RECEIVED

SEP 03 1998

OSTI

DOE/ER/61679--T1

Cloning and Sequencing of the *trpE* Gene from *Arthrobacter globiformis* ATCC
8010 and Several Related Subsurface *Arthrobacter* Isolates

Tatiana Chernova¹, V.K. Viswanathan², Natasha Austria and Brian P. Nichols*

Laboratory for Molecular Biology

Department of Biological Sciences MC067

University of Illinois at Chicago

900 S. Ashland Ave.

Chicago, IL 60607

Phone: (312) 996-5064

FAX: (312) 413-2691

E-mail: brian.p.nichols@uic.edu

*Corresponding Author

¹Present address: Winship Cancer Center, Emory University, GA 30322

²Present address: Department of Microbiology-Immunology, Northwestern
University, Chicago, IL 60611

Author Box:

This work was undertaken as a part of the Deep Microbiology subprogram of the
US Department of Energy Subsurface Science Program. Drs. Chernova and
Viswanathan contributed equally to this work.

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

MASTER

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.

DISCLAIMER

**Portions of this document may be illegible
electronic image products. Images are
produced from the best available original
document.**

ABSTRACT

Tryptophan dependent mutants of *Arthrobacter globiformis* ATCC 8010 were isolated and *trp* genes were cloned by complementation and marker rescue of the auxotrophic strains. Rescue studies and preliminary sequence analysis reveal that at least the genes *trpE*, *trpC*, and *trpB* are clustered together in this organism. In addition, sequence analysis of the entire *trpE* gene, which encodes component I of anthranilate synthase, is described. Segments of the *trpE* gene from 17 subsurface isolates of *Arthrobacter* sp. were amplified by PCR and sequenced. The partial *trpE* sequences from the various strains were aligned and subjected to phylogenetic analysis. The data suggest that in addition to single base changes, recombination and genetic exchange play a major role in the evolution of the *Arthrobacter* genome.

INTRODUCTION

Arthrobacter globiformis is a gram-positive aerobic bacterium belonging to the coryneform-actinomycete family of the high G+C subdivision of bacteria. *Arthrobacter* spp. are the most numerous bacteria in many soils (Paul & Clark 1989), and because it has effective survival capabilities, it is one of the most numerous among bacterial types recovered from deep subsurface samples, being found at depths up to several hundred meters (Haldeman et al. 1993; Stim 1994).

In addition to their ubiquitous presence in soil samples, *Arthrobacter* sp. have been shown to contain degradative pathways capable of catabolizing a wide variety of substituted aromatic compounds, including halogenated aromatic compounds, halogenated alkanes and aromatic hydrocarbons (Schenk et al. 1989; Scholtz et al. 1987; Fredrickson et al. 1991). A better understanding of the various *Arthrobacter* sp. occurring in the natural environment, their relationship to one another, and evolutionary processes contributing to their heterogeneity should also aid in more efficient methods of bioremediation.

The subsurface bacterial isolates used in this study were derived from the Yakima Barricade borehole site at Hanford, Washington described by Kieft et al. (1995). Strains related to *Arthrobacter* were isolated and characterized (Crocker, Reeves & Balkwill 1996) and distributed by the Subsurface Microbial Culture Collection, Gainesville FL. A subset of the *Arthrobacter* strains have been characterized by analysis of 16S rRNA sequences (Crocker, Reeves & Balkwill 1996), *recA* sequences (van Waasbergen El-Masri & Miller 1996), and as we report here, *trpE* sequences.

The relatively slow fixation of nucleotide differences in 16S rRNA makes this sequence of unsurpassed utility for constructing bacterial phylogenies and identifying and typing new or unknown bacterial samples (Olsen Woese & Overbeek 1994). In addition, the study of intraspecies structural and metabolic gene sequences has led to the characterization of additional

evolutionary processes such as recombinational events, that are active within a species (reviewed in Milkman 1996). Similarly, horizontal gene transfer between very divergent bacterial groups, including the high G+C coryneform bacteria, have been suggested by analysis of genes encoding metabolic enzymes such as *trp* (Crawford 1989).

The tryptophan biosynthetic pathway is comprised of five reactions starting from chorismate, and the pathway appears to be biochemically invariant in all organisms studied to date, although the organization and regulation of the *trp* biosynthetic genes varies considerably among different organisms (Nichols 1996; Crawford 1975; Crawford 1989). On the basis of these observations, it has been suggested that the tryptophan biosynthetic pathway evolved only once, and the myriad of contemporary variations in the gene organization, gene fusions, regulation, and gene sequences reflects divergence from a single ancestral origin (Crawford 1989).

Here we describe the isolation of various *trp* genes by complementation of auxotrophic strains of *Arthrobacter globiformis* ATCC 8010, and sequence analysis of *trpE*. We also describe the amplification of segments of *trpE* from subsurface isolates of *Arthrobacter* sp. using PCR primers designed on the *A. globiformis* ATCC 8010 sequence. Analysis of the *trpE* segments and comparison with 16S rRNA tree topology suggests that intra- and inter-genic recombinational exchange has played a large role in generating diversity in subsurface *Arthrobacter* species.

MATERIALS AND METHODS

Strains, plasmids and media. *A. globiformis* ATCC 8010 from the American Type Culture Collection (ATCC) was the parent for construction of Trp⁻ mutants and the source of DNA for library construction. Subsurface *Arthrobacter*-related isolates from the Yakima Barricade borehole at the Hanford site were obtained from the Subsurface Microbial Culture Collection (SMCC) and provided by D. Balkwill. *E. coli* DH5 α (*supE44 hsdR17 recA1 endA1 gyrA96 thi-1 relA1*)

(Sambrook Fritsch & Maniatis 1989) was used for plasmid amplification and purification. A coryneform-*E. coli* shuttle vector p21350, based on the *Brevibacterium lactofermentum* pBL1 minimal replicon (Miwa et al., 1985; Trautwetter & Blanco, 1991) and containing tetracycline and kanamycin resistance determinants was obtained from C. Blanco. Bluescript SKII(+) (Stratagene) was used for DNA subcloning and plasmid preparation from *E. coli* hosts.

LB (Miller 1972) was used as the rich growth medium. M9 minimal medium (Sambrook Fritsch & Maniatis 1989) was supplemented where appropriate with tryptophan (40 µg/ml), anthranilate (20 µg/ml), indole (20 µg/ml), thiamin (5 µg/ml), kanamycin (100 µg/ml) and tetracycline (12.5 µg/ml). *Arthrobacter* strains were grown at 30°C, while *E. coli* strains were grown at 37°C.

Isolation of *A. globiformis* tryptophan auxotrophs. Auxotrophic derivatives of *A. globiformis* ATCC 8010 were obtained by N-methyl-N'-nitro-N-nitrosoguanidine (NTG) mutagenesis. Cells were grown at 30°C in 5 ml LB to exponential phase, harvested by centrifugation, washed once with 0.1 M sodium phosphate pH 7.0, resuspended in 2 ml of the same buffer, and mutagenized for 30 min at 30°C without shaking in the presence of 50 µg/ml NTG. Mutagenized cells were harvested by centrifugation, washed in phosphate buffer, resuspended in 2 ml of LB medium and diluted 1/100 in 10 ml aliquots of fresh LB. Following growth to stationary phase, cells were diluted and plated on M9 medium supplemented with tryptophan. After 3-4 days of incubation colonies were replicated onto plates of both M9 and M9 supplemented with tryptophan. Colonies which grew on supplemented but not on unsupplemented media were picked as tryptophan auxotrophs.

Plasmid DNA isolation and transformation. Plasmid DNA was isolated from mid-logarithmic cultures of *A. globiformis* by alkaline lysis (Birnboim & Doly 1979) following freezing of the harvested cell pellet.

E. coli transformation was performed as described by Inoue Nojima &

Okayama (1990). *A. globiformis* was electrotransformed using a Gene Pulser equipped with the Pulse Controller attachment (Bio-Rad). A modification of the recommended protocol (Bio-Rad) was used. Mid-logarithmic cultures of *A. globiformis* were harvested by centrifugation and washed 3 times with one volume, 0.5 volume and 0.02 volume of ice-cold 10% glycerol. Cells were resuspended in 0.01 volume of 10% glycerol and either used directly for electroporation or stored at -70°C for later use. Concentrated cells (80 µl) were mixed with DNA, immediately transferred into a pre-cooled 2 mm cuvette and pulsed at 2.5 kV and 400 Ω. The electroporated cells were immediately diluted with 1 ml of SOC medium (Hanahan 1983) and incubated with shaking at 30°C for 2 hours before plating on LB + kanamycin plates for selection of transformants.

Isolation of chromosomal DNA. *Arthrobacter* chromosomal DNA was prepared using a modification of the protocol developed for *Bacillus subtilis* (Harwood & Cutting, 1990). Late-log phase cultures of *A. globiformis* were harvested from 100 ml LB, washed in 10 ml of lysis buffer (50 mM EDTA pH 8.0, 0.1 M NaCl) and frozen at -20°C for at least 1 hour. The pellet was resuspended in 4 ml of lysis buffer with the subsequent addition of 10 mg/ml of lysozyme, and incubated at 37°C for 20 min with gentle agitation. Cells were lysed by addition of 20% N-laurylsarcosine for 10 min at 37°C, followed by 10 min incubation at 65°C. Chromosomal DNA was purified by phenol-chloroform extraction and precipitated with ice cold ethanol.

Construction of genomic library of *A. globiformis*. *A. globiformis* ATCC 8010 chromosomal DNA was partially digested with *Sau3AI*, and DNA fragments of 4 to 15 kb were isolated from a 0.8% agarose gel, following the method of Zhen & Swank (1993). DNA fragments were ligated with *Bam*HI-digested and dephosphorylated p21530 leading to inactivation of the Tc^r gene by insertion. *E. coli* DH5α was transformed with the ligation mixture and plasmid DNA was isolated from 40 Km^r transformants. Restriction analysis revealed that 85% of

the transformants contained plasmids with an insert of *A. globiformis* DNA of average size of 7.6 kb. Transformants were pooled and plasmid DNA was isolated. Seven pools contained about 1.4×10^3 individual Km^r Tc^s transformants. This DNA library was used for the selection of *trpE* plasmids as described below.

PCR amplification of DNA from subsurface *Arthrobacter* isolates. DNA primers (Integrated DNA Technologies, Coralville, Iowa) were used to amplify 10 ng samples of *Arthrobacter* DNA according to the protocol of Bej Mahbubani & Atlas (1991). The primers AGCon3A and AGCon5 (Figure 2) were designed on the basis of the *A. globiformis* ATCC 8010 sequence, and used to amplify a 304 bp fragment from subsurface *Arthrobacter* isolates. The amplified DNA fragments were sequenced, and an invariant sequence in this region was used to design the primer AGCon6 (Figure 2). AGCon6 was used in conjunction with AG6 (Figure 2) to amplify a ~1070 bp DNA segment. For strains that did not yield a product with AG6-AGCon6, a downstream primer, AGV1, was used in conjunction with AGCon6 to amplify a relatively shorter region.

DNA sequence analysis. Specific restriction fragments were cloned in Bluescript SKII(+) (Stratagene). Short-track DNA sequencing was done by the dideoxy chain termination method of Sanger Nicklen & Coulson (1977), using Sequenase version 2.0 DNA sequencing kits (United states Biochemical, Cleveland, Ohio). Large-scale sequencing was done by the Genetic Engineering Facility of University of Illinois at Urbana-Champaign. Sequencing of PCR products was performed using the Sequenase PCR product sequencing kit (United States Biochemical, Cleveland, Ohio). All sequences have been deposited with GenBank. *A. globiformis* ATCC 8010 *trpE* has been assigned accession number U89690.

Phylogenetic analysis. Construction of phylogenies was accomplished using MacClade version 3 (Maddison & Maddison, 1992) and PAUP 3.0s (Swofford 1991). Parsimonius trees were checked by bootstrap analysis.

RESULTS

Identification of *trp* mutants. Thirty seven tryptophan auxotrophic strains were obtained by NTG treatment of *A. globiformis* strain ATCC 8010. On the basis of growth on M9 medium, and on M9 supplemented with anthranilate, indole, or tryptophan (Table 1), three different groups of *trp*⁻ mutants were identified. The first group, represented by strain M6 and 18 other 19 *trp*⁻ isolates, grew on medium containing anthranilate, characteristic of mutants lacking anthranilate synthase activity, encoded by *trpE* and *trpG*. The second group, represented by strain M1 and 5 other *trp*⁻ mutants, failed to grow on medium containing indole, characteristic of mutants deficient in tryptophan synthase β subunit activity, encoded by *trpB*. The third group, represented by M22 and 14 other mutants, grew on indole supplemented medium but not anthranilate supplemented medium, and were apparently deficient in one of the other enzymes of tryptophan biosynthesis encoded by *trpD*, *trpF*, *trpC* or *trpA*.

Isolation of the *trp* gene cluster of *A. globiformis*. *A. globiformis* strain M6, deficient in anthranilate synthase activity, was transformed with *A. globiformis* sub-genomic library pools, constructed as described in Materials and Methods. Kn^r transformants were selected and screened for complementation of the tryptophan auxotrophy on M9 medium. Two Kn^r, Trp-independent transformants were isolated, and shown to harbor identical plasmids by restriction analysis. This plasmid was designated pTC1, and contained a 7kb *A. globiformis* DNA insert (Figure 1). pTC1 also complemented the auxotrophy of M1 and M22, suggesting that it contained *trpE* or *trpG*, *trpB*, and at least one other *trp* gene.

Individual *trp* genes were localized by subcloning fragments from pTC1. pTC2 and pTC4 were able to complement the anthranilate synthase deficient strain M6, but not any other auxotrophs tested, suggesting that *trpE* or *trpG* lay on the 1.6 kb *EcoRI*-*Bam*HI fragment of pTC4. pTC3 complemented strain M1,

suggesting that *trpB* was located within another 1.6 kb *EcoRI*-*Bam*HI fragment of pTC1. Limited DNA sequence analysis of the subclones confirmed the locations of *trpE* and *trpB*, and further identified and localized *trpC*, as indicated in Figure 1. It became apparent by sequence analysis that complete genes (e.g. *trpB*) were not present on the "complementing" plasmids. Therefore, growth must have been due to recombinational marker rescue, rather than complementation. Although there are gaps of sufficient size that other genes may be present, it is clear that at least *trpE*, *trpC* and *trpB* are present in this gene cluster.

Features of the DNA sequence of *A. globiformis trpE*. *A. globiformis* chromosomal DNA from the plasmid pTC4 and pTC5 was subcloned into the vector Bluescript SK(+) to obtain pTC6 and pTC7, respectively. Determination of a 2017 nucleotide region of pTC6 and a portion of pTC7 yielded the complete nucleotide sequence of *trpE*, and 400 bp of upstream flanking region (Figure 2). The deduced open reading frame encoded a 531 amino acid protein with strong similarity to anthranilate synthase component I of other organisms. The overall G+C content of the coding region was 67.2%, while 86.7% of the codons ended in G or C. Thirty bp upstream of the first nucleotide of the Met codon, appears a 91 bp sequence containing 6 tandemly arranged 15 bp repeats. Although the function of these sequences has not been established, it seems likely that they may play a role in the regulation of *trpE* expression.

***trpE* sequences of subsurface isolates related to *A. globiformis*.** PCR primers were designed on the basis of the *A. globiformis* ATCC 8010 *trpE* sequence to amplify segments of *trpE* from subsurface isolates of *Arthrobacter* sp. One set of primers (Agcon5 and Agcon3) were designed to amplify a 304 bp region from the 3'-terminal segment of *trpE*, and additional primers (Agcon6, AG6, and AGV1) were designed to amplify the 5'-segment.

In all, partial *trpE* sequences were determined from 17 subsurface *Arthrobacter* strains (Table 2). Several of the sequences were identical, or

nearly so (< 1.5% divergent), and further analysis was carried out using only a single representative of each identity group. The sequences are indicated in Table 1 as belonging to group 1, 2, 3 or as unique (U). Interestingly, the 5'-terminal sequence of ZAT012 belongs to identity group 2, while the 3'-terminal sequence belongs to identity group 1. The remainder belonged to a single group regardless of the portion of sequence compared, or were unique.

A total of 844 nucleotide positions could be compared among 2 or more strains. As expected, the large majority (78%) of the nucleotide substitutions were synonymous with respect to amino acid coding. Thirty seven of 281 amino acid positions (13%) suffered replacements. Slightly more than half (54%) of codons without amino acid replacement substitutions contained synonymous nucleotide substitutions. The patterns of nucleotide and amino acid sequence divergence were not significantly different in the 5'- and 3'-terminal sequences.

Relationships among deep subsurface *Arthrobacter* strains. The DNA sequences from subsurface isolates were co-linear with the *A. globiformis* ATCC 8010 *trpE* sequence, requiring no gaps for alignment. The sequences were used to investigate the phylogenetic relationships among the type and subsurface *Arthrobacter* strains, and the results are shown in Figure 3. The identity groups and phylogenetic relationships were independent of the depth at which the samples were taken, and only partially correspondent with the relationships determined by 16S rRNA sequence analysis. Two pairs of strains (G966 and G979; G959 and G982) which were identical in 16S rRNA sequences were also identical in *trpE* sequences. However, important discrepancies existed in examining other pairwise relationships. For example, ZAT031 16S rRNA had been found to be most similar to *Micrococcus agilis* 16S rRNA, but we found the *trpE* sequence to be identical to the *A. globiformis* ATCC 8010 type strain. At the 16S rRNA level, strain G986 had been determined to have the closest relationship to the *A. globiformis* ATCC8010 type strain, but we found that

these two *trpE* sequences differ by 14% indicating a much larger divergence in this portion of the genome.

DISCUSSION

The *A. globiformis trpE* sequence reported here is the second from the high G+C gram positive grouping of eubacteria, and as such would be expected to be most similar in amino acid sequence to that reported for *Corynebacterium glutamicum* (Matsui Sano & Ohtsubo 1986; Heery & Dunican 1990). This has not proved to be the case however, as analysis has suggested a surprisingly close relationship between *trp* genes of *C. glutamicum* and the enteric bacteria (Crawford 1989). This relationship crosses the systematic boundaries between two major eubacterial groups, well established according to 16S rRNA-based systematics and other approaches, and the most probable explanation is that the *trp* genes of these organisms have had a much more recent common ancestry than their 16S rRNA genes, presumably the result of horizontal transfer (Crawford 1989). If this is the case, then the *A. globiformis trpE* sequence is more likely to represent the high G+C gram positive group of organisms.

When compared across a large number of widely divergent species, the amino acid sequences encoded by *trpE* show a biphasic pattern of divergence. Sequences in the N-terminal half of the protein are highly divergent, and difficult to align, requiring the insertion of several, and sometimes somewhat arbitrary, gaps. The C-terminal portions, however, are more conserved in sequence, and require the introduction of few gaps for alignment (Crawford, 1989; Nichols 1996). Of 31 invariant residues in known TrpE amino acid sequences, only one occurs in the N-terminal portion. Nevertheless, the numbers and patterns of both nucleotide and amino acid substitutions in the closely related *Arthrobacter* sp. are quite similar in both portions of the sequence, with the exception of the apparent ZAT012 recombinant. This trend is also true for comparisons between other closely related genera such as *E. coli*-*Salmonella* (Nichols van Cleemput & Yanofsky 1981; Yanofsky & van Cleemput

1982), and the Pseudomonads (Essar Eberly & Crawford 1990). The extreme variability observed at the TrpE N-terminus may reflect either adaptation to a yet unknown species-specific function, or may represent the continued accumulation of replacements over time in a non-essential region of the protein.

The investigation of *trpE* sequences from *Arthrobacter* isolates from subsurface core samples yielded a collection of nucleotide sequences differing from one another by up to 22%, and the collection contained members that were both identical to and differed from the sequence of the contemporary surface *A. globiformis* ATCC 8010 type strain. Several of the *trpE* alleles were widely dispersed in *Arthrobacter*-like isolates, independent of both depth of isolation and 16S rRNA phylogeny. The lack of correspondence in relationships determined by 16S rRNA and *trpE* sequence phylogenies may have been due to genetic exchange of sequences under environmental conditions. These data were consistent with the evolution of bacterial genomes by clonal segments and clonal frames, driven by recombination, as has been described in detail for *E. coli* (Milkman & Stoltzfus 1988; Milkman & McKane-Bridges 1990; Milkman 1996). The active recombination system in *A. globiformis* has been observed in the laboratory (El-Masri et al. 1995) and was evident in this work in our marker rescue experiments designed to test the presence of various *trp* genes on several plasmid clones. Similarly, the mosaic structure of the *trpE* sequence from subsurface strain ZAT012 also suggests a relatively recent recombination event between two *trpE* alleles. Given these considerations, it would not be surprising to find differences in phylogenies for any DNA sequence segment studied, each representing the ancestry of a different clonal segment of DNA, rather than the organism as a whole. This phenomenon is likely to confound attempts at tracing lineages of bacterial species throughout the subsurface environment.

ACKNOWLEDGMENTS

We would like to thank the SMCC for providing strains and D. Balkwill and R. Miller for communicating results prior to publication. We thank Mary Ashley for many helpful discussions. This work was supported by the US Department of Energy grant DEFG0293ER61679.

REFERENCES

- Bej AK, Mahbubani MH, Atlas RM (1991) Amplification of nucleic acid by polymerase chain reaction (PCR) and other methods and their application. *Critical Reviews in Biochemistry and Molecular Biology*, 26, 301-334.
- Birnboim HC, Doly J (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Research*, 7, 1513-1523.
- Crawford IP (1975) Gene rearrangements in the evolution of the tryptophan synthetic pathway. *Bacteriological Reviews*, 39, 87-120.
- Crawford IP (1989) Evolution of a biosynthetic pathway: the tryptophan paradigm. *Annual Review of Microbiology*, 43, 567-600.
- Crocker FH, Reeves RH, Balkwill DL (1996) Phylogenetic and physiological diversity of subsurface *Arthrobacter* strains. *Abstracts of the General Meeting of the American Society for Microbiology*, 96, 474.
- El-Masri A, Walls SD, Van Waasbergen LG, Brown JT, Miller RV (1995) Analysis of *recA* expression and evolution in deep-subsurface *Arthrobacter* species. *Abstracts of the General Meeting of the American Society for Microbiology*, 95, 443.
- Essar DW, Eberly L, Crawford IP (1990) Evolutionary differences in chromosomal locations of four early genes of the tryptophan pathway in fluorescent pseudomonads: DNA sequences and characterization of *Pseudomonas putida* *trpE* and *trpGDC*. *Journal of Bacteriology* 172, 867-883.

Fredrickson JK, Brockman FJ, Workman DJ, Li SW, and Stevens TO (1991) Isolation and characterization of a subsurface bacterium capable of growth on toluene, naphthalene, and other aromatic compounds. *Applied and Environmental Microbiology*, 57, 763-803.

Haldeman DL, Amy PS, Ringleberg D, White DC (1993) Characterization of the microbiology within a 21 m³ section of rock from the deep subsurface. *Microbial Ecology*, 26, 145-159.

Hanahan D (1983) Studies on transformation of *Escherichia coli* with plasmids. *Journal of Molecular Biology*, 166, 557-580.

Harwood CR, Cutting SM (1990) *Molecular Biological Methods for Bacillus*, John Wiley & Sons, Chichester, England.

Heery DM, Dunican LK (1990) Nucleotide sequence of the *Corynebacterium glutamicum trpE* gene. *Nucleic Acids Research*, 18, 7138.

Inoue H, Nojima H, Okayama H (1990) High efficiency transformation of *Escherichia coli* with plasmids. *Gene*, 96, 23-28.

Kieft TL, Fredrickson JK, McKinley JP, Bjornstad BN, Rawson SA, Phelps TJ, Brockman FJ, Pfiffner SM (1995) Microbiological comparisons within and across contiguous lacustrine, paleosol, and fluvial subsurface sediments. *Applied and Environmental Microbiology*, 61, 749-757.

Maddison WP, Maddison DR (1992) *MacClade: Analysis of Phylogeny and Character Evolution*, Sinauer Associates, Sunderland, Massachusetts.

Matsui K, Sano K, Ohtsubo E (1986) Complete nucleotide and deduced amino acid sequences of the *Brevibacterium lactofermentum* tryptophan operon. *Nucleic Acids Research*, 14, 10113-10114.

Milkman R (1996) Recombinational exchange among clonal populations. In: *Escherichia coli and Salmonella* (ed. Neidhardt FC) pp 2663-2684, ASM Press, Washington DC.

Milkman R, McKane Bridges M (1990) Molecular evolution of the *E. coli* chromosome. II. Clonal segments. *Genetics*, 120, 359-366.

Milkman R, Stoltzfus A (1988) Molecular evolution of the *E. coli* chromosome. III. Clonal frames. *Genetics*, 126, 505-517.

Miller J (1972) *Experiments in Molecular Genetics*, Cold Spring Harbor Laboratory Press, New York.

Miwa K, Matsui K, Terabe M, Ito K, Ishida M, Takagi H, Nakamori S, Sano K, (1985) Construction of novel shuttle vectors and a cosmid vector for the glutamic acid-producing bacteria *Brevibacterium lactofermentum* and *Corynebacterium glutamicum*. *Gene*, 39, 281-286.

Nichols BP (1996) Evolution of genes and enzymes of tryptophan biosynthesis. In: *Escherichia coli and Salmonella* (ed. Neidhardt FC) pp 2638-2648, ASM Press, Washington DC.

Nichols BP, van Cleemput M, Yanofsky C (1981) Nucleotide sequence of *Escherichia coli* *trpE*: anthranilate synthetase component I contains no tryptophan residues. *Journal of Molecular Biology*, 146, 45-54.

Olsen, GJ, Woese CR, Overbeek R, (1994) The winds of (evolutionary) change: breathing new life into microbiology. *Journal of Bacteriology*, 176, 1-6.

Paul EA, Clark FE (1989) *Soil Microbiology and Biochemistry*, Academic Press, San Diego.

Sambrook J, Fritsch EF, and Maniatis T (1989) *Molecular Cloning - a Laboratory Manual*, 2nd edn. Cold Spring Harbor Laboratory Press, New York.

Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the USA*, 74, 5463-5467, 1977)

Schenk T, Müller R, Mörsberger F, Otto MK, Lingens F (1989) Enzymatic dehalogenation of pentachlorophenol by extracts from *Arthrobacter* sp. strain ATCC 33790. *Journal of Bacteriology*, 171, 5487-5491.

Scholtz R, Schmuckle A, Cook AM, Leisinger T (1987) Degradation of eighteen 1-monohaloalkanes by *Arthrobacter* sp. strain HA1. *Journal of General Microbiology*, 133, 267-274.

Stim KP (1994) A phylogenetic analysis of micro-organisms isolated from subsurface environments. *Molecular Ecology*, 3, 1-10.

Swofford DL (1991) *Phylogenetic analysis using parsimony (PAUP) version 3.0s*. Illinois Natural History Survey, Champaign, Illinois.

Trautwetter, A, Blanco, C (1991) Structural organization of the *Corynebacterium glutamicum* plasmid pCG100. *Journal of General Microbiology*,

137, 2093-2101.

Van Waasbergen L, El Masri A, Miller RV (1996) Deducing the phylogenetic relationships of deep-subsurface *Arthrobacter* species through evolution of the *recA* gene. *Abstracts of the General Meeting of the American Society for Microbiology*, 96, 404.

Yanofsky C, van Cleemput M (1982) Nucleotide sequence of *trpE* of *Salmonella typhimurium* and its homology with the corresponding sequence of *Escherichia coli*. *Journal of Molecular Biology*, 155, 235-246.

Zhen L, Swank RT (1993) A simple and high yield method for recovering DNA from agarose gels. *Biotechniques*, 14, 894-898.

Table 1
Growth response of *A. globiformis* ATCC 8010 *trp* mutants to
various supplements

Strain	Supplement				Gene Affected
	none	anthranilate	indole	tryptophan	
M6	-	+	+	+	<i>trpE</i> and/or <i>trpG</i>
M1	-	-	-	+	<i>trpB</i>
M22	-	-	+	+	<i>trpD, F, C</i> or <i>A</i>

- = No growth

+ = Growth

Table 2

Partial *trpE* sequences from subsurface *Arthrobacter* strains

Strain	Depth	Sequence 1	Length	Sequence 2	Length	Group
ATCC 8010	0	491-1005	514	1251-1629	378	1
G915	580	521-666	145	1407-1603	196	2
G919	648	491-727	236	1251-1620	369	U
G950	713	522-668	146	1407-1604	197	2
G954	643	519-727	208			3
G959	713	511-668	157	1407-1604	197	2
G961	567	525-726	201	1329-1606	277	2
G962	578	652-1005	353	1402-1629	227	U
G963	582	548-1005	457	1402-1582	180	U
G966	588	703-1003	300	1295-1500	205	U
G968	591	521-802	281	1402-1587	185	1
G979	591	702-1005	303	1398-1629	231	U
G982	648	527-726	199	1380-1573	193	2
G986	591	520-727	207			3
ZAT002	567	521-672	151	1415-1586	171	2
ZAT012	567	521-676	155	1406-1586	180	2, 1
ZAT031	619	520-756	236	1399-1629	230	1
ZAT104	615	653-1004	351	1294-1464	171	U

FIGURE LEGENDS

Figure 1. Restriction maps and marker rescue patterns of *A. globiformis* ATCC 8010 *trp* clones used in this study. The thick line indicates vector DNA, the thin line *A. globiformis* DNA.

Figure 2. Nucleotide and deduced amino acid sequences of *A. globiformis* ATCC 8010 *trpE*. Underlined sequences represent six repeated sequences upstream of the *trpE* coding segment. Labeled arrows illustrate primers used to amplify DNA from subsurface isolates.

Figure 3. Phylogenetic relationships of *A. globiformis* ATCC 8010 and subsurface *Arthrobacter* sp. *trpE* sequences. The three trees represent different portion of *trpE* sequences: a) residues 491-727; b) residues 652-1005; c) residues 1251-1629.

Figure 1

1 AACAGGTCTANTGCTGGGCTGGATGACAAACGAGNCCCTCAANGACCATACCANCGGCCCGGGTNTCGTTCTACTCCCCTCCCAGGAGTACTGGC
 101 GCAAGGGTGACACGTCGGCCTCGTGCAGTGGGTCAAGTCGGTGGCCTGGACTGCGACGGCGACGCCCTCCTCGTGGGGTGGACCAGGTGGGTGCAGC
 201 ATGCCACACCGGTACGCGGACCTGCTTTGACTTCCGGGCTTCGACGTAGTGGCAGGTCCGGCCAACAGCGGCCCCNCCCCAACAGGCCCCACC
 301 CAACAGGCCCTCACCAACAGGCCCTCACACTCAACAGGCCCCACCAACAGGCCCCACCCACTGGCACTGGCCTAGAACAGACGGAGAACCAATAGC
 401 AATGCAGGACCTCGGAATCATCAGCCCTGGCCTGGAAGAATTCGGGAACCTCGCGCCACAGCCGCGTCATCCCTGTCCGGCTCAAGGTGCTGGCGGAC
 M Q D L G I I S P G L E E F R E L A A H S R V I P V R L K V L A D
 501 GCTGAAACGCCGATCGGGCTGTACCGGAAGCTGGCGCAGGGCCAGCCAGGCACGTTCTCATGGAGTCGGCAGCTGTGGCGGGCCTGGTCCCCTACT
 A E T P I G L Y R K L A Q G Q P G T F L M E S A A V G G A W S R Y S
 601 CCTTCATCGGTTCCCGGTCCCGGCCACGCTGACCACCAAGGACGGACAGGCGCACTGGCTGGGCGAGCCGCGGGGGGTGCCCGTCGACGGAAATCC
 F I G S R S R A T L T T K D G Q A H W L G E P P A G V P V D G N P
 701 GGTGATGCCATCCGCGACCATCGAGGCGCTGCGCACCGACCGGTTGACGGGCTTCGCGGCTTACCTCCGGCTGGTGGGCTTCTGGGCTGGGAA
 V D A I R D T I E A L R T D R F D G L P P F T S G L V G F L G W E
 801 ACGGTCCGCCACTGGGAGAAGCTCACCAGACCGCCCGAGGACGACCTGCAGCTGCCGAACCTGGCGCTGAACCTCGTCACGGACATGGCCGTCCACGACA
 T V R H W E K L T R P P E D D L Q L P E L A L N L V T D M A V H D N
 901 ACATGGACGGCACGGTGTGCTCATCGCCAACGCCATCAACTTTGACAACAGCTCCGAACCGGTTGACGAGGCGCTGGCATGACGCCGTGCCCCGCGTGAA
 M D G T V L L I A N A I N F D N S S E R V D E A W H D A V A R V K
 1001 GGAGCTGCTGGCACGCATCAGCACCCCGGTGCTCAACCGGTGTCCGTGCTGGAGCCGGCTGCCCTCGATTTTGTGCCAGCGTCCAGGAACGCTGGAAC
 E L L A R I S T P V L Q P V S V L E P A A L D F A A S V Q E R W N
 1101 GAGCCCGAGTACCTTGGCCCGCTCGACCGGGCAAGGAGGCGATCGTGGACGGGAAGTCTTCCAGGTGGTCATCTCCCGCCGCTTCGAATGGAGTGGC
 E P E Y L A A L D R G K E A I V D G E V F Q V V I S R R F E M E C G
 1201 GGGCGTCGCCCTTGGATGTCTACCGGGTGTGCGCAACACCAACCCGAGCCGTCATGTACATTTTACGCTCGAGGACGCCGCCGGCCGGCAGTATTC
 A S P L D V Y R V L R N T N P S P Y M Y I F S L E D A A G R Q Y S
 1301 GATCGTCGGTTCGTGCGCGGAGGCGCTGGTGACGGTCACCGGCGAGGACGTATCAGCACCCCTATCGCCGGCTCCCGCCGCGCGGCAAGACCGTGGAC
 I V G S S P E A L V T V T G E D V I T H P I A G S R P R G K T V D
 1401 GCGGACAAGGCATTGGCCGAGGAACCTCTCGCAGACCAGAAGGAACGGGCGGAACACCTTATGCTGGTGGACCTGTCCCGGAACGATCTTTGGAAGGTCT
 A D K A L A E E L L A D Q K E R A E H L M L V D L S R N D L S K V C
 1501 GCGTGGCGGGTTCGCTGGACGTCACCCAGTTTATGGAGGTGGAGCGCTTACGCCACATCATGCACCTGGTCTCCACCGTCTGGGCGGCTGGTCCGAC
 V A G S V D V T Q F M E V E R F S H I M H L V S T V V G R L A P T
 1601 GGCGAAAGCCTATGACGTCCTGAAGGCCACCTTCCCGGCCGGAACCTTTCCGGCGCCCCGAAGCCCCGGGCGCTGCGGCTCCTGGATGAACCTCGAGCCG
 A K A Y D V L K A T F P A G T L S G A P K P R A L R L L D E L E P
 1701 CACCGGCGCGGCATCTACGGCGGCGTGGTGGGCTACCTGGACTTCGCGGCGGACATGGACATGGCCATTGCCATCCGTTCCGCACTGCTGCGCGAAGGAC
 H R R G I Y G G V V G Y L D F A G D M D M A I A I R S A L L R E G R
 1801 GGGCCTACGTCCAGCGGGCGGCGGCGCATCGTGGCGGATTGAGCAATCCGGCCGAGGCGCAGGAACCGTCAACAAGGCGGCGGCCCCGCTGCGTGTGT
 A Y V Q A G G G I V A D S S N P A E A Q E T V N K A A A P L R A V
 1901 CCACACGGCCGGATCGCTGCACAACATCAGCGCGGACTCCGTGAGCGCTCCCGATTCCGTGAGCAGCCCCGATTCCGTACGGAGGCAAACTCCTGATGG
 H T A G S L H N I T P D S V S A P D S V S S P D S V T E A N S *
 2001 CGGTTCCCGCAGGATCC

Figure 2

Chernova, Viswanathan, Austria and Nichols

Figure 3

Chemova, Viswanathan, Austria and Nichols