ACTA BIOLOGICA SLOVENICA LJUBLJANA 2002 Vol. 45, Št. 1: 9 - 15

tra region of the natural conjugative *Escherichia coli* plasmid pRK100 is F-like

Regija *tra* naravnega konjugativnega plazmida pRK100 bakterije *Escherichia coli* je podobna plazmidu F

Marjanca STARČIČ ERJAVEC^{1,2}, Wim GAASTRA², Darja ŽGUR-BERTOK¹

¹Department of Biology, Biotechnical Faculty, University of Ljubljana, Večna pot 111, 1000 Ljubljana, Slovenia, and

² Department of Infectious Diseases and Immunology, Utrecht University, P.O. Box 80.165, 3508 TD Utrecht, The Netherlands; E-mail: darja.zgur@uni-lj.si

Abstract. The aim of the presented study was to identify the similarity of the pRK100 *tra* region with *tra* regions of other conjugative plasmids of *Enterobacteriaceae*. For this purpose several *tra* genes were amplified with PCR and the nucleotide sequences of the obtained PCR products were determined. The pRK100's nucleotide sequences were compared to the nucleotide sequences deposited in GenBank and the nucleotide divergence between them was calculated. The obtained results clearly demonstrated, that the *tra* region of pRK100 is the most similar to the *tra* region of plasmid F.

Key words: conjugative plasmid, tra region, Enterobacteriaceae.

Izvleček. Cilj raziskave je bil ugotoviti podobnost regije *tra* plazmida pRK100 z regijami *tra* drugih konjugativnih plazmidov enterobakterij. V ta namen smo več genov regije *tra* plazmida pRK100 pomnožili v PCR in določili nukleotidno zaporedje dobljenih produktov PCR. Nukleotidna zaporedja plazmida pRK100 smo primerjali z drugimi, deponiranimi, nukleotidnimi zaporedji in izračunali nukleotidno divergenco. Dobljeni rezultati so jasno pokazali, da je regija *tra* plazmida pRK100 najbolj podobna regiji *tra* plazmida F.

Ključne besede: konjugativen plazmid, regija tra, Enterobacteriaceae.

Introduction

Conjugative plasmids are extrachromosomal elements, that can promote their own DNA transfer, as well as of co-resident plasmids or even chromosomal DNA from a donor to a recipient cell in a process named conjugation (Firth & al. 1996). All information needed for this process is encoded in a large, approximately 30 kb long, plasmid region, often denoted as the *tra* region. Conjugative plasmids facilitate the exchange and spread of resistances to antibiotics, chemicals, virulence factors and metabolic properties.

One of the best studied *tra* regions is that of the *Escherichia* coli plasmid F (Fig. 1) The *tra* region of plasmid F contains approximately 40 genes that are organised into three operons with a complex network of regulation of gene expression (Frost & al. 1994). The genes of the F-plasmid *tra* region can be divided, according to function, into 5 groups: i) regulatory genes (*finP*, *traJ*, *traY*, and *finO*); ii) genes for pilus biogenesis (*traA*, *traL*, *traE*, *traK*, *traB*, *traV*, *traC*, *trbI*, *traW*, *traU*, *trbC*, *traF*, *traQ*,

traG, traH, traX); iii) genes for DNA metabolism (traM, traY, traD, traI); iv) genes for aggregate stabilisation (traN, traG); and v) genes for surface exclusion (traS, traT) (Firth & al. 1996).

Several conjugative plasmids of *Enterobacteriaceae* have been found to harbour F-like *tra* regions: pCoIV-K30, P307, R100, R1, pSLT. Even though these plasmids have many similarities at the level of nucleotide sequences, differences in regulation of conjugation are observed. For example: plasmid R1 has two promoters upstream of the *traJ* gene, while plasmid F has only one promoter (Dempsey 1994). Therefore, characterisation of different conjugative plasmids and their *tra* regions is of interest.



Figure 1: Physical and genetic map of the F-plasmid *tra* region. Genes of the F *tra* region are depicted. To clearly show the region of a gene, some boxes representing genes are offline. Capital letters are shortened abbreviations for the genes named *tra*. The boxes designating the pRK100 genes analysed in this study are filled. The figure is based on Frost & al. 1994.

Slika 1: Razpored genov *tra* plazmida F. Prikazani so geni *tra* plazmida F. Kvadrati, ki označujejo gene, katerih nukleotidna zaporedja smo analizirali, so zapolnjeni. Slika je pripravljena po Frost & al. 1994.

In the presented study the *tra* region of the natural conjugative plasmid pRK100 was characterised with regard to similarity with *tra* regions of related plasmids. pRK100 is a ~145-kb plasmid isolated from a uropathogenic *Escherichia coli* strain and it has been partially characterised (Ambrožič & al. 1998). It is a member of the IncF incompatibility group and encodes two antibiotic resistances, ampicillin and tetracycline, two colicins, CoIV and CoIIa, and the aerobactin iron uptake system (Žgur-Bertok & al. 1990).

For the purpose of defining the similarity of pRK100's *tra* region, PCR products of several *tra* genes were amplified and their nucleotide sequences were determined. The obtained (partial) nucleotide sequences of genes *traM*, *finP*, *traJ*, *traY*, *traD*, *traI*, and *finO* were compared with the nucleotide sequences of the same genes of plasmids F, pColV-K30, P307, R100, R1 and pSLT and the nucleotide divergence between pRK100 and the other plasmids was determined. From the obtained results it can be concluded that the *tra* region of pRK100 is most similar to the *tra* region of plasmid F.

Methods

Bacterial strains, plasmids and growth conditions

Bacterial strains and plasmids used in this study are presented in Tab. 1. Bacteria were grown in Luria-Bertani (LB) medium with aeration at 37°C. Ampicillin (Ap, 100 μ g/ml) and tetracycline (Tc, 10 μ g/ml) were added to the growth media, when appropriate.

Strain or DNA	Relevant features	Reference or source*
Strains HB101 CL225 DH5α	hsdR hsdM recA13 supE44 leuB6 lacZ proA2 HB101 harbouring pRK100 Φ80dlacZΔM15 Δ(lacZYA-argF)U169 endA1 recA1 hsdR17 deoR thi-1 supE44 gyrA96 relA1	D. Ehrlich ¹ Ambrožič & al. 1998 BRL Life Technologies
Plasmids pGEM-T Easy pRK100	T-vector for cloning of PCR products; Ap' natural plasmid; Ap', Tc'	Promega Ambrožič & al. 1998

Table 1: Bacterial strains and plasmids Tabela 1: Bakterijski sevi in plazmidi

*1 Laboratoire Genetique Microbienne, INRA, Domain de Vilvert, Jouy en Josas, France

General DNA manipulation techniques

Plasmid DNA isolation, ligation and transformation experiments were performed using standard methods (Sambrook & al. 1989). Restriction endonuclease digestions were carried out as specified by the manufacturer (Promega, Boehringer). DNA fragments were isolated from agarose gels using the QIAquick Gel Extraction kit (Qiagen). DNA sequencing was performed using a dye rhodamine terminator cycling reaction and an ABI PRISM[™] 310 Genetic Analyzer automated sequencer and ABI PRISM[™] software.

Polymerase chain reaction

The primers used for the various PCR reactions are presented in Tab. 2. The polymerase chain reactions (PCR) were performed in a 50 μ l PCR reaction mixture with 20 pmol of the two primers, 5 μ l ligation mixture, 0.2 mM of dNTP mixture (Pharmacia), 0.625 U Taq DNA polymerase (Promega) and 1× PCR buffer (Promega). The PCR programs used in this study are listed in Tab. 2. Each PCR amplification program started with a prolonged denaturation (94°C – 4 min) step before 1st cycle, and ended with a prolonged extension step (72°C – 10 min) after the last cycle. The pGEM-T Easy system (Promega) was used for cloning of PCR products.

Table 2: Oligonucleotide primers and PCR programs

Tabela 2: Oligonukleotidni začetniki in programi verižne reakcije s polimerazo (PCR)

Primer	Nucleotide sequence	PCR program
Oligonucleotide primers for traM, finP, traJ PCR: FinP-1 FinP-2	5'-TATTGAGAAGCGTCGACAGG-3' 5'-TGACGAACATGAGCAGCATC-3'	(94°C-1:00, 55°C-1:00, 72°C-1:00) 30×
for traY PCR: TraY-f ¹ TraY-r ¹	5'-GGAATTCAAGATTTGGTACACGTTCTGC-3' 5'-GGAATTCCTTCCTCTTTATCTGCCTCCC-3'	(94°C-1:00, 63°C- 1:00, 72°C-2:00) 30×
for traD PCR: TraD-f ¹ TraD-r ¹	5'-GGAATTCCAGATTGCGTCCATGCGTATCC-3' 5'-GGAATTCATCACCACACATATCACCGCGC-3'	(94°C-1:00, 63°C-1:00, 72°C-1:00) 30×
for <i>tral</i> PCR: TraI-1 FinO-4	5'-ACAGCGAATATACGTGACGG-3' 5'-CGTGGTGACATTGATGATGG-3'	(94°C-0:30, 57°C-0:30, 72°C-3:00) 30×
for <i>finO</i> PCR: FinO-f ¹ FinO-r ¹	5'-GGAATTCGAAGCGACCGGTACTGACACTG-3' 5'-GGAATTCGCCTGAAGTTCTGCCTTTATCCG-3'	(94°C-1:00, 63°C-1:00, 72°C-2:00) 30×

¹ Nucleotide sequence is based on Boyd & al. 1996

¹ Nukleotidno zaporedje z ozirom na Boyd & al. 1996

Sequence analysis

DNA sequences were compiled and analysed using the CLUSTAL W (Thompson & al. 1994) program for sequence alignment and the program DNADIST in the PHYLIP package (Felsenstein 1993, Felsenstein 1989) for calculating distance matrixes.

Results

The compared nucleotide sequence of pRK100 *traM* is most similar to *traM* of plasmids F and pColV-K30

The *traM* gene of F-like plasmids is approximately 380 bp long. Our obtained nucleotide sequence of the pRK100 *traM* gene is partial, covering 308 nucleotides (GenBank accession number AF237698). Subsequently, in nucleotide sequence analysis the obtained 308 nucleotides were compared with

corresponding sequences of related plasmids. The phylogenetic distance between the obtained *traM* nucleotide sequence and *traM* of pColV-K30 and plasmid F is the smallest, namely 0.0098 (Fig. 2A). On the basis of the calculated nucleotide divergence, we can conclude, that the *traM* nucleotide sequence of pRK100 is most similar to *traM* of F and pColV-K30.

The compared nucleotide sequence of pRK100 finP is identical to finP of plasmid F

The *finP* gene of the F-like plasmids is approximately 80 bp long and the full length pRK100 *finP* gene sequence (GenBank accession number AF237698) was compared with corresponding sequences of related plasmids. Since there is no nucleotide divergence between pRK100 *finP* and F *finP* (Fig. 2B), the *finP* of pRK100 is completely the same as the *finP* of plasmid F.

The compared nucleotide sequence of pRK100 traJ is most similar to traJ of plasmid F

The *traJ* gene of F-like plasmids is 600-700 bp long. Our obtained nucleotide sequence of pRK100 *traJ* gene is partial, covering approximately one half of the gene (363 bp) (GenBank accession number AF237698). Therefore, in nucleotide sequence analysis the obtained 363 bp were compared with corresponding sequences of related plasmids. The phylogenetic distance between the obtained *traJ* nucleotide sequence and *traJ* of plasmid F is the smallest, only 0.0028. The distance to *traJ* of plasmid pColV-K30 is also very small, only 0.0055 (Fig. 2C). We can conclude that, the *traJ* nucleotide sequence of pRK100 is most similar to *traJ* of plasmid F.

The compared nucleotide sequence of pRK100 traY is most similar to traY of plasmids F and pColV-K30

The *traY* genes of F-like plasmids can be one of two lengths, approximately 230 bp, and probably due to duplication (Maneewannakul & al. 1996), approximately 400 bp. Our obtained nucleotide sequence of the pRK100 *traY* gene is partial, covering 326 nucleotides (GenBank accession number AF237695). Due to the larger size, we can conclude that, the pRK100 *traY* is of the duplicated type. In our sequence analysis the obtained 326 bp were compared with corresponding sequences of related plasmids. The compared nucleotide sequences of pRK100, pColV-K30 and plasmid F were completely identical with no nucleotide divergence among them (Fig. 2D).

The compared nucleotide sequence of pRK100 traD is most similar to traD of plasmid F

The *traD* gene of F-like plasmids is approximately 2200 bp long. Our obtained nucleotide sequence of pRK100 *traD* gene is partial, covering 574 bp (GenBank accession number AF237693). Therefore, in nucleotide sequence analysis the obtained 574 nucleotides were compared with corresponding sequences of related plasmids. The phylogenetic distance between the obtained *traD* nucleotide sequence and *traD* of plasmid F is the smallest, only 0.0359 (Fig. 2E). We can conclude, that the *traD* nucleotide sequence of pRK100 is most similar to *traD* of plasmid F.

The compared nucleotide sequence of pRK100 tral is most similar to tral of plasmid R100

The *traI* gene of F-like plasmids is approximately 5200 bp long. Our obtained nucleotide sequence of pRK100 *traI* gene is partial, covering only 258 bp (GenBank accession number AY230887). In the nucleotide sequence analysis the obtained 258 bp were compared with corresponding sequences of related plasmids. The phylogenetic distance between the obtained *traI* nucleotide sequence and *traI* of plasmid R100 is the smallest, the nucleotide divergence is 0.0237 (Fig. 2F). The *traI* nucleotide sequence of pRK100 is therefore, most similar to *traI* of plasmid R100.

The compared nucleotide sequence of pRK100 finO is most similar to finO of plasmid F

The *finO* gene of the F-like plasmids is approximately 560 bp long. The obtained pRK100 *finO* gene sequence is partial, encompassing 486 nucleotides (GenBank accession number AF237696) and these were compared with corresponding sequences of related plasmids. The smallest nucleotide divergence, 0.0316, was found to be between pRK100 *finO* and F *finO* (Fig. 2G). Therefore, we can conclude that *finO* of pRK100 is most similar to *finO* of plasmid F.



pSLT

P307

R1

0

0.0847

0.088

0 1142

0,2

0,4

0.6

PSLT

F

0

0.2222

0.4

0.2

0.6

Nucleotide divergence

0.8

1

1.2



1

1,2

0.8

Figure 2: Nucleotide divergence of the obtained pRK100 tra sequences and related plasmid sequences. The nucleotide sequences of pRK100 related plasmids used in the sequence analysis of tra pRK100 nucleotide sequences are deposited in GenBank under the following accession numbers: F-plasmid -U01159; pCoIV-K30 - AF237697 (traM, finP, traJ), AF237694 (traY), AF237692 (traD); P307 -M62986; R100 - AP000342; R1 - M19710 and pSLT - AE006471.

Nucleotide divergence, as the measurement of phylogenetic nucleotide distance, was calculated according to the Kimura-2 parameter. Nucleotide divergences between pRK100 sequences and sequences of other related plasmids are plotted: Panel A – nucleotide divergence of traM, panel B – nucleotide divergence of *finP*, panel C – nucleotide divergence of *traJ*, panel D – nucleotide divergence of *traY*, panel E - nucleotide divergence of traD, panel F - nucleotide divergence of traI and panel G nucleotide divergence of finO.

Slika 2: Divergenca nukleotidnih zaporedij genov tra plazmida pRK100 in nukleotidnih zaporedij sorodnih plazmidov. Nukleotidna zaporedja plazmidov sorodnih pRK100, katera smo uporabili v analizi zaporedij tra genov pRK100, so shranjena v GenBank in označena s sledečimi številkami: Plazmid F - U01159; pColV-K30 - AF237697 (traM, finP, traJ), AF237694 (traY), AF237692

(traD); P307 – M62986; R100 – AP000342; R1 – M19710 in pSLT – AE006471.

Nukleotidna divergenca kot izračun filogenetske drugačnosti nukleotidov je bila preračunana z uporabo parametra Kimura-2. Divergence nukleotidnih zaporedij plazmida pRK100 in nukleotidnih zaporedij sorodnih plazmidov so prikazane: panel A – nukleotidna divergenca gena *traM*, panel B – nukleotidna divergenca gena *traM*, panel C – nukleotidna divergenca gena *traJ*, panel D – nukleotidna divergenca gena *traY*, panel E – nukleotidna divergence gena *traD*, panel F – nukleotidna divergenca gena *traI* in panel G – nukleotidna divergence gena *finO*.

Discussion

In order to characterise the *tra* region of pRK100, nucleotide sequences of seven different *tra* region genes, amplified by PCR, were determined and compared with nucleotide sequences of related plasmids. The chosen *tra* region genes were: *traM*, *finP*, *traJ*, *traY* from one end the and *traD*, *traI*, *finO* from the other end of the pRK100 *tra* region. The obtained pRK100 nucleotide sequences were compared with nucleotide sequences of the following F-like plasmids: plasmid F, pCoIV-K30, P307 and R1 from *Escherichia coli*, R100 from *Shigella flexneri* and pSLT from *Salmonella typhimurium*.

The choice of related plasmids was based on the fact, that pRK100 seems to be an F-like plasmid, which evolved from at least two F-like plasmids. In our previous work it was discovered, that pRK100 is an IncFI plasmid harbouring a RepFIB (Ambrožič & al. 1998) and RepFIIA replication region. The RepFIB replication region is similar to the RepFIB replication region of plasmid F, pColV-K30, P307, and the RepFIIA replication region is similar to R100 (our unpublished data). pRK100 also encodes colicin V and the aerobactin uptake system as does plasmid pColV-K30 (Ambrožič & al. 1998). Further, it also carries IS1, which is also present on pColV-K30 and on plasmid R100. Apart from IS1, pRK100 also encodes IS2 and IS3 insertion sequences (our unpublished data), which are also present on plasmid F, but not on R100.

A similar melange also emerges from results presented in this paper. The *traM* gene of pRK100 has the same nucleotide divergence with *traM* of F and *traM* of pColV-K30; the *finP* gene of pRK100 is exactly the same as *finP* of F; *traY* of pRK100 has the same nucleotide sequence as *traY* of F but also the same sequence as *traY* of pColV-K30; pRK100 *traJ*, *traD* and *finO* are most similar to *traJ*, *traD* and *finO* of plasmid F; but *traI* of pRK100 is most similar to *traI* from R100. On the basis of our results we can therefore conclude that, the *tra* region of pRK100 is F-like and it is most similar to the *tra* region of plasmid F.

Even though all the plasmids, incorporated into this study, are known to be F-like plasmids, nucleotide sequence divergence among the different genes is not the same, some genes are more conserved than others. For example, the regulatory genes *traJ* and *traY* exhibit greater differences, than for example the *traD* gene, whose product is involved in transmembrane conveyance of nucleic acids (Firth & al. 1996). Plasmid genes are mosaic in structure due to multiple recombination events between diverse ancestral genes (Boyd & al. 1996). Conjugation by introducing horizontally transferred DNA into cells increases the opportunity of different plasmids to meet and exchange genetic information. Knowledge of plasmid structure and transfer genes is necessary to develop efficient means to reduce plasmid transfer and dissemination of antibiotic resistances as well as bacterial virulence factors.

Conclusions

To summarise and conclude:

- 1. different *tra* genes of pRK100 show different levels of nucleotide divergence with different related plasmids;
- 2. the tra region of pRK100 is F-like;
- 3. the pRK100 tra region is pronouncedly most similar to the tra region of plasmid F
- 4. the pRK100 tra region is mosaic.

Acknowledgement

The authors are thankful to Peter Trontelj for his help with computer programs CLUSTAL W and PHYLIP.

Literature

- AMBROŽIČ J., OSTROVERŠNIK A., STARČIČ M., KUHAR I., GRABNAR M. & D. ŽGUR-BEERTOK 1998: Escherichia coli ColV plasmid pRK100: genetic organisation, stability and conjugal transfer. Microbiology 144: 343-352.
- BOYD E. F., HILL C. W., RICH S. M. & D. L. HARTL 1996: Mosaic structure of plasmids from natural populations of *Escherichia coli*. Genetics 143:1091-100.
- DEMPSEY W. B. 1994: *traJ* sense RNA initiates at two different promoters in R100-1 and forms two stable hybrids with antisense *finP* RNA. Mol. Microbiol. 13:313-326
- FELSENSTEIN J. 1993: PHYLIP (Phylogeny Inference Package) version 3.5c. Distributed by the author. Department of Genetics, University of Washington, Seattle.
- FELSENSTEIN J. 1989: PHYLIP Phylogeny Inference Package (Version 3.2) Cladistics 5:164-166.
- FIRTH N., IPPEN-IHLER K. & R. A. ŠKŪRŘAY; NEIDHARDT F. C., CURTIS R. I., INGRAHAM J. L., & al. (ed.) 1996: Structure and function of the F factor and mechanism of conjugation. *Escherichia coli* and *Salmonella*. 2nd ed. Washington: ASM Press pp. 2377-401.
- FROST L., IPPEN-IHLER K. & R. A. SKURRAY 1994: Analysis of the sequence and gene products of the transfer region of the F sex factor. Microbiol. Rev. 58:162-210.
- MANEEWANNAKUL K., KATHIR P., ENDLEY S., MOORE D., MANCHAK J., FROST L. & K. IPPEN-IHLER 1996: Construction of derivatives of the F plasmid pOX-*tra715*: characterization of *traY* and *traD* mutants that can be complemented *in trans*. Mol Microbiol 22:197-205.
- SAMBROOK J., FRITSCH E. F. & T. MANIATIS 1989: Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- THOMPSON, J.D., HIGGINS, D.G. & T. J. GIBSON 1994: CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Research 22:4673-4680.
- ŽGUR-BERTOK D., MODRIČ E. & M. GRABNAR 1990: Aerobactin uptake system, ColV production, and drug resistance encoded by a plasmid from an urinary tract infection *Escherichia* coli strain of human origin. Can J Microbiol. 36:297-299.