JOURNAL OF BACTERIOLOGY, Apr. 1997, p. 2433–2434 Vol. 179, No. 7
0021-9193/97/$04.00+0

The *rtn* Gene of *Proteus vulgaris* Is Actually from *Escherichia coli*

BARRY G. HALL*

University of Rochester, Rochester, New York 14627

Received 29 October 1996/Accepted 16 January 1997

The *rtn* gene, identified as coming from *Proteus vulgaris* ATCC 13315, is present in *Escherichia coli* K-12, and over a 440-bp region of *rtn* is identical to the published *Proteus* sequence, with the exception of a single G insertion. It was not possible to verify the presence of *rtn* in *P. vulgaris*.

Escherichia coli K-12 strain MH034 was isolated as a mini-Tn*10*(Cam)-induced mutant of strain SJ134 (F^- Δ*lacZ4680 ebgA51*) during a hunt for genes in which mini-Tn*10* insertions alter the adaptive mutation rate. Details of that mutant hunt will be reported elsewhere. MH034 has a modestly increased rate of adaptive mutations to *ebgR* (details reported elsewhere). To identify the gene that had been disrupted, genomic DNA from strain MH034 was digested with restriction endonuclease *Kpn*I, which does not cut within mini-Tn*10*(Cam), and the resulting fragments were ligated into the *Kpn*I-digested vector pUF2000. *E. coli* JM109 was transformed with the ligation mixture by electroporation, and after an hour for expression of drug resistance, the transformed culture was spread onto L broth (3) chloramphenicol plates to select for hybrid plasmids carrying mini-Tn*10*(Cam) and the region flanking the insertion. The resulting plasmid was designated pMH034.

Plasmid pMH034 was purified with the QiaQuick kit (Qiagen) according to the manufacturer's instructions, and the regions flanking mini-Tn*10*(Cam) were sequenced on an ABI model 370A automated DNA sequencer with primers corresponding to bp 1359 to 1383 and the complement of bp 78 to 102 of mini-Tn*10*(Cam). The resulting sequences were used to screen the non-redundant nucleotide databases with the BLAST program as implemented by the National Center for Biotechnology Information. The results of that search showed that the region flanking mini-Tn*10*(Cam) corresponded to the *rtn* gene of *Proteus vulgaris* ATCC 13315 (GenBank accession number U12780). The alignments of the flanking regions with the *rtn* gene showed a virtually perfect match with *rtn* and indicated that mini-Tn*10*(Cam) had inserted at bp 1573 of the reported *rtn* gene sequence.

The similarity between the *E. coli* region flanking mini-Tn*10*(Cam) and *rtn* of *P. vulgaris* was astonishingly high: 95 and 97% identity for the two flanking regions. Although *P. vulgaris* is not used in my laboratory, I considered the possibility that MH034 was a *P. vulgaris* contaminant. To test that possibility, oligonucleotide primers corresponding to bp 1320 to 1345 and the complement of bp 1797 to 1821 of the published *rtn* sequence were used in PCR amplifications with genomic DNA from *E. coli* MH034, its parent strain, SJ134, and a classic wild-type *E. coli* K-12 strain, MG1655. Analysis of the amplification products on 1% agarose gels showed that strains MG1655 and SJ134 had produced the expected 0.5-kb fragments, and strain MH034 had produced a 1.9-kb fragment corresponding to the region expected if it contained mini-Tn*10*(Cam).

The amplified fragment from the wild-type *E. coli* strain, MG1655, was purified with a Qiagen PCR purification kit according to the manufacturer's instructions and was sequenced with the same two primers used to amplify that fragment. The region from bp 1351 to 1790 was sequenced on both strands and over that 440 bp was identical to the published sequence, with the single exception of an additional G that was between published positions 1584 and 1585.

The *rtn* gene was originally isolated by shotgun cloning into plasmid pUC9 with selection of *E. coli* HB101 transformants that were resistant to bacteriophage λ (1). The resulting plasmid was shown to encode a protein that conferred resistance to both λ and to the *E. coli* phage N4, and on the basis of sodium dodecyl sulfate-polyacrylamide gel electrophoresis analysis, it was concluded that the Rtn protein was probably a membrane protein (1). The sequence of *rtn* was reported, and *rtn* was shown to encode a 484-amino-acid polypeptide that includes hydrophobic regions that are consistent with its being a membrane-associated protein (2).

Among members of the family *Enterobacteriaceae*, *P. vulgaris* is one of the species most distantly related to *E. coli* (4). *P. vulgaris* is unusual in that the G+C content of its DNA is only 37%, while most members of the *Enterobacteriaceae* have G+C contents in the 50 to 55% range (4). Excluding the *rtn* sequence, among the *P. vulgaris* sequences in the non-redundant nucleotide databases that are longer than 200 bp and that are not from plasmids, the G+C content is 41%. In contrast, the G+C content of the *rtn* sequence is 50.3%, and the G+C content of the third positions of the codons, which are the positions most free to drift toward the overall genomic content, is 56.6%. On the basis of G+C content, the *rtn* gene is much more like an *E. coli* gene than like a typical *P. vulgaris* gene.

The divergence rate at synonymous (silent) sites between *E. coli* and *Salmonella* spp. is estimated to be about 1% per million years (4). *E. coli* and *Salmonella* spp. are estimated to have diverged between 120 and 140 million years ago, while *E. coli* and *P. vulgaris* diverged about four times that long ago (4); thus, it is extremely unlikely that 440 bp of the *rtn* genes of *E. coli* and *Proteus* would have accumulated no base substitutions in that time. If the *rtn* gene is present in *P. vulgaris*, it must have been transferred from *E. coli* in the very recent past.

To determine whether *rtn* is present in *P. vulgaris*, genomic DNA prepared from *P. vulgaris* ATCC 13315 was used as a template in a PCR with the same primers that efficiently amplified the *rtn* gene from the *E. coli* K-12 strain. It has not proven possible to amplify *rtn* from *P. vulgaris* DNA. I con-

* Mailing address: Biology Department, Hutchison Hall, River Campus, University of Rochester, Rochester, NY 14627. Phone: (716) 275-0721. Fax: (716) 275-2070. E-mail: bhall@bgh.biology.rochester.edu.

clude that the published sequence did not originate from *P. vulgaris*, but instead originated from the *E. coli* K-12 strain, probably as a contaminant in the preparation of the cloning vector.

Nucleotide sequence accession number. The region sequenced from *E. coli* K-12 strain MH034 has been deposited in the GenBank database under accession number U83404.

I am grateful to Joe Yanni for expert technical assistance.

This study was supported by grant NP-932 from the American Cancer Society.

REFERENCES

1. **Chae, K.-S., and O. J. Yoo.** 1986. Cloning of lambda resistant genes from *Brevibacterium albidum* and *Proteus vulgaris* into *Escherichia coli*. Biochem. Biophys. Res. Commun. **140:**1101–1105.
2. **Hwang, U. W., D. W. Lee, C. S. Kim, and K.-S. Chae.** 1994. Structural analysis of an *rtn* gene isolated from *Proteus vulgaris*. Mol. Cells **4:**387–391.
3. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
4. **Ochman, H., and J. G. Lawrence.** 1996. Phylogenetics and the amelioration of bacteria genomes, p. 2627–2637. *In* F. C. Neidhardt, R. Curtiss III, J. L. Ingraham, E. C. C. Lin, K. B. Low, B. Magasanik, W. S. Reznikoff, M. Riley, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella*: cellular and molecular biology. ASM Press, Washington, D.C.