, no clonal expansion has been seen to have occurred in C. cronae to date (Fig. 2C).
In *C. freundii*, the plasmid underwent further evolution resulting in the fusion of acquired plasmid A and resident plasmid B to the megaplasmid C ultimately recovered in *C. cronae*. We hypothesized that this happened by plasmid fusion, since plasmid C contains regions with genetic homology of close to 100% across the full length of plasmid A and plasmid B. In addition, plasmid C contained regions harboring parts of transposons which were also present in the chromosome of *C. freundii* cluster Cf2 strains, indicating that this organism was most likely the host of the plasmid fusion. However, the possibility of a plasmid fusion in *C. cronae* cannot be ruled out (Fig. 5, gray area).

We speculate that *C. freundii* Cf2 strains “distributed” plasmid A to *C. freundii* Cf1 and plasmid C to *C. cronae*. However, it is also possible that Cf1 and Cf2 independently acquired plasmid A from *P. aeruginosa* or that Cf2 acquired plasmid A from Cf1. Although less likely, the plasmid fusion resulting in plasmid C might have occurred in *C. cronae* after independent transfer of plasmids A and B from any of the other three bacteria. However, *C. cronae* is also lacking a copy of the transposon region in its chromosome which is present in cluster Cf2, making a fusion in *C. cronae* highly unlikely (Fig. S1). The data presented in Fig. 5 depict all possible trajectories of the adaptation processes mediated by plasmid HGT leading to three bacterial species and four clones with multiple antibiotic resistances in a single hospital within a few years.

In parallel, the class 1 integron in *P. aeruginosa* harboring the antimicrobial resistance genes, including *bla*<sub>IMP-8</sub>, underwent various rearrangements such as deletions and integration of AMR genes. In 12 of the *P. aeruginosa* isolates, one or more AMR genes were lost (Table S4), and at least 9 strains show evidence of gene cassette shuffling (Fig. S2).

---

**FIG 5** Concept of plasmid evolution and transmission across three bacterial species. *P. aeruginosa* *bla*<sub>IMP-8</sub> was isolated approximately 2.5 years prior to the first isolation of *Citrobacter* species harboring *bla*<sub>IMP-8</sub>, leading to the hypothesis of a transfer of plasmid A from *P. aeruginosa* to *Citrobacter* species. Occurrences of the *C. freundii* and *C. cronae* *bla*<sub>IMP-8</sub> genes started at the same time; thus, the timeline does not suggest a specific direction of the transfer. However, the existence of the transposon region in the chromosome of *C. freundii* cluster Cf2 (marked with a black triangle) makes it the most likely host of the merging of plasmid A and B, which are linked by two copies of the transposon region. Solid arrows represent the transmission sequence resulting from this hypothesis.
In conclusion, we demonstrated the successful application of Nanopore sequencing to track the transmission and rapid evolution of an antibiotic resistance plasmid(s) within and between multiple bacterial species in a comprehensive and systematic collection of multidrug-resistant Gram-negative bacteria obtained from a large cohort of high-risk patients and corresponding environment samples.

**DISCUSSION**

Understanding the evolution and spread of multidrug-resistant organisms has become a major challenge in the medical field, necessitating the development of novel diagnostic methods in order to effectively combat increasing numbers of infections with these organisms. The clinical importance of an antimicrobial resistance gene is determined by (i) the class of antibiotics that is rendered resistant, (ii) the pathogenicity of the bacterium, and (iii) the genetic location of the AMR gene. The localization of one or more AMR genes on a mobile genetic element, e.g., a plasmid, strongly increases the risk of resistance spreading between different bacterial genera, including well-adapted and successful human pathogens established in the hospital environment.

In several studies, the spread of carbapenemase gene-harboring plasmids has been demonstrated to happen in the hospital environment (see, e.g., reference 10). For example, Conlan et al. examined plasmids harboring bla\textsubscript{KPC-2} and bla\textsubscript{KPC-3} and provided evidence for horizontal gene transfer between *Klebsiella pneumoniae*, *Enterobacter* sp., and *Citrobacter* sp. (10). Interestingly, the *Citrobacter* strain described in their study (CFNIH1), which was isolated from the hospital environment, contained a 272-kb KPC-encoding plasmid and clustered very closely (core SNP distance of 28 nucleotides) with our study’s *C. cronae* P28 isolate (Fig. 2C), which harbored the large plasmid C. This might indicate that the genetic background of this *Citrobacter* strain enables large plasmid uptake or formation of megaplasmids in this species. The formation of megaplasmids conferring multidrug resistance has also been noted in other *Enterobacteriales*. For example, Desmet et al. analyzed two clinical isolates (a *Klebsiella pneumoniae* strain and an *Enterobacter cloacae* complex strain) harboring bla\textsubscript{OXA-427} carbapenemase and identified a 321-kb megaplasmid which resulted from a cointegration of the MDR plasmid in another plasmid background (24). A further study demonstrated that a fusion plasmid had occurred as a result of recombination in a clinical *Escherichia coli* isolate containing the bla\textsubscript{NDM-5} carbapenemase gene (25). Similarly to the results seen with the C plasmid that we isolated from *C. cronae*, this megaplasmid also harbored duplicated transposon-containing regions, likely as a result of the fusion event. However, the fusion plasmid was not stable when transferred to an *E. coli* recipient strain (25). This is in line with our observations. Fusion plasmid C contained a duplicated region, most likely as a result of recombination. While plasmid C was stable within our study isolates, it was never detected afterwards, suggesting that the plasmid was not positively selected in the hospital environment. However, further studies are needed to elucidate the factors involved in megaplasmid evolution dynamics.

Although the importance of plasmid evolution and horizontal gene transfer for the spread of MDR bacteria has clearly been documented, the epidemiological surveillance of HGT within hospitals is not commonly performed on a routine basis and remains limited to few centers. Short-read sequencing technology, which is available in many hospitals, cannot reliably distinguish between plasmids and chromosomes and such analyses often lead to the occurrence of fragmented genome and plasmid assemblies. Long-read sequencing technologies, on the other hand, enable high-quality, finished assemblies of plasmids. With the emergence of Nanopore sequencing, a fast and inexpensive alternative technology for *de novo* assembly of multidrug-resistant bacteria isolates became available (11, 13, 26). Here, we demonstrated that the application of Nanopore sequencing in combination with Illumina short reads and epidemiological data enabled detailed tracking of plasmid evolution in a comprehensive consecutive collection of bla\textsubscript{IMP-8}-harboring multidrug-resistant Gram-negative bacteria.
to multiple plasmid-based horizontal gene transfers, we were able to detect rearrange-
ments within the multidrug resistance gene cassette, as well as fusion of two plasmids
to a megaplasmid. While the presence and absence of antimicrobial resistance genes
can be postulated based on Illumina short-read assemblies, identification of their
locations on mobile elements and determination of the structure of multidrug resis-
tance gene cassettes remain challenging due to difficulties with assembling repetitive
regions. In the *P. aeruginosa* genomes assembled using Nanopore data, we were readily
able to detect continuous reads confirming the circularity of the plasmid and the exact
order of the resistance gene cassette and were able to distinguish between the bacteria
harboring the megaplasmid and those harboring the two independent plasmids,
further emphasizing the power of long reads for determination of structures of mobile
genetic elements.

**Conclusion.** The application of Nanopore sequencing and the establishment of a
computational pipeline for genome and plasmid assembly, annotation, and compara-
tive analysis (termed “pathoLogic,” including the novel plasmid analysis method plas-
miDent) enabled us to investigate plasmid-driven adaptation and emergence of
multidrug-resistant bacteria using a comprehensive strain collection that included
patient and environment isolates. Using Nanopore-based de novo assemblies, we
demonstrated that horizontal gene transfer between *P. aeruginosa*, *C. freundii*, and
*C. cronae* via a multidrug resistance plasmid (plasmid fusion), resulting in a mega-
plasmid and evolution of the multidrug resistance gene cassette, had occurred
within the short period of 3 years within our hospital. The chosen method for
tracking of MDR plasmids and their evolutionary dynamics represents a powerful
approach which could be applied for real-time infection control surveillance,
thereby contributing to successful countermeasures and efficient containment of
hospital outbreaks. In summary, we developed and showcased a novel pipeline for
de novo bacterial genome assembly, AMR gene and plasmid characterization, and
comparative analysis across species, enabling rapid tracking of AMR transmission
via plasmids in hospital settings.

**MATERIALS AND METHODS**

**Study isolates.** In total, 54 hospital strains were included in the study, comprising *P. aeruginosa*
(*n* = 45), *C. freundii* (*n* = 8), and *C. cronae* (*n* = 1) strains. The strains were obtained from patient
specimens, including rectal screening culture sources (*n* = 40) and water-related environment sources
(toilet or sink; *n* = 14). All isolates were cultured and identified following standard microbiology
protocols as described before (27) and were positive for the *bla*<sub>IMP-8</sub> gene as determined by PCR (28). All
isolates were recovered from samples processed in the hemato-oncology department between July 2009
and July 2015. During this time, the sampling strategy for screening cultures and environmental
surveillance was adjusted as a consequence of the *P. aeruginosa* *bla*<sub>IMP-8</sub> outbreak. Between July 2009 and
October 2010, only clinical specimens were obtained. Weekly rectal screening programs of all hemato-
oncology patients and environment screening of toilets, sinks, and showers in a 14-day cycle were
introduced in October 2010.

**Nanopore and Illumina sequencing.** Nanopore sequencing was performed on an Oxford Nanopore
Technologies MinION device with three different chemistries (versions 6, 7, and 8) and flow cell versions
(FLO-MAP103 version Pk.1, FLO-MIN105 version R9, and FLO-MIN106 version R9:4). An overview of the
chemistry and flow cell versions used for each sample is shown in Table S2 in the supplemental material.

(i) **ONT chemistry version 6.** Sequencing libraries were prepared with a Genomic DNA Sequencing
SQK–MAP006 kit using 1.5 µg of genomic DNA (gDNA) as starting material. Briefly, nick-repaired DNA
(NEBNext FFPE DNA Repair Mix; NEB) was sheared in a Covaris g-TUBE (Covaris, Inc.), followed by end
repair and dA tailing (NEBNext Ultra II End Repair/dA-tailing module; NEB). The leader and hairpin
sequencing adapters (ONT) were ligated using blunt TA ligase (NEB). After tether addition, the final library
was purified with MyOne streptavidin C1 beads (Thermo Fisher). The MinION flow cell (FLO-MAP103,
ONT) was primed and loaded with the library for a 48-h run with 24-h intervals for adding new
presequencing mix, running buffer, and Fuel Mix (ONT).

(ii) **ONT chemistry version 7 and 8.** Libraries were prepared with Genomic DNA Sequencing Kit
SQK–NSK007 and SQK–LSK108, starting with 1.5 µg of gDNA sheared in a Covaris g-TUBE (Covaris, Inc.)
and nick-repaired with N E B N e x t F F P E DNA repair mix (NEB). Subsequently, DNA was end-repaired and
adenylated (NEBNext Ultra II End-Repair/dA-tailing module; NEB) followed by ligation of adaptor (ONT)
using NEB Blunt/TA master mix (NEB). After priming of the flow cells, FLO-MIN105 libraries for kit
SQK–NSK007 and FLO-MIN106 libraries for kit SQK–LSK108 were loaded and run for 48 h following the
protocols of the manufacturer (ONT).
(iii) Illumina sequencing. Due to the advances in sequencing technology that became available over the study period, different protocols were used to obtain short-read sequences, as described before (17, 27, 29). In brief, early isolates were sequenced using 2 × 50 bp on an Illumina HiSeq 2000 sequencer (17) or using 2 × 300 bp on an Illumina MiSeq sequencer (29) or using 2 × 250 bp on an Illumina MiSeq sequencer (27). Table S2 provides a detailed overview of the sequencing protocols applied.

Hybrid de novo assembly pipeline using long and short reads. To achieve complete de novo genome assemblies, we developed a custom pipeline (termed pathoLogic; see Fig. 1) consisting of individual steps for read preprocessing, hybrid de novo assembly, quality control, and generation of assembly statistics. First, long Nanopore reads are subjected to adapter trimming with Porechop (https://github.com/rwick/Porechop), quality filtering with Filtlong (https://github.com/rwick/Filtlong), and quality control (QC) using Nanoplot (30). Adapter trimming and QC for short reads is performed using SeqPurge (31). We benchmarked multiple assembly approaches implemented in pathoLogic. Unicycler, a hybrid assembler using short and long reads (16), produced the longest contigs at high and low read coverage and was therefore used in this study. Finally, assembly statistics are calculated and contigs shorter than 2,000 bp are removed. Application-specific parameters are documented in the published source code and configuration file. All tools are included in the provided Docker image (release v1.0) available on github (plasmIDent, https://github.com/imag/plasmIDent; pathoLogic, https://github.com/imag/pathoLogic).

Phylogenetic analysis. Assembly of the short-read Illumina data for all studied isolates was performed using Spades version 3.7.0 (32), followed by alignment using ProgressiveMauve (version 3.2.1) (33) with a locally collinear block size of 1,000 bp. Phage content was removed using Phast (34). The obtained alignment was used for phylogeny calculation, applying IQ tree version 1.6.3 in UFboot mode (http://tree.bio.ed.ac.uk/software/figtree/). In brief, early isolates were sequenced using 2

Plasmid detection and annotation. For most assemblies, the produced one or a few large chromosomal scaffolds along with several shorter contigs (between 10 kb and 200 kb in length). The latter might have stemmed either from complete circular plasmids or from fragments of the chromosome or plasmids. We therefore developed the plasmIDent tool, which uses long reads to ascertain whether a scaffold is circular, identifies all antibiotic resistance genes, and calculates characteristic metrics such as GC content and read coverage. PlasmIDent takes assembled genomes in fasta format and Nanopore reads in fastq format as input. First, contig ends are fused in order to mimic a circular layout. Next, minimap2 is used to align Nanopore reads to the putative plasmid and the end-to-end fusion site. In cases in which long reads continuously cover the scaffold and the artificially closed gap, we assume that the sequence originated from a circular plasmid. Furthermore, sudden changes of median GC content within the plasmid are used to predict ancestral fusions of two or multiple plasmids. Finally, plasmIDent supports discovery of resistance genes using the CARD database.

Genome annotations. Assembled FASTA files were uploaded to the ResFinder tool (https://cge.cbs.dtu.dk/services/ResFinder/), applying a 98% identity threshold and a minimum overlapping length of 60%.

For the P. aeruginosa blaVIM-2 gene, the sequence was obtained from the ResFinder database. Additionally, CARD-based annotations automatically generated by plasmIDent were merged with the ResFinder results. Finally, we used the RAST Web server to obtain complete genome and plasmid annotations for all isolates and the ISFinder Web server to specifically identify transposons and insertion sequences. We displayed the best hits and annotated the transposons or IS elements to the family level in the duplicated regions.

Comparative genome and plasmid analysis across species. (i) Whole-genome alignment (WGA).

Multiple whole-genome alignments of all assembled plasmids were generated with progressiveMauve in order to find highly similar regions. Plasmids with highly homologous regions were additionally compared by pairwise sequence alignment using nucmer (see, e.g., Fig. S1 in the Supplemental material), resulting in a pairwise identity score and the annotation of homologous regions. We used dot plots (pathoLogic utility scripts) of the pairwise alignments to visually identify rearrangements in plasmids. Homologous regions between plasmids and chromosomal scaffolds were identified using pairwise alignment (nucmer) between a plasmid of interest and the concatenated sequence representing all scaffolds in an isolate’s genome assembly. More specifically, we identified homologous sequences of the transposon-containing region found in plasmid C but not in plasmid A and plasmid B in order to ascertain whether a Citrobacter isolate contained only plasmids A and B and the transposon-containing region inserted in the chromosome or contained plasmid C with the transposon-containing region in the plasmid.

(ii) Read coverage (density) analysis. We chose megaplasmid C of isolate 28_P_CC as the reference plasmid, as it integrates both plasmid A and plasmid B involved in the studied horizontal gene transfer of AMR genes. We used bwa-mem to realign Illumina short reads of each isolate to the reference plasmid,
thereby determining the presence or absence of specific regions based on read density (i.e., whether regions without read coverage were absent in a studied isolate; see Fig. 4 [see also Fig. S2]). We identified breakpoints, indicating structural variants or the end of plasmids, based on clip or split reads. We defined deletions as regions with very-low-density read coverage, with split or paired reads spanning the two breakpoints. (Plasmid ends were identified by circularization as described before.) Furthermore, we evaluated whether putatively deleted resistance genes were also absent from the plasmid AMR gene annotations by ResFinder and CARD.

(iii) AMR gene rearrangements. WGA of the resistance gene cassette of all isolates assembled with Nanopore reads identified two haplotypes, termed RSC1 and RSC2, distinguished by two translocations of AMR genes. In order to assign all sequenced isolates to one or the other cassette configuration, we aligned Illumina short reads to the 4 breakpoints per haplotype (two breakpoints for each translocation event per cassette configuration). Then, we compared the numbers of aligned reads spanning the four breakpoints in RSC1 versus RSC2 and computed the log-transformed fraction of breaks in RSC1 and RSC2, each normalized by the corresponding amount of total reads. Isolates showing log values above 1 were assigned to RSC1 and those showing log values below –1 to RSC2, while the other isolates remained unassigned.

Data availability. All sequence data have been deposited at the European Nucleotide Archive (study accession number PRJEB31907). pathoLogic (https://github.com/imgag/pathoLogic) and plasmIDent (https://github.com/imgag/plasmIDent) were developed in this study and are freely available on GitHub.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

FIG S1, TIF file, 0.1 MB.

FIG S2, TIF file, 0.4 MB.

FIG S3, TIF file, 0.2 MB.

TABLE S1, PDF file, 0.2 MB.

TABLE S2, PDF file, 0.2 MB.

TABLE S3, PDF file, 0.3 MB.

TABLE S4, PDF file, 0.2 MB.

ACKNOWLEDGMENTS

We thank Nadine Hoffmann and Baris Bader for expert technical assistance. The study was conducted in accordance with the local ethic committee (Ethic Committee Medical Faculty, University of Tübingen, No. 741/2016BO2 and 407/2013R).

We declare that we have no competing interests. S.P. received a speaker’s honorarium from bioMérieux, Deutschland, and consultancy honoraria from IDbyDNA, San Francisco, CA, USA, and Illumina, Cambridge, United Kingdom, which had no role in the design, analysis, interpretation, and writing of the manuscript or in the decision to publish the data.

The study was funded by the Faculty of Medicine of the University of Tübingen, the Spanish Ministry of Economy and Competitiveness, the Centro de Excelencia Severo Ochoa, the CERCA Program/Generalitat de Catalunya and the “la Caixa” Foundation. The funders had no role in the design, analysis, interpretation, and writing of the manuscript.

Software development and bioinformatics analysis were performed by M.B. during the first half of the 3-year project and by C.G. during the second half. S.P., D.B., P.O., J.G., M.M., M.W., I.G., and M.G. generated the laboratory and sequencing data. J.L., W.V., and D.D. gathered epidemiological data. S.P., M.B., C.G., and S.O. performed the data analysis. C.G., M.B., and S.O. developed the bioinformatics methods and pipelines, and L.B. developed the GUI. S.P., I.A., and S.O. designed the study. S.P., C.G., and S.O. wrote the manuscript.

REFERENCES

1. Roca I, Akova M, Baquero F, Carlet J, Cavalieri M, Coenen S, Cohen J, Findlay D, Gysens I, Heure OE, Kahlmeter G, Kruse H, Laxminarayan R, Liébana E, López-Cerero L, MacGowan A, Martins M, Rodriguez-Baño J, Rolain J-M, Segovia C, Sigauque B, Tacconelli E, Wellington E, Vila J. 2015. The global threat of antimicrobial resistance: science for intervention. New Microbes New Infect 6:22–29. https://doi.org/10.1016/j.jmni.2015.02.007.

2. Marcola AE, Cascella M, Corcione S, Cuomo A, Nappa S, Borgia G, De Rosa FG, Gentile I. 2017. Management of multidrug-resistant *Pseudomonas aeruginosa* in the intensive care unit: state of the art. Expert Rev Anti Infect Ther 15:861–871. https://doi.org/10.1080/14787210.2017.1367666.

3. Kengkla K, Kongpakwattana K, Saokaew S, Apisarnthanarak A, Chaiyakunapruk N. 2018. Comparative efficacy and safety of treatment options for MDR and XDR *Acinetobacter baumannii* infections: a systematic review and network meta-analysis. J Antimicrob Chemother 73:22–32. https://doi.org/10.1093/jac/dcx368.

4. Tacconelli E, Cataldo MA, Dancer SJ, De Angelis G, Falcone M, Frank U, Kahlmeter G, Pan A, Petrosillo N, Rodriguez-Bano J, Singh N, Venditti M,
Monitoring AMR Plasmid Transfer by Nanopore Sequencing

Yokae DS, Cookson B, European Society of Clinical M. 2016. ESMID guidelines for the management of the infection control measures to reduce transmission of multidrug-resistant Gram-negative bacteria in hospitalized patients. Clin Microbiol Infect 20(Suppl 1):1–55. https://doi.org/10.1111/1469-0691.12427.

Naylor NR, Atun R, Zhu N, Kulasabathan K, Silva S, Chatterjee A, Knight GM, Robotham JV. 2018. Estimating the burden of antimicrobial resistance: a systematic literature review. Antimicrob Resist Infect Control 7:58. https://doi.org/10.1186/s13756-018-0336-y.

Gilchrist CA, Cotton JA, Burker C, Aruj T, Gilmartin A, Lin Y, Ahmed E, Steiner K, Alam M, Ahmed S, Robinson G, Zaman SU, Kabir M, Sanders M, Chalmers RM, Ahmed T, Ma JZ, Haque R, Faruque ASG, Berriman M, Petri WA. 2018. Genetic diversity of *Cystoseira* *paramoena* in a Bangladeshi community as revealed by whole-genome sequencing. J Infect Dis 218:259–264. https://doi.org/10.1093/infdis/jiy121.

Partridge SR, Kwong SM, Firth N, Jensen SO. 2018. Mobile genetic elements associated with antimicrobial resistance. Clin Microbiol Rev 31:e00088-17. https://doi.org/10.1128/CMR.00088-17.

Arredondo-Alonso S, Willems RJ, van Schaik W, Schurch AC. 2017. On the (im)possibility of reconstructing plasmids from whole-genome short-read sequencing data. Microb Genom 3:e000128. https://doi.org/10.1099/mgen.0.000128.

Conlan S, Park SB, DeAngelis D, Thomas PJ, Young AC, Coleman H, Sison C, NISC Comparative Sequencing Program, Weingarten RA, Lau AF, Dekker JP, Palmore TN, Frank KM, Segre JA. 2016. Plasmid dynamics in KPC-positive *Klebsiella pneumoniae* during long-term patient colonization. mBio 7:e00742-16. https://doi.org/10.1128/mBio.00742-16.

Conlan S, Thomas PJ, Deming C, Park M, Lau AF, Dekker JP, Snitkin ES, Clark TA, Luong K, Song Y, Tsai YC, Boitano D, Mayal J, Brookes SY, Schmidt B, Young AC, Thomas JW, Bouffard GG, Blakesley RW, NISC Comparative Sequencing Program, Mullikin JC, Korch J, Henderson DK, Frank KM, Palmore TN, Segre JA. 2014. Single-molecule sequencing to track plasmid diversity of hospital-associated carbapenemase-producing Enterobacteriaceae. Sci Transl Med 6:254ra126. https://doi.org/10.1126/scitranslmed.3009845.

Lemon JK, Rhi PF, Frank KM, Dekker JP. 2017. Rapid Nanopore sequencing of plasmids as an early warning system. Nature 535:3530–3543. https://doi.org/10.1038/nature19067.

George S, Pankhurst L, Hubbard A, Votintseva A, Stoesser N, Sheppard AE, Mathers AE, Norris RV, Navickaite I, Eaton CA, Iqbal Z, Crook DW, Phan HT, HTT. 2017. Resolving plasmid structures in Enterobacteriaceae using the MiniON nanopore sequencer: assessment of MiniON and MiniON/ Illumina hybrid data assembly approaches. Microb Genom 3:e000118. https://doi.org/10.1099/mgen.0.000118.

Dong N, Yang X, Zhang R, Chan EW, Chen S. 2018. Tracking microevolution events among ST11 carbapenemase-producing hypervirulent *Klebsiella pneumoniae* outbreak strains. Emerg Microb Infect 7:146. https://doi.org/10.1038/s41426-018-0146-6.

Chin CS, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, Turner SW, Korlach J. 2013. Non-hybrid, finished microbial genome assemblies from long-read SMRT sequencing data. Nat Methods 10:563–569. https://doi.org/10.1038/nmeth.2474.

Antipov D, Korobeynikov A, McLean JS, Pevzner PA. 2016. hybridSPAdes: an algorithm for hybrid assembly of short and long reads. Bioinformatics 32:1009–1015. https://doi.org/10.1093/bioinformatics/btw688.

Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput Biol 13:e1005595. https://doi.org/10.1371/journal.pcbi.1005595.

Willmann M, Bezdan D, Zapata L, Susak H, Vogel W, Schroppel C, Liese J, Weidenmaier C, Autenrieth I, Ossowski S, Willmann M. 2015. Genomic characterisation of clinical and environmental *Pseudomonas putida* group strains and determination of their role in the transfer of antimicrobial resistance genes to *Pseudomonas aeruginosa*. BMC Genomics 16:1889. https://doi.org/10.1186/s12864-017-4216-2.

Pitout JD, Gregson DB, Poirel L, McClure JA, Defrise M, Liese J, Weidenmaier C, Willmann M. 2017. Genomic characterisation of clinical and environmental *Pseudomonas aeruginosa* producing metallo-beta-lactamases in a large centralized laboratory. J Clin Microbiol 43:3129–3135. https://doi.org/10.1128/JCM.03129-17.

Pfennigwerth N. 2018. Bericht des Nationalen Referenzzentrums (NRZ) für gramnegative Krankenhauserreger, Dezember 2017. Epid Bull 28:263–267.

Pfennigwerth N. 2019. Bericht des Nationalen Referenzzentrums für gramnegativen Krankenhauserreger – Zeitraum 1. Januar 2017 – 31. Dezember 2017. Epid Bull 28:263–267.

July/August 2020 Volume 5 Issue 4 e00025-20 msphere.asm.org 15
C. 2017. Nextflow enables reproducible computational workflows. Nat Biotechnol 35:316–319. https://doi.org/10.1038/nbt.3820.
40. Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. 2017. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. Genome Res 27:722–736. https://doi.org/10.1101/gr.215087.116.
41. Li H. 2016. Minimap and miniasm: fast mapping and de novo assembly for noisy long sequences. Bioinformatics 32:2103–2110. https://doi.org/10.1093/bioinformatics/btw152.
42. Kolmogorov M, Yuan J, Lin Y, Pevzner PA. 2018. Assembly of long error-prone reads using repeat graphs. bioRxiv https://doi.org/10.1101/247148.
43. Lin Y, Yuan J, Kolmogorov M, Shen MW, Chaisson M, Pevzner PA. 2016. Assembly of long error-prone reads using de Bruijn graphs. Proc Natl Acad Sci U S A 113:E8396–E8405. https://doi.org/10.1073/pnas.1604560113.
44. Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: quality assessment tool for genome assemblies. Bioinformatics 29:1072–1075. https://doi.org/10.1093/bioinformatics/btt086.