

Genetic and physical map of broad host range cosmid pRG930cm

Andrew Chibeu[#]

Laboratory of Microbial Interactions
Department of Cellular and Molecular Interactions
Flanders Interuniversity Institute of Biotechnology
Vrije Universiteit Brussel
Pleinlaan 2, B-1050
Brussels, Belgium

Rob Lavigne*

Division of Gene Technology
Katholieke Universiteit Leuven
Kasteelpark Arenberg 21
Leuven, B-3001, Belgium
Fax: 32 016 32 19 65
E-mail: rob.lavigne@biw.kuleuven.be

Wesley Mattheus

Division of Gene Technology
Katholieke Universiteit Leuven
Kasteelpark Arenberg 21
Leuven, B-3001, Belgium

Sandra Matthijs

Laboratory of Microbial Interactions
Department of Cellular and Molecular Interactions
Flanders Interuniversity Institute of Biotechnology
Vrije Universiteit Brussel
Pleinlaan 2, B-1050
Brussels, Belgium

Pierre Cornelis

Laboratory of Microbial Interactions
Department of Cellular and Molecular Interactions
Flanders Interuniversity Institute of Biotechnology
Vrije Universiteit Brussel
Pleinlaan 2, B-1050
Brussels, Belgium

Guido Volckaert

Division of Gene Technology
Katholieke Universiteit Leuven
Kasteelpark Arenberg 21
Leuven, B-300, Belgium

Financial support: AC was initially supported by a joint grant of the laboratories involved and currently holds an IRO scholarship from KU Leuven. RL holds a postdoctoral fellowship from the FWO (Fonds voor Wetenschappelijk Onderzoek), Vlaanderen.

Keywords: cosmid, pRG930cm, *Pseudomonas*, sequencing.

Present address: [#]Division of Gene Technology, Katholieke Universiteit Leuven, Kasteelpark Arenberg 21, Leuven, B-3001, Belgium.

Abbreviations: Cm^R: chloramphenicol resistance cassette
IncPa: incompatibility group P alpha
MCS: multiple cloning site
ORF: open reading frame

We hereby present the complete sequence and annotation of pRG930cm, a spectinomycin/streptomycin/chloramphenicol-resistant cosmid vector.

pRG930cm (17,256 bp; GenBank Accession No.: FM174471) has a broad host range, and is stably maintained by a number of Gram-negative bacteria

*Corresponding author

including *Pseudomonas spp.*, *Escherichia coli*, *Agrobacterium tumefaciens* and *Azorhizobium caulinodans* ORS571. pRG930cm is already widely used and its sequence will aid efficient construction and analysis of cosmid libraries.

Incompatibility group P alpha (IncP α) low-copy-number plasmids have a wide host range and replicate/maintain themselves stably in most Gram-negative bacteria (Pansegrau et al. 1994). pVS1 plasmid, an IncP plasmid isolated from *Pseudomonas aeruginosa* replicates in a wide range of Gram-negative bacteria but not in *E. coli* (Itoh et al. 1984). The plasmid's regions for replication, stability and mobilization are clustered within an 8 kb region which was ligated into pBR325 resulting in plasmid pGV910, a broad host range cloning vector with pVS1 regions essential for maintenance, stability and mobilization. pGV910 is 15.6 kb in size and has both ColE1 and pVS1 origins of replication (Van den Eede et al. 1992). Cosmid vector pRG930 was constructed by ligating the 1.6 kb *Bgl*III *cos* fragment from cosmid pLAFR1 (Friedman et al. 1982; Vanbleu et al. 2004) into the *Bam*H1 site of pGV910 and inserting part of pUC8 into the unique *Eco*RI site. To increase the scale of antibiotic resistance, and therefore the host range, the Cm^R cassette of pBR325 was inserted into the unique *Hind*III site of pRG930 resulting in pRG930cm (Matthijs et al. 2004).

pGV910 and its descendants pRG930 and pRG930cm are not self-transmissible but can be mobilized by pRK2013 (Figurski and Helinski, 1979). pRK2013 is a ColE1 plasmid carrying its own *mob* genes plus the RK2 *tra* genes (Comai et al. 1983).

Knowledge of the genes and restriction sites present on plasmid pGV910 or on cosmids pRG930 and pRG930cm was lacking and could not be reconstructed from available sequences of the ancestor plasmids due to their mode of construction. However, pRG930cm has been used in the construction of a number of stable clone libraries in *Pseudomonas*, *Burkholderia*, and *Actetobacter* (Aendekerck et al. 2002; de Chial et al. 2003; Matthijs et al. 2004; Plesa et al. 2004; Denayer et al. 2007). Subcloning and further analysis of these clone libraries has been tedious due to the lack of the full DNA sequence of the cosmid.

MATERIALS AND METHODS

We sequenced pRG930cm with average five-fold coverage by shotgun sequencing, for which each nucleotide position was sequenced at least three times. pRG930cm sequence is available under GenBank Accession Number FM174471. The sequencing procedure involved sonication of the entire cosmid DNA, end repair of the fragments and ligation of 1- to 2-kb DNA fragments into the *Sma*I site of pUC18 and transformation with *E. coli* XL1-Blue cells. Plasmids extracted from single colonies were used directly as template for sequencing (Ganguly et al. 2005). After standard ethanol precipitation of the samples, they were

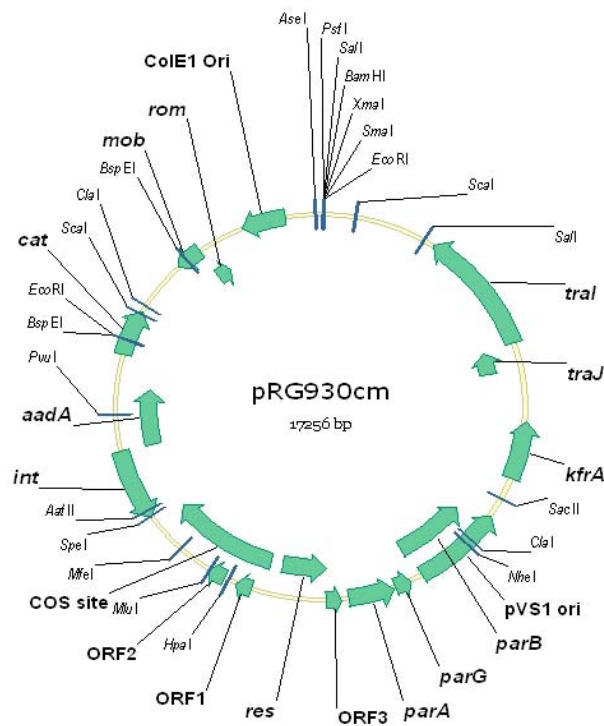


Figure 1. Genetic map of cosmid vector pRG930cm including the restriction sites of enzymes cutting at not more two locations. Position 1 on the vector is the cut site of *PstI* enzyme.

fractionated and analyzed on an ABI 3130 capillary sequencing device (Applied Biosystems, Foster City, CA). Sequence assembly was performed with Sequencher 4.1 software (Genecodes, Ann Arbor, MI). Both DNA strands were sequenced using shotgun sequencing while uncertain positions were verified by direct sequencing using pRG930cm-specific primers. The complete and truncated ORFs present on the vector were predicted using GeneMark hmm (Besemer and Borodovsky, 1999) and verified manually. The translated ORFs were compared to known proteins using standard protein-protein BLASTP (Altschul et al. 1990). Annotation of identified ORFs was done by manual curation and interpretation of outputs from a BLASTP search. Information obtained is summarized in Figure 1 and Table 1.

RESULTS AND DISCUSSION

pRG930cm is a 17,256 bp cosmid with the ColE1 and pVS1 origins of replication. BLASTN similarity search shows that the 599 bp region (from nucleotide coordinates 16178 to 16776) corresponds to ColE1-ori. Nucleotide coordinates 5930 to 7195 correspond to the pVS1 ori based on similarity to the ori of replication of plasmid pZU634. The 1.7 kb *cos* region derived from cosmid pLAFR1 contains the *cos* site and two phage 80 genes, ORF1 and ORF2. ORF3 is a hypothetical protein with no significant homology to proteins in the databases.

Table 1. ORFs encoded on the cosmid vector pRG930cm.

| Position | Gene name | Function/characteristic | Source |
|-------------|-------------|---|--|
| 1570-3357 | <i>tral</i> | DNA relaxase | (Thorsted et al. 1998) |
| 3354-3728 | <i>traJ</i> | OriT-recognizing protein | (Fürste et al. 1989; Ziegelin et al. 1989) |
| 4506-5336 | <i>kfrA</i> | Transcription regulation | (Jagura-Burdzy et al. 1992) |
| 6108-7283 | <i>parB</i> | pVS1 replication protein ParB | (Austin and Abeles, 1983) |
| 7374-7589 | <i>parG</i> | pVS1 partitioning protein Par G | (Hayes, 2000) |
| 7610-8239 | <i>parA</i> | Partitioning protein ParA | (Austin and Abeles, 1983) |
| 8326-8541 | ORF 3 | Hypothetical protein | |
| 8538-9224 | | Resolvase | (Swinfield et al. 1991) |
| 9560-9790 | ORF 1 | Bacteriophage phi 80 unknown gene | |
| 9952-10680 | ORF 2 | Bacteriophage phi 80 unknown gene | |
| 11288-12301 | <i>IntI</i> | Integrase | (Esposito and Scocca, 1997) |
| 12300-13241 | <i>aadA</i> | streptomycin/spectinomycin adenylyltransferase | (Goldschmidt-Clermont, 1991) |
| 13702-14361 | <i>cat</i> | Chloramphenicol acetyltransferase | (Shaw, 1983) |
| 15161-15529 | <i>mob</i> | Mob-like protein | |
| 15528-15752 | <i>rom</i> | Rom/Rop protein. RNA I inhibition modulator protein | (Cesareni et al. 1982) |
| 16178-16776 | | ColE1 <i>ori</i> | (Hashimoto-Gotoh and Timmis, 1981) |

In addition, regions conferring chloramphenicol (*cat*) and spectinomycin/streptomycin (*aadA*) resistance are present. The genes responsible for conferring tetracycline (*tetA* and *tetR*; positions 14503-14922 and 1195 bp-1412 respectively) and ampicillin resistances (*bla*; position 16938-17256) are non-functional truncations. pRG930cm has *Pst*I, *Sal*I, *Bam*HI, *Xma*I, *Sma* I and *Eco*RI restriction sites in its multiple cloning site (MCS). Restriction enzyme *Sal*I is however not unique and cannot be used for cloning. The presence of common unique restriction enzyme sites in the MCS enable efficient cloning into this vector. The truncated *cat* gene (position 18-464) is located adjacent to the MCS. An identical sequence of the complete *cat* gene occurs at position 13702-14361. This has made it difficult to design primers annealing to a unique sequence flanking the MCS. This has led to difficulties in analyzing cosmid inserts in the past. With the complete pRG930cm sequence determined, primers designed upstream (5'-GATACCGCGAGACCCACGCT-3') and downstream (5'-CATTGCCATACGGAATTCCC-3') of the MCS,

respectively, have successfully been used in establishing the flanking sequences of cosmid inserts.

The *parA*, *parG* and *parB* gene cluster (6108-8239 bps) is involved in partitioning and stabilization of the pRG930cm. ParB protein is responsible for localizing to opposite poles DNA sequences adjacent to the pVS1 *ori*. ParA ATPase activity is regulated by ParB (Abeles et al. 1985). The cosmid encoded pVS1 *parG* is homologous to the *parG* gene of plasmid TP288 partitioning system. The positioning of the genes upstream from *parF* in pVS1 has suggested that *parG* also plays a role in partitioning (Hayes, 2000). A pVS1 resolvase of the Serine Recombinase (SR) family is encoded on vector backbone positions 8538 to 9224. The protein is also likely to contribute to the segregational stability of pRG930cm by facilitating efficient partitioning through conversion of plasmid multimers into monomers (Swinfield et al. 1991). An integrase gene homologous to the *P. aeruginosa* integrase gene is also encoded in pRG930cm. Such integron-integrases and are found to

occur naturally in most mobile elements including transposons and conjugative plasmids (Esposito and Scocca, 1997).

pRG930cm also has a *mob* gene which encodes a Mob-like protein similar to that of *E. coli* species Sflu5. The *rom* gene which is adjacent to the *mob* gene but transcribed in opposite direction codes for the Rom/Rop protein which plays a role in maintaining the cosmid copy number by inhibiting plasmid replication. This is by enhancing the pairing between RNA I and RNA II so that processing of the primer can be inhibited even at relatively low concentrations of RNA I (Cesareni et al. 1982).

The complete characterization of pRG930cm will enhance its role in restriction enzyme mapping and rapid chromosomal walking. The sequence may also be used to screen out components when "shotgun" sequencing inserts are cloned into this cosmid or into its predecessor pRG930.

REFERENCES

- ABELES, A.L.; FRIEDMAN, S.A. and AUSTIN, S.J. Partition of unit-copy miniplasmids to daughter cells. III. The DNA sequence and functional organization of the P1 partition region. *Journal of Molecular Biology*, September 1985, vol. 185, no. 2, p. 261-272.
- AENDEKERK, Séverine; GHYSELS, Bart; CORNELIS, Pierre and BAYSSE, Christine. Characterization of a new efflux pump, MexGHI-OpmD, from *Pseudomonas aeruginosa* that confers resistance to vanadium. *Microbiology*, August 2002, vol. 148, no. 8, p. 2371-2381.
- ALTSCHUL, Stephen F.; GISH, Warren; MILLER, Webb; MYERS, Eugene W. and LIPMAN, David J. Basic local alignment search tool. *Journal of Molecular Biology*, October 1990, vol. 215, no. 3, p. 403-410.
- AUSTIN, Stuart and ABELES, Ann. Partition of unit-copy miniplasmids to daughter cells. I. P1 and F miniplasmids contain discrete, interchangeable sequences sufficient to promote equipartition. *Journal of Molecular Biology*, September 1983, vol. 169, no. 2, p. 353-372.
- BESEMER, John and BORODOVSKY, Mark. Heuristic approach to deriving models for gene finding. *Nucleic Acid Research*, October 1999, vol. 27, no. 19, p. 3911-3920.
- CESARENI, G.; MUESING, M.A. and POLISKY, B. Control of COLE1 DNA replication: the *rop* gene product negatively affects transcription from the replication primer promoter. *Proceedings of the National Academy of Sciences of the United States of America*, October 1982, vol. 79, no. 20, p. 6313-6317.
- COMAI, Luca; SCHILLING-CORDARO, Christine; MERGIA, Ayalew and HOUCK, Catherine M. A new technique for genetic engineering of *Agrobacterium* Ti plasmid. *Plasmid*, July 1983, vol. 10, no. 1, p. 21-30.
- DE CHIAL, Magaly; GHYSELS, Bart; BEATSON, Scott A.; GEOFFROY, Valérie; MEYER, Jean Marie; PATTERY, Theresa; BAYSSE, Christine; CHABLAIN, Patrice; PARSONS, Yasmin N.; WINSTANLEY, Craig; CORDWELL, Stuart J. and CORNELIS, Pierre. Identification of type II and type III pyoverdine receptors from *Pseudomonas aeruginosa*. *Microbiology*, April 2003, vol. 149, no. 4, p. 821-831.
- DENAYER, Sarah; MATTHIJS, Sandra and CORNELIS, Pierre. Pyocin S2 (Sa) kills *Pseudomonas aeruginosa* strains via the FpvA type I ferripyoverdine receptor. *Journal of Bacteriology*, November 2007, vol. 189, no. 21, p. 7663-7668.
- ESPOSITO, D. and SCOCCA, J.J. The integrase family of tyrosine recombinases: evolution of a conserved active site domain. *Nucleic Acids Research*, September 1997, vol. 25, no. 18, p. 3605-3614.
- FIGURSKI, D.H. and HELINSKI, D.R. Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided in *trans*. *Proceedings of the National Academy of Sciences of the United States of America*, April 1979, vol. 76, no. 4, p. 1648-1652.
- FRIEDMAN, Alan M.; LONG, Sharon R.; BROWN, Susan E.; BUIKEMA, William J. and AUSUBEL, Frederick M. Construction of a broad host range cosmid cloning vector and its use in the genetic analysis of *Rhizobium* mutants. *Gene*, June 1982, vol. 18, no. 3, p. 289-296.
- FÜRSTE, Jens Peter; PANSEGRAU, Werner; ZIEGELIN, Gunter; KRÖGER, Manfred and LANKA, Erich. Conjugative transfer of promiscuous IncP plasmids: interaction of plasmid-encoded products with the transfer origin. *Proceedings of the National Academy of Sciences of the United States of America*, March 1989, vol. 86, no. 6, p. 1771-1775.
- GANGULY, Tapan; CHEN, Peiqin; TEETSEL, Rebecca; ZHANG, Lan Ping; PAPAIOANNOU, Elias and CIANCIARULO, Joseph. High-throughput sequencing of high copy number plasmids from bacterial cultures by heat lysis. *Biotechniques*, September 2005, vol. 39, no. 3, p. 304-308.
- GOLDSCHMIDT-CLERMONT, Michel. Transgenic expression of aminoglycoside adenine transferase in the chloroplast: a selectable marker of site-directed transformation of *Chlamydomonas*. *Nucleic Acids Research*, August 1991, vol. 19, no. 15, p. 4083-4089.
- HASHIMOTO-GOTOH, T. and TIMMIS, K.N. Incompatibility properties of Col E1 and pMB1 derivative plasmids: random replication of multicopy replicons. *Cell*, January 1981, vol. 23, no. 1, p. 229-238.
- HAYES, Finbarr. The partition system of multidrug resistance plasmid TP228 includes a novel protein that

- epitomizes an evolutionarily distinct subgroup of the ParA superfamily. *Molecular Microbiology*, January 2000, vol. 37, no. 3, p. 528-41.
- ITOH, Yoshifumi; WATSON, John M.; HAAS, Dieter and LEISINGER, Thomas. Genetic and molecular characterization of the *Pseudomonas* plasmid pVS1. *Plasmid*, May 1984, vol. 11, no. 3, p. 206-220.
- JAGURA-BURDZY, Grazyna; KHANIM, Farhat; SMITH, Christopher A. and THOMAS, Christopher M. Crosstalk between plasmid vegetative replication and conjugative transfer: repression of the *trfA* operon by *trbA* of broad host range plasmid RK2. *Nucleic Acids Research*, August 1992, vol. 20, no. 15, p. 3939-3944.
- MATTHIJS, Sandra; BAYSSE, Christine; KOEDAM, Nico; TEHRANI ABBASPOUR, Kouros; VERHEYDEN, Lieve; BUDZIKIEWICZ, Herbert; SCHÄFER, Mathias; HOORELBEKE, Bart; MEYER, Jean-Marie; DE GREVE, Henri and CORNELIS, Pierre. The *Pseudomonas* siderophore quinolobactin is synthesized from xanthurenic acid, an intermediate of the kynurenine pathway. *Molecular Microbiology*, April 2004, vol. 52, no. 2, p. 371-384.
- PANSEGRAU, W.; LANKA, E.; BARTH, P.T.; FIGURSKI, D.H.; GUINEY, D.G.; HAAS, D.; HELINSKI, D.R.; SCHWAB, H.; STANISICH, V.A. and THOMAS, C.M. Complete nucleotide sequence of Birmingham IncPa plasmids. Compilation and comparative analysis. *Journal of Molecular Biology*, June 1994, vol. 239, no. 5, p. 623-663.
- PLESA, Maria; KHOLTI, Abdelaziz; VERMIS, Karen; VANDAMME, Peter; PANAGEA, Stavroula; WINSTANLEY, Craig and CORNELIS, Pierre. Conservation of the *opcL* gene encoding the peptidoglycan-associated outer-membrane lipoprotein among representatives of the *Burkholderia cepacia* complex. *Journal of Medical Microbiology*, May 2004, vol. 53, no. 5, p. 389-398.
- SHAW, William V. Chloramphenicol acetyltransferase: enzymology and molecular biology. *Critical Reviews in Biochemistry and Molecular Biology*, 1983, vol. 14, no. 1, p. 1-47.
- SWINFIELD, Tract-Jane; JANNIERE, Laurent; EHRLICH, Dusko S. and MINTON, Nigel P. Characterization of a region of the *Enterococcus faecalis* plasmid pAM β 1 which enhances the segregational stability of pAM β 1-derived cloning vectors in *Bacillus subtilis*. *Plasmid*, November 1991, vol. 26, no. 3, p. 209-221.
- THORSTED, Peter B.; MACARTNEY, Donia P.; AKHTAR, Parveen; HAINES, Anthony S.; ALI, Nasima; DAVIDSON, Philip; STAFFORD, Theresa; POCKLINGTON, Michael J.; PANSEGRAU, Werner; WILKINS, Brian M.; LANKA, Erich and THOMAS, Christopher M. Complete sequence of the IncP β plasmid R751: implications for evolution and organisation of the IncP backbone. *Journal of Molecular Biology*, October 1998, vol. 282, no. 5, p. 969-990.
- VAN DEN EEDE, G.; DEBLAERE, R.; GOETHALS, K.; VAN MONTAGU, M. and HOLSTERS, M. Broad host range and promoter selection vectors for bacteria that interact with plants. *Molecular Plant-Microbe Interactions*, May-June 1992, vol. 5, no. 3, p. 228-234.
- VANBLEU, Els; MARCHAL, Kathleen and VANDERLEYDEN, Jos. Genetic and physical map of the pLAFR1 vector. *DNA sequence*, 2004, vol. 15, no. 3, p. 225-227.
- ZIEGELIN, G.; FÜRSTE, J.P. and LANKA, E. TraJ protein of plasmid RP4 binds to a 19-base pair invert sequence repetition within the transfer origin. *Journal of Biological Chemistry*, July 1989, vol. 264, no. 20, p. 11989-11994.