Characterization of two multidrug-resistant IncA/C plasmids from the 1960s by using Oxford Nanopore MinION sequencer device

Mónika Szabó, Tibor Nagy, Tímea Wilk, Tibor Farkas, Anna Hegyi, Ferenc Olasz, János Kiss#

Agricultural Biotechnology Institute, National Agricultural Research and Innovation Centre, Gödöllő, Hungary

Running Title: Early IncA/C plasmids analyzed by using MinION device

Address correspondence to János Kiss, kiss.janos@abc.naik.hu

T. N. and T. W. contributed equally to this work.

Two IncA/C family plasmids from the 1960s have been sequenced and classified into A/C_2 Type 1 group. R16a and IP40a contain novel antibiotic resistance islands and a complete GIsul2 island not previously found in the family. In the 173.1 kb R16a the 29.9 kb ARI is located in a unique backbone position not utilized by ARIs. ARI_{R16a} consists of Tn*1*, Tn*6020*, Tn*6333* harboring the resistance genes bla_{TEM-1D}, *aphA1b* and a *mer* module, respectively, a truncated Tn*5393* copy and a gene cluster with unknown function. Plasmid IP40a is 170.4 kb in size and contains a 5.6 kb ARI inserted into *kfrA* gene. ARI_{IP40a} carrying bla_{TEM-1D} and *aphA1b* genes is composed of Tn*1* with Tn*6023* insertion. Additionally, IP40a harbors single IS*2*, IS*186* and Tn*1000* insertions scattered in the backbone, an IS*150* copy in GIsul2 and a complete Tn*6333* encoding a *mer* module at the position of ARI_{R16a}. Loss of resistance markers in R16a, IP40a and R55 was observed during stability tests. Every phenotypic change proved to be the result of recombination events involving mobile elements. Intramolecular transposition of IS copies that generated IP40a derivatives lacking large parts of the backbone could account for the formation of other family members, too. MinION platform proved to be a valuable tool in bacterial genome sequencing since it generates long reads that span repetitive elements and facilitates full length plasmid or chromosome assembly. Nanopore technology enables rapid characterization of large, low-copy plasmids and their rearrangement products.

INTRODUCTION

Plasmids belonging to the A/C incompatibility group (IncA/C) are large, low-copy, conjugative extrachromosomal elements that often encode resistance genes(1)(2)(3)(4). The efficient conjugation system, broad host range of IncA/C plasmids and their ability to mobilize multidrug-resistant genomic islands (GIs) are presumably responsible for the rapid dissemination of resistance genes among Gram-negative enteric bacteria(5)(6). IncA/C plasmids are classified into two distinct subgroups(2)(7) designated as A/C_1 (RA1 as sole sequenced member) and A/C_2 (all the other plasmids sequenced to date). Historically, the A/C complex was created by combining the IncA group, consisting only of RA1 and RA2(8), with the IncC group, which has recently been suggested to correspond to the A/C_2 subgroup(9). A/C_2 plasmids are further characterized as Type 1 or Type 2 based on the alleles of *orf1832/1847* and *rhs1/2* genes they possess in loci R1 and R2, respectively, and the presence/absence of two short insertions (i1 and i2)(10)(2).

Comparative genomics of IncA/C plasmids revealed their modular structure: a conserved backbone contains genes required for maintenance (*rep*, *par*) and conjugative transfer (*tra*, *aca*), while the variable accessory modules are often identified as antibiotic resistance islands (ARIs) that harbor various resistance determinants associated with complex arrays of transposons (Tn*3*-family, Tn*7*-like and IS*CR* elements, IS*26*) and integrons(11)(12)(2). IncA/C plasmids contain ARIs (ARI-A in Type 1 and ARI-B in Type 1 and 2) at two specific positions, though in some family members ARIs can also be found in the *rhs-kfrA* region(2). Relatively little is known about the origin of ARIs, but ARI-Bs presumably evolved by incorporation of GIsul2(13) into the IncA/C backbone in the early stage of evolution and subsequent internal replacements and rearrangements of the island(2). However, a direct evidence for this assumption, namely a family member containing a complete GIsul2, has not been reported so far.

In the course of sequencing plasmids and bacterial chromosomes that harbor repetitive elements (e.g. IS elements and transposons) in multiple copies, it is challenging to assemble the short reads generated by new generation sequencing (NGS) platforms into a single contig. Third generation (i.e. single molecule, real-time and nanopore) sequencing produces long reads that can span repetitive elements and therefore facilitate more contiguous assembly of NGS contigs. Oxford Nanopore Technologies' (ONT) MinION, a portable DNA sequencer device, detects bases of ssDNA passing through a nanopore(14). Despite the relatively high error rate, MinION reads are suitable for de novo

assembly of complete genomes(15)(16)(17)(18), scaffolding NGS contigs(19), metagenomic studies(20)(21) and real-time epidemiological investigations(22)(23). The main advantages of MinION platform over NGS technologies are long reads (there is no theoretical limit of read length), low investment cost per device, portability, flexible run and reduced turnaround time(22)(24).

Plasmids R16a and IP40a (R40a)(25) were isolated in the Pasteur Institute from abscess and urine samples collected in 1966 (St-Antoine Hospital, Paris) and 1969 (Necker Hospital, Paris), respectively. The nucleotide sequences of the two IncA/C plasmids and their spontaneous rearrangement products have been determined in the framework of the ONT MinION Access Programme (MAP). Both plasmids belong to A/C_2 Type 1 group and carry *aphA1b*, $bla_{TEM\text{-}1D}$ and *sul2* genes conferring the previously determined resistance to kanamycin, ampicillin and sulphonamides(26), respectively. In addition, both plasmids provide resistance to mercury. The rearrangements detected in R16a, IP40a and R55(25) seem to be associated with transposons. The presence of relatively few and archaic antibiotic resistance genes and the complete GIsul2 island, which is absent from all the known IncA/C plasmids, suggests that these plasmids represent an ancestral stage in Type 1 lineage.

MATERIALS AND METHODS

DNA and microbial techniques. Plasmid DNA was extracted by using QIAGEN Plasmid Midi kit (Qiagen). R16a and IP40a were purified from their *E. coli* K-12 TG90F$^-$ transconjugants obtained from crosses with *E. coli* K-12 J53 donor strains (gifts from Benoît Doublet). The recipient strain TG90F$^-$ derived from TG90(27) by curing F'. Deletion derivatives were isolated by replica plating following 1-5 passages of TG1Nal(28) transconjugants harboring R16a, IP40a or R55 under non-selective conditions(28). Mercury resistance of TG1Nal, TG1Nal/R16a and TG1Nal/IP40a strains was tested as follows: cells were grown in LB until OD_{600}~0.8, serially diluted tenfold to 10^7, then each dilution was dropped onto LB plates supplemented with 0, 2.5, 5, 10, 15 or 20 µg/ml $HgCl_2$ and incubated overnight at 30 ºC. Bacterial strains were routinely grown at 37 ºC in LB supplemented (if applicable) with ampicillin (150 µg/ml), chloramphenicol (20 µg/ml), kanamycin (30 µg/ml), nalidixic acid (20 µg/ml), gentamicin (25 µg/ml), tetracycline (10 µg/ml).

Sequencing and analysis. R16a and IP40a were sequenced on Illumina platform at the Department of Biochemistry and Molecular Biology, Faculty of Science and Informatics, University of Szeged,

Hungary. MiSeq 2×300 bp paired-end reads were de novo assembled using MIRA 4.0.2(29). Gap closure and exact determination of deletion endpoints were performed by Sanger sequencing of PCR amplicons covering these regions on ABI 3130xl instrument. PCRs were carried out using Pwo (Roche) or Phusion polymerases and Long PCR enzyme mix (Thermo Scientific) according to the manufacturer's recommendations. For oligonucleotide primers and detailed workflow see Table S1 and Text S1.

MinION libraries were prepared using SQK-MAP005 (deletion derivatives) and SQK-MAP006 (original plasmids) kits according to the manufacturer's instructions and sequenced on R7.3 flow cells. Fast5 read files were base-called via 2D workflow of ONT's Metrichor software. Earlier versions of 2D workflow (v.1.12-1.14) yielded 1D and 2D reads for deletion derivatives, while higher versions (v.1.16) resulted only 2D reads for R16a and IP40a (Table 1).

R16a and IP40a MinION reads were de novo assembled into a single full length contig by Minimap and Miniasm(30). LAST v. 393(31) was used to align MinION reads from the deletion derivatives to the complete R16a, IP40a and R55 reference sequences, and to map MiSeq MIRA-contigs to the corresponding single Miniasm-contigs of R16a and IP40a. Final sequences of plasmids were annotated using RAST version 2.0(32) and corrected manually.

Libraries of deletion derivatives prepared by SQK-MAP005 kit and protocol yielded less favorable read length distribution (Figure S1), lower number of 2D reads (Table 1), and therefore significantly lower coverage and accuracy compared to R16a and IP40a libraries generated by SQK-MAP006. Thus, for determination of deletions all 1D and 2D reads were mapped to the complete reference sequences. Endpoints were identified by sequencing the PCR amplicons spanning the deleted regions (Text S1).

Accession numbers. Complete sequences of R16a and IP40a were deposited in GenBank under the accession numbers KX156773 and KX156772, respectively. MiSeq Fastq and base-called MinION Fast5 read files are available under the BioProject PRJNA318408 (https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA318408), while the Miniasm-contigs (g00003c and g00001c) and Mira-contigs generated for the two plasmids are available at http://emboss.abc.hu/minionarticle/. Accession numbers of all the other sequenced IncA/C plasmids are listed in Table S2.

RESULTS AND DISCUSSION

Plasmid sequencing. The loss of resistance markers was observed when plasmids R55, R16a and IP40a were passaged without selection. To study the background of rearrangements, the yet unknown plasmids R16a and IP40a were sequenced on Illumina platform. After filtering *E. coli* chromosomal sequences, 22 and 8 contigs were obtained from R16a (N50: 158234; coverage 115; 68555 mapped reads; Phred scores, r1: 36, r2: 31) and IP40a (N50: 31122; coverage 120; 69652 mapped reads; Phred scores, r1: 36, r2: 29), respectively. Due to the abundance of repetitive elements MiSeq reads alone could not be assembled unambiguously into single contigs. However, de novo assembly of long MinION reads (Figure S1) by using Miniasm software, which was developed for the efficient mapping and assembling of low accuracy reads without error correction, yielded full-length contigs, even for IP40a where relatively low coverage was achieved (Table 1). Single Miniasm-contigs were used for aligning the corresponding MIRA-contigs (Figure S2). Deduced plasmid sequences were confirmed by PCR and direct sequencing of gap-bridging amplicons (Text S1).

Comparative analysis of R16a and IP40a. Sequence analysis revealed that R16a and IP40a possess all distinctive traits of A/C_2 Type 1 group(2) as they carry *orf1832* and *rhs*1 alleles at R1 and R2 loci, and lack i1 and i2 insertions (Table 2). Both plasmids harbor an intact GIsul2, the presumptive ancestor of ARI-Bs, in the standard ARI-B site and unique ARIs (ARI_{R16a} and ARI_{IP40a}) at unconventional sites, while neither contains ARI-A. A newly identified Tn*3*-family transposon Tn*6333* was found in the same position of R16a and IP40a. While it is intact in IP40a, the insertion of a transposon-like gene cluster in Tn*6333* led to the formation of ARI_{R16a} in R16a.

The two plasmids are apparently very close relatives (Figure 1A) as their 127.7 kb backbones only differ in 12 SNPs and they both contain a 131-bp deletion unique among the known family members. The deletion causes frameshift after the Thr61 codon of a putative nitrite reductase gene downstream of *dsbA* (starting at 11670 bp in R16a). Furthermore R16a carries an 11-bp insertion upstream of *mobI* and a single base insertion upstream of a DNA primase gene near to *rhs1*. Compared to the ca. 40 years later isolated R148(33) the backbones of R16a and IP40a differ in 51 and 53 SNPs, respectively, and have the same 3-bp insertion and 1-bp deletion (the latter is probably a sequencing mistake in a 7-8-bp oligoT-stretch). The IP40a backbone contains an additional 1-bp insertion. The

presence of complete GIsul2 is exceptional among the known IncA/C plasmids and suggests that these plasmids represent ancestral forms in Type 1 lineage.

The major differences between the plasmids are in the structure of ARIs and the presence of four additional transposons in IP40a. These Tns are typical for *E. coli*, though IS*186* and Tn*1000* ($\gamma\delta$) can be found in strains of *P. aeruginosa*, the original host of IP40a, while IS*2* and IS*150* have also been described from several species (*Shigella, Salmonella*, *Proteus, Morganella*, *Yersinia*, etc.). The question whether IP40a had captured these elements before its isolation in 1969 or after its transfer to *E. coli* K-12 J53, during the ca. 45 years of laboratory maintenance, is difficult to answer. The first scenario is supported, however, by the fact that R16a did not acquire these elements even though it has undergone similar procedures (Benoît Doublet, pers. comm.). Furthermore, Tn*1000*, a regular resident of the F plasmid in *E. coli* K-12, can derive neither from the original donor J53 strain, which is F^- and has no genomic Tn*1000* copy (NZ_AICK01000000), nor the F'-cured recipient TG90F^- strain (see. Materials and methods).

Structure of R16a ARIs. The 173094 bp R16a contains the ARI_{R16a} integrated into the backbone at a site (127177 bp) not utilized by ARIs in other family members (Figure 1B). ARI_{R16a} is composed of the IS*26*-based compound transposon Tn*6020*(34) and three different Tn*3*-family members, i.e. Tn*1*(35), Tn*5393*(36) and the newly identified Tn*6333*. All resistance genes of ARI_{R16a} are located in transposons Tn*1* ($bla_{TEM\text{-}1D}$), Tn*6020* (*aphA1b*) and Tn*6333* (*mer* module). Interestingly, integrons, often accumulating resistance genes in other family members, were not found either on R16a or IP40a. ARI_{R16a} is delimited by IRs of Tn*6333* inserted at position 124178 bp of R16a (145801 bp in R148). The Tn*6333* transposase (Tpase) gene is interrupted by a transposon-like cluster flanked by an intact and a truncated copy of Tn*1* elements in direct orientation (Figure 1B). The full length Tn*1* shows 99% homology to Tn*1* of IncP plasmid RK2(37) (4936/4950 bp identity and 2 gaps) and its closest relatives were found in *Shigella dysenteriae* plasmids pA5468 and pBU53M1 (4942/4949 bp identity). The truncated Tn*1* has 7 SNPs compared to the corresponding region of the intact copy, and its closest relatives, differing in 6 SNPs and an 1-bp insertion, can be found in several plasmids and genomic islands. The Tn*1*-delimited cluster, which cannot be found in GenBank as a unit, includes a yet undescribed 5757 bp region carrying 12 orfs, a truncated copy of Tn*5393* and a compound transposon Tn*6020* consisting of two identical, directly oriented IS*26* elements. The Tn*5393* fragment

has 100% homologues in numerous plasmids and bacterial chromosomes, but it differs from the corresponding region of Tn*5393* prototype described from *Erwinia amylovora* plasmid pEa34(36) in 7 SNPs. The sequences of IS*26* elements in Tn*6020* are identical to IS*15*DII (https://www-is.biotoul.fr//)(38) and contain 1 SNP compared to IS*26* (GenBank: X00011).

The lack of 8-bp target duplication (TD) typical for IS*6* family (http://www-is.biotoul.fr) indicates that after insertion Tn*6020* generated a deletion via intramolecular transposition leading to the actual structure, where Tn*6020* is bracketed by the two truncated Tn*3*-family elements. On the other hand, the IRs of Tn*6333* and the outer IRs of the Tn*1*-based cluster are delimited by 5-bp TD (Figure 1B) characteristic for Tn*3* family transposons. This arrangement of TDs and the inactivation of Tn*6333* Tpase by Tn*1* insertion suggest a scenario in which Tn*6333* integration preceded the insertion of Tn*1*-based cluster.

The other island in R16a, the complete GIsul2 located in the previously described site of ARI-B, contains the *sul2* gene and an arsenic resistance operon (*ars*^{R}). GIsul2 has been found in the chromosome and/or plasmids of several species, such as *Providencia stuartii*, *Escherichia coli*, *Sphingopyxis granuli*, *Morganella morganii*, *Shigella flexneri*, *Achromobacter xilosoxidans*, *Enterobacter cloacae*, *Acinetobacter baumannii*, but not in other IncA/C plasmids. Compared to the copy described from *A. baumannii* ATCC 17978(13) GIsul2 in R16a is 278 bp longer and it contains five SNPs and six 1-bp indels, but it is almost identical to the variants found in *P. stuartii* ATCC 33672 chromosome and pHUSEC411-like plasmid of *E. coli* PMV-1 (2 SNPs). GIsul2 in R16a, similarly to other intact GIsul2 copies found in databases, is bordered by GGGA direct repeats (Table S3). The chromosomal GIsul2 islands are all integrated into the 3' end of a GMP synthase gene *guaA*, while the plasmid-borne copies are located at different sites, such as the intergenic region near *tolA* gene in pHUSEC411-like plasmid, the topoisomerase gene in R485, or the 3' end of a hypothetical gene in R16a. Although the short direct repeats are reminiscent of TDs generated via transposition, the preference for the *guaA* integration site and the lack of Tpase able to produce TDs may support the idea that GIsul2 is an integrating element(13).

Structure of IP40a ARIs. The 170404 bp IP40a has the same resistance genes in similar components as R16a, but arranged into three blocks (Figure 1B). In the position of ARI_{R16a} IP40a harbors an intact Tn*6333* copy including the *mer* module. This situation may represent the initial state of ARI_{R16a}

formation as it was suggested above. *Bla*$_{TEM-1D}$ and *aphA1b* genes are carried by a unique transposon-in-transposon structure ARI$_{IP40a}$ inserted into the *kfrA* gene at a position not utilized by other ARIs in the IncA/C family (Table S2). The first step of ARI$_{IP40a}$ evolution was most likely the simple insertion of the *bla*$_{TEM-1D}$-bearing Tn*1* into *kfrA*. The presence of a 5-bp TD delimiting Tn*1* IRs clearly supports this idea. The second step must have been the Tn*6023* insertion that was presumably followed by an intramolecular transposition event. This removed the inner part (1507-3441 bp) of Tn*1* (the remaining parts differ from the full length copy of R16a in 3 SNPs), inactivated the Tpase gene by deleting its 5' half and led to the present state, where no TD flanks the outer IRs of Tn*6023*. *AphA1b* is carried by Tn*6023* composed of two identical, inversely oriented IS*26* elements (Figure 1B). These IS*26* copies are identical to IS*15*DI and carry 3 SNPs compared to the reference sequence of IS*26* (https://www-is.biotoul.fr//). The IP40a-borne Tn*6023* is not identical to the prototype found in pSRC125,(39) in which the *aphA1* cassette is 21 bp longer and the second IS*26* copy contains 3 SNPs. It also differs from Tn*6020* of R16a in the orientation and sequence of the IS elements (4 SNPs), and in the intervening segment that contains the resistance gene without any additional orf. The other resistance genes (*sul2* and *ars*^{R} operon) reside in GIsul2, which has 2 SNPs compared to the R16a-borne version and an IS*150* insertion in the integrase gene, but resides at the same position as in R16a and is also flanked by GGGA direct repeats.

Description of Tn*6333* The novel 11514 bp transposon, Tn*6333*, is delimited by 37-bp inverted repeats and generates 5-bp target duplication (Figure 1C). The transposon shares the highest homology with two unidentified transposons found in the genome of *Providencia stuartii* strain ATCC 33672 (GenBank: CP008920) and *Shewanella* sp. ANA-3 plasmid 1 (GenBank: CP000470), respectively. Compared to Tn*6333* the Tn copy in *P. stuartii* is identical except it has an IS*26* insertion in the Tpase gene, while the one on *Shewanella* sp. plasmid has 21 SNPs, a 12-bp internal deletion and 451-bp deletion effecting the left end. Based on the amino acid sequence of the transposase, Tn*6333* was classified into the Tn*3* family. The transposon also contains a resolvase gene (*res*), a putative partitioning protein gene (*parA*), 4 hypothetical orfs with unknown function and the 7 orfs of *mer* module. The *mer* module, which is located between the Tpase and resolvase genes, contains a mercuric reductase gene (*merA*), mercuric transport protein genes (*merC,P,T*), a Hg(II)-responsive transcriptional regulator gene (*merR*) and a short orf encoding a putative MerR-family transcription

coregulator domain. The module also includes an orf encoding a putative signaling protein with the conserved EAL domain (Figure 1C). The m*er* module provides equal level of resistance to mercury for both plasmids. Five µg/ml $HgCl_2$ that was completely inhibitory for the growth of the negative control TG1Nal strain had no significant effect on the titer of plasmid bearing derivatives (LB+5 µg/ml $HgCl_2$: 4.0×10^8; LB:1.7×10^9 CFU/ml), and single colonies were obtained even on 15 µg/ml concentration, but with 4 orders of magnitude less frequency than on LB or LB+5 µg/ml (LB+15 µg/ml $HgCl_2$: 7.0×10^4 CFU/ml).

Analysis of spontaneous rearrangements detected in three IncA/C plasmids.

R55, a previously sequenced Type 2 plasmid isolated simultaneously with R16a and IP40a, was included in the stability tests, where the loss of resistance markers was examined. Frequency of plasmid species that lost one or several markers was in the range of $1.8\text{-}38.0\times10^{-3}$ in the passaged cell populations (Table S4). We first examined whether homologous recombination between directly repeated elements of the original plasmids could lead to the observed phenotypes as it was described for SGI1 variants(28). Considering the structure of ARI_{R16a}, Km^SAp^R R16a derivatives can arise by recombination between the directly oriented IS*26* or Tn*1* copies. Similarly, the observed $Km^SGm^SAp^SFlo^RCm^R$ phenotype in R55 can be the result of deletions between the directly oriented IS*5075* copies bracketing Tn*6187*. Km^SAp^S IP40a derivatives, however, could not be formed this way due to the lack of extensive direct repeats around the resistance region.

R16a and R55 derivatives were tested by PCRs indicative for the presumed deletions (Text S1). Type R16a_d1, R16a_d3 and R55_d11 deletions corresponded to the three predicted rearrangements (Figure 2), while deletion types R16a_d2 and R55_d20 could not be mapped this way. The structure of R16a_d1-type deletion, which was independently isolated three times, was consistent with the putative product of homologous recombination between IS*26* elements of Tn*6020*. The single isolate R16a_d3 and R55_d11, which were detected in four independent assays, appeared to derive in a similar way involving two directly repeated Tn*1* segments and IS*5075* elements, respectively (Figure 2). Although the most probable explanation for the formation of the three deletion-types is homologous recombination, R16a_d1 could be the result of IS*26* transposition (40)(41)(42), too.

R16a_d2 and R55_d20, which could not be determined in the first round, were further analyzed by PCR and sequencing, together with IP40a derivatives IP40a_d5 and IP40a_d8, in which the

rearrangements were not predictable from the plasmid sequence. Subsequent PCR test revealed that R16a_d2 carries a deletion between the IRR of the first IS*26* element and the resolvase gene of Tn*6333*, indicating that it has been formed by intramolecular transposition. Rearrangements in R55_d20 and the two IP40a deletions were identified by MinION sequencing. Alignment of MinION reads from R55_d20 to the reference showed that the deletion occurred inside the ARI and started from the IRR of a single IS*1* element located next to the *catA1* gene. The same method revealed that the transposition of the right IS*26* copy of Tn*6023* eliminated considerable part of IP40a backbone in IP40a_d5 and IP40a_d8 as the endpoints of deletions are in the Tn*1000* Tpase gene and the intergenic region near to the left end of Tn*1000,* respectively (Figure 2). Highly truncated versions of the 127 kb conserved backbone, similar to the IP40a derivatives, can be found in numerous IncA/C family members. In these plasmids the remaining backbone segments are often delimited by IS elements (Figure S3) suggesting that their origin is analogous to that of IP40a_d5 and IP40a_d8, and supporting the idea that transposons are key players in remodeling of IncA/C plasmids(43).

Conclusions. Two early IncA/C isolates containing novel ARIs and complete GIsul2 have been described in this work. Compared to other ARIs in the family (e.g. ARI_{R55}), ARI_{R16a} and ARI_{IP40a} have a relatively simple structure that can be explained by a few transposition events. In R16a and IP40a the absence of integrons and “modern” antibiotic resistance genes, and the presence of a complete GIsul2 suggest that they represent an early stage of IncA/C evolution. However, due to the unique 131-bp deletion in both backbones, it is unlikely that these plasmids are direct ancestors of present family members. The analysis of further IncA/C plasmids collected in the pre- and early antibiotic era could contribute to the better understanding of the lineages. The current taxonomy of family members(2), that is based on 4 backbone markers, appears to be adequate for the majority of the ca. 110 plasmid species identified to date. However, classification of emerging deletion derivatives and hybrids (pNDMCFuy, pHM881QN), presumably originating from recombination between Type 1 and 2 species, is ambiguous and further loci should be considered in the classification scheme. The third generation MinION sequencer proved to be a suitable tool for mapping deletions and generating complete sequences from plasmids abundant with repetitive elements. The IS-mediated rearrangements identified in three IncA/C plasmids may elucidate the evolution of ARIs and the origin of family members having truncated backbone sequences.

ACKNOWLEDGEMENTS

We are grateful to Axel Cloeckaert and Benoît Doublet (INRA, Nouzilly, France) for providing us the IncA/C plasmids R16a, IP40a and R55. We thank Oxford Nanopore for the opportunity to participate in the MinION Access Program, and the staff of ONT Support team, in particular Margherita Coccia, for their help. Finally, special thanks to Erika Sztánáné-Keresztúri and Mária Turai for the excellent technical assistance.

FUNDING

This study was funded by a grant from Hungarian Scientific Research Fund (K 105635 to J.K.). MinION device and kits supplied by Oxford Nanopore were received in the frame of ONT's MinION Access Program.

TRANSPARENCY DECLARATIONS

None to declare.

SUPPLEMENTAL MATERIALS

Figure S1-*S3*, Table S1-S4 and Supplemental methods Text S1.

References

1. **Johnson TJ**, **Lang KS**. 2012. IncA/C plasmids An emerging threat to human and animal health? Mob Genet Elements **2:1**:55–58.
2. **Harmer CJ**, **Hall RM**. 2015. The A to Z of A/C plasmids. Plasmid **80**:63–82.
3. **Carattoli A**. 2009. Resistance plasmid families in Enterobacteriaceae. Antimicrob Agents Chemother **53**:2227–2238.
4. **Carattoli A**. 2013. Plasmids and the spread of resistance. Int J Med Microbiol **303**:298–304.
5. **Doublet B**, **Boyd D**, **Mulvey MR**, **Cloeckaert A**. 2005. The Salmonella genomic island 1 is an

integrative mobilizable element. Mol Microbiol **55**:1911–1924.

6. **Carraro N**, **Matteau D**, **Luo P**, **Rodrigue S**, **Burrus V**. 2014. The Master Activator of IncA/C Conjugative Plasmids Stimulates Genomic Islands and Multidrug Resistance Dissemination. PLoS Genet **10**:e1004714.

7. **Carattoli A**, **Miriagou V**, **Bertini A**, **Loli A**, **Colinon C**, **Villa L**, **Whichard JM**, **Rossolini GM**. 2006. Replicon typing of plasmids encoding resistance to newer β-lactams. Emerg Infect Dis **12**:1145–1148.

8. **Datta N**, **Hedges RW**. 1973. R factors of compatibility group A. J Gen Microbiol **74**:335–7.

9. **Harmer CJ**, **Partridge SR**, **Hall RM**. 2016. pDGO100, a type 1 IncC plasmid from 1981 carrying ARI-A and a Tn1696-like transposon in a novel integrating element. Plasmid.

10. **Harmer CJ**, **Hall RM**. 2014. pRMH760, a Precursor of A/C2 Plasmids Carrying blaCMY and blaNDM Genes. Microb Drug Resist **20**:416–423.

11. **Welch TJ**, **Fricke WF**, **McDermott PF**, **White DG**, **Rosso M-L**, **Rasko D a**, **Mammel MK**, **Eppinger M**, **Rosovitz MJ**, **Wagner D**, **Rahalison L**, **LeClerc JE**, **Hinshaw JM**, **Lindler LE**, **Cebula TA**, **Carniel E**, **Ravel J**. 2007. Multiple Antimicrobial Resistance in Plague: An Emerging Public Health Risk. PLoS One **2**:e309.

12. **Fricke WF**, **Welch TJ**, **McDermott PF**, **Mammel MK**, **LeClerc JE**, **White DG**, **Cebula TA**, **Ravel J**. 2009. Comparative Genomics of the IncA/C Multidrug Resistance Plasmid Family. J Bacteriol **191**:4750–4757.

13. **Nigro SJ**, **Hall RM**. 2011. GIsul2, a genomic island carrying the sul2 sulphonamide resistance gene and the small mobile element CR2 found in the Enterobacter cloacae subspecies cloacae type strain ATCC 13047 from 1890, Shigella flexneri ATCC 700930 from 1954 and Acinetobacter bauma. J Antimicrob Chemother **66**:2175–2176.

14. **Ward AC**, **Kim W**. 2015. Minlon^{TM}: New, long read, portable nucleic acid sequencing device. J Bacteriol Virol **45**:285–303.

15. **Goodwin S**, **Gurtowski J**, **Ethe-Sayers S**, **Deshpande P**, **Schatz MC**, **McCombie WR**. 2015. Oxford Nanopore sequencing, hybrid error correction, and de novo assembly of a eukaryotic genome. Genome Res **25**:1750–1756.

16. **Loman NJ**, **Quick J**, **Simpson JT**. 2015. A complete bacterial genome assembled de novo using only nanopore sequencing data. Nat Methods **12**:733–735.

17. **Quick J**, **Quinlan AR**, **Loman NJ**. 2014. A reference bacterial genome dataset generated on the MinION(TM) portable single-molecule nanopore sequencer. Gigascience **3**:22.

18. **Wang J**, **Moore NE**, **Deng YM**, **Eccles DA**, **Hall RJ**. 2015. MinION nanopore sequencing of an influenza genome. Front Microbiol **6**:1–7.

19. **Ashton PM**, **Nair S**, **Dallman T**, **Rubino S**, **Rabsch W**, **Mwaigwisya S**, **Wain J**, **O'Grady J**. 2014. MinION nanopore sequencing identifies the position and structure of a bacterial antibiotic resistance island. Nat Biotechnol **33**:296–300.

20. **Kilianski A**, **Haas JL**, **Corriveau EJ**, **Liem AT**, **Willis KL**, **Kadavy DR**, **Rosenzweig CN**, **Minot SS**. 2015. Bacterial and viral identification and differentiation by amplicon sequencing on the MinION nanopore sequencer. Gigascience **4**:12.

21. **Benitez-Paez A**, **Portune K**, **Sanz Y**. 2015. Species level resolution of 16S rRNA gene amplicons sequenced through MinIONTM portable nanopore sequencer. bioRxiv 021758.

22. **Greninger AL**, **Naccache SN**, **Federman S**, **Yu G**, **Mbala P**, **Bres V**, **Stryke D**, **Bouquet J**, **Somasekar S**, **Linnen JM**, **Dodd R**, **Mulembakani P**, **Schneider BS**, **Muyembe-Tamfum J-J**, **Stramer SL**, **Chiu CY**. 2015. Rapid metagenomic identification of viral pathogens in clinical samples by real-time nanopore sequencing analysis. Genome Med **7**:99.

23. **Quick J**, **Ashton P**, **Calus S**, **Chatt C**, **Gossain S**, **Hawker J**, **Nair S**, **Neal K**, **Nye K**, **Peters T**, **De Pinna E**, **Robinson E**, **Struthers K**, **Webber M**, **Catto A**, **Dallman TJ**, **Hawkey P**, **Loman NJ**. 2015. Rapid draft sequencing and real-time nanopore sequencing in a hospital outbreak of Salmonella. Genome Biol **16**:114.

24. **Judge K**, **Harris SR**, **Reuter S**, **Parkhill J**, **Peacock SJ**. 2015. Early insights into the potential of the Oxford Nanopore MinION for the detection of antimicrobial resistance genes. J Antimicrob Chemother **70**:2775–2778.

25. **Chabbert YA**, **Scavizzi MR**, **Witchitz JL**, **Gerbaud GR**, **Bouanchaud DH**. 1972. Incompatibility Groups and the Classification of f- Resistance Factors. J Bacteriol **112**:666–675.

26. **Douard G**, **Praud K**, **Cloeckaert A**, **Doublet B**. 2010. The Salmonella genomic island 1 is specifically mobilized in trans by the IncA/C multidrug resistance plasmid family. PLoS One **5**:e15302.

27. **Gonzy-Treboul G**, **Karmazyn-Campelli C**, **Stragier P**. 1992. Developmental regulation of transcription of the Bacillus subtilis ftsAZ operon. J Mol Biol **224**:967–979.

28. **Kiss J**, **Nagy B**, **Olasz F**. 2012. Stability, entrapment and variant formation of Salmonella genomic island 1. PLoS One **7**:e32497.

29. **Chevreux B**, **Wetter T**, **Suhai S**. 1999. Genome Sequence Assembly Using Trace Signals and Additional Sequence Information. Ger Conf Bioinforma 45–56.

30. **Li H**. 2015. Minimap and miniasm: fast mapping and de novo assembly for noisy long sequences. arXiv 1–7.

31. **Kielbasa SM**, **Wan R**, **Sato K**, **Horton P**, **Frith MC**. 2011. Adaptive seeds tame genomic sequence comparison. Genome Res **21**:487–493.

32. **Aziz RK**, **Bartels D**, **Best A a**, **DeJongh M**, **Disz T**, **Edwards R a**, **Formsma K**, **Gerdes S**, **Glass EM**, **Kubal M**, **Meyer F**, **Olsen GJ**, **Olson R**, **Osterman AL**, **Overbeek R a**, **McNeil LK**, **Paarmann D**, **Paczian T**, **Parrello B**, **Pusch GD**, **Reich C**, **Stevens R**, **Vassieva O**, **Vonstein V**, **Wilke A**, **Zagnitko O**. 2008. The RAST Server: rapid annotations using subsystems technology. BMC Genomics **9**:75.

33. **Tipmongkolsilp N**, **Castillo CS del**, **Hikima J**, **Jung T-S**, **Kondo H**, **Hirono I**, **Aoki T**. 2012. Multiple Drug-resistant Strains of Aeromonas hydrophila Isolated from Tilapia Farms in Thailan. Fish Pathol **47**:56–63.

34. **Nigro SJ**, **Post V**, **Hall RM**. 2011. Aminoglycoside resistance in multiply antibiotic-resistant Acinetobacter baumannii belonging to global clone 2 from Australian hospitals. J Antimicrob Chemother **66**:1504–1509.

35. **Hedges RW**, **Jacob AE**. 1974. Transposition of ampicillin resistance from RP4 to other replicons. Mol Gen Genet **132**:31–40.

36. **Chiou CS**, **Jones AL**. 1993. Nucleotide sequence analysis of a transposon (Tn5393) carrying streptomycin resistance genes in Erwinia amylovora and other gram-negative bacteria. J Bacteriol **175**:732–740.

37. **Pansegrau W**, **Lanka E**, **Barth PT**, **Figurski DH**, **Guiney DG**, **Haas D**, **Helinski DR**, **Schwab H**, **Stanisich V a**, **Thomas CM**. 1994. Complete Nucleotide Sequence of Birmingham IncPα Plasmids. J Mol Biol **239**:623–663.

38. **Siguier P**. 2006. ISfinder: the reference centre for bacterial insertion sequences. Nucleic Acids Res **34**:D32–D36.

39. **Cain AK**, **Hall RM**. 2011. Transposon Tn 5393 e Carrying the aphA1 -Containing Transposon

Tn 6023 Upstream of strAB Does Not Confer Resistance to Streptomycin. Microb Drug Resist **17**:389–394.

40. **Iida S**, **Mollet B**, **Meyer J**, **Arber W**. 1984. Functional characterization of the prokaryotic mobile genetic element IS26. Mol Gen Genet **198**:84–89.

41. **He S**, **Hickman B**, **Varani AM**, **Siguier P**, **Chandler M**, **Dekker JP**. 2015. Insertion Sequence IS 26 Reorganizes Plasmids in Clinically Isolated Multidrug-Resistant Bacteria by Replicative Transposition. MBio **6**:15.

42. **Harmer CJ**, **Hall RM**. 2015. IS26-Mediated precise excision of the IS26-apha1a translocatable unit. MBio **6**:1–9.

43. **Meinersmann RJ**, **Lindsey RL**, **Bono JL**, **Smith TP**, **Oakley BB**. 2013. Proposed Model for the High Rate of Rearrangement and Rapid Migration Observed in Some IncA/C Plasmid Lineages. Appl Environ Microbiol **79**:4806–4814.

44. **Doublet B**, **Boyd D**, **Douard G**, **Praud K**, **Cloeckaert A**, **Mulvey MR**. 2012. Complete nucleotide sequence of the multidrug resistance IncA/C plasmid pR55 from Klebsiella pneumoniae isolated in 1969. J Antimicrob Chemother **67**:2354–2360.

Legends to Figures

Figure 1 (A) Comparison of R16a and IP40a to reference IncA/C_2 Type 1 plasmid R148. Major backbone genes and the site of ARIs are indicated. The intact transposons and IS elements located outside of ARI_{IP40a} are shown below the graph (B) Detailed structure of ARIs in R16a and IP40a. The horizontal lines with the major backbone genes represent the whole plasmid sequences. Coordinates below the maps show the ends of ARIs and the standard position of ARI-A. Orfs of ARIs encoding >50 amino acids are shown by arrows. Transposons and IS elements are represented by color-coded rectangles with white (IRL) and black (IRR) arrowheads indicating the inverted repeats of the element. Direct repeats bracketing the mobile elements are shown in capitals. IS*26* variants are marked with '26'. Antibiotic resistance genes are marked as red, the IS*CR2* element is light brown, metal-resistance operons, compound transposons and prophages are indicated. Insertion site of *E. coli*-related IS elements and Tn*1000* are indicated below the graph of IP40a. Figures are drawn not to scale. (C) Schematic representation of Tn*6333*. Orfs are color coded, depending on functional annotations: green, transposition/recombination; orange, mer resistance, grey, partitioning; white, unknown function. EAL (glutamate-alanine-leucine) domain: gene encoding a putative signaling protein with the conserved diguanilate phosphodiesterase EAL domain. Linear maps are drown to scale except the enlarged regions showing ARIs in detail.

Figure 2 Deletion derivatives of three IncA/C plasmids. The name and endpoints of deletions are shown below the plasmid maps. Plasmid backbones are represented by horizontal lines. ARIs involved in rearrangements are shown in more details, otherwise only mobile elements (color-coded) and antibiotic resistance genes (red arrows) are indicated. Coordinates of GIs, ARIs and some transposable elements involved in the formation of deletions are indicated below the graphs. Plasmid map of R55 was created according to the published R55 sequence JQ010984.1.(44) Figures are drawn not to scale.

456 **Table 1** MinION reads and coverage data

	R16a [a]	IP40a [a]	IP40a_d5	IP40a_d8	R55_d20
Library preparation protocol and kit	SQK-MAP006		SQK-MAP005		
Total reads	-	-	52308	4065	5126
Reads mapped to ref.	-	-	4785	774	716
Mean length of mapped reads	-	-	2158	3263	2330
Median length of mapped reads	-	-	1194	1999	1230
Coverage of mapped reads	-	-	73	29	10
Mean Sequence Quality (Phred Score)	-	-	5	5	7
2D reads	19966	12702	10554	344	804
2D reads mapped to ref.	13685	3119	1304	75	158
Mean length of mapped 2D reads	6383	4340	1557	3985	2113
Median length of mapped 2D reads	6349	3067	810	2914	996
Coverage of mapped 2D reads	305	55	19	4	3
Mean Sequence Quality (Phred Score)	10	10	8	9	10

457 [a] 2D basecalling work-flow v.1.16 provided only 2D reads for R16a and IP40a.

458

459 **Table 2** Major characteristics of R16a and IP40a compared to reference plasmids

	R16a	IP40a	R148 (Type 1)	R55 (Type 2)
i1	-	-	-	+
i2	-	-	-	+
R1	*orf1832*	*orf1832*	*orf1832*	*orf1847*
R2	*rhs1*	*rhs1*	*rhs1*	*rhs2*
ARI-A	-	-	+	-
ARI-B	GIsul2	GIsul2	-	+
other ARI	+ (unique position)	+ (unique position in *kfrA*)	-	+ (unique position in *kfrA*)
IS*Ecp1*	-	-	-	-

460