



OPEN

## Genomic analysis and phylogenetic position of the complex IncC plasmid found in the Spanish monophasic clone of *Salmonella enterica* serovar Typhimurium

Xenia Vázquez<sup>1,2,11</sup>, Patricia García<sup>1,9,11</sup>, Vanesa García<sup>1,10</sup>, María de Toro<sup>3</sup>, Víctor Ladero<sup>4,5</sup>, Jürgen J. Heinisch<sup>6</sup>, Javier Fernández<sup>2,7</sup>, Rosaura Rodicio<sup>2,8</sup> & M. Rosario Rodicio<sup>1,2</sup>✉

pUO-STmRV1 is an IncC plasmid discovered in the Spanish clone of the emergent monophasic variant of *Salmonella enterica* serovar Typhimurium, which has probably contributed to its epidemiological success. The sequence of the entire plasmid determined herein revealed a largely degenerated backbone with accessory DNA incorporated at four different locations. The acquired DNA constitutes more than two-thirds of the pUO-STmRV1 genome and originates from plasmids of different incompatibility groups, including IncF (such as R100 and pSLT, the virulence plasmid specific of *S. Typhimurium*), IncN and IncI, from the integrative element *Glsu2*, or from yet unknown sources. In addition to pSLT virulence genes, the plasmid carries genes conferring resistance to widely-used antibiotics and heavy metals, together with a wealth of genetic elements involved in DNA mobility. The latter comprise class 1 integrons, transposons, pseudo-transposons, and insertion sequences, strikingly with 14 copies of IS26, which could have played a crucial role in the assembly of the complex plasmid. Typing of pUO-STmRV1 revealed backbone features characteristically associated with type 1 and type 2 IncC plasmids and could therefore be regarded as a hybrid plasmid. However, a rooted phylogenetic tree based on core genes indicates that it rather belongs to an ancient lineage which diverged at an early stage from the branch leading to most extant IncC plasmids detected so far. pUO-STmRV1 may have evolved at a time when uncontrolled use of antibiotics and biocides favored the accumulation of multiple resistance genes within an IncC backbone. The resulting plasmid thus allowed the Spanish clone to withstand a wide variety of adverse conditions, while simultaneously promoting its own propagation through vertical transmission.

*Salmonella enterica* serovar 4,[5],12:i:- is a monophasic variant of *S. Typhimurium*, which emerged as a main food-borne pathogen causing outbreaks and sporadic cases of salmonellosis worldwide<sup>1–3</sup>. Several multidrug resistant (MDR) clones of this variant have been described, with the resistance genes being located either on the

<sup>1</sup>Departamento de Biología Funcional, Área de Microbiología, Universidad de Oviedo, 33006 Oviedo, Spain. <sup>2</sup>Grupo de Microbiología Traslacional, Instituto de Investigación Sanitaria del Principado de Asturias (ISPA), 33011 Oviedo, Spain. <sup>3</sup>Plataforma de Genómica y Bioinformática, Centro de Investigación Biomédica de La Rioja (CIBIR), 26006 Logroño, Spain. <sup>4</sup>Instituto de Productos Lácteos de Asturias, Consejo Superior de Investigaciones Científicas (IPLA-CSIC), 33300 Villaviciosa, Spain. <sup>5</sup>Grupo de Microbiología Molecular, Instituto de Investigación Sanitaria del Principado de Asturias (ISPA), 33011 Oviedo, Spain. <sup>6</sup>Department of Genetics, Faculty of Biology and Chemistry, University of Osnabrück, Barbarastrasse 11, 49076 Osnabrück, Germany. <sup>7</sup>Servicio de Microbiología, Hospital Universitario Central de Asturias, 33011 Oviedo, Spain. <sup>8</sup>Departamento de Bioquímica y Biología Molecular, Universidad de Oviedo, 33006 Oviedo, Spain. <sup>9</sup>Present address: Department of Microbiology, University Hospital A Coruña (CHUAC)-Biomedical Research Institute A Coruña (INIBIC), 15006 A Coruña, Spain. <sup>10</sup>Present address: Laboratorio de Referencia de Escherichia coli (LREC), Departamento de Microbiología e Parasitología, Facultad de Veterinaria; Instituto de Investigación Sanitaria de Santiago de Compostela (IDIS), Universidade de Santiago de Compostela (USC), 27002 Lug, Spain. <sup>11</sup>These authors contributed equally: Xenia Vázquez and Patricia García. ✉email: rrodicio@uniovi.es

bacterial chromosome<sup>4–6</sup> or on plasmids of different incompatibility groups<sup>7,8</sup>. In the so-called “Spanish clone”, the MDR phenotype is conferred by large (ca. 150–200 kb), non-conjugative IncA/C plasmids<sup>7</sup>, here renamed as IncC. This nomenclature follows a proposal based on the compatibility between IncA (formerly IncA/C<sub>1</sub>, represented by plasmid RA1<sup>9</sup>) and IncC (formerly IncA/C<sub>2</sub>) plasmids<sup>10,11</sup>.

Members of the IncC group are large, low-copy number plasmids with broad-host range, and variable capacity for conjugation. The majority of the genes required for conjugation are arranged in two separate regions, namely transfer region 1 and transfer region 2<sup>10,12</sup>. Based on single nucleotide polymorphisms (SNP) in the backbone, IncC plasmids were separated into two major groups, called type 1 and type 2. These two types also differ in Regions R1 and R2, containing alternative versions of an open reading frame located within transfer region 1 (*orf1832* in type 1 and *orf1847* in type 2), and in the large *rhs* (recombinant hot spot) gene (*rhs1* in type 1 and *rhs2* in type 2). Additionally, two small insertions, i1 (428 bp) and i2 (462 bp), are present in type 2 but not in type 1 plasmids<sup>10,12</sup>. Type 1 plasmids were further separated into type 1a and type 1b, based on SNP accumulations in a 14.5 kb region (Type 1a patch), which starts at the 3'-end of the replication initiation gene *repA* and extends to the region downstream of the *ant-tox* antitoxin-toxin genes<sup>13</sup>. Despite incompatibility and entry exclusion, hybrid plasmids with backbone features from both type 1 and type 2, and also with some additional features (like *orf1854* and *rhs3*), have occasionally been found<sup>10</sup>.

IncC plasmids occur in a number of pathogenic bacteria from diverse environments, where they serve as vehicles for classical and emergent resistance genes. Apparently, these genes were acquired multiple times during evolution, usually as part of clearly recognizable mobile genetic elements inserted at particular locations within the IncC backbone<sup>10</sup>. IncC plasmids of the monophasic Spanish clone were previously shown to contain the *bla*<sub>TEM-1</sub>, *cmlA1*, *aac(3)-IV*, [*aadA1*, *aadA2*], [*sul1*, *sul2*, *sul3*], *tet(A)* and *dfrA12* genes which confer resistance to ampicillin, chloramphenicol, gentamicin, streptomycin-spectinomycin, sulfonamides, tetracycline and trimethoprim, respectively<sup>7,14</sup>. Genes potentially encoding resistance to biocides; multiple insertion sequences, transposons and integrons; or derived from the serovar-specific virulence plasmid of *S. Typhimurium*, pSLT, were also detected in the IncC plasmids of the Spanish clone<sup>7,14</sup>.

Given the interest in IncC plasmids as a model to understand the evolutionary mechanisms leading to the acquisition and dissemination of accessory DNA, including not only resistance but also virulence genes, we here decided to determine the complete nucleotide sequence of a plasmid with the IncC backbone (pUO-STmRV1), which was obtained from a clinical isolate selected as a representative of the monophasic Spanish clone of *S. Typhimurium*. Sequence analysis confirmed experimentally obtained results, detected additional resistance genes, revealed novel features related to the structure of the plasmid, and disclosed its phylogenetic position within the IncC group.

## Results and discussion

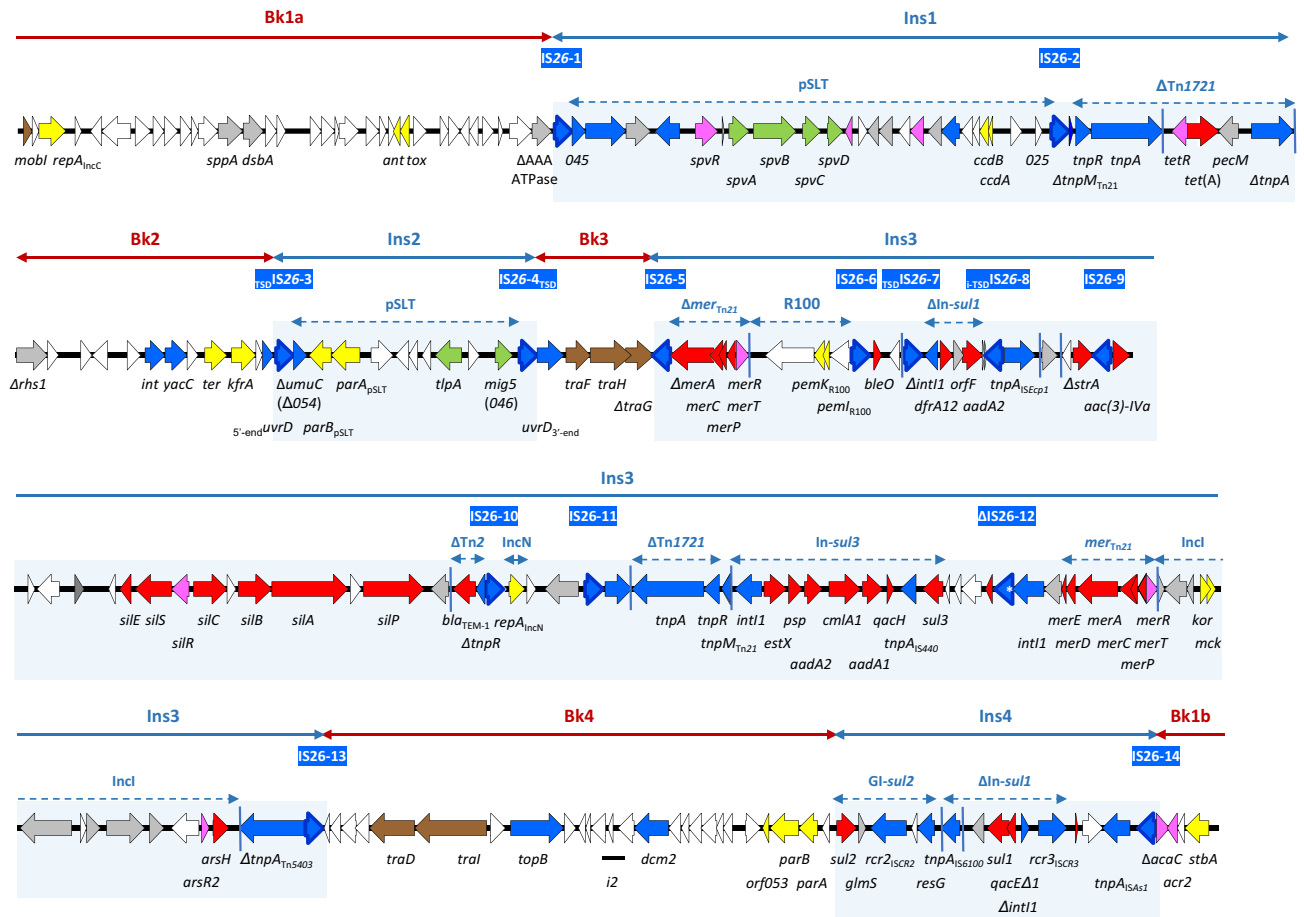
### pUO-STmRV1 is a complex plasmid shaped by a wealth of genetic elements involved in DNA mobility.

Plasmid pUO-STmRV1 (Fig. 1) consists of 197,365 bp with a mean GC content of 52.3%. Approximately 170 open reading frames (*orf*) could be identified, comprising all *orfs* of more than 100 nt and some selected smaller *orfs*. Those with predicted functions (62%) were mainly associated with plasmid propagation (replication, maintenance and conjugative transfer), with virulence and resistance, and with multiple genetic elements, including intact or defective insertion sequences (IS26, *ISEcp1*, IS440, ISCR2, IS6100, ISCR3, and ISAs1), transposon remnants (Tn1721, Tn2, Tn21 and Tn5403), and class 1 integrons of the *sul1* and *sul3* types (Fig. 1). Notably, pUO-STmRV1 harbors 14 copies of IS26 (IS26-1 to IS26-14; all intact except IS26-12 that has a frameshift mutation in the *tnpA* gene). This insertion sequence has probably played a major role in the reductive evolution observed in the pUO-STmRV1 backbone, as well as in the acquisition of the accessory DNA. IS26 can provoke these events by (1) generating pseudo-compound transposons (consisting of a central segment flanked by two copies of IS26 in direct orientation) and translocatable units (formed by one copy of IS26 and the adjacent DNA), and by (2) using two different mechanisms of movement: (a) the copy-in mechanism, which is replicative since both the IS as well as 8 bp originally present at a randomly selected target site are duplicated, and (b) the targeted conservative mechanism involving two copies of IS26, in which IS26 is neither replicated, nor does the target site duplication (TSD) occur<sup>15–17</sup>. Together with IS26, the many other genetic elements found in pUO-STmRV1, and homologous recombination between repeated DNA (like parts of class 1 integrons, of the *mer* operon of Tn21, and of Tn1721, in addition to IS26), could have contributed to the complexity, and also to the high plasticity and intrinsic instability of the IncC plasmid of the monophasic Spanish clone, indicated by the large number of variants discovered so far<sup>7,14</sup>.

### Extensive deletions and loss of synteny altered the IncC backbone in pUO-STmRV1.

Three extensive deletions were detected in the pUO-STmRV1 backbone (Bk), which is considerably smaller than the typical IncC backbone (60,875 bp compared to 127.8 kb and 129.2 kb reported for type 1 and type 2 plasmids, respectively<sup>10</sup>). The remaining backbone is scattered into four segments, here designated as Bk1 to Bk4 (Figs. 1 and 2). The expected order and orientation is conserved for all these segments except Bk4, which corresponds to the *parAB* region<sup>10</sup> (see below). The 26,987 bp sequence of Bk1 (designated as Bk1a and Bk1b in the linear maps shown in Figs. 1 and 2, but contiguous in the circular plasmid), is flanked by two oppositely oriented copies of IS26 (IS26-14 and IS26-1). In addition to *repA*, Bk1a contains the *ant* and *tox* genes (also known as *ata* and *tad*) that encode the antitoxin and toxin of a plasmid addition system, respectively<sup>12,18</sup>. An extensive deletion in Bk1a includes the location of the 428 bp i1 insertion found in type 2 but not in type 1 IncC plasmids. Upstream of *repA*, in Bk1b, the plasmid also lacks sequences of the master regulator required for transcriptional activation of transfer genes<sup>19</sup>, retaining only *acr2* and  $\Delta$ *acaC*.

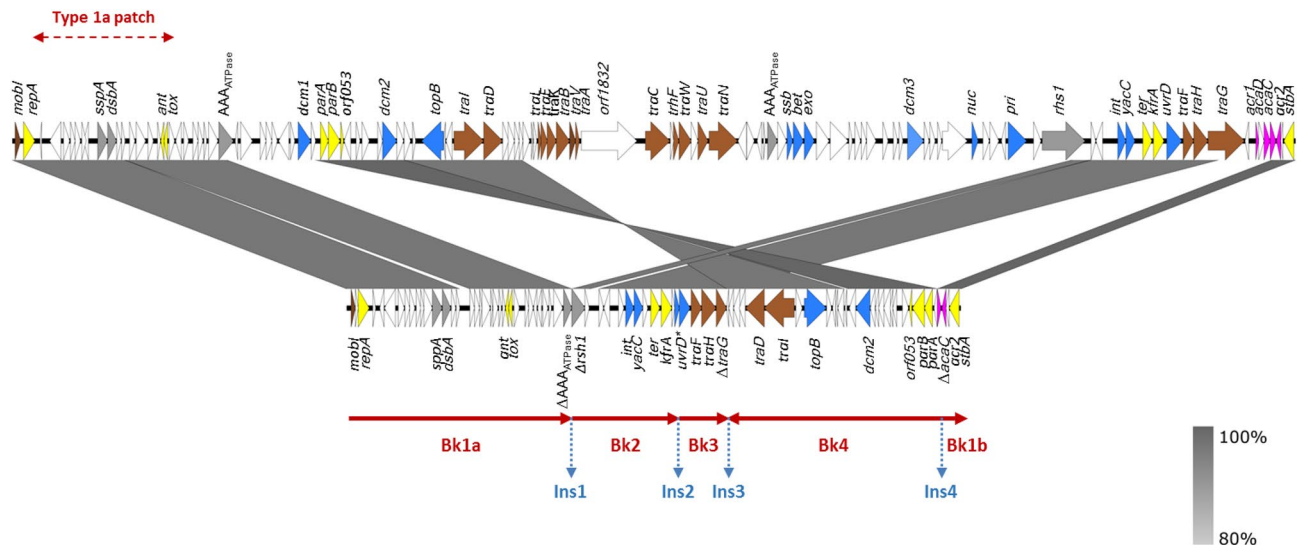
With regard to the 10,671 bp Bk2 (Figs. 1 and 2), it is of note that only 1,317 bp of the 3'-end of the *rhs* gene are retained in pUO-STmRV1, and that they share 98% identity to the corresponding end of the *rhs1* gene of the



**Figure 1.** Gene organization of pUO-STmRV1 (drawn to scale from accession number CP018220). Open reading frames are represented by arrows indicating the direction of transcription with different colors based on the predicted functions: yellow, plasmid replication, maintenance and segregation; brown, conjugative transfer; blue, DNA metabolism; light orange, regulation of gene expression; red, resistance; green, virulence; grey, other roles; white, unknown function. The 14 copies of IS26 detected in pUO-STmRV1 are labelled with white letters on dark blue background. All IS26 are delineated by inverted repeats (IR; not shown), and all except IS26-12 are intact. The frameshift mutation detected in the *tnpA* gene of the latter is indicated with a white asterisk inside the gene. IR in other transposable elements are represented by vertical blue lines. None of the individual IS, including the 14 copies of IS26, or Tn are surrounded by target site duplications. The position of the *i2* insert (between *topB* and *dcm2*) is indicated by a horizontal black line. The DNA segments (Ins1 to Ins4) inserted into the pUO-STmRV1 backbone are indicated and highlighted within blue boxes. The segments in which the plasmid backbone is distributed (Bk1a and Bk1b, which are contiguous in the circular plasmid, to Bk4), are also indicated.

type 1 reference plasmid pR148<sup>10</sup> (26 mismatches in the common sequence of 1,317 bp). In contrast, only the first 725 bp of the *rhs* remnant of pUO-STmRV1 coincided in the *rhs* genes of R55 and pYR1, showing 93.5% and 94.9% identities, respectively (47 and 37 differences). Deletions in *rhs1* have previously been detected, though they mostly affected the 3'-end of the gene. The insertion site of ARI-A, the resistance island located upstream of *rhs1* in most type 1 plasmids, is missing in pUO-STmRV1, since the deletion affecting the 5'-end of *rhs* extends into the adjacent DNA. Downstream of  $\Delta rhts$ , the *int*, *yacC*, *ter* and *kfrA* genes are followed by the *uvrD* gene, interrupted by an insertion of foreign DNA (see below).

The 4,667 bp Bk3 corresponds to part of transfer region 2, carrying *orf*s homologous to *traF*, *traH* and *traG* (required for assembly of the conjugative type IV secretion system), with *traG* truncated at the 3'-end. Finally, Bk4 comprises 20,890 bp that are translocated and inverted as compared to the typical IncC backbone, causing a loss of synteny (Figs. 1 and 2). Bk4 together with Ins4 (ARI-B; see below) could have been released from Bk1 as part of a larger segment flanked by copies of IS26 in direct orientation. This will enable the segment to move as a translocatable unit<sup>15,16</sup>, which would have targeted an IS26 already present at the new location. Intramolecular copy-in transposition by the *trans* pathway<sup>17</sup>, could then have led to the observed inversion, but an additional deletion will be also required to explain the current configuration. Bk4 contains the plasmid partition genes *parA* and *parB*, as well as a conserved *orf* (*orf053*), previously shown to be required for stable plasmid maintenance<sup>18</sup>. This *orf*, together with *repA*, *parA* and *parB*, was selected as part of the pMLST (plasmid multilocus sequence typing) scheme proposed for IncC plasmids<sup>18</sup>. The *traI* and *traD* genes of transfer region 1, which encode a relaxase of the MOB<sub>H</sub> group<sup>20</sup> and a coupling protein, respectively, are also located in Bk4. Yet, other transfer genes



**Figure 2.** Comparison of the pUO-STmRV1 backbone (below) with that of plasmid pR148 used as reference of the IncC group (above). The alignment was created with EasyFig BLASTn, based on accession no. JX141473 and CP018220. Open reading frames are represented by arrows pointing in the direction of transcription and having different colors based on the predicted functions: yellow, plasmid replication, maintenance and segregation; brown, conjugative transfer; blue, DNA metabolism; light orange, regulation of gene expression; grey, other roles; white, unknown function. Gray shading between the backbones connects regions of nucleotide sequence identity ranging from 80 to 100%, according to the scale shown in the lower right part of the figure. The extent of the four segments of the pUO-STmRV1 backbone (Bk1a and Bk1b, contiguous in the circular plasmid, Bk2, Bk3 and Bk4; the later translocated and inverted with respect to the corresponding segment in pR148), and the position of the insertions located between them (Ins1 to Ins4) are indicated. The *uvrD* gene, interrupted by Ins2 and distributed between Bk2 and Bk3, is marked with an asterisk and represented by two arrows, corresponding to the 5'- and 3'-ends. The type 1a patch used to differentiate type 1a and type 2b IncC plasmids is shown above the pR148 backbone.

of this region are absent, as well as any of the two large *orf*s placed downstream of *traA* in type 1 (*orf1832*) and type 2 (*orf1847*) IncC plasmids<sup>10,12</sup>. However, the 462 bp *i2* insertion characteristically found in type 2 plasmids is present. Consistent with the extensive deletions affecting the two transfer regions, pUO-STmRV1 failed to be conjugated into *E. coli*<sup>7</sup>. It is finally of note that the RepA, ParA, ParB and Orf053 proteins of pUO-STmRV1, known to be essential for replication and maintenance of IncC plasmids, are all more than 99% identical to the equivalent proteins of pR148, pSN254, R55 and pYR1.

**Nearly two-thirds of pUO-STmRV1 comprises accessory DNA providing virulence genes and genes conferring resistance to antimicrobial agents and heavy metals.** Apart from a reduced IncC backbone, pUO-STmRV1 contains four exogenous regions, termed Ins1 to Ins4, which account for 136,525 bp out of the 197,365 bp (69.17%) of the plasmid genome (Figs. 1 and 2). They are located upstream of the truncated *rhs* gene (Ins1), within the *uvrD* gene (Ins2), downstream of  $\Delta traG$  (Ins3), and downstream of *parA* (Ins4). All four are either flanked by or adjacent to IS26, with additional copies of the latter appearing interspersed within Ins1 and Ins3. Comparative analysis reveals that large portions of the acquired DNA originated from plasmids of different incompatibility groups, including IncF (such as R100 and pSLT), IncN and IncI, and from the integrative element *GI**sul2*<sup>21</sup>. Other sequences are of unknown origin.

Ins1 (30,516 bp) consists of a segment of pSLT-DNA flanked by two copies of IS26 (IS26-1 and IS26-2), which is followed by a  $\Delta tnpM$  gene of Tn21 and a truncated Tn1721 that supplies the *tetR* and *tet(A)* for tetracycline resistance (Fig. 1). The pSLT DNA (from pSLT045 to pSLT025) comprises the *spv* region, which encodes the main virulence factors associated with serovar-specific virulence plasmids in *S. enterica*<sup>22</sup>, as well as the toxin-antitoxin *ccdAB* genes, which could further enhance the stability of pUO-STmRV1.

Ins2 (10,908 kb), flanked by IS26-3 and IS26-4, consists of a second segment of pSLT-DNA (from pSLT $\Delta$ 054 to pSLT046), which comprises the *parAB*<sub>pSLT</sub> partition genes and the macrophage-induced virulence gene *mig5*, coding for a putative carbonic anhydrase that gets induced inside macrophages. Interestingly, none of the individual copies of IS26 found in pUO-STmRV1 displays the 8 bp TSD. However, such 8 bp TSD (GTCGAAGG, which belongs to the targeted *uvrD* gene of the IncC backbone) are located at the 5'-end of IS26-3 and the 3'-end of IS26-4. A two-step mechanism would explain the observed configuration: (1) copy-in transposition of IS26 into *uvrD* generating the TSD, followed by (2) conservative targeting of this IS element by a translocatable unit<sup>15,16</sup> consisting of a single copy of IS26 and the pSLT segment spanning from  $\Delta umuC$  (pSLT $\Delta$ 054) to *mig5* (pSLT046). This would have occurred after the initial acquisition of a contiguous pSLT segment (from pSLT $\Delta$ 054 to pSLT025; already carrying an internal copy of IS26), which is now separated into Ins1 and Ins2 (Fig. 1).

The large (81,581 bp) and highly complex Ins3 region is also surrounded by two copies of IS26 (IS26-5 and IS26-13) and carries seven internal copies of this element (IS26-6 to IS26-12). Eight bp TSD (TATCTTTA and TAAAGATA; Fig. 1) are only found in the segment flanked by IS26-7 and IS26-8, although the position and orientation of one of the repeats has been altered. This could have originated from an intramolecular copy-in transposition event by the *trans* pathway<sup>17</sup>, accompanied by inversion of the segment carrying a defective class 1 integron of the *sul1* type. However, inversion of the segment and the TSD through homologous recombination between the oppositely oriented copies of IS26, cannot be ruled out. Many other genetic elements and resistance genes are also located within Ins3. They comprise genes involved in resistance to heavy metals (see below), such as arsenic (*arsR2* and *arsH* genes), mercury (with intact: *merRTPCADE* and truncated: *merRTPC-ΔmerA* copies of the *mer* region of Tn21), and silver (*silESRCBAP* genes), as well as additional resistance genes, most of them carried by conventional (*sul1*) and atypical (*sul3*) class 1 integrons, or transposon remnants [ $\Delta$ Tn2 and  $\Delta$ Tn1721]. The *pemI* and *pemK* genes responsible for stable maintenance of plasmid R100, the *repA* gene of the IncN incompatibility group (supplied by a 965 nt segment that is 99.17% identical to the corresponding DNA of R46), and a segment homologous to sequences of IncI plasmids (showing 99.1% identity with pR64 spanning 11.2 kb), including the *kor* and *mck* genes for plasmid stability, as well as the *arsR2* and *arsH* genes, are also located within Ins3. Altogether, four plasmid addition systems are carried by pUO-STmRV1 (*ant-tox*, *ccdA-ccdB*, *pemK-pemI* and *mck-kor*), which are likely to ensure maintenance of the plasmid even in the absence of selective pressure. However, this will not prevent further evolution through loss or acquisition of accessory DNA, as a means of adaptation to changing conditions.

Finally, Ins4 (13,520 bp), located downstream of *parA*, is a *sul2*-containing region of the ARI-B type. ARI-B is a resistant island found in both type 1 and type 2 IncC plasmids, which has most likely originated from the integrative element GI*sul2*<sup>21,23</sup>. It covers the region from the *sul2*-end to *resG*, but lacks the other genes of GI*sul2* including the *arsHCB* operon. In contrast, a second class 1 integron of the *sul1* type, can be detected downstream of *resG* which lacks gene cassettes in the variable region but carries *qacEΔ1* and *sul1* in the 3'-conserved segment (Fig. 1).

**pUO-STmRV1 confers resistance to heavy metals.** Although many genes carried by pUO-STmRV1 were previously associated with the resistance of LSP 389/97 to multiple antibiotics<sup>7,14</sup>, resistance towards heavy metals has not been experimentally confirmed. According to the presence of the *mer* and *sil* clusters in pUO-STmRV1, LSP 389/97 was resistant to HgCl<sub>2</sub> and AgNO<sub>3</sub>, with MIC values of 32 μg/ml and 125 μM, respectively, which are 4- to 8- and 4- fold higher than those of the control strains (S. Typhimurium LT2 and ATCC 14028; Table 1). In the case of arsenic, toxicity is strongly dependent on the chemical form and oxidation state<sup>24</sup>. It has been demonstrated that the *arsH* gene, which is controlled by the product of *arsR2* (both carried by pUO-STmRV1), encodes an organoarsenical oxidase that detoxifies trivalent methylated and aromatic arsenicals by oxidation to the relatively innocuous pentavalent species<sup>24,25</sup>. Therefore, phenylarsine oxide, an aromatic As(III) compound, was tested in the present study, using the organic As(V) roxarsone, and two inorganic compounds: NaAsO<sub>2</sub> [As(III)] and Na<sub>2</sub>AsHO<sub>4</sub> [As(V)], as controls. As expected, the MIC of LSP 389/97 to the inorganic forms (64 μg/ml and 128 μg/ml for NaAsO<sub>2</sub> and Na<sub>2</sub>AsHO<sub>4</sub>, respectively) coincided with those of the control strains. In contrast, the MIC of the phenylarsine oxide (4 μg/ml), was 16 times higher. In the case of roxarsone, a largely harmless pentavalent organic form, the MIC values were extremely high for the three strains tested. This compound is used in the poultry industry, and to a lesser extent also in the pork industry, as a feed additive to promote growth and prevent coccidial infections<sup>26</sup>. Interestingly, the entire integrative GI*sul2* element carried by plasmid pIP40a and reported as the progenitor of ARI-B, includes an apparently intact *arsRHCB* cluster<sup>23</sup>, which is not present in the deleted ARI-B of pUO-STmRV1. The *Pseudomonas aeruginosa* strain carrying pIP40a was susceptible to NaAsO<sub>2</sub> and Na<sub>2</sub>HAsO<sub>4</sub>, but resistance to trivalent organic arsenicals was not tested. The products of the *arsR2* and *arsH* genes of pUO-STmRV1 are closely related to those reported in plasmids of the IncI incompatibility group, like R64<sup>27</sup>, but only distantly related to those carried by GI*sul2*.

**Phylogenetic analysis reveals that pUO-STmRV1 belongs to an ancient lineage which diverged early from the main clade of IncC plasmids.** Sequence analysis combined with in silico PCR typing<sup>28</sup> identified pUO-STmRV1 as a novel hybrid plasmid, sharing features of type 1 IncC plasmids (such as the R2 region containing *rhs1*, although only 1,317 bp of the 3' end of the gene are conserved in pUO-STmRV1; see above), and type 2 IncC plasmids (i2 insertion). Other specific traits used to define IncC types are missing in pUO-STmRV1, due to the extensive deletions affecting the plasmid backbone (Table 2).

To precisely establish the evolutionary position of pUO-STmRV1, a phylogenetic tree was constructed based on SNP detected in a total of 28 core genes conserved in 67 IncC plasmids and RA1 (IncA), separated in the tree as outgroup (Fig. 3; see Tables S1, S2 and S3 for details). pUO-STmRV1 is most closely related to pBML2526, a 204,791 bp plasmid from *Providencia retggeri*. Interestingly, these two plasmids appear to belong to an ancient lineage which diverged at the root of the IncC tree from pYR1 and the main clade including all other IncC plasmids used in the tree. The latter are in turn distributed into two sub-clades. One of them comprises type 1a plasmids and the distantly related pCFSAN001921, previously proposed as a potential new subtype<sup>10</sup>, while the second contains type 1b and type 2 plasmids. The relationship between the latter two groups is not only supported by the numbers of SNP (Table S2), but also by comparisons of the type 1a patch (see Fig. 2) used to discriminate type 1a and type 1b plasmids<sup>10,12</sup>, which showed 99.96% identity between pSN254 (type 1b) and R55 (type 2), but only 96.38% identity between pR148 (type 1a) and pSN254. The pUO-STmRV1 region corresponding to the type 1a patch presented 90.16%, 92.94%, 92.90% and 96.91% identity with the equivalent regions in pR148, pSN254, R55 and pYR1, respectively. It should be noted that with the exception of pYR1, all other IncC plasmids previously reported as hybrids and included in the tree (Table S1), grouped together with either type 1b or type 2 plasmids.

Strain	MIC					
	HgCl <sub>2</sub> (µg/ml)	AgNO <sub>3</sub> (µM)	NaAsO <sub>2</sub>	Na <sub>2</sub> HAsO <sub>4</sub> (µg/ml)	Phenylarsine oxide (µg/ml)	Roxarsone (µg/ml)
LSP 389/97	32	125	64	128	4	8,192
LT2	4	31	64	128	0.25	2,048
ATCC 14028	8	31	64	128	0.25	2,048

**Table 1.** Minimum inhibitory concentration (MIC) of heavy metals for *Salmonella enterica* serovar 4,5,12:i:- strain LSP 389/97 carrying plasmid pUO-STmRV1. *S. enterica* serovar Typhimurium strains LT2 and ATCC 14028 were included as negative controls.

Plasmid	IncC type	i1	i2	orf	rhs	ARI-B	Accession number	References
pR148	Type 1a	–	–	1,832	<i>rhs1</i>	–	JX141473	<sup>31</sup>
pSN254	Type 1b	–	–	1,832	<i>rhs1</i>	+	CP000604	<sup>32</sup>
R55	Type 2	+	+	1,847	<i>rhs2</i>	+	JQ010984	<sup>33</sup>
pYR1	Hybrid	+	+	1,847	<i>rhs3</i>	+	CP000602	<sup>32</sup>
pUO-STmRV1	Hybrid	– (NP)	+	– (NP)	$\Delta$ <i>rhs1</i>	+	CP018220	This study

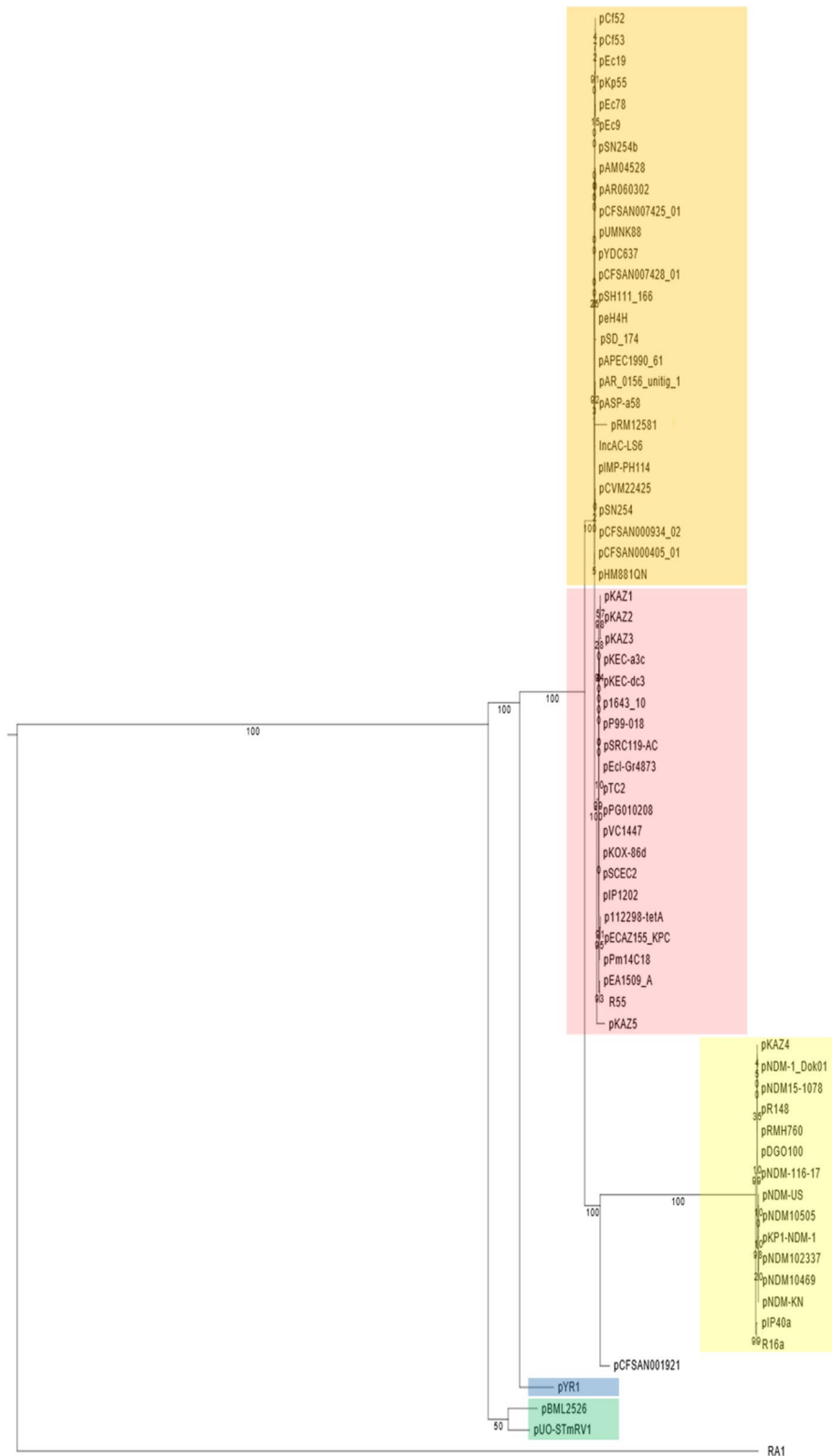
**Table 2.** In silico PCR-typing of pUO-STmRV1. NP, regions surrounding the i1 and orf locations are not present;  $\Delta$ *rhs1*, only 1,317 bp of the 3'-end of *rhs1* (4,254 bp) are present. Based on<sup>28</sup>.

The overall topology of the tree shown in Fig. 3 coincides with that reported by Hancock et al.<sup>18</sup>, built in a similar fashion from SNP in core genes using plasmid RA1 as outgroup. Both trees failed to separate type 1 and type 2 plasmids, in contrast to the unrooted tree generated from plasmid backbones, and published by Lei et al.<sup>29</sup>. Plasmid backbones are particularly well suited to disclose diversity, since intergenic regions in addition to core genes are used in the alignments, while rooted trees based on SNP differences in core genes are better suited to disclose evolutionary distances.

In conclusion, typing of pUO-STmRV1 reveals backbone features characteristically associated with type 1 and type 2 IncC plasmids, and could thus be regarded as a new hybrid plasmid. Such plasmids are traditionally assumed to originate from homologous recombination between type 1a and type 1b plasmids, despite their incompatibility and entry exclusion mechanisms<sup>11</sup>. However, a phylogenetic analysis based on SNP in core genes suggests that pUO-STmRV1, and also pYR1, belong to ancient lineages which have separated at an early stage from the main branch leading to most extant IncC plasmids detected thus far. pUO-STmRV1 appears to have evolved at a time when uncontrolled use of antibiotics and biocides propitiated the accumulation of multiple resistance genes into an IncC platform. This was facilitated and mediated by a wealth of potentially mobile genetic elements that also provided virulence genes, at the expense of significantly shortening the plasmid backbone. The resulting pUO-STmRV1 allowed the Spanish monophasic clone of *S. Typhimurium* to withstand a variety of adverse conditions, while simultaneously promoting its own maintenance and propagation through accumulation of plasmid addiction systems and by vertical transmission within the clone.

## Methods

**Bacterial isolate, genomic DNA extraction and whole genome sequencing.** LSP 389/97 4,5,12:i:- (phage type U302) was the monophasic isolate used in the present study. It was recovered from feces of a patient with gastroenteritis in 1997, and assigned to the Spanish clone<sup>7,14,30</sup>. Genomic DNA of the isolate was extracted with the “GenElute Bacterial Genomic DNA Kit” (Sigma Aldrich) following the manufacturer’s instructions. WGS was performed in parallel with Illumina (short-reads) and PacBio (long-reads) technologies. Illumina sequencing was carried out at Era7 Bioinformatics (Madrid, Spain), using paired-end reads of 90 nt from a fragment library of 500 bp. The reads were assembled with the VelvetOptimizer.pl script implemented in the “on line” version of PLACNETw (<https://castillo.dicom.unican.es/upload/>). PacBio sequencing was performed at Expression Analysis Inc. (Durham, NC, USA), using the Pacific Biosciences RS II platform, from a library of 6.5 kb DNA fragments on three single-molecule real-time (SMRT) cells (Pacific Biosciences, Menlo Park, CA, USA). The reads were assembled with the Hierarchical Genome Assembly Process (HGAP 3) version 3.0. The assembly comprised two contigs, one corresponding to the entire *Salmonella* chromosome and the other to pUO-STmRV1, which was manually circularized after removing the terminal repeats in the assembled sequence. Errors in the PacBio sequence of pUO-STmRV1 were manually corrected by comparison with Illumina contigs of plasmid origin that were identified with PLACNETw in conjunction with BLASTn comparisons (<http://blast.ncbi.nlm.nih.gov/>). Moreover, before approaching the sequence of the entire genome of LSP 389/97, plasmid fragments were cloned and subjected to Sanger sequencing (data not shown). In this way, more than 30% of the plasmid sequence (about 60 kb) was already generated, and it was also used to correct errors. The corrected sequence of pUO-STmRV1 is available in GenBank under accession no. CP018220.



**Figure 3.** Phylogenetic relationships of pUO-STmRV1. The nucleotide sequences of 67 IncC plasmids and of RA1 (IncA) were extracted from GenBank-NCBI (Table S1). The tree was based on a  $17,713 \pm 37$  bp core genome including 28 orthologous common genes (Table S2) with at least 80% identity and 80% coverage, using 100 bootstrapping replicates. IncC types and subtypes are highlighted in yellow, type 1a; orange, type 1b; pink, type 2; blue, pYR1; green, pUO-STmRV1 and pBML2526. As shown in the pairwise SNP distance matrix used to build the tree (Table S3), the number of SNP separating pUO-STmRV1 from pBML2526, pYR1, pR148 (type 1a), pSN254 (type 1b) and R55 (type 2), which are all IncC, and from the IncA plasmid RA1 are 78, 143, 410, 212, 215 and 1,336, respectively.

**Plasmid annotation, typing and phylogenetic analysis.** Annotation of pUO-STmRV1 was automatically accomplished by the NCBI Prokaryotic Genome Annotation Pipeline ([http://www.ncbi.nlm.nih.gov/genome/annotation\\_prok/](http://www.ncbi.nlm.nih.gov/genome/annotation_prok/)), and manually refined using sequence comparisons with BLASTn and BLASTp. Typing of the plasmid was performed in silico following the PCR scheme of Harmer and Hall<sup>28</sup>. Plasmids pR148<sup>31</sup> (accession no. JX141473), pSN254<sup>32</sup> (CP000604), R55<sup>33</sup> (JQ010984) and pYR1<sup>32</sup> (CP000602) were included as representatives of type 1a, type 1b, type 2 and hybrid IncC plasmids, respectively. A phylogenetic tree was constructed based on variable positions (i.e. SNP) in genes encoding core proteins of pUO-STmRV1 and another 67 fully sequenced IncC plasmids, in addition to the IncA plasmid RA1. Plasmid sequences were retrieved from GenBank-NCBI (<ftp://ftp.ncbi.nlm.nih.gov/genomes/genbank/bacteria/>) and their accession numbers are shown in Table S1. The core genome was defined by the collection of 28 orthologous genes common to all plasmids, which shared at least 80% identity and 80% coverage<sup>34</sup>, and accounted for a core length of 17,713 ± 37 bp (Table S2). The tree was built with RAXML (Randomized Axelerated Maximum Likelihood)<sup>35</sup>, with bootstrap support based on 100 replicates<sup>36</sup>. The pairwise SNP distance matrix used to build the tree is shown in Table S3.

**MICs of heavy metals.** To gather additional information on the resistance properties of the monophasic Spanish clone, the minimum inhibitory concentration (MIC) of several heavy metals were determined by microdilution tests using the following compounds and concentrations: HgCl<sub>2</sub> (0–256 µg/ml); AgNO<sub>3</sub> (0–125 µM); NaAsO<sub>2</sub> (0–256 µg/ml); Na<sub>2</sub>HAsO<sub>4</sub> × 7H<sub>2</sub>O (0–256 µg/ml); phenylarsine oxide [0–16 µg/ml], and roxarsone (4-hydroxy-3-nitrobenzenearsonic acid; 0–8192 µg/ml). All these compounds were purchased from Sigma. *S. Typhimurium* strains LT2 and ATCC 14028 were included as negative controls.

Received: 10 July 2020; Accepted: 7 May 2021

Published online: 01 June 2021

## References

- EFSA. EFSA Panel on Biological Hazards (BIOHAZ); Scientific Opinion on monitoring and assessment of the public health risk of “*Salmonella* Typhimurium-like” strains. *EFSA J.* **8**, 1826. <https://doi.org/10.2903/j.efsa.2010.1826> (2010).
- EFSA. EFSA and ECDC (European Food Safety Authority and European Centre for Disease Prevention and Control). The European Union One Health 2018 Zoonoses Report. *EFSA J.* **17**(5926), 2018. <https://doi.org/10.2903/j.efsa.2019.5926> (2019).
- Switt, A. I., Soyer, Y., Warnick, L. D. & Wiedmann, M. Emergence, distribution, and molecular and phenotypic characteristics of *Salmonella enterica* serotype 4,5,12:i. *Foodborne Pathog. Dis.* **6**, 407–415. <https://doi.org/10.1089/fpd.2008.0213> (2009).
- Boland, C. *et al.* Extensive genetic variability linked to IS26 insertions in the *fliB* promoter region of atypical monophasic variants of *Salmonella enterica* serovar Typhimurium. *Appl. Environ. Microbiol.* **81**, 3169–3175. <https://doi.org/10.1128/AEM.00270-15> (2015).
- Garcia, P. *et al.* Horizontal acquisition of a multidrug-resistance module (R-type ASSuT) is responsible for the monophasic phenotype in a widespread clone of *Salmonella* Serovar 4,[5],12:i. *Front. Microbiol.* **7**, 680. <https://doi.org/10.3389/fmicb.2016.00680> (2016).
- Lucarelli, C. *et al.* Nucleotide sequence of the chromosomal region conferring multidrug resistance (R-type ASSuT) in *Salmonella* Typhimurium and monophasic *Salmonella* Typhimurium strains. *J. Antimicrob. Chemother.* **67**, 111–114. <https://doi.org/10.1093/jac/dkr391> (2012).
- Garcia, P., Guerra, B., Bances, M., Mendoza, M. C. & Rodicio, M. R. IncA/C plasmids mediate antimicrobial resistance linked to virulence genes in the Spanish clone of the emerging *Salmonella enterica* serotype 4,[5],12:i. *J. Antimicrob. Chemother.* **66**, 543–549. <https://doi.org/10.1093/jac/dkq481> (2011).
- Garcia, P. *et al.* Diversity of plasmids encoding virulence and resistance functions in *Salmonella enterica* subsp. *enterica* serovar Typhimurium monophasic variant 4,[5],12:i:- strains circulating in Europe. *PLoS One* **9**, e89635. <https://doi.org/10.1371/journal.pone.0089635> (2014).
- Fricke, W. F. *et al.* Comparative genomics of the IncA/C multidrug resistance plasmid family. *J. Bacteriol.* **191**, 4750–4757. <https://doi.org/10.1128/JB.00189-09> (2009).
- Ambrose, S. J., Harmer, C. J. & Hall, R. M. Evolution and typing of IncC plasmids contributing to antibiotic resistance in Gram-negative bacteria. *Plasmid* **99**, 40–55. <https://doi.org/10.1016/j.plasmid.2018.08.001> (2018).
- Ambrose, S. J., Harmer, C. J. & Hall, R. M. Compatibility and entry exclusion of IncA and IncC plasmids revisited: IncA and IncC plasmids are compatible. *Plasmid* **96–97**, 7–12. <https://doi.org/10.1016/j.plasmid.2018.02.002> (2018).
- Harmer, C. J. & Hall, R. M. The A to Z of A/C plasmids. *Plasmid* **80**, 63–82. <https://doi.org/10.1016/j.plasmid.2015.04.003> (2015).
- Harmer, C. J. & Hall, R. M. Evolution in situ of ARI-A in p B2–1, a type 1 IncC plasmid recovered from *Klebsiella pneumoniae*, and stability of Tn4352B. *Plasmid* **94**, 7–14. <https://doi.org/10.1016/j.plasmid.2017.10.001> (2017).
- Guerra, B., Soto, S. M., Arguelles, J. M. & Mendoza, M. C. Multidrug resistance is mediated by large plasmids carrying a class 1 integron in the emergent *Salmonella enterica* serotype [4,5,12:i:-]. *Antimicrob. Agents Chemother.* **45**, 1305–1308. <https://doi.org/10.1128/AAC.45.4.1305-1308.2001> (2001).
- Harmer, C. J., Moran, R. A. & Hall, R. M. Movement of IS26-associated antibiotic resistance genes occurs via a translocatable unit that includes a single IS26 and preferentially inserts adjacent to another IS26. *mBio* **5**, e01801–01814. <https://doi.org/10.1128/mBio.01801-14> (2014).
- Harmer, C. J., Pong, C. H. & Hall, R. M. Structures bounded by directly-oriented members of the IS26 family are pseudo-compound transposons. *Plasmid* **111**, 102530. <https://doi.org/10.1016/j.plasmid.2020.102530> (2020).
- He, S. *et al.* Insertion sequence IS26 reorganizes plasmids in clinically isolated multidrug-resistant bacteria by replicative transposition. *mBio* **6**, e00762. <https://doi.org/10.1128/mBio.00762-15> (2015).
- Hancock, S. J. *et al.* Identification of IncA/C plasmid replication and maintenance genes and development of a plasmid multilocus sequence typing scheme. *Antimicrob. Agents Chemother.* **61**. doi:<https://doi.org/10.1128/AAC.01740-16> (2017).
- Carraro, N., Matteau, D., Luo, P., Rodrigue, S. & Burrus, V. The master activator of IncA/C conjugative plasmids stimulates genomic islands and multidrug resistance dissemination. *PLoS Genet* **10**, e1004714. <https://doi.org/10.1371/journal.pgen.1004714> (2014).
- Garcillan-Barcia, M. P., Francia, M. V. & de la Cruz, F. The diversity of conjugative relaxases and its application in plasmid classification. *FEMS Microbiol. Rev.* **33**, 657–687. <https://doi.org/10.1111/j.1574-6976.2009.00168.x> (2009).
- Nigro, S. J. & Hall, R. M. 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 baumannii* ATCC 17978 from 1951. *J. Antimicrob. Chemother.* **66**, 2175–2176. <https://doi.org/10.1093/jac/dkr230> (2011).
22. Guiney, D. G. & Fierer, J. The role of the *spv* genes in *Salmonella* pathogenesis. *Front. Microbiol.* **2**, 129. <https://doi.org/10.3389/fmicb.2011.00129> (2011).
  23. Harmer, C. J., Hamidian, M. & Hall, R. M. pIP40a, a type 1 IncC plasmid from 1969 carries the integrative element *Glusl2* and a novel class II mercury resistance transposon. *Plasmid* **92**, 17–25. <https://doi.org/10.1016/j.plasmid.2017.05.004> (2017).
  24. Chen, J., Bhattacharjee, H. & Rosen, B. P. ArsH is an organoarsenical oxidase that confers resistance to trivalent forms of the herbicide monosodium methylarsenate and the poultry growth promoter roxarsone. *Mol. Microbiol.* **96**, 1042–1052. <https://doi.org/10.1111/mmi.12988> (2015).
  25. Yang, H. C. & Rosen, B. P. New mechanisms of bacterial arsenic resistance. *Biomed. J.* **39**, 5–13. <https://doi.org/10.1016/j.bj.2015.08.003> (2016).
  26. Konkel, L. Organoarsenic drugs over time: The pharmacokinetics of roxarsone in chicken meat. *Environ. Health Perspect.* **124**, A150. <https://doi.org/10.1289/ehp.124-A150> (2016).
  27. Sampei, G. *et al.* Complete genome sequence of the incompatibility group I1 plasmid R64. *Plasmid* **64**, 92–103. <https://doi.org/10.1016/j.plasmid.2010.05.005> (2010).
  28. Harmer, C. J. & Hall, R. M. PCR-based typing of IncC plasmids. *Plasmid* **87–88**, 37–42. <https://doi.org/10.1016/j.plasmid.2016.08.002> (2016).
  29. Lei, C. W. *et al.* A novel type 1/2 hybrid IncC plasmid carrying fifteen antimicrobial resistance genes recovered from *Proteus mirabilis* in China. *Plasmid* **93**, 1–5. <https://doi.org/10.1016/j.plasmid.2017.07.002> (2017).
  30. Garcia, P., Malorny, B., Hauser, E., Mendoza, M. C. & Rodicio, M. R. Genetic types, gene repertoire, and evolution of isolates of the *Salmonella enterica* serovar 4,5,12:i:- Spanish clone assigned to different phage types. *J. Clin. Microbiol.* **51**, 973–978. <https://doi.org/10.1128/JCM.02777-12> (2013).
  31. Del Castillo, C. S. *et al.* Comparative sequence analysis of a multidrug-resistant plasmid from *Aeromonas hydrophila*. *Antimicrob. Agents Chemother.* **57**, 120–129. <https://doi.org/10.1128/AAC.01239-12> (2013).
  32. Welch, T. J. *et al.* Multiple antimicrobial resistance in plague: An emerging public health risk. *PLoS ONE* **2**, e309. <https://doi.org/10.1371/journal.pone.0000309> (2007).
  33. Doublet, B. *et al.* Complete nucleotide sequence of the multidrug resistance IncA/C plasmid pR55 from *Klebsiella pneumoniae* isolated in 1969. *J. Antimicrob. Chemother.* **67**, 2354–2360. <https://doi.org/10.1093/jac/dks251> (2012).
  34. Lanza, V. F. *et al.* Plasmid flux in *Escherichia coli* ST131 sublineages, analyzed by plasmid constellation network (PLACNET), a new method for plasmid reconstruction from whole genome sequences. *PLoS Genet* **10**, e1004766. <https://doi.org/10.1371/journal.pgen.1004766> (2014).
  35. Stamatakis, A. RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22**, 2688–2690. <https://doi.org/10.1093/bioinformatics/btl446> (2006).
  36. Felsenstein, J. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**, 783–791. <https://doi.org/10.1111/j.1558-5646.1985.tb00420.x> (1985).

## Acknowledgements

This work was supported by the “Fondo de Investigación Sanitaria, Instituto de Salud Carlos III, Ministerio de Economía y Competitividad”, Spain, cofunded by European Regional Development Fund of the European Union: a way to making Europe, project FIS PI17/00474. XV and PG were recipients of grants of the Program “Severo Ochoa” (ref. BP17-008 and BP08-031, respectively) for support of Research and Teaching in the Principality of Asturias (Spain). V.G. acknowledges the “Consellería de Cultura, Educación e Ordenación Universitaria, Xunta de Galicia” for her post-doctoral grant (Grant number ED481B-2018/018).

## Author contributions

M.R.R. conceived and designed the research. X.V., P.G., V.G. and J.F. performed the research. X.V., P.G., V.G., V.L., M.T., R.R. and M.R.R. performed bioinformatic and phylogenetic analyses. X.V., P.G., V.G., V.L., M.T., J.H., J.F., R.R. and M.R.R. prepared figures and tables. X.V., P.G., J.H., J.F., R.R. and M.R.R. drafted the main manuscript text and the supplementary material. All authors contributed to analysis and interpretation of the results, reviewed the manuscript, and gave final approval for publication.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-90299-z>.

**Correspondence** and requests for materials should be addressed to M.R.R.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher’s note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021