








## Combining sequencing approaches to fully resolve a carbapenemase-encoding megaplasmid in a *Pseudomonas shirazica* clinical strain

João Botelho <sup>a</sup>, Cédric Lood <sup>b,c</sup>, Sally R. Partridge <sup>d</sup>, Vera van Noort <sup>b,e</sup>, Rob Lavigne <sup>c</sup>,  
Filipa Grosso <sup>a</sup> and Luísa Peixe <sup>a</sup>

<sup>a</sup>UCIBIO/REQUIMTE, Laboratório de Microbiologia, Faculdade de Farmácia, Universidade do Porto, Porto, Portugal; <sup>b</sup>Centre of Microbial and Plant Genetics, Department of Microbial and Molecular Systems, KU Leuven, Leuven, Belgium; <sup>c</sup>Laboratory of Gene Technology, Department of Biosystems, KU Leuven, Leuven, Belgium; <sup>d</sup>Centre for Microbiology and Infectious Diseases, The Westmead Institute for Medical Research, The University of Sydney, Westmead Hospital, Sydney, Australia; <sup>e</sup>Institute of Biology, Leiden University, Leiden, The Netherlands

### ABSTRACT

Horizontal transfer of plasmids plays a pivotal role in dissemination of antibiotic resistance genes and emergence of multidrug-resistant bacteria. Plasmid sequencing is thus paramount for accurate epidemiological tracking in hospitals and routine surveillance. Combining Nanopore and Illumina sequencing allowed full assembly of a carbapenemase-encoding megaplasmid carried by multidrug-resistant clinical isolate FFUP\_PS\_41. Average nucleotide identity analyses revealed that FFUP\_PS\_41 belongs to the recently proposed new species *Pseudomonas shirazica*, related to the *P. putida* phylogenetic group. FFUP\_PS\_41 harbours a 498,516-bp megaplasmid (pJBCL41) with limited similarity to publicly-available plasmids. pJBCL41 contains genes predicted to encode replication, conjugation, partitioning and maintenance functions and heavy metal resistance. The [aacA7|blaVIM-2|aacA4] cassette array (resistance to carbapenems and aminoglycosides) is located within a class 1 integron that is a defective Tn402 derivative. This transposon lies within a 50,273-bp region bound by Tn3-family 38-bp inverted repeats and flanked by 5-bp direct repeats (DR) that composes additional transposon fragments, five insertion sequences and a Tn3-Derived Inverted-Repeat Miniature Element. The hybrid Nanopore/Illumina approach allowed full resolution of a carbapenemase-encoding megaplasmid from *P. shirazica*. Identification of novel megaplasmids sheds new light on the evolutionary effects of gene transfer and the selective forces driving antibiotic resistance.





**ARTICLE HISTORY** Received 28 March 2019; Revised 17 July 2019; Accepted 19 July 2019

**KEYWORDS** *Pseudomonas*; megaplasmids; Nanopore; Illumina; antibiotic resistance


### Introduction

Bacteria can become resistant to antibiotics through chromosomal mutations and/or by the acquisition of resistance genes carried on mobile genetic elements, including plasmids and integrative and conjugative elements [1]. Plasmids are autonomous self-replicating elements of which some are capable to drive horizontal transfer (HGT) of antibiotic resistance genes by conjugation [2–5]. The mobility of a plasmid depends on the set of genes that it carries, and these extrachromosomal elements may be conjugative, mobilisable or non-transmissible [2,3]. Conjugative plasmids carry all the machinery necessary for self-transfer: i) a relaxase, a key protein in conjugation; ii) an origin of transfer (*oriT*); iii) a set of genes encoding for the type-IV secretion system

(T4SS); and iv) a gene encoding a type-IV coupling protein (T4CP) [2,3]. Mobilisable plasmids lack the complete set of genes encoding the T4SS and may use the conjugative apparatus of a helper plasmid present in the cell to be successfully transferred. Conjugative plasmids tend to be low copy number and large, whereas mobilisable plasmids are frequently high copy number and smaller (<30 kb) [2,3]. The term megaplasmids [6] has been used for very large replicons (>350 kb) which, in contrast to chromids [7], do not carry essential core genes. Megaplasmids frequently have mosaic structures, carrying genetic modules that originate from different ancestral sources [8]. The formation of mosaic plasmids may be influenced by several factors, such as the abundance of conjugative plasmids and transposons, selection pressures,

**CONTACT** João Botelho  botelho@evolbio.mpg.de  Antibiotic Resistance Evolution Group, Max-Planck-Institute for Evolutionary Biology, 24306 Plön, Germany; Department of Evolutionary Ecology and Genetics, Zoological Institute, Christian-Albrechts-Universität zu Kiel, Am Botanischen Garten 1-9, 24118 Kiel, Germany; Luísa Peixe  lpeixe@ff.up.pt  Laboratório de Microbiologia, Faculdade de Farmácia da Universidade do Porto, Rua Jorge Viterbo Ferreira n° 228, 4050-313 Porto, Portugal

\*Present address: Antibiotic Resistance Evolution Group, Max-Planck-Institute for Evolutionary Biology, Plön, Germany; Department of Evolutionary Ecology and Genetics, Zoological Institute, Christian-Albrechts-Universität zu Kiel, Kiel, Germany

 Supplemental data for this article can be accessed <https://doi.org/10.1080/22221751.2019.1648182>

© 2019 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group on behalf of Shanghai Shangyixun Cultural Communication Co., Ltd  
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

incompatibility groups and the host's tolerance of foreign DNA. According to the plasmid hypothesis, megaplasmids are the evolutionary precursors of chromids, due to the amelioration of genomic signatures to those of the host's chromosome and the acquisition of essential genes [7].

To date, fourteen incompatibility groups (IncP-1 to IncP-14) have been characterised amongst *Pseudomonas* plasmids [9,10]. Narrow host range plasmids comprise IncP types -2, -5, -7, -10, -12 and -13 and cannot be transferred into *Escherichia coli*. In contrast, other groups display a broad host range, as they are also included in the typing scheme for *Enterobacteriaceae* plasmids: IncP-1 (IncP), IncP-4 (IncQ) and IncP-6 (IncG) [9,10]. Unlike *Enterobacteriaceae* plasmids, no replicon-based PCR typing of other *Pseudomonas* plasmids has been created yet. Even though a few reports have characterised large plasmids among pseudomonads [11–13], the role of these elements in the spread of antibiotic resistance in this genus remains poorly understood.

Plasmids may harbour accessory module(s) that provide adaptive advantage(s) for their host, such as virulence-encoding factors and antibiotic resistance genes [9,14–16]. These elements frequently harbour carbapenemase-encoding genes, which confer resistance to  $\beta$ -lactams, including carbapenems, frequently last resort antibiotics for infections caused by multi-drug resistant bacteria [9,17]. Sequencing of plasmids is thus paramount to the success of accurate epidemiological tracking strategies in the hospital setting and routine surveillance, helping to identify transmission routes and to prevent future outbreaks [18–23]. The advent of WGS has enabled the *in silico* analysis of a wide array of plasmids, most of them from assembly of short-read sequencing data [11,24–27]. However, fully resolving plasmids with short-read sequencing technologies remains challenging due to the presence of numerous long repeated regions [28], and currently the most accurate approach to assemble these plasmids is to use a combination of short-read and long-read methods [18–23,29,30].

Here, we combined Nanopore and Illumina sequencing to fully assemble a carbapenemase-encoding megaplasmid carried by a clinical isolate belonging to the recently proposed *Pseudomonas shirazica* species [31].

## Material and methods

### Bacterial isolate

Isolate FFUP\_PS\_41 was obtained in 2008 from endotracheal tube secretions of a patient with pneumonia admitted to the Neonatal/Pediatric Intensive Care unit of Centro Hospitalar do Porto – Hospital de

Santo António, in Porto, Portugal, as part of regular surveillance of carbapenemase-producers among clinical isolates.

FFUP\_PS\_41 was initially identified at the hospital as *Pseudomonas putida* by VITEK-2 (bioMérieux), a routine phenotypic based method for bacterial identification. In this study we re-classified the strain by pairwise average nucleotide identity based on BLAST+ (ANIb) using PyANI v0.2.7 (<https://github.com/widdowquinn/pyani>) [32,33]. Antimicrobial susceptibility testing was conducted by standard disc diffusion and broth microdilution (for colistin) methods, according to EUCAST guidelines (<http://www.eucast.org/>).

### Whole-plasmid sequencing and bioinformatics

Genomic DNA from FFUP\_PS\_41 was extracted using a QIAamp DNA Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. Sequencing libraries were prepared using Illumina Nextera and the 1D ligation library approach from Oxford Nanopore Technology (ONT) where we used a Covaris gTube to fragment the gDNA around 10 kbp. Libraries were sequenced on the Illumina HiSeq 2500 sequencer or the MinION sequencer from ONT equipped with a flowcell of chemistry type R9.4, respectively.

Illumina reads were verified for quality using FastQC and Trimmomatic [34,35] yielding 5.9 M paired-end reads of 125 bp in length totalling 1.4 B bases (245x est. coverage), while the MinION reads were processed with ONT's albacore v2.3.0 followed by demultiplexing using porechop v0.2.3, yielding 62.6 k reads totalling 504 M bases (84x est. coverage). Both datasets were then combined using the Unicycler assembly pipeline [36] with a finishing step of Pilon v1.22. The assemblies were visually inspected using the assembly graph tool Bandage v0.8.1 [37]. Annotation of the megaplasmid was performed with Prokka v1.13 using default parameters [38]. To improve annotation, we downloaded additional files of trusted proteins from NCBI RefSeq plasmids (<ftp://ftp.ncbi.nih.gov/refseq/release/plasmid/>), the NCBI Bacterial Antimicrobial Resistance Reference Gene Database ([ftp://ftp.ncbi.nlm.nih.gov/pathogen/Antimicrobial\\_resistance/](ftp://ftp.ncbi.nlm.nih.gov/pathogen/Antimicrobial_resistance/)) and the Antibacterial Biocide- and Metal-Resistance Genes database (Bac-Met, <http://bacmet.biomedicine.gu.se/index.html>, all accessed on the 01/10/18). EggNOG mapper v4.5.1 and NCBI's Conserved Domain Database CDSEARCH/cdd v3.16 were used for functional annotation and conserved domain search of protein sequences, respectively [39–41]. Inference of orthologous groups (OGs) was achieved with OrthoFinder v2.2.6 [42]. The coding sequence (CDS) annotations of the megaplasmid were visualised with Circos v0.69–6 [43]. We used ISfinder [44] to look for insertion sequences (IS). Antimicrobial resistance genes and associated mobile elements were annotated

using Galileo™ AMR (<https://galileoamr.arcbio.com/mara/>, Arc Bio, Cambridge, MA) [45]. Plasmid copy number was estimated based on coverage of the Illumina dataset. GenSkew (<http://genskew.csb.univie.ac.at/>) was used to compute and plot nucleotide skew data to predict the origin of replication.

### Plasmid transfer and S1/I-CeuI-PFGE

Plasmid transfer by conjugation assays was attempted as previously described [46], using a spontaneous rifampicin-resistant mutant of *P. aeruginosa* PAO1 as recipient strain. Transconjugant selection was performed using Mueller–Hinton agar containing rifampicin (100 mg/L) and imipenem (2 mg/L).

S1 and I-CeuI-PFGE was performed as previously described [47] to confirm the presence of extrachromosomal elements.

### Accession number

The sequence of plasmid pJBCL41 was deposited in GenBank accession number MK496050.

## Results

### Antimicrobial susceptibility and taxonomy testing

Clinical isolate FFUP\_PS\_41 has a multidrug resistance (MDR) phenotype, showing resistance to imipenem, meropenem, ceftazidime, cefepime, aztreonam, piperacilin + tazobactam, gentamicin, tobramycin, amikacin, ciprofloxacin but remains susceptible to colistin (MIC = 1 mg/L). FFUP\_PS\_41 was initially identified as *P. putida* by VITEK-2. However, it displays an ANIb value of 99.1% (above the 95% cut-off for species identification [32]) when compared with the *P. shirazica* type strain genome [31], suggesting that it belongs to this species related to the *P. putida* phylogenetic group.

### Comparative megaplasmidomics between pJBCL41 and related *Pseudomonas* plasmids

Using a hybrid assembly approach, we were able to fully resolve a single extrachromosomal element carried by *Pseudomonas* sp. FFUP\_PS\_41 (Figure S1). This mosaic megaplasmid (named pJBCL41) is 498,516 bp long and a total of 608 predicted CDS were annotated (Figure 1). It has an average GC content of 56.0%, which is lower than that observed for the chromosome (62.6%) and the mean content for strains identified as *P. putida* (62.0%, according to information retrieved on the 08/03/2019 on <https://www.ezbiocloud.net/taxon?tn=Pseudomonas%20putida>).

NCBI's conserved domains database (CDD) calls 42.1% (256) of the predicted CDS for pJBCL41 (Table S1), indicating that most genes encode proteins of unknown function. The backbone of this megaplasmid harbours genes predicted to be responsible for plasmid replication and heavy metal resistance and carries two predicted type-II toxin-antitoxin (TA) systems and genes encoding for partition systems (Figure 1) [48]. Several genes encoding transport and metabolic processes, as well as transposable elements and CDS associated with transcription, regulatory, chemotaxis signal transduction and mobility functions could be identified (Table S2). These traits are frequently overrepresented on large plasmids (Figure 2) [6,49]. Also, pJBCL41 harbours several genes coding for the synthesis of DNA precursors, which may promote replication and transcription processes to help alleviate the burden that this acquired element may impose on the host cell.

pJBCL41 has low nucleotide sequence identity with *Pseudomonas* megaplasmids deposited in public databases (Table 1 and Figure S2). OrthoFinder assigned 59.4% of proteins encoded by pJBCL41 and the most closely-related plasmid, pQBR103 from *Pseudomonas fluorescens* [50], to 335 OGs (Table S3). pQBR103 was found in *Pseudomonas* populations colonising the leaf and root surfaces of sugar beet plants growing at Wytham, United Kingdom and carries no antimicrobial resistance genes [50]. Curiously, a blastp analysis between the proteins encoded by these megaplasmids revealed that the average amino acid sequence identity is 72.8% among sequences producing significant alignments.

Large plasmids identified among the *Pseudomonas* genus usually belong to the IncP-2 incompatibility group [10,11,27]. However, the IncP-2-type stability/replication/conjugal transfer system is absent from pJBCL41 as previously observed for other megaplasmids carried by different *Pseudomonas* species [51,52]. Two replication initiation genes could be identified here. One replicase gene is located at positions 458,679–457,813 on the plasmid (locus\_tag: pJBCL41\_00568), in close proximity to the predicted origin of replication (Figure S3). pJBCL41 is estimated to be present as a single copy, from read coverage vs. the chromosome. Like many megaplasmids, pJBCL41 appears to possess a full set of genes for self-transmission [2,3]. We identified a cluster of genes encoding an F-type T4SS, encompassing i) a gene encoding a TraD homolog (locus\_tag: pJBCL41\_00295), an AAA + ATPase of the pfamVirD4 type, known as the T4CP and which is a key protein in conjugation; ii) a gene encoding a TraI (locus\_tag: pJBCL41\_00297) relaxase homolog, which together with accessory proteins is responsible for cleaving the plasmid in a site-specific manner to initiate DNA transfer and iii) a set of genes (*traEFGKNV* homologues, positions



**Table 1.** Megaplasmiids identified by BLASTn search with pJBCL41.

Plasmid identifier	Max score	Total score	Query cover	E value	Ident (%)	Species	Genbank accession	Size (bp)	Year of isolation	Source	Country
pOBR103	11073	1.06E + 05	44%	0	72.99%	<i>P. fluorescens</i>	NC_009444.1	425094	2008	Environment	UK
XWY-1	10059	1.22E + 05	9%	0	99.59%	<i>P. sp.</i>	NZ_CP026333.1	394537	2016	Environment	China
pJB37	7285	43666	17%	0	99.90%	<i>P. aeruginosa</i>	KY494864.1	464804	2008	Clinical	Portugal
pSY153	6131	1.41E + 05	20%	0	99.85%	<i>P. putida</i>	KY883660.1	468170	2012	Clinical	China
pOZI176	6129	77698	19%	0	99.71%	<i>P. aeruginosa</i>	KC543497.1	500839	2000	Clinical	China
pBM413	5306	51713	17%	0	99.93%	<i>P. aeruginosa</i>	CP016215.1	423017	2012	Clinical	China
RW109	4728	41531	17%	0	71.75%	<i>P. aeruginosa</i>	NZ_L1969519.1	552265	NA	Industrial	NA
P19E3	4715	28494	15%	0	71.79%	<i>P. koreensis</i>	NZ_CP027478.1	467568	2014	Environment	Switzerland
AR439	4697	39223	16%	0	71.79%	<i>P. aeruginosa</i>	NZ_CP029096.1	437392	NA	Clinical	NA

NA indicates no data available.

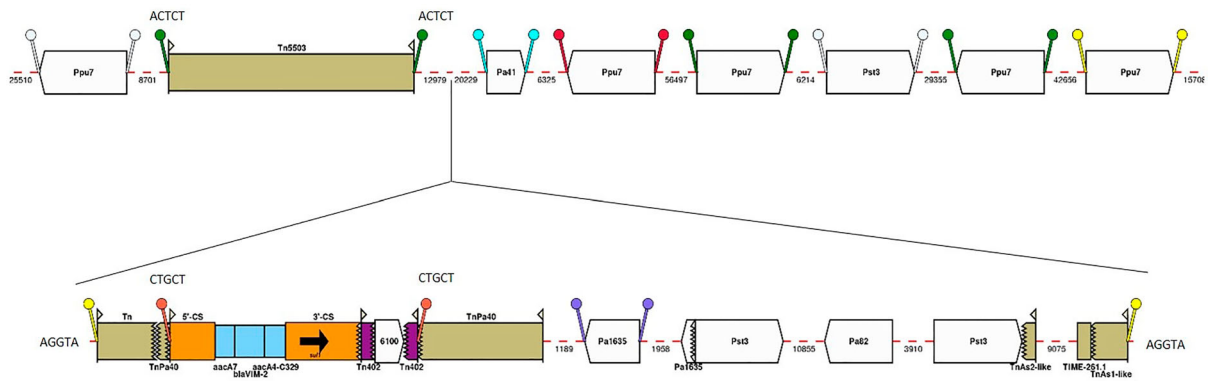
182,497–203,751) coding for a mating pair formation system responsible for pilus assembly and retraction (Figure 1) [2,3,53].

We were unable to transfer the pJBCL41 *in vitro* to a spontaneous rifampicin-resistant mutant of *P. aeruginosa* PAO1, under tested conditions. S1/I-CeuI-PFGE confirmed the presence of a ~500 kb extra-chromosomal element.

### **pJBCL41 carries a complex 50 kb multidrug resistance region**

pJBCL41 carries genes typically found on IncP-2 plasmids encoding resistance to tellurite, which could allow co-selection and enrichment of bacteria with MDR plasmids [54]. It also harbours a class 1 integron with the  $[aacA7|bla_{VIM-2}|aacA4]$  cassette array (named In103 by INTEGRALL [55]) (Figure 3): *aacA7* confers resistance to aminoglycosides (amikacin, netilmicin and tobramycin) and *bla\_{VIM-2}* encodes resistance to  $\beta$ -lactams (including carbapenems). The *bla\_{VIM-2}* gene is by far the most frequently described carbapenemase-encoding gene, both geographically and phylogenetically (across *Pseudomonas* spp.) [56,57]. The *aacA4* gene cassette has a C residue at nucleotide position 329 corresponding to a serine residue associated with gentamicin resistance [58]. The same cassette array has been observed previously among isolates from Portuguese hospitals [25]. The integron is of the In4 type, with a complete 5'-CS bounded by the 25 bp inverted repeat IRi, 2,239 bp of the 3'-CS and IS6100 flanked by two fragments of the IRT end of Tn402 [9,59]. As the region between IRi and IRT lacks *tni* transposition genes, this constitutes a Tn402-like transposon that would be defective in self-transposition.

This defective Tn402-like transposon is flanked by 5-bp direct repeats (DR) (5'-CTGCT-3') (Figure 3), suggesting integration by transposition close to the predicted resolution (*res*) site of a Tn3-family transposon. About 300 bp at the IR<sub>L</sub> end of the transposon are related (~86% identical) to TnAs1 (ISfinder), followed by a region containing a gene which may encode a methyl-accepting chemotaxis protein. From the predicted recombination crossover point in the *res* site the sequence matches TnPa40 (ISfinder). This "hybrid" transposon is not flanked by characteristic 5 bp DR but the 5 bp adjacent to IR<sub>L</sub> (5'-AGGTA-3') are repeated 50,273 bp away, immediately adjacent to the 38 bp repeat of a 1,100 bp transposon fragment ~97% identical to part of both Tn1721 (GenBank accession no. X61367.1, [60]) and TnAs1 (Figure 3). This transposon is truncated by 261 bp region that apparently corresponds to a Tn3-Derived Inverted-Repeat Miniature Element (designated TIME-261.1 here). TIMEs are non-autonomous mobile elements commonly found in *Pseudomonas* spp. [61].



**Figure 3.** Map of resistance genes and mobile genetic elements inserted in the backbone of pJBCL41. Gene cassettes are shown as blue boxes labelled with the cassette name and are oriented in the 5'-CS to 3'-CS direction. IS are shown as block arrows labelled with the IS name/number, with the pointed end corresponding to IR<sub>R</sub>. TIME-261.1 and fragments of Tn3-family transposons are shown as beige boxes with 38 bp IR represented by flags. The fragment annotated as "TnAs1-like" is ~97% identical to a region in common between Tn1721 (GenBank accession no. X61367.1) and TnAs1 in ISfinder. The fragment annotated as "TnAs2-like" is ~94% identical to TnAs2 in ISfinder. The integron is inserted in a proposed hybrid transposon, apparently created by *res*-mediated recombination between a *tnp* region matching TnPa40 and another transposon, labelled "Tn", that is ~86% identical to TnAs1 over the ~300 bp at the IR<sub>L</sub> end only. DR are shown as a pair of "lollipops" of the same colour flanking an IS or a pair of IRs (but note that the same colour may be used to indicate more than one pair of DR), with sequences indicated for DR of transposons. Mobile elements are shown to scale and numbers below dashed red lines indicate the lengths of intervening regions in bp. This figure was constructed from diagrams generated using Galileo™ AMR.

Most of the region between these transposon elements consists of a 16,782 bp segment flanked by directly oriented copies of ISPst3 (IS21 family). This region, except for insertion of ISPa82 (IS66 family) and an adjacent deletion in pJBCL41, matches several *Pseudomonas* chromosomes (e.g. *P. aeruginosa* PA7 in Figure S4) and different parts of it are found in plasmids in *Pseudomonas*, *Acinetobacter* and *Enterobacteriaceae*, sometimes also flanked by IS. The sequence between TnPa40 and the left-hand ISPst3 in pJBCL41 is a duplication of part of the 16,782 bp region, with ISPa1635 (IS4 family) inserted, flanked by characteristic 8 bp DR, instead of ISPa82 and ends with a partial ISPa1635. The right-hand ISPst3 truncates a transposon related to TnAs2 [62], which is separated from TIME-261.1 by a 9,075 bp region that also matches *Pseudomonas* chromosomes and includes a putative aminoglycoside phosphotransferase gene.

Blast searches with the complete 50 kb region identified a 59 kb region in the chromosome of *P. aeruginosa* AR\_0440 (GenBank accession no. CP029148.1) that has similar ends, but lacks an integron, with an additional Tn5393 insertion and a different region in place of the ISPst3-bounded segment (Figure S4). This 59 kb region is flanked by 5 bp DR (5'-AATGA-3') and an uninterrupted version of the flanking sequence matches other *Pseudomonas* chromosomes.

A Tn5503-like transposon encoding a type-II TA system and two metal dependent phosphohydrolases is also inserted in pJBCL41 [63] and is flanked by 5-bp DR (5'-ACTCT-3'), indicating that this element transposed independently of the 50 kb region (Figure 3). It has only 10 nucleotide differences from

the original Tn5503 on plasmid Rms149, the archetype of *Pseudomonas* plasmid incompatibility group IncP-6 [63], and additional copies of short repeats in a GC-rich region within a gene encoding an ATP-utilizing enzyme. An additional ISPst3, five ISPpu7 (IS21 family) and one ISPa41 (IS5 family) - all flanked by DR of characteristic length, are also inserted in the pJBCL41 backbone (Figures 1 and 3).

## Discussion

In this study, we took advantage of a hybrid assembly approach to fully resolve and characterise a carbapenemase-encoding megaplasmid and to compare it with related *Pseudomonas* megaplasmids. The lower GC content of pJBCL41 compared with the FFUP\_PS\_41 chromosome and strains belonging to the *P. putida* phylogenetic group may be related to a more relaxed selection acting on these secondary replicons, as the maintenance of GC-rich genomes is energetically more demanding [64,65]. Ongoing studies will help to characterise the biology and genomic signatures related to this newly characterised *P. shirazica* species (Botelho *et al*, unpublished data). Even though we were unable to transfer pJBCL41 by conjugation to a *P. aeruginosa* strain under the conditions used, we hypothesise that it may be transferrable to other strains belonging to the *P. putida* phylogenetic group. Strains belonging to this group display a GC content lower than those of *P. aeruginosa*, and differences in GC content are a known biological barrier for HGT [66].

Since secondary replicons are under strong pressure to undergo genomic reshuffling [64], the observed low nucleotide sequence identity between pJBCL41 and

large *Pseudomonas* plasmids deposited in public databases might be expected. Even though pJBCL41 and pQBR103 are similar in size and functionalities, there is a high level of divergence between genes encoding related proteins. Indeed, it is rare to identify megaplasmids with a similar nucleotide sequence in strains belonging to different species within the same genus [6,52]. These results suggest that pJBCL41 and pQBR103 may share a common ancestor, but independent evolutionary trajectories have led to significant diversification among related genes.

The presence of different replicons suggests that pJBCL41 may have resulted from co-integration of distinct plasmid modules. The replication module defines plasmid copy number and plasmid survival in different hosts. Low copy-number plasmids are more frequently lost, due to random assortment at cell division [2,3] and extra stability modules, such as TA and partition systems, may be required to ensure that large plasmids such as pJBCL41 are maintained [48,67].

The DR flanking the 50 kb region in pJBCL41 and the related 59 kb region in the *P. aeruginosa* AR\_0440 chromosome could reflect insertion of each region by transposition, possibly mediated by the intact transposase and resolvase of TnPa40. However, the size, complexity and differences between the internal parts of these related regions may be more consistent with initial insertion of a simple transposon followed by further insertions, deletions and rearrangements. A similar situation is seen in plasmid pCTX-M360, which carries a complete Tn2 flanked by the 5 bp DR, and the highly-related pCTX-M3, in which the ends of Tn2 are present in the same position but the central part of the transposon has undergone extensive rearrangements [68]. The identification of all or part of the 16,782 bp segment found within the 50 kb region in pJBCL41 in other locations also suggests that some of the genes it carries may encode advantageous functions, but this needs further analysis. Identification of other sequences related to parts of these 50 and 59 kb region segments may also shed light on how they have arisen and evolved.

In summary, we show that a hybrid Nanopore/Illumina approach is useful for producing contiguous assemblies and allowed full resolution of a carbapenemase-encoding *Pseudomonas* megaplasmid. The presence of this large plasmid may provide a selective advantage to the host cell. However, given their size and gene content, acquisition of these secondary replicons may pose a significant cost [69–71]. The high level of gene variation when compared to publicly available megaplasmids suggests that these secondary replicons frequently undergo gene loss and gain through HGT. The reduced purifying selection and the high prevalence of transposable elements frequently observed on megaplasmids may help to explain why these elements readily acquire foreign DNA [6,64,72]. In fact, mosaic

plasmids such as pJBCL41 and the majority of megaplasmids have a high proportion of mobile genetic elements [73]. The identification of novel megaplasmids may shed light on the evolutionary effects of gene transfer and the selective forces driving antibiotic resistance.

## Disclosure statement

SRP responsible for the Galileo<sup>TM</sup> AMR database for Arc Bio.

## Funding

This work was supported by the Applied Molecular Biosciences Unit-UCIBIO which is financed by national funds from FCT/MCTES (UID/Multi/04378/2019). JB and FG were supported by grants from Fundação para a Ciência e a Tecnologia (SFRH/BD/104095/2014 and SFRH/BPD/95556/2013, respectively). CL is supported by an SB PhD fellowship from FWO Vlaanderen (1S64718N).

## ORCID

João Botelho  <http://orcid.org/0000-0002-2771-2345>  
 Cédric Lood  <http://orcid.org/0000-0001-7826-3378>  
 Sally R. Partridge  <http://orcid.org/0000-0002-0666-8330>  
 Vera van Noort  <http://orcid.org/0000-0002-8436-6602>  
 Rob Lavigne  <http://orcid.org/0000-0001-7377-1314>  
 Filipa Grosso  <http://orcid.org/0000-0002-4345-2723>  
 Luísa Peixe  <http://orcid.org/0000-0001-5810-8215>

## References

- [1] Alekshun MN, Levy SB. Molecular mechanisms of antibacterial multidrug resistance. *Cell*. 2007;128:1037–1050.
- [2] Smillie C, Garcillán-Barcia MP, Francia MV, et al. Mobility of plasmids. *Microbiol Mol Biol Rev*. 2010;74:434–452.
- [3] Garcillán-Barcia MP, Alvarado A, de la Cruz F. Identification of bacterial plasmids based on mobility and plasmid population biology. *FEMS Microbiol Rev*. 2011;35:936–956.
- [4] Shintani M, Sanchez ZK, Kimbara K. Genomics of microbial plasmids: classification and identification based on replication and transfer systems and host taxonomy. *Front Microbiol*. 2015;6:242.
- [5] Orlek A, Stoesser N, Anjum MF, et al. Plasmid classification in an era of whole-genome sequencing: application in studies of antibiotic resistance epidemiology. *Front Microbiol*. 2017;8:182.
- [6] diCenzo GC, Finan TM. The divided bacterial genome: structure, function, and evolution. *Microbiol Mol Biol Rev*. 2017;81:e00019–17.
- [7] Harrison PW, Lower RPJ, Kim NKD, et al. Introducing the bacterial ‘chromid’: not a chromosome, not a plasmid. *Trends Microbiol*. 2010;18:141–148.
- [8] Pesesky MW, Tilley R, Beck DAC. Mosaic plasmids are abundant and unevenly distributed across prokaryotic taxa. *Plasmid*. 2019;102:10–18.
- [9] Partridge SR, Kwong SM, Firth N, et al. Mobile genetic elements associated with antimicrobial resistance. *Clin Microbiol Rev*. 2018;31:e00088–17.

- [10] Thomas CM, Haines AS. Plasmids of the genus *Pseudomonas*. *Pseudomonas*. 2004. Springer US, Boston, MA, 197–231.
- [11] Botelho J, Grosso F, Quinteira S, et al. The complete nucleotide sequence of an IncP-2 megaplasmid unveils a mosaic architecture comprising a putative novel *bla*<sub>VIM-2</sub>-harbouring transposon in *Pseudomonas aeruginosa*. *J Antimicrob Chemother*. 2017;72:2225–2229.
- [12] Sun F, Zhou D, Wang Q, et al. The first report of detecting the *bla*<sub>SIM-2</sub> gene and determining the complete sequence of the SIM-encoding plasmid. *Clin Microbiol Infect*. 2016;22:347–351.
- [13] Cazares A, Moore MP, Grimes M, et al. A megaplasmid family responsible for dissemination of multidrug resistance in *Pseudomonas*. *bioRxiv*. 2019:630780. Available from: <https://www.biorxiv.org/content/10.1101/630780v1>.
- [14] San Millan A. Evolution of plasmid-mediated antibiotic resistance in the clinical context. *Trends Microbiol*. 2018;26:978–985.
- [15] Diene SM, Rolain J-M. Carbapenemase genes and genetic platforms in gram-negative bacilli: *Enterobacteriaceae*, *Pseudomonas* and *Acinetobacter* species. *Clin Microbiol Infect*. 2014;20:831–838.
- [16] Rozwandowicz M, Brouwer MSM, Fischer J, et al. Plasmids carrying antimicrobial resistance genes in *Enterobacteriaceae*. *J Antimicrob Chemother*. 2018;73:1121–1137.
- [17] EFSA Panel on Biological Hazards (BIOHAZ). Scientific opinion on carbapenem resistance in food animal ecosystems. *EFSA J*. 2013;11:3501. Available from: <https://efsa.onlinelibrary.wiley.com/doi/abs/10.2903/j.efsa.2013.3501>.
- [18] Greig DR, Dallman TJ, Hopkins KL, et al. MinION nanopore sequencing identifies the position and structure of bacterial antibiotic resistance determinants in a multidrug-resistant strain of enteroaggregative *Escherichia coli*. *Microb Genomics*. 2018;4:e000213.
- [19] Phan HTT, Stoesser N, Maciucă IE, et al. Illumina short-read and MinION long-read WGS to characterize the molecular epidemiology of an NDM-1 *Serratia marcescens* outbreak in Romania. *J Antimicrob Chemother*. 2018;73:672–679.
- [20] Dong N, Lin D, Zhang R, et al. Carriage of *bla*<sub>KPC-2</sub> by a virulence plasmid in hypervirulent *Klebsiella pneumoniae*. *J Antimicrob Chemother*. 2018;73:3317–3321.
- [21] Ludden C, Reuter S, Judge K, et al. Sharing of carbapenemase-encoding plasmids between *Enterobacteriaceae* in UK sewage uncovered by MinION sequencing. *Microb Genomics*. 2017;3:e000114.
- [22] George S, Pankhurst L, Hubbard A, et al. Resolving plasmid structures in *Enterobacteriaceae* using the MinION nanopore sequencer: assessment of MinION and MinION/Illumina hybrid data assembly approaches. *Microb Genomics*. 2017;3:e000118.
- [23] Lemon JK, Khil PP, Frank KM, et al. Rapid nanopore sequencing of plasmids and resistance gene detection in clinical isolates. *J Clin Microbiol*. 2017;55:3530–3543.
- [24] Botelho J, Grosso F, Peixe L. Characterization of the pJB12 plasmid from *Pseudomonas aeruginosa* reveals Tn6352, a novel putative transposon associated with mobilization of the *bla*<sub>VIM-2</sub>-harboring In58 integron. *Antimicrob Agents Chemother*. 2017;61:e02532–16.
- [25] Botelho J, Grosso F, Quinteira S, et al. Two decades of *bla*<sub>VIM-2</sub>-producing *Pseudomonas aeruginosa* dissemination: an interplay between mobile genetic elements and successful clones. *J Antimicrob Chemother*. 2018;73:873–882.
- [26] Bonnin RA, Poirel L, Nordmann P, et al. Complete sequence of broad-host-range plasmid pNOR-2000 harbouring the metallo-β-lactamase gene *bla*<sub>VIM-2</sub> from *Pseudomonas aeruginosa*. *J Antimicrob Chemother*. 2013;68:1060–1065.
- [27] Xiong J, Alexander DC, Ma JH, et al. Complete sequence of pOZ176, a 500-kilobase IncP-2 plasmid encoding IMP-9-mediated carbapenem resistance, from outbreak isolate *Pseudomonas aeruginosa* 96. *Antimicrob Agents Chemother*. 2013;57:3775–3782.
- [28] Arredondo-Alonso S, Willems RJ, van Schaik W, et al. On the (im)possibility of reconstructing plasmids from whole-genome short-read sequencing data. *Microb Genomics*. 2017;3:e000128.
- [29] T'Syen J, Raes B, Horemans B, et al. Catabolism of the groundwater micropollutant 2,6-dichlorobenzamide beyond 2,6-dichlorobenzoate is plasmid encoded in *Aminobacter* sp. MSH1. *Appl Microbiol Biotechnol*. 2018;102:7963–7979.
- [30] Albers P, Lood C, Öztürk B, et al. Catabolic task division between two near-isogenic subpopulations co-existing in a herbicide-degrading bacterial consortium: consequences for the interspecies consortium metabolic model. *Environ Microbiol*. 2018;20:85–96.
- [31] Keshavarz-Tohid V, Vacheron J, Dubost A, et al. Genomic, phylogenetic and catabolic re-assessment of the *Pseudomonas putida* clade supports the delineation of *Pseudomonas alloputida* sp. nov., *Pseudomonas inefficax* sp. nov., *Pseudomonas persica* sp. nov., and *Pseudomonas shirazica* sp. nov. *Syst Appl Microbiol*. 2019;42:468–480.
- [32] Varghese NJ, Mukherjee S, Ivanova N, et al. Microbial species delineation using whole genome sequences. *Nucleic Acids Res*. 2015;43:6761–6771.
- [33] Pritchard L, Glover RH, Humphris S, et al. Genomics and taxonomy in diagnostics for food security: soft-rotting enterobacterial plant pathogens. *Anal Methods*. 2016;8:12–24.
- [34] Andrews S. FastQC: a quality control tool for high throughput sequence data. 2010. Available from: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.
- [35] Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014;30:2114–2120.
- [36] Wick RR, Judd LM, Gorrie CL, et al. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. *PLOS Comput Biol*. 2017;13:e1005595.
- [37] Wick RR, Schultz MB, Zobel J, et al. Bandage: interactive visualization of *de novo* genome assemblies: Fig. 1. *Bioinformatics*. 2015;31:3350–3352.
- [38] Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics*. 2014;30:2068–2069.
- [39] Marchler-Bauer A, Derbyshire MK, Gonzales NR, et al. CDD: NCBI's conserved domain database. *Nucleic Acids Res*. 2015;43:D222–D226.
- [40] Marchler-Bauer A, Bo Y, Han L, et al. CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. *Nucleic Acids Res*. 2017;45:D200–D203.
- [41] Huerta-Cepas J, Szklarczyk D, Forslund K, et al. eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. *Nucleic Acids Res*. 2016;44:D286–D293.



- [42] Emms DM, Kelly S. Orthofinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biol.* 2015;16:157.
- [43] Krzywinski M, Schein J, Birol I, et al. Circos: an information aesthetic for comparative genomics. *Genome Res.* 2009;19:1639–1645.
- [44] Siguier P, Perochon J, Lestrade L, et al. ISfinder: the reference centre for bacterial insertion sequences. *Nucleic Acids Res.* 2006;34:D32–D36.
- [45] Partridge SR, Tsafnat G. Automated annotation of mobile antibiotic resistance in Gram-negative bacteria: the Multiple antibiotic resistance Annotator (MARA) and database. *J Antimicrob Ch.* 2018;73:883–890.
- [46] Quinteira S, Sousa JC, Peixe L. Characterization of In100, a new integron carrying a metallo- $\beta$ -lactamase and a carbenicillinase, from *Pseudomonas aeruginosa*. *Antimicrob Agents Ch.* 2005;49:451–453.
- [47] Quinteira S, Peixe L. Multiniche screening reveals the clinically relevant metallo-beta-lactamase VIM-2 in *Pseudomonas aeruginosa* far from the hospital setting: an ongoing dispersion process? *Appl Environ Microbiol.* 2006;72:3743–3745.
- [48] Diaz-Orejas R, Espinosa M, Yeo CC. The importance of the expendable: toxin–antitoxin genes in plasmids and chromosomes. *Front Microbiol.* 2017;8:1479.
- [49] Janssen PJ, Van Houdt R, Moors H, et al. The complete genome sequence of *Cupriavidus metallidurans* strain CH34, a master survivalist in harsh and anthropogenic environments. *PLoS One.* 2010;5:e10433.
- [50] Tett A, Spiers AJ, Crossman LC, et al. Sequence-based analysis of pQBR103; a representative of a unique, transfer-proficient mega plasmid resident in the microbial community of sugar beet. *ISME J.* 2007;1:331–340.
- [51] Sun F, Zhou D, Wang Q, et al. Genetic characterization of a novel *bla*<sub>DIM-2</sub>-carrying megaplasmid p12969-DIM from clinical *Pseudomonas putida*. *J Antimicrob Ch.* 2016;71:909–912.
- [52] Smith BA, Leligdon C, Baltrus DA. Just the two of us? A family of *Pseudomonas* megaplasmids offers a rare glimpse into the evolution of large mobile elements. *Genome Biol Evol.* 2019;11:1192–1206.
- [53] Thomson NR, Cerdeño-Tárraga AM, Brown CJ, et al. Annotation of plasmid genes. *Plasmid.* 2017;91:61–67.
- [54] Gullberg E, Albrecht LM, Karlsson C, et al. Selection of a multidrug resistance plasmid by sublethal levels of antibiotics and heavy metals. *MBio.* 2014;5:e01918–14.
- [55] Moura A, Soares M, Pereira C, et al. INTEGRALL: a database and search engine for integrons, integrases and gene cassettes. *Bioinformatics.* 2009;25:1096–1098.
- [56] Botelho J, Roberts AP, León-Sampedro R, et al. Carbapenemases on the move: it's good to be on ICEs. *Mob DNA.* 2018;9:37.
- [57] Oliver A, Mulet X, López-Causapé C, et al. The increasing threat of *Pseudomonas aeruginosa* high-risk clones. *Drug Resist Updat.* 2015;21–22:41–59.
- [58] Tsafnat G, Copty J, Partridge SR. RAC: repository of antibiotic resistance cassettes. *Database.* 2011;2011:bar054.
- [59] Partridge SR, Brown HJ, Stokes HW, et al. Transposons Tn1696 and Tn21 and their integrons In4 and In2 have independent origins. *Antimicrob Agents Ch.* 2001;45:1263–1270.
- [60] Allmeier H, Cresnar B, Greck M, et al. Complete nucleotide sequence of Tn1721: gene organization and a novel gene product with features of a chemotaxis protein. *Gene.* 1992;111:11–20.
- [61] Szuplewska M, Czarnecki J, Bartosik D. Autonomous and non-autonomous Tn3-family transposons and their role in the evolution of mobile genetic elements. *Mob Genet Elements.* 2014;4:1–4.
- [62] Pfeiffer F, Zamora-Lagos M-A, Blettinger M, et al. The complete and fully assembled genome sequence of *Aeromonas salmonicida* subsp. *pectinolytica* and its comparative analysis with other *Aeromonas* species: investigation of the mobilome in environmental and pathogenic strains. *BMC Genomics.* 2018;19:20.
- [63] Haines AS, Jones K, Cheung M, et al. The IncP-6 plasmid Rms149 consists of a small mobilizable backbone with multiple large insertions. *J Bacteriol.* 2005;187:4728–4738.
- [64] Cooper VS, Vohr SH, Wrocklage SC, et al. Why genes evolve faster on secondary chromosomes in bacteria. *PLoS Comput Biol.* 2010;6:e1000732.
- [65] Rocha EPC, Danchin A. Base composition bias might result from competition for metabolic resources. *Trends Genet.* 2002;18:291–294.
- [66] Popa O, Hazkani-Covo E, Landan G, et al. Directed networks reveal genomic barriers and DNA repair bypasses to lateral gene transfer among prokaryotes. *Genome Res.* 2011;21:599–609.
- [67] Baxter JC, Funnell BE. Plasmid partition mechanisms. *Microbiol Spectr.* 2014;2. Available from: <https://www.asmscience.org/content/journal/microbiolspec/10.1128/microbiolspec.PLAS-0023-2014>.
- [68] Zhu W, Luo L, Wang J, et al. Complete nucleotide sequence of pCTX-M360, an intermediate plasmid between pEL60 and pCTX-M3, from a multidrug-resistant *Klebsiella pneumoniae* strain isolated in China. *Antimicrob Agents Ch.* 2009;53:5291–5293.
- [69] Romanchuk A, Jones CD, Karkare K, et al. Bigger is not always better: transmission and fitness burden of ~1 MB *Pseudomonas syringae* megaplasmid pMPPla107. *Plasmid.* 2014;73:16–25.
- [70] Dougherty K, Smith BA, Moore AF, et al. Multiple phenotypic changes associated with large-scale horizontal gene transfer. *PLoS One.* 2014;9:e102170.
- [71] San Millan A, MacLean RC. Fitness costs of plasmids: a limit to plasmid transmission. *Microbiol Spectr.* 2017;5. Available from: <https://www.asmscience.org/content/journal/microbiolspec/10.1128/microbiolspec.MTBP-0016-2017>.
- [72] Rankin DJ, Rocha EPC, Brown SP. What traits are carried on mobile genetic elements and why? *Heredity (Edinb).* 2011;106:1–10.
- [73] Pesesky MW, Tilley R, Beck DAC. Mosaic plasmids are abundant and unevenly distributed across prokaryotic taxa. *Plasmid.* 2019;102:10–18.