Genomic Plasticity and Catabolic Potential of *Pseudomonas cepacia*
T.G. Lessie, Department of Microbiology
University of Massachusetts, Amherst, MA 01003

Project Summary (Grant No. DE-FG02-91 ER20051)

The primary goal of our project was to gain information about the size and organization of the genome of *Burkholderia cepacia* (formerly *Pseudomonas cepacia*), a microbe which continues to attract attention because of its extraordinary degradative abilities and potential as an agent of bioremediation. This bacterium is no longer considered to be a member of genus *Pseudomonas* nor does it belong in the gamma-subclass of the proteobacteria, in which the authentic pseudomonads are grouped. It belongs in the less well characterized beta-subclass of the proteobacteria (Palleroni, 1992; Yabuuchi *et al.*, 1992). We used the technology for manipulation of large DNA fragments developed by Cantor *et al.* (1988) to demonstrate that chromosomal multiplicity, a characteristic yet to be observed in a gamma-subclass bacterium, is common among *B. cepacia* strains. We also constructed a derivative of Tn5 suitable for determining the chromosomal locations of various *B. cepacia* genes.

The biochemically well-characterized strain *B. cepacia* 249 (ATCC 17616) contained three chromosomes of 3.4, 2.5, and 0.9-Mb and a cryptic plasmid of 170 kb (Cheng and Lessie, 1994). Other strains we analysed contained between two and four large replicons of 1.1 to 5.8 Mb. Overall genome sizes varied between 4.6 and 8.5-Mb. The number and distribution of *rnm* genes in strain 17616 differed from that in other *B. cepacia* isolates. Major genomic rearrangements were observed in different stocks of strain 249 and of another *B. cepacia* isolate (ATCC 29424). Various insertion sequences we have detected in *B. cepacia* presumably contribute to such genomic plasticity.

In an attempt to gain further insight into mechanisms of IS-element-dependent adaptation of *B. cepacia* to utilize novel substrates, we investigated the activation of a cryptic D-serine deaminase (*dsd*) gene in strain 249. Our working hypothesis was that expression of this gene might be a consequence of insertional activation. It therefore seemed reasonable that factors influencing IS element transposition might affect the frequency of *dsd* gene activation and likewise affect overall genomic plasticity. We were particularly interested in the *B. cepacia* element IS407, a member of the IS3 family, and the regulation of its transposition by translational frameshifting. In this context we undertook *i*) to isolate the *dsd* gene and investigate the mechanism(s) of its activation and *ii*) to construct a variant of IS407, carrying a drug resistance marker and rare restriction enzyme site, whose transposition could be monitored under conditions predicted to inhibit or enhance translational frameshifting.

Progress Report

Genomic organization of *B. cepacia* 249 (ATCC 17616). We have constructed a physical map of the genome of *B. cepacia* 249 which shows the distribution of *Swa*I, *Pac*I, and *Pme*I sites on the three chromosomes of 3.4, 2.5, and 0.9 Mb comprising the genome of this bacterium (Cheng and Lessie, 1994). We also constructed a Tn5 variant containing a trimethoprim resistance gene and a *Swa*I site, which we have used to determine the map coordinates of selected genes on the 3.4- and 2.5-Mb chromosomes. Southern hybridization experiments indicated that the 3.4-Mb chromosome contained at least three sets ribosomal RNA (*rrn*) genes. The 2.5- and 0.9-Mb replicons each appeared to contain at least one set of *rrn* genes. *Ceu*I is an enzyme which appears to cleave bacterial DNA only within genes encoding 23S r-RNA (Lieu and Sanderson, 1995). *Ceu*I digests of DNA from strain 249 contained five-fragments whose sizes were consistent with the presence of three sets of *rrn* genes on the 3.4-Mb chromosome and one set on each of the remaining two chromosomes.

Genomic complexity of other *B. cepacia* isolates. In our studies of *B. cepacia* 249 we demonstrated that preparations of large DNA contained linearized forms of each of the large circular replicons present in this strain. Derivatives of strain 249 in which DNA had been deleted from the 2.5- and 0.9-Mb replicons contained correspondingly smaller linearized replicons. Similar results have been reported for other bacteria containing multiple chromosomes (Michaux *et al.*, 1993; Rodley *et al.*, 1995). This observation afforded a strategy for estimating the number and sizes of replicons in other *B. cepacia* isolates. We prepared large DNA from over a dozen different *B. cepacia* isolates and resolved the linear DNA species present by CHEF gel electrophoresis. For these experiments we substituted pronase E for proteinase K during the preparation of large DNA. This substitution gave significantly better yields of linearized species of the circular replicons present in strain 249. The linear DNA species were resolved at low voltages for approximately 160 h. The results of our survey of various *B. cepacia* strains are summarized in Table 1. All of the *B. cepacia* isolates examined contained two or more large replicons with sizes greater than 1 Mb. Certain of the strains contained as many as four replicons in this size range. A few strains contained several replicons in the size range of 0.3 to 1 Mb. The overall genome sizes varied from 4.6 to ca. 8.5 Mb of DNA.

Based on differences in two phenotypes, ability to utilize phthalate and ability D-serine, we were able to sort the strains examined into two major groups. Members of one group utilized phthalate but failed to utilize D-serine. These included strains 249, DB01, CRE-7, G4, and 67-46. Members of the other group utilized D-serine, but failed to utilize phthalate. These included strains 25416, NTC 10661, 382, 383, 542, and SW3. The majority of the phthalate utilizers contained three chromosomes similar in size to those in strain 249 (see Table 1). The D-serine utilizers were a more diverse group. They contained between two and four chromosomes. Two of the strains containing only two chromosomes, 383 and SW3, had relatively small

genomes equivalent to ca 4.7Mb of DNA. Other D-serine utilizers such as strains 25416 and NTC 10661 contained three replicons, and had significantly larger genomes equivalent to ca 8 Mb of DNA. Recently Rodley *et al.* (1995) published a macrorestriction map of the genome of strain 25416. The size estimates we obtained for the linearized replicons from preparations of large DNA were in general agreement with the sizes of the chromosomes defined by macrorestriction fragment mapping.

The marked differences in size and number of replicons among *B. cepacia* isolates raises questions about what constitutes a core *B. cepacia* genome. One possibility is that the majority of strains contain a similar primary chromosome (perhaps a replicon of ca 3.5-Mb) as well as accessory chromosomes added by lateral transfer from diverse organisms. One approach that might be taken to test this possibility would be to access the extent of DNA sequence homology between different replicons in a given strain such as 17616 or 25416, and then to carry out cross-hybridizations with linearized DNA species from the less well characterised strains. A complementary approach would be to compare the nucleotide sequences of *rrn* genes from different replicons within the same strain.

To obtain information about the chromosomal distribution of *rrn* operons we compared the sizes and numbers of fragments generated by *Ceul* treatment of DNA from different strains. We expected that the *Ceul* fragment patterns of DNA from phthalate-utilizing strains (which contained similar numbers and sizes of linearized replicons) might be the same. However our experiments revealed significant differences between the distribution of *Ceul* sites strain 17616 and the other strains (see Table 2). Digests of DNA from strain 17616 contained only 5 *Ceul* fragments. DNA from the other phthalate utilizers and from the different D-serine utilizers contained six fragments. The sizes as well as the number of *Ceul* fragments from strain 17616 were distinct compared to the other strains, indicating a marked lack of conservation of the distribution of *rrn* operons (Lessie and Manning, 1995).

In the course of surveying various isolates the presence of multiple replicons we discovered cases where different stocks of the same isolate had undergone major genomic rearrangements. One such case involved *B. cepacia* PHK, a phthalate utilizer, which has been the focus of studies of the evolution of pathways related to aromatic compound utilization (Zylstra *et al.*, 1989). This strain was originally isolated by P. Kaiser in the lab of D. Ribbons, as strain PHK, and deposited in the American Type Culture Collection where it was assigned the ATTC number and deposited in the American Type Culture Collection where it was assigned the ATTC number 29424. Subsequently D. Ballou obtained a stock of strain PHK from the Ribbons Lab and then maintained it at the University of Michigan as strain DB01. Another stock of strain PHK was maintained by R. Eaton at the U.S. EPA Environmental Research Lab in Gulf Breeze, FL with its original designation. In the course of our survey we discovered that the PHK variant maintained at Gulf Breeze, which we designated PHK-GB, two replicons of 2.4 and 3.4 Mb had fused to form a

stable 5.8-Mb replicon. Both the ATCC stock and strain DB01 contained three replicons of 1.1, 2.4, and 3.4 Mb. Strain PHK-GB contained only the 1.4 and 5.8-Mb replicons. *Xba*I digests of DNA from the different strains were similar but not identical. The results indicate that the 5.8-Mb replicon of strain PHK-GB was derived from the two smaller replicons. Subsequently we have identified a similar case of replicon fusion involving the 2.5 and 3.4-Mb replicons of strain 249.

Isolation and characterization of the *dsd* gene of *B. cepacia* 249. D-serine is ordinarily toxic for *B. cepacia* 249 (ATCC 17616). This strain not only fails to utilize D-serine; its growth on alternative substrates such as D-mannitol is also inhibited in the presence of D-serine. However D-serine resistant variants which can utilize this amino acid as sole carbon and energy source inhibition by D-serine are isolated readily. For example when ca 2×10^8 bacteria were plated to medium containing 0.2% D-serine, ca 50 colonies appeared on the plates over a several day incubation at 37° C. The D-serine utilizing mutants had acquired to ability to form high levels of a D-hydroxyamino acid deaminase not detected in extracts of the wild type.

In an attempt to isolate the activated *dsd* gene, bacteria from a plate containing ca 100 Dsd positive colonies were collected, and DNA was isolated and subjected to limited digestion with *Sau*3a. Fragments of ca 5-kb in size were ligated with *Bam*HI-digested pACYC184, and the recombinant plasmids were transformed into the DsdA negative *E. coli* mutant AC6082. D-serine resistant mutants were selected on plates containing 0.2% each of D-mannitol and D-serine, and then tested for ability to utilize D-serine as sole carbon and energy source. One such D-serine utilizer, containing a plasmid insert of ca 6 kb, was chosen for subsequent analysis of its DNA. This included subcloning experiments to identify the region of the cloned DNA conferring D-serine resistance and ability to utilize D-serine, as well as subsequent DNA sequence analysis. The *dsd* gene was identified as an open reading frame of 1095 starting ca 700 nt downstream from one end of the cloned fragment. SDS/PAGE analysis of bacterial extracts indicated that derivatives of strain AC6082 containing the *B. cepacia dsd* gene contained a new peptide of ca 45 kD, consistent with the size to the predicted gene product. A similar size gene product was detected in extracts of Dsd positive derivatives of strain 249, but not in the parent strain. The nucleotide sequence of the cloned *B. cepacia dsd* gene did not resemble that of the *dsd* gene of *E. coli*. It also failed to exhibit a significant degree of similarity with other sequences deposited in GenBank/EMBL. Our results suggest that the *dsd* gene we have identified might be a suitable probe for phylogenetic identification of *B. cepacia* isolates.

Mechanism of activation of the *B. cepacia dsd* gene. A working hypothesis, based on our previous studies of the activation of foreign genes introduced into *B. cepacia* on broad-host-range plasmids, was that expression of the cryptic *dsd* gene might depend upon insertional activation by IS elements. To test this possibility we used two approaches. 1) We used the isolated *dsd* gene as a probe in Southern

hybridization experiments to compare homologous DNA sequences in wild type and Dsd positive variants of strain 17616 and 2) we used PCR technology to amplify sequences upstream of the *dsd* gene in the wild type and Dsd positive variants. The results of the Southern hybridization experiments indicated that there were no major rearrangements within a ca 700-bp region immediately upstream of the *dsd* gene. Similar results were obtained using PCR primers that would amplify a ca 700 nt region including 500-nt immediately upstream of the *dsd* gene. A PCR product of this size was detected with template DNA from both the wild type and mutants. The simplest interpretation of the results is that some mechanism other than insertional activation is responsible for the majority of cases of *dsd* gene activation. We are in the process of using PCR technology to amplify DNA upstream of the *dsd* gene in various Dsd positive mutants, and will attempt to define the mutations responsible for *dsd* gene activation.

Transposition of IS3 family elements. One of the reasons that led us to explore the possible involvement of IS elements in *dsd* gene activation was the discovery that two insertion sequences we had isolated in *B. cepacia*, IS401 and IS407, were members of the IS3 family (Wood *et al.*, 1992; Byrne and Lessie, 1994). The transposases of IS3 family elements appear to be transframe proteins formed by a translational frameshift event (Chandler and Fayet, 1993). Translational frameshifting appears to depend upon ribosome pausing at "slippery" sequences, where ribosomes exhibit significant potential to shift to a new reading frame. Such an event must shift the stop codon of the upstream ORF out-of-phase and the downstream ORF in phase. One factor likely to influence the extent of translational frameshifting is the abundance of charged t-RNAs recognizing codons at the 3' end of the slippery sequence. The putative slippery sequences in IS401 and IS407 are AAAG and TTTT.TG, respectively. In the case of IS401, our working hypothesis is that formation of the IS401 transframe protein involves a rightward shift near the end of the first ORF, that brings into the correct phase the TTG codon (specifying trp) in the sequence TTTT.TG. We expect that this type of frameshift is enhanced in the presence of trp. In the case of IS407, we believe that the appropriate shift is leftward shift that places the AAG codon (specifying lys) in the sequence AAAG out of phase. Here we would expect the shift to be favored under conditions of lysine deprivation.

In the context of *dsd* gene activation, it should be noted that when we tested the influence of various amino acids on adaptation of strain 17616 to utilize D-serine, we found that lysine decreased the frequency of Dsd positive mutants about 10-fold. In contrast, tryptophan increased the frequency of Dsd positive mutants nearly 50-fold. No other amino acids had such a dramatic effect on the incidence of Dsd positive variants. These findings would have been consistent with the possibility that lys and trp decreased or increased the extent of *dsd* gene activation by influencing extent of insertional activation by elements such as IS401 and IS407. However our failure to demonstrate insertional activation in Dsd positive variants isolated in absence of lysine suggests that the influence of this amino acid on *dsd* activation is not related to

its effect on transposition.

There are two approaches will be used to test further whether translational frameshifting is important in regulating the extent of transposition of IS401 and IS407. One, which is indirect, will be to examine whether tryptophan increases insertional activation of the *B. cepacia dsd* gene. The other which is more direct is to examine the influence of lysine on transposition of an IS407 variant constructed in our lab by Ning Li. This element, IS407.2S, was constructed by insertion of a trimethoprim resistance gene and a *Swal* site into the unique *DraI* site in IS407. Precautions were taken to avoid interrupting regions of the element important for transposition. We then constructed, pTGL914, a derivative of pME305 carrying IS407.2S. Plasmid pTGL914, like its parent pME305, is temperature sensitive with respect to its replication and carries a non-transposable tetracycline resistance marker (Rella and Haas, 1985). We have introduced pTGL914 into a lysine requiring strain of *B. Cepacia* 17616, and now plan to select T^pR Tc^S transposants in which IS407.2S has transposed to various chromosomal sites. The experiments will be carried out using populations that have been limited for lysine or propagated that have been grown under conditions of lysine limitation or lysine excess, and the extent of element transposition under the two conditions will be compared. The aim of the ongoing experiments is to understand how IS elements contribute to the adaptability of *B. cepacia*, and whether various conditions of environmental and physiological stress might influence gene activation and genomic plasticity by triggering element transposition.

Literature Cited:

- Byrne, A.M. and T.G. Lessie 1994. Characteristics of IS401, a new member of the IS3 family implicated in plasmid rearrangements in *Pseudomonas cepacia*. Plasmid 31: 138-147.
- Cantor C.R., C.L. Smith, and M.K. Mathew. 1988. Pulsed-field gel electrophoresis of very large DNA molecules. Ann. Rev. Biophys. Biophys. Chem. 17:287-304.
- Chandler, M. And O. Fayet. 1993. Translational frameshifting in the control of transposition in bacteria. Mol. Microbiol. 7: 497-503.
- Cheng, H.-P. and T.G. Lessie. 1994. Multiple replicons constituting the genome of *Pseudomonas cepacia* 17616. J. Bacteriol. 176:4034-4042.
- Lieu, S.L. and K.E. Sanderson 1995. I-Ceul reveals conservation of the genome of independent strains of *Salmonella typhimurium*. J. Bacteriol. 177:3355-3357.
- Michaux S., J. Paillison. M.J. Carles-Nurit, G. Bourg, A. Allard-Servent, and M. Ramuz. 1993. Presence of two independent chromosomes in the *Brucella melitensis* 16M genome. J. Bacteriol. 175:701-705.
- Palleroni, N.J. 1992. Present situation in the taxonomy of aerobic pseudomonads, p. 105-115. In E. Galli, S. Silver, and B. Witholt (ed.), *Pseudomonas: molecular biology and biotechnology*. American Society for Microbiology, Washington, DC.

- Rodley P.D, U. Romling, and B. Tummier . 1995. A physical genome map of the *Burkholderia cepacia* type strain. *Mol. Microbiol.* 17:57-67.
- Wood, M.S., A.M. Byrne, and T.G. Lessie. 1991. IS406 and IS407, two gene-activating insertion sequences from *Pseudomonas cepacia*. *Gene* **105**: 105-105.
- Yakabuuchi E, Y. Kosako, H. Oyaizu, I. Yano, H. Hotta, Y. Hashimoto, T. Ezaki, and M. Arakawa. 1992. Proposal of *Burkholderia* gen. nov. and transfer of seven species of the genus *Pseudomonas* homology group II to the new genus, with the type species *Burkholderia cepacia* (Palleroni and Holmes 1981) comb. nov. *Microbiol. Immunol.* 36:1251-1275.

TABLE 1. Genomic complexity of selected *B. cepacia* isolates.

Strain	Large Replicons (Mb)	Genome Size ^a (Mb)
NTC10661	3.4, 3.4, 1.6	8.5
G4 ^b	3.4, 2.4, 1.2, 1.1	8.1
ATCC 25416	3.6, 3.4, 1.0	8.0
ATCC 17759 (382)	3.5, 3.3, 1.1	7.9
ATCC 29424 (PHK/DB01) ^c	3.4, 2.4, 1.3	7.1
PHK-GB ^c	5.8, 1.3	7.1
CRE-7	3.4, 3.2, 1.3	7.0
ATCC 17616 (249)	3.4, 2.5, 0.9	6.8
542	3.4, 3.0	6.4
ATCC 53867 ^d (AC1100)	3.7, 2.6	6.3
67-46	3.6, 1.8	5.4
ATCC 17760 (383)	3.4, 1.3	4.7
SW3	3.5, 1.1	4.6

^a The estimates of genome size do not include replicons smaller than 0.9-Mb. For example, the 170-kb cryptic plasmid of strain 17616 (Gaffney and Lessie, 1995) and the 200-kb plasmid of strain 25416 (Rodley *et al.*, 1995) are not included.

^b Unrestricted DNA from strain G4 also contained DNA species corresponding to replicons of ca 0.4, 0.3, 0.2, and 0.1 Mb.

^c Strains DB01 and PHK-GB represent different stocks of the phthalate-utilizing strain PHK isolated by P. Kaiser and deposited in the American culture collection as ATCC 29424.

^d Unrestricted DNA from strain AC1100 also contained DNA species corresponding to replicons of ca 0.5 and 0.3 Mb.

Table 2. Fragments^a generated by treatment of genomic DNA with *CeuI*^b

<i>E. coli</i> / <i>S. typhimurium</i>		<i>Burkholderia cepacia</i>					
		17616 ^c	25416	G4	CRE-7	AC1100	29424
2460	2460	2500	3000	2400	2300	2200	2400
700	775	1900	2200	2100	1900	1900	2200
670	740	900	1200	1400	1400	1700	1400
530	550	900	1100	1200	1200	500	1100
130	145	600	300	350	350	300	300
92	92		250	250	250	200	250
44	44						

^a Size estimates are in kilobase pairs.

^b *CeuI* recognizes a 26-bp sequence present in genes encoding 23S RNA (Lieu and Sanderson, 1994).

^c The 3.4-Mb replicon of strain 17616 contained three *CeuI* fragments of 1.9, 0.9, and 0.6 Mb. The 2.5 and 0.9-Mb replicons each contained a single *CeuI* site. The total number of *CeuI* fragments was equivalent to the number of r-RNA operons estimated from Southern hybridization experiments in which DNA fragments obtained by digestion of genomic DNA with *SwaI* and *PacI* were probed with radiolabelled 16S or 23S r-RNA (Cheng and Lessie, 1994). Table 2. Fragments^a generated by treatment of genomic DNA with *CeuI*^b

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

Student training:

The project has resulted in the training of two Ph.D. Candidates, Dr. Hai-Ping Cheng and Ms. Stacy Montgomery. Dr. Cheng was awarded a Ph.D. Degree by the University of Massachusetts in 1994. Ms. Montgomery expects to complete her Ph.D. Degree requirements in the academic year 1995/96. Two other students, Brendan Manning, who is currently a Ph.D. Candidate in the Biology Program at Yale University and Ning Li, who is currently a graduate student in our program., were also involved in aspects of the project.

Publications:

- Byrne, A.M. And T.G. Lessie. 1994. Characteristics of IS401, a new member of the IS3 family implicated in plasmid rearrangements in *Pseudomonas cepacia*. Plasmid **31**: 138-147.
- Cheng, H.-P. And T.G. Lessie. 1994. Multiple replicons constituting the genome of *Pseudomonas cepacia* 17616. J. Bacteriol. **176**: 4034-4042.

Abstracts/Symposium Reports:

- Cheng, H.-P. and T.G. Lessie. 1992. Macrorestriction mapping of the *Pseudomonas cepacia* 17616 genome. Abstr. Annu. Meet. Amer. Soc. Microbiol. H178, p. 212.
- Cheng, H.-P. and T.G. Lessie. 1994. Characterization of the multiple replicons comprising the genome of *Pseudomonas cepacia*. Abstr. Annu. Meet. Amer. Soc. Microbiol. H188, p. 233.
- Montgomery, S.O. And T.G. Lessie. 1995. Activation of a cryptic D-serine deaminase (*dsd*) gene from *Pseudomonas cepacia* 17616. Abstr. Annu. Meet. Amer. Soc. Microbiol. H160, p. 520.
- Lessie, T.G. And B.D. Manning. 1995. Chromosomal multiplicity in *Burkholderia cepacia*. Proceed. EPA Symp. Environmental Releases of Biotechnology Products. Clarion Res. Center, Pensacola, FL.

In preparation:

- Montgomery, S.O. And T.G. Lessie. 1995. Identification and characterization of a cryptic D-serine deaminase (*dsd*) gene from *Burkholderia cepacia* 17616. In preparation for submission to J. Bacteriol.
- Lessie, T.G. And B.D. Manning. 1995. Multiple large replicons in environmental isolates of *Burkholderia cepacia*. In preparation for submission to Appl. Environ. Microbiol.
- Lessie, T.G. 1996. Genomic plasticity and catabolic potential of *Burkholderia cepacia*. Invited Mini-review to appear in FEMS Lett. Microbiol.

Reprints and preprint removed to
for separate cycling

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.