The complete nucleotide sequence of cosmid vector pTL5: location and origin of its genetic components

(Recombinant DNA; pBR322; phage λ; phage φ80; bacteriophage cohesive ends; Charon 4A)

Jerry L. Slightom[a,b], David R. Siemieniak[b], Ben F. Koop[c] and Leroy Hood[d]

[a]*Molecular Biology Unit 7242, The Upjohn Company, Kalamazoo, MI 49007, USA;* [b]*Human Genome Center, The University of Michigan, Ann Arbor, MI 48109-0650, USA. Tel. (1-313) 747-4783;* [c]*Department of Biology-Center for Environmental Health, University of Victoria, Victoria, BC, V8W2Y2, Canada. Tel. (1-604) 656-8146; and* [d]*Department of Molecular Biotechnology, University of Washington, Seattle, WA 98105, USA. Tel. (1-206) 685-7367*

SUMMARY

The complete nucleotide sequence (5793 bp) of the cosmid vector pTL5 and the origin of its genetic components has been determined. Cosmid pTL5, a derivative of cosmid vector pHC79, is composed of genetic components from pBR322, bacteriophage λ and the hybrid lambdoid bacteriophage Charon (Ch) 4A cohesive ends (*cos*) region. The Ch4A *cos* region contains genetic components from two bacteriophages, the λ *cos*-left arm and the φ80 *cos*-right arm regions. The Ch4A *cos* region has been used in the construction of many other cosmid-type vectors, some of which have been sequenced and entered into the GenBank database.

INTRODUCTION

Cosmid-type cloning vectors are generally composed of an *E. coli* plasmid origin of replication (in many cases the ColE1 *ori*), at least one antibiotic resistance gene and a bacteriophage *cos* region. Hohn and Collins (1980) constructed the versatile cosmid, pHC79; it contains pBR322 (ColE1 *ori* and the *tet* and *bla* genes that encoding Tc[R] and Ap[R], respectively), a 0.65-kb fragment from the λ *cro-cII* region and the chimeric bacteriophage *cos* region contained within a 1.74-kb *Bgl*II fragment isolated from Ch4A. Cosmid vector pTL5 is essentially pHC79 in which

---

Correspondence to: Dr. J.L. Slightom, Molecular Biology Unit 7242, The Upjohn Co., Kalamazoo, MI 49007, USA. Tel. (1-616) 385-5304; Fax (1-616) 385-4599; e-mail: jslighto@upj.com

Abbreviations: aa, amino acids(s); Ap, ampicillin; *bla*, gene encoding β-lactamase; bp, base pair(s); Ch, Charon; *cos, cohesive end(s)* of λ or φ80 bacteriophage; *E., Escherichia*; GCG, Genetics Computer Group (Madison, WI, USA); kb, kilobase(s) or 1000 bp; nt, nucleotide(s); *ori*, origin of DNA replication; [R], resistance/resistant; Tc, tetracycline; *tet*, gene encoding Tc[R].

the *bla* gene and the *Bgl*II site adjacent to the λ *cos*-left arm were removed (Lund et al., 1982).

EXPERIMENTAL AND DISCUSSION

(a) Genetic components of pTL5

Knowledge of the nt sequence of cosmid vectors is important to cloning, mapping and nt sequencing projects, especially those that utilize the 'shotgun' method (Koop et al., 1993), because some of the nt sequencing runs will include vector-derived sequences. While sequencing parts of the mouse and human T-cell receptor variable gene clusters (Lai et al., 1988; Koop et al., 1994; Slightom et al., 1994) we also determined the complete double-stranded sequence of the cosmid vector pTL5. The sequence of pTL5 (GenBank accession No. U07340) was determined by the manual primer-walking method using ^{32}P-radioisotope labeling (Siemieniak et al., 1991) and the shotgun sequencing method using dye-labeled primers and the ABI 373A instrument (Koop et al., 1993).

The identification of pTL5 genetic components is for

TABLE I

Components of pTL5

Component	pTL5 location[a]	Origin (Ori)	GenBank No.	Reference(s)
Tc[R](tet)	1/1460 (99.93)[b]	pBR322	J01749	Bolivar et al. (1977); Sutcliffe (1978)
λ cro-cII	1461/2113 (100)	λ	J02459/M17233	Daniels et al. (1982)
cos-right arm	2114/3448 (98.5)[c]	φ80	none	D.L. Daniels and F.R. Blattner (unpublished)
cos-left arm	3437/3854 (99.8)[d]	λ	J02459/M17233	Daniels et al. (1982)
ColE1 ori	3855/5793 (100)	pBR322	J01749	Bolivar et al. (1977); Sutcliffe (1978); Watson (1988)

[a] Numbers in parentheses indicate the % nt identity between the pTL5 nt sequence and the indicated GenBank entry.

[b] Note that GenBank entry No. J01749 differs at position 1134(C) of tet vs. pTL5 (T), which encodes the aa replacement Thr$^{350} \rightarrow$ Ile. Other GenBank entries for tet share the identical nt at position 1134 with the pTL5 sequence; see Ahmed (1989) and GenBank accessions M20189 and L08817.

[c] 100% identity between the φ80 and pTL5 sequences is not expected because the sequence of φ80 is preliminary (D.L. Daniels and F.R. Blattner, personal communication).

[d] These nt sequences differ at λ position 138, where the λ sequence contains an extra G that is not found in pTL5.

the most part straightforward (Table I), except for the 1.3-kb cos-right arm region. The Ch4A cos region is composed of the φ80 cos-right and λ cos-left arm regions (Williams and Blattner, 1979). Interestingly, there are only limited nt sequence homologies between bacteriophage λ and φ80 (Fiandt et al., 1971). This was confirmed by comparing the pTL5 nt sequences (positions 2114 to 3448) and λ right arm nt sequences (Sanger et al., 1982; Daniels et al., 1982) which reveals only a small region of nt identity (63 bp that share 89% identity which includes the identical 12-bp cos sequence). Confirmation that the pTL5 cos-right arm was derived from φ80 was obtained by comparing this nt sequence with that from the φ80 cos-right arm (Table I). These sequences share 98.5% identity over the compared 1.3 kb, the small degree of non-identity being compatible ambiguities in the preliminary sequence of the φ80 cos-right arm.

(b) Structure and sequence of pTL5-related cosmids

Many other cosmid vectors use all or part of the chimeric Ch4A cos region and some of these have been completely sequenced and entered into GenBank, as revealed by a GenBank search (release 80, using the GCG FASTA

computer program). Table II lists these pTL5-related cosmid vectors. Cosmid pHC79 was not included in the FASTA results, because it is a computer assembled sequence, rather than experimentally determined sequence, in which the λ cos-right arm sequence was used in its assembly, see GenBank accession No. VB0100.

ACKNOWLEDGEMENTS

TABLE II

Sequenced cosmid vectors that contain the φ80 cos-right arm genetic component

Name	φ80 identity region[a]	GenBank No.
pIB8/pHomerI	3284-4620 (100)	U00003
pTCF	7657-6321 (100)	L19899
svPHEP	5962-7298 (100)	L19900
pWE15	5986-7322 (100)	X65279/M99569
pWE15A	6026-6362 (100)	Z12112
sCOS-1	1147-299 (97.9)	M99566

[a] Numbers in parentheses indicate the % nt identity between GenBank entry sequences and the φ80 cos-right arm sequence of pTL5.

REFERENCES

Ahmed, A.: A vector for sequencing long (40-kb) DNA fragments. Gene 75 (1989) 315–321.

Bolivar, F., Rodriguez, R.L. Greene, P.J., Betlach, M.C., Heyneker, H.L., Boyer, H.W., Crosa, J.H. and Falkow, S.: Construction and characterization of new cloning vehicles. II. A multipurpose cloning system. Gene 2 (1977) 95–113.

Daniels, D.L., Schroeder, J.L., Szybalski, W., Sanger, F., Coulson, A.R., Hong, G.-F., Hill, D.F., Peterson, G.B. and Blattner, F.R.: Complete annotated lambda sequence. In: Hendrix, R.W., Roberts, J.W., Stahl, F.W. and Weisberg, R.A. (Eds.), Lambda II. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1983, pp. 519–676.

Fiandt, M., Hradecna, Z., Lozeron, H.A. and Szybalski, W.: Electron micrographic mapping of deletions, insertions, inversions, and homologies in the DNAs of coliphages lambda and phi80. In: Hershey, A.D. (Ed.), The Bacteriophage Lambda, Cold Spring Harbor Laboratories, Cold Spring Harbor, NY, 1971, pp. 329–354.

Hohn, B. and Collins, J.: A small cosmid for efficient cloning of large DNA fragments. Gene 11 (1980) 291–298.

Koop, B.F., Rowen, L., Chen, W.-Q., Deshpande, P., Lee, H. and Hood, L.: Sequence length and error analysis of sequences and automated *Taq* cycle sequencing methods. BioTechniques 14 (1993) 442–447.

Koop, B.F., Rowen, L., Wang, K., Kuo, C.L., Seto, D., Lenstra, J.A., Howard, S., Shan, W., Deshpande, P. and Hood, L.: The human T-cell receptor TCRAC/TCRDC Cα/Cδ region: organization, sequence, and evolution of 97.6 kb of DNA. Genomics 19 (1994) 478–493.

Lai, E., Concannon, P. and Hood, L.:. Conserved organization of the human and murine T-cell receptor β-gene families. Nature 331 (1988) 543–546.

Lund, T., Grosveld, F.G. and Flavell, R.A.: Isolation of transforming DNA by cosmid rescue. Proc. Natl. Acad. Sci. USA 79 (1982) 520–524.

Sanger, F., Coulson, A.R., Hong, G.F., Hill, D.G. and Petersen, G.B.: Nucleotide sequence of bacteriophage λ DNA. J. Mol. Biol. 162 (1982) 729–773.

Siemieniak, D.R., Sieu, L.C. and Slightom, J.L.: Strategy and methods for directly sequencing cosmid clones. Anal. Biochem. 192 (1991) 441–448.

Slightom, J.L., Siemieniak, D.R., Sieu, L.C., Koop, B.F. and Hood, L.: Nucleotide sequence analysis of 77.7 kb of the human V_β T-cell receptor gene locus. Genomics 20 (1994) 149–168.

Sutcliffe, J.G.: Complete nucleotide sequence of the *Escherichia coli* plasmid pBR322. Cold Spring Harbor Symp. Quant. Biol. 43 (1978) 77–90.

Watson, N.: A new revision of the sequence of plasmid pBR322. Gene 70 (1988) 399–403.

Williams, F.B. and Blattner, F.R.: Construction and characterization of the hybrid bacteriophage lambda Charon vectors for DNA cloning. J. Virol. (1979) 555–575.