J Mol Evol (2005) 61:682–690
DOI: 10.1007/s00239-004-0349-4

JOURNAL OF MOLECULAR EVOLUTION

The Evolution of Microbial Phosphonate Degradative Pathways

Jinling Huang,[1,*] Zhengchang Su,[1,2] Ying Xu[1,2]

[1] Department of Biochemistry and Molecular Biology, University of Georgia, Athens, GA 30602, USA
[2] Computational Biology Institute, Oak Ridge National Laboratory, Oak Ridge, TN 27831, USA

Received: 8 December 2004 / Accepted: 26 May 2005 [*Reviewing Editor:* Dr. Martin Kreitman]

Abstract. Phosphonate utilization by microbes provides a potential source of phosphorus for their growth. Homologous genes for both C–P lyase and phosphonatase degradative pathways are distributed in distantly related bacterial species. The phn gene clusters for the C–P lyase pathway show great structural and compositional variation among organisms, but all contain phnG–phnM genes that are essential for C–P bond cleavage. In the γ-proteobacterium *Erwinia carotovora*, genes common to phosphonate biosyntheses were found in neighboring positions of those for the C–P lyase degradative pathway and in the same transcriptional direction. A gene encoding a hypothetical protein DUF1045 was found predominantly associated with the phn gene cluster and was predicted functionally related to C–P bond cleavage. Genes for phosphonate degradation are frequently located in close proximity of genes encoding transposases or other mobile elements. Phylogenetic analyses suggest that both degradative pathways have been subject to extensive lateral gene transfers during their evolution. The implications of plasmids and transposition in the evolution of phosphonate degradation are also discussed.

Key words: Phosphonate degradation — phn operon — Lateral gene transfer — Phylogenetic analyses

*Current address: Department of Biology, East Carolina University, Greenville, NC 27858, USA
Correspondence to: Ying Xu; *email:* xyn@bmb.uga.edu

Introduction

Phosphorus is an essential element for many biomolecules and plays an important role in many biological processes. Phosphonates are a class of organic phosphorus compounds characterized by a stable carbon–phosphorus (C–P) bond, which resists enzymatic hydrolysis, thermal decomposition, and other biochemical degradations (Ternan et al. 1998). The presence of phosphonates in both prokaryotes and eukaryotes is often linked to antibiotic activities or components of cellular macromolecules such as glycoproteins, glycolipids, and phosphonolipids (Hilderbrand and Henderson 1983). Phosphonates have also been artificially made for a wide range of industrial uses as pesticides, lubricant additives, and fire retardant agents etc. However, the increasing abundance of artificial phosphonates in the ecosystem has become an issue of environmental concern because of the inert nature of C–P bonds (Ternan et al. 1998).

The ability of microbes to break down phosphonates has attracted particular attention because of its potential applications in biodegradation. In natural systems, the degradation of phosphonates relies solely on microorganisms. The wide distribution of phosphonates makes them a valuable phosphorus source for microbes to sustain survival and growth. In some natural systems, phosphonates provide a significant fraction of the phosphorus supply required by microbes. The utilization of phosphonates by microbes may be realized by cleaving the C–P bond with several enzymes, including C–P lyase complex, phosphonatase, phosphonoacetate hydrolase, and phosphonopyruvate hydrolase (Ternan et al. 1998;

Kononova and Nesmeyanova 2002); the activities of the latter two enzymes were reported only in strains of *Curtobacterium*, *Pseudomonas*, and *Burkholderia* (Kulakova et al. 1997, 2001; Ternan et al. 1998; McGrath et al. 1999). It is generally believed that C–P lyase complex and phosphonatase, which are expressed under the condition of phosphate starvation, are responsible for the degradation of most phosphonates in nature (Kononova and Nesmeyanova 2002; Obojska and Lejczak 2003). The phosphonatase pathway acts mainly on 2-aminoethylphosphonate (AEP), whereas the C–P lyase pathway can break down a wide variety of structurally diverse phosphonates (Metcalf and Wanner 1993). The C–P lyase pathway has been extensively studied in *Escherichia coli* (Wanner and Boline 1990; Wanner and Metcalf 1992; Metcalf and Wanner 1993) and other bacteria such as *Sinorhizobium* (Parker et al. 1999), *Enterobacter* (Lee et al. 1992), and *Pseudomonas* (White and Metcalf 2004). In *E. coli*, 14 genes (phnC–phnP) related to phosphonate uptake and breakdown are arranged in tandem in the genome and cotranscribed from a single promoter (Metcalf and Wanner 1993). Experimental studies suggested that phnG–phnM genes are likely involved in C–P bond cleavage, and phnN and phnP are possibly accessory proteins (Metcalf and Wanner 1993). The phosphonatase pathway involves two steps. The AEP transaminase (encoded by the phnW gene) first catalyzes the reaction of AEP and pyruvate to form L-alanine and phosphonoacetaldehyde, and the latter is further converted to acetaldehyde by phosphonoacetaldehyde hydrolase (phosphonatase, encoded by the phnX gene). The first reaction is reversible, and the AEP transaminase is also used in the biosynthesis of AEP, where phosphonoacetaldehyde acts as a precursor.

The past decade has seen an increasing number of studies on phosphonate degradation; however, the distribution of the degradative pathways in living organisms is still less clear and no study on the evolution of phosphonate degradation has been reported so far. In this study, we have mined the available DNA and protein databases and performed comparative and phylogenetic analyses of the two major phosphonate degradative pathways. Our data suggest that these pathways scatter in distantly related bacterial groups and the spread of the phosphonate utilization capability among bacteria is significantly affected by events of lateral gene transfer.

Materials and Methods

Data Sources

One hundred ninety-seven prokaryotic genomes were downloaded from the NCBI GenBank database. Each of these genomes was fed into the program *formatdb* of the NCBI BLAST package (Altschul et al. 1990) to create an organism-specific database. Reciprocal similarity search was performed between *E. coli* K12 and each of the other organisms, and the best reciprocal hits were identified as orthologues in the species. Results of sequence similarity searches were parsed and the orthologues of phn genes (phnA–phnQ) and phnX were extracted for each species. Positions of phn and phnX orthologues on the chromosome/plasmid were manually inspected for each species. To verify the presence/absence of the two degradative pathways in eukaryotes and other prokaryotes, we also used phnA–phnQ protein sequences from *E. coli* and phnX sequence from *Salmonella typhimurium* to search the GenBank nonredundant (nr) protein database, GeneDB (www.genedb.org), and the eukaryotic gene orthologues (EGO) database (www.tigr.org/tdb/tgi/ego/index.shtml). Similarity searches of the above databases also led to identification of extra copies of the phn gene cluster in some bacterial species.

Phylogenetic Analyses

Multiple protein sequence alignment was performed using ClustalX (Thompson et al. 1997) for phnX of the phosphonatase pathway and each of the phnG–phnM genes of the C–P lyase pathway, followed by manual refinement. These genes were chosen for phylogenetic analyses since they are likely the indispensable components of the two pathways. Only unambiguously aligned sequence portions were used. For phnG–phnM, which are often conserved in gene order, unambiguously aligned segments of each gene were concatenated. Phylogenetic analyses were performed with a maximum likelihood method using TREE-PUZZLE version 5.1 for Unix (Schmidt et al. 2002) and a distance method using the program *neighbor* of the PHYLIP version 3.6a package (Felsenstein 2004). Bootstrap support was estimated using 1000 replicates for distance analyses and quartet puzzling values were calculated using 10,000 puzzling steps for maximum likelihood analyses. Distance calculation used the Jones–Taylor–Thornton (JTT) substitution matrix (Jones et al. 1992), and site-substitution variation was modeled with a gamma-distribution whose shape parameter was estimated from the data. For maximum likelihood analyses, a mixed model of eight gamma-distributed rates and one invariable rate was used to calculate the pairwise maximum likelihood distances.

Results

Similarity Search of the Phosphonate Degradative Pathways

Similarity search of *E. coli* phn genes against available DNA and protein sequence databases found significant hits only in bacteria, mostly the proteobacteria. Within the proteobacteria, the phn gene clusters were identified only in α-, γ-, β-, and δ-proteobacteria. Only two cyanobacteria (*Nostoc* sp. and *Trichodesmium erythraeum*) and one species from each of the firmicutes (*Oceanobacillus iheyensis*) and the green nonsulfur bacteria (*Chloroflexus aurantiacus*) were found to contain homologous phn gene clusters. No phn gene clusters were identified in available eukaryotic sequence data in our analyses.

The phosphonatase degradative pathway also seems limited to a number of bacterial groups. Homologues of phnX, which is specific to the phosphonatase pathway, were found in only a few

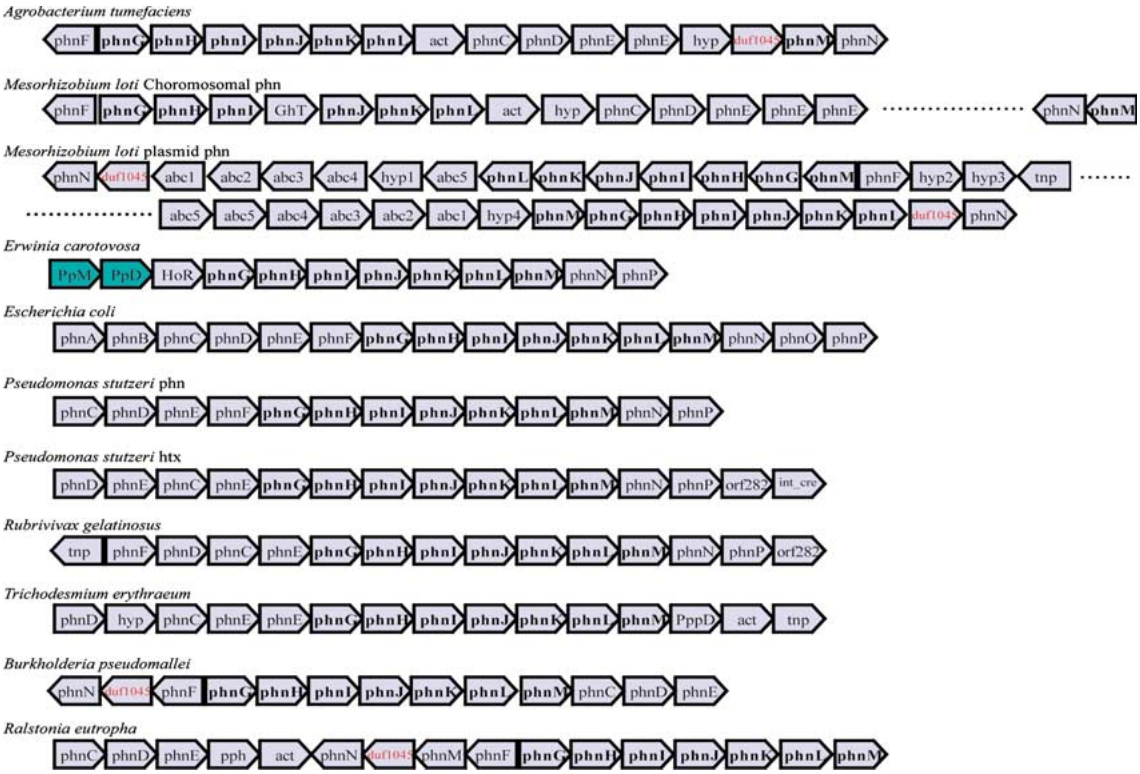

Fig. 1. Structure of bacterial phn gene clusters. Green boxes indicate genes for phosphonate biosynthesis, whereas blue boxes indicate genes for the C–P lyase degradative pathway and others. The genes are not drawn to scale. hyp, hypothetical protein; abc, ABC transporter; tnp, transposase; int_cre, Cre recombinase; PpM, phosphoenolpyruvate mutase; PpD, phosphonopyruvate decarboxylase; HoR, 2-hydroxy-3-oxopropionate reductase; GhT, glutathione transferase; PppD, phenylpropionate dioxygenase; pph, phosphohydrolase; act; acetyltransferase.

firmicutes (*Bacillus*, *Lactobacillus*, and *Enterococcus*), proteobacteria (*Burkholderia*, *Salmonella*, *Pseudomonas*), and a bacteroidete (*Bacteroides fragilis*). In many cases, phnW and phnX are located adjacently on the chromosome.

Structure of the phn Gene Cluster

All phn clusters identified in this study contain phnG–phnM genes (Fig. 1), which are essential for C–P bond cleavage. However, the structure of the phn gene cluster varies greatly among species. In γ-proteobacteria *E. coli*, *Yersinia pestis*, and other nonproteobacterial species such as *C. aurantiacus*, *T. erythraeum*, and *O. iheyensis*, the phn gene cluster contains ordered phnG–phnM genes in a consecutive arrangement. In most α-proteobacteria investigated in the study (e.g., *Agrobacterium*, *Mesorhizobium*, and *Sinorhizobium*), these phn genes are separated by those encoding hypothetical or other proteins (Fig. 1). One of these genes encodes the hypothetical protein DUF1045 and is mainly associated with phn genes in a few proteobacteria. DUF1045 is often located adjacent to phnN and appears not to be correlated with the absence of any specific phn gene. In addition, many of the phn gene clusters are found adjacent to mobile elements. For example, genes encoding transposases (or inactivated transposases) are located in proximity of the phn clusters in *T. erythraeum*, *Rubrivivax gelatinosus*, *Sinorhizobium meliloti*, and *Desulfovibrio desulfuricans*. Similarly, a gene downstream (separated by one gene) of a phn operon homologue (htx) in *Pseudomonas stutzeri* encodes a homologue of Cre recombinase (Fig. 1).

Although genes of the C–P lyase pathway tend to cluster on the chromosome (or plasmid), they are not always transcribed in the same direction. The phnF gene is often positioned next to other phn components, but except for a few instances, they are transcribed in opposite directions. Similar cases were also found in the β-proteobacteria *Ralstonia*, *Burkholderia*, and *Chromobacterium*, where phnF, DUF1045, and phnN form an ordered cluster 150–300 bp upstream of the phnG–phnM genes on the opposite strand.

Multiple copies of the phn gene clusters were identified in *Mesorhizobium loti* and *Pseudomonas stutzeri*. *Mesorhizobium loti* contains a phn gene cluster on the chromosome and two additional clusters on a plasmid. The chromosomal phn gene cluster of *M. loti* is similar to those of other α-proteobacteria (e.g., *S. meliloti* and *Agrobacterium tumefaciens*) in

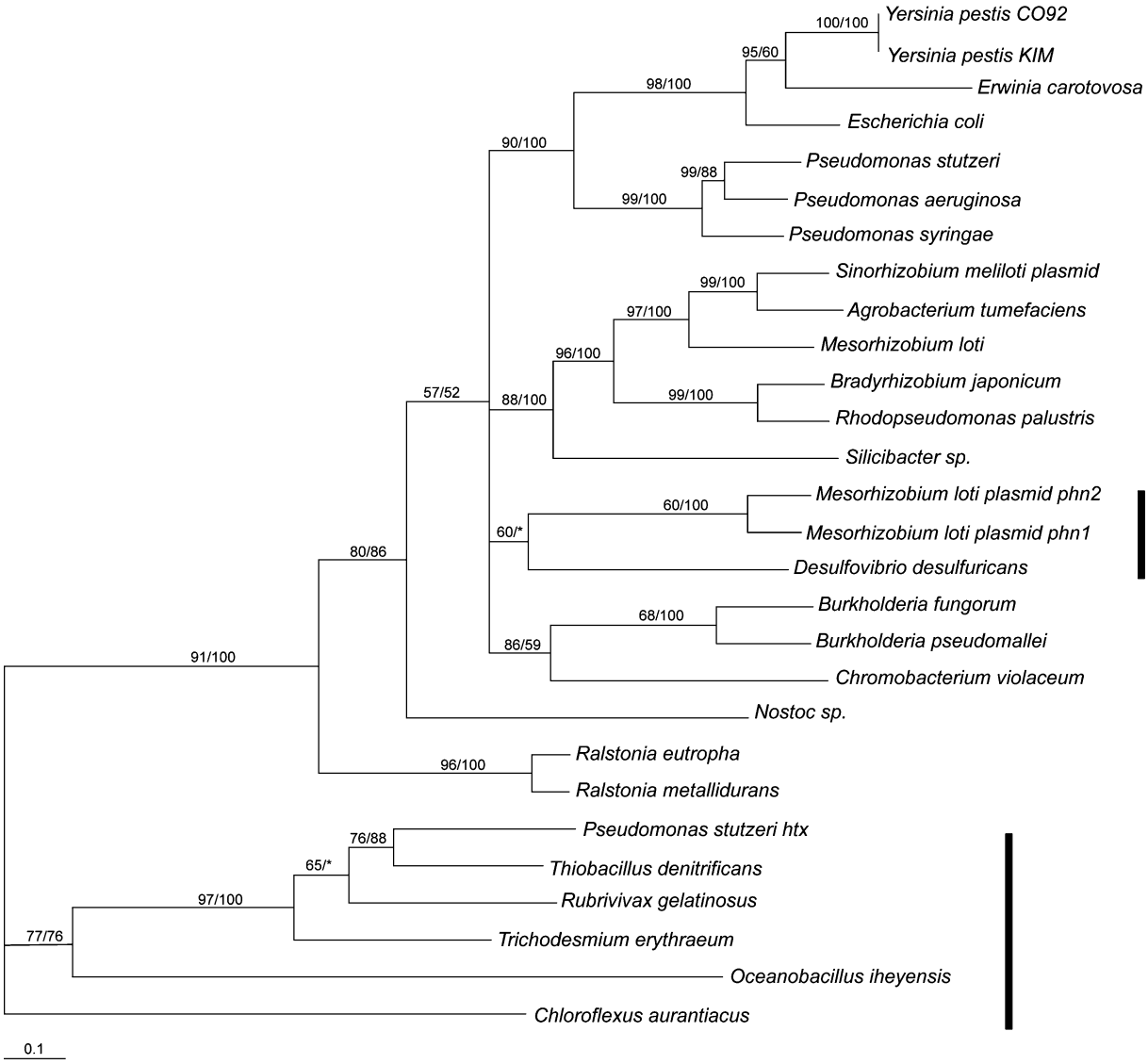

Fig. 2. Phylogenetic analyses of phnG–phnM protein sequences of the C–P lyase pathway. The tree topology is based on maximum likelihood analyses. Numbers above the branches show quartet puzzling frequency for maximum likelihood analyses and bootstrap support for neighbor-joining analyses, respectively. Asterisks indicate that support for the branch is below 50%. The scale is as indicated. Vertical bars show where lateral gene transfers are involved.

gene composition, cluster structure, and sequence identity. The two plasmid clusters are large inverted repeats located on opposite strands and are about 110 kb away. Each of them consists of 14–16 genes, including an ordered phnMG–phnL cluster and several other genes encoding ABC transporters and one encoding the hypothetical protein DUF1045 (Fig. 1). In addition, homologous genes from the two plasmid phn clusters often share the highest protein sequence identity. *Pseudomonas stutzeri* contains two chromosomal phn operon homologues. The regular phn operon of *P. stutzeri* shares the highest protein sequence identity with those from congeneric species *P. syringae* and *P. aeruginosa*, whereas the other operon (htx) is more similar to homologues from *Thiobacillus denitrificans*, *R. gelatinosus*, and *T. erythraeum* in both structure and sequence identity. The *P. stutzeri* htx operon also contains a gene encoding a hypothetical protein (orf282) and an inactivated transposase immediately downstream. Orf282 is found in only a limited number of bacterial species and has the highest sequences identity (55%) with its homologue from *R. gelatinosus*.

Phylogenetic Analyses

Multiple protein sequence alignment for seven genes (phnG–phnM) of the C–P lyase pathway provided a total of 1633 unambiguously aligned amino acid residues, which were further concatenated and used for phylogenetic analyses. The unrooted tree from the analyses is shown in Fig. 2. In this tree, the phn lin-

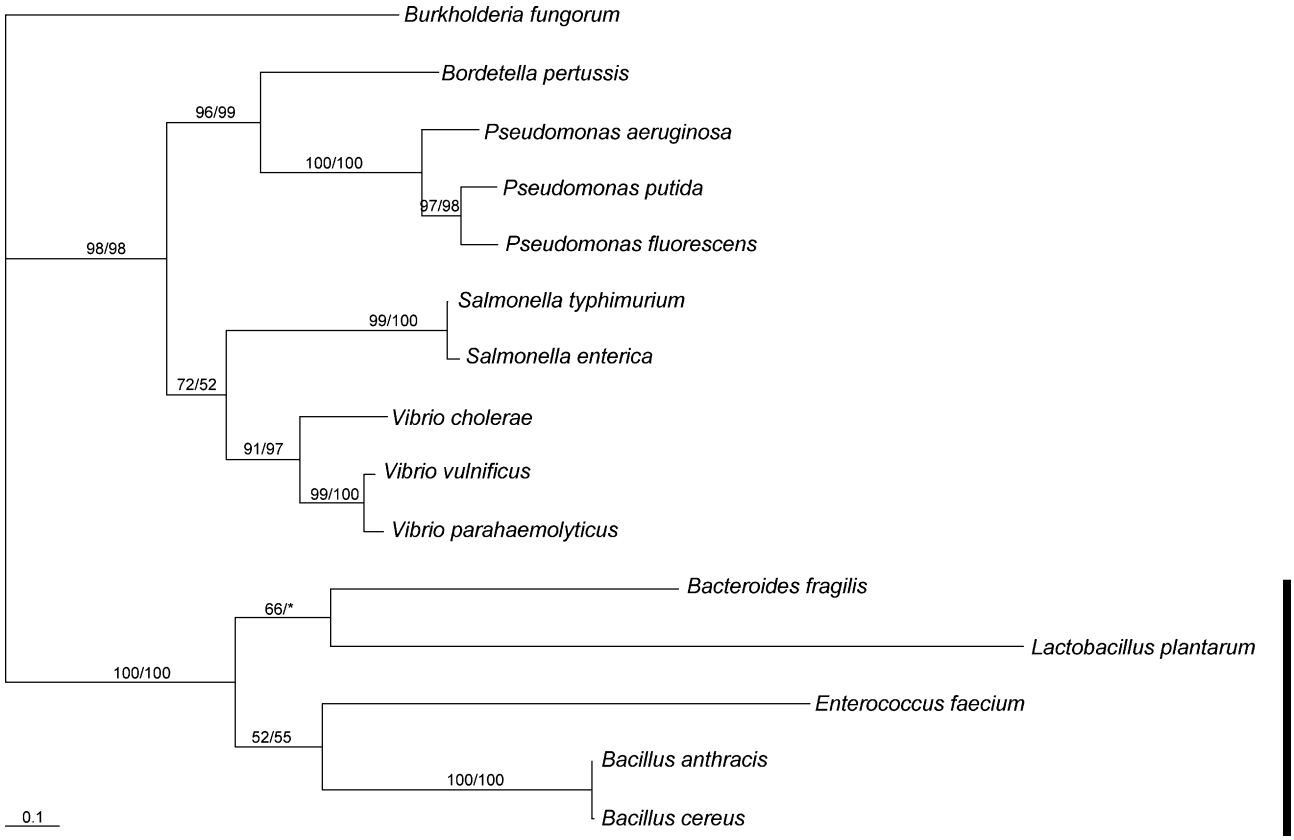

Fig. 3. Phylogenetic analyses of phnX protein sequences of the phosphonatase pathway. The tree topology is based on maximum likelihood analyses. Numbers above the branches show quartet puzzling frequency for maximum likelihood analyses and bootstrap support for neighbor-joining analyses, respectively. The asterisk indicates that support for that branch is below 50%. The scale is as indicated. Vertical bar shows branches where lateral gene transfers are involved.

eages from the green nonsulfur bacterium *C. aurantiacus*, cyanobacterium *T. erythraeum*, and firmicute *O. iheyensis* form a well-supported group with their homologues of the β-proteobacteria *R. gelatinosus*, *T. denitrificans*, and the htx operon of the γ-proteobacterium *P. stutzeri*. The phn genes in the *δ*-proteobacterium *D. desulfuricans* group with two plasmid homologues from the α-proteobacterium *M. loti* with modest support. The phn genes from the two cyanobacteria, *Nostoc* sp. and *T. erythraeum*, are of different origins in our analyses.

Results of the phylogenetic analyses of phnX gene from the phosphonatase pathway are shown in Fig. 3. The phnX sequences from firmicutes *Bacillus cereus* and *Ba. thuringiensis* and the bacteroidete *B. fragilis* have the shortest pairwise distances and form a well-supported group with those from other firmicutes. The position of the β-proteobacterium *Burkholderia fungorum* phnX sequence appears isolated from others.

Discussion

In this study, we have mined the available sequence databases for the distribution of two major phosphonate degradative pathways and performed comparisons and phylogenetic analyses on their component genes. Our study revealed the conserved phnG–phnM genes in the phn gene cluster, supporting the previous finding that these genes are essential for C–P bond cleavage in the C–P lyase degradative pathway (Metcalf and Wanner 1993).

Distribution of the Phosphonatase and the C–P Lyase Pathways

Similarity searches of the available sequence databases indicated that the two major phosphonate degradative pathways exist in a number of bacterial groups. A few species also contain both pathways. However, it should be noted that available microbial genome sequencing data are largely biased toward pathogens and a few groups of agriculturally or ecologically important microbes. Failure to identify homologues of the two degradative pathways in certain groups of microbes in our analyses does not necessarily suggest that these pathways are absent from these organisms, since their genomes may have not yet been sequenced. For example, C–P lyase activities were detected by experimental studies in

actinobacteria *Saccharopolyspora erythraea* (Obojska and Lejczak 2003), *Arthrobacter* sp. (Schowanek and Verstraete 1990), the firmicute *Bacillus megaterium* (Kononova and Nesmeyanova 2002), and a few soil-borne fungi (Krzysko-Lupicka et al. 1997), but we were unable to identify their phn gene homologues because of the lack of genome sequences for these species.

The capability of microbes to cleave the C–P bond is often restricted to strains instead of being species-specific or lineage-specific (Schowanek and Verstraete 1990). This strain-specific nature may also account for the observation that phn clusters are not identified in the genome even though C–P bond cleavage activities have been reported in related species. For example, C–P lyase activities were previously detected in several actinobacteria including *Streptomyces lusitanus*, *S. fradiae*, *S. morookaensis*, and *S. vinaceus* (Obojska and Lejczak 2003), but our database search failed to identify the phn clusters in two sequenced congeneric species, *S. avermitilis* and *S. coelicolor*. Therefore, the distribution of the two major phosphonate degradative pathways presented in this study should only be considered our best understanding from available genome sequence data rather than a conclusive list of their existence in all microbes.

Structure of the phn Operon

The phn gene cluster shows great variation in structure and gene composition. Except for a few organisms (e.g., *E. coli*), most phn clusters contain genes encoding hypothetical proteins or other functions. The gene encoding the hypothetical protein DUF1045 is limited to a few proteobacteria, but its presence in the genome is almost always associated with the phn gene cluster. Since component genes in an operon are often functionally related, it may be useful to carry out further experimental investigations to see whether DUF1045 is involved in phosphonate degradation activities.

Operons are pervasive in prokaryotes. Although the mechanism underlying the evolution of an operon is not yet fully understood, it has been suggested that selection for more efficient coregulation or lateral gene transfer is the driving force (Lawrence and Roth 1996). If these hypotheses hold up, the great variation of the phn gene cluster as revealed in our data would showcase the dynamics of the operon evolution. The phn operons in *E. coli, Pseudomonas,* and related species, where 13 or 14 genes involved in phosphonate transport and breakdown are cotranscribed, are apparently more effective for both coordinated expression and lateral gene transfer. On the other hand, the phn genes in the α-proteobacteria and other lineages (e.g., *Nostoc*) are often more loosely arranged. In some species (e.g., *Ralstonia*, *Chromobacterium*, and *Burkholderia*), component genes in the cluster may also be divergently transcribed. It should also be noted that the phnF gene, except in the γ-proteobacteria, is almost always transcribed in the opposite direction from most of the other phn genes. These loosely organized operon structures are sometimes considered "destructed" due to genome rearrangement, and secondary in origin (Itoh et al. 1999). While genome rearrangement and even gene displacement can be common during operon evolution (Omelchenko et al. 2003), fragmentation of a well-adapted operon will at least require the evolution of regulatory elements for newly generated gene clusters but may not confer selective advantages to the organism. Disruption of an essential functional module such as the phnG–phnM suboperon will more likely be deleterious and, thus, eliminated from the population. Therefore, these loosely organized phn gene clusters may actually represent various phases toward the formation of a more sophisticated operon.

Interestingly, genes for both phosphonate degradation and biosynthesis form a cluster in the γ-proteobacterium *Erwinia carotovora*. In this species, the phn genes (phnG–phnP) are located immediately downstream of those encoding phosphoenolpyruvate mutase (PpM), phosphonopyruvate decarboxylase (PpD), and 2-hydroxy-3-oxopropionate reductase (Fig. 1). PpM and PpD are the first two enzymes used in AEP biosynthesis and they are also likely the only enzymes common to all biosyntheses of phosphonates (Schwartz et al. 1998). Whether these genes for different pathways are cotranscribed and form a true operon warrants further experimental investigations.

Lateral Gene Transfer of Phosphonate Degradation Among Bacteria

The scattered distribution of the phn gene cluster and the phnX gene among a broad spectrum of bacteria serves as an instant reminder of lateral gene transfers across distantly related organisms, even though the evolutionary scenario that an ancestral pathway, either C–P lyase or phosphonatase, was independently lost among numerous bacterial groups remains theoretically possible. Lateral gene transfer is widely accepted as a major force for bacterial genome evolution (Ochman et al. 2000). Methods used for detecting gene transfers in prokaryotes are often based on sequence similarity search, codon bias, and uncharacteristic base composition. These methods are useful in initial screening for lateral transfer events, but they are less accurate and provide no information about the evolutionary origin of the subject DNA or protein sequence (Logsdon and Faguy 1999; Koski et al. 2001; Ragan 2001). In this

study, we have performed rigorous phylogenetic analyses, in addition to comparisons of sequence identity, operon structure, and taxonomic distribution. Our phylogenetic analyses show that sequences from various major bacterial lineages often form well-supported monophyletic groups (Figs. 2 and 3), suggesting that lateral gene transfer indeed plays an important role in the evolution of phosphonate degradation in bacteria.

The γ-proteobacterium *P. stutzeri* contains another homologous htx operon in addition to a regular phn operon. The htx operon is currently identified in certain strains of *P. stutzeri* (White and Metcalf 2004) and the β-proteobacterium *Alcaligenes faecalis* (Wilson and Metcalf 2005; only htxABCD reported); its distribution in other microbes is yet to be investigated. In *P. stutzeri*, this htx operon is only distantly related to its infraspecific homologue, excluding the possibility of arising from a recent intragenomic duplication event. Instead, the htx genes are highly similar to their homologues from the β-proteobacteria *R. gelatinosus* and *T. denitrificans*, and the cyanobacterium *T. erythraeum*, with protein sequence identities ranging from 37% to 79%. Given its strain-specific nature (White and Metcalf 2004), the htx operon in *P. stutzeri* was likely recently acquired from other organisms. This hypothesis of recent acquisition of the htx operon in *P. stutzeri* is further supported by its shared indels with *R. gelatinosus* and *T. erythraeum* in multiple protein sequence alignment (data not shown) and the unique distribution of the gene encoding orf282 (Fig. 1). Most importantly, our phylogenetic analyses show a strong monophyletic group of *P. stutzeri* htx and its homologues from *T. denitrificans*, *R. gelatinosus*, *T. erythraeum, O. iheyensis*, and *C. aurantiacus*, with *Ralstonia* as their immediate outgroup (Fig. 2). Since these species are members of different bacterial lineages, our data also point to more rampant lateral gene transfers across distantly related bacterial groups.

Another apparent case of lateral gene transfer involves the phnX gene of *B. fragilis*, an endogenous anaerobic microbe commonly found in intestinal tracks. The phnX gene is identified in the strain *B. fragilis* YCH46, but not in another strain ATCC 25285, whose genome sequence is currently incomplete, and the congeneric species *B. thetaiotaomicron*. The phnX gene from *B. fragilis* shares the highest protein sequence identity with those from *B. cereus* and *B. thuringiensis*. Phylogenetic analyses further support a monophyletic group of these sequences with those from other firmicutes, suggesting that *B. fragilis* likely obtained its phnX gene from firmicutes.

Previous experimental data suggested that C–P lyase activities exist in a rather broad spectrum of bacteria (Cook et al. 1978; Schowanek and Verstraete 1990), many of which have not yet been sequenced. In our sequence analyses, *Nostoc*, *Trichodesmium*, *Chloroflexus*, and *Oceanobacillus* are the only taxa outside the proteobacteria found to contain a phn gene cluster. Nevertheless, the phn clusters in these taxa all appear to have been acquired from other organisms. However, lateral gene transfers may not be limited to these species. Even within the proteobacteria, the phn gene cluster is far from being universally distributed. The phn gene cluster homologues are found in only 4 of 14 α-proteobacteria and 11 of 43 γ-proteobacteria whose complete genomes are available, and appear to be lacking in the ε-proteobacteria. This pattern of taxonomic distribution implies either multiple species-specific losses or, alternatively, involvement of widespread lateral gene transfer in the evolution of the C–P lyase pathway within the phylum. Similar scenarios also happened in the evolution of the phosphonatase degradative pathway, which is found sparsely distributed in firmicutes, δ- and γ-proteobacteria, and bacteroidetes in our sequence analyses.

The evolution of prokaryotic genomes is a dynamic process characterized by constant exchanges of genetic materials among organisms (Ochman and Moran 2001). It is conceivable that lateral gene transfer of large DNA fragments containing C–P bond cleavage genes could confer host organisms immediate abilities to utilize a wide range of phosphonates as nutrients that were previously unavailable. This may particularly be true for bacteria in ecosystems such as marine sediments and open seas, where sources for inorganic phosphorus are limited (Sundareshwar et al. 2003). Indeed, the widespread distribution of phosphonates in nature, together with the presence of both C–P lyase and phosphonatase degradative pathways in *Pseudomonas* and *Burkholderia* and multiple phn gene clusters in some other organisms (e.g., *M. loti* and *P. stutzeri*), may all imply the importance of phosphonates as potential nutrients for certain bacterial groups.

The introduction of foreign genes or genetic materials does not guarantee their successful integration and function in the host genome. Newly acquired genes can frequently be deleted from the host genome or decay into pseudogenes (Ochman and Jones 2000; Liu et al. 2004) unless benefits are conferred to the host. The phnX and phn genes appear to be intact in all species investigated in this study, suggesting that these genes, including those acquired from other organisms, are likely functional in their host organisms. In fact, the functionality of these genes has been demonstrated by experimental studies on several organisms, including *P. stutzeri*, whose regular phn and recently acquired htx operons act in

concert to promote phosphite and phosphonate utilization (White and Metcalf 2004).

Mechanisms Underlying the Lateral Gene Transfer

Given the finding that lateral gene transfer plays a significant role in the evolution of phosphonate degradation among a broad taxonomic group of bacteria, it would be useful to speculate the process mediating these transfer events. The presence of phn gene clusters in plasmids of both *M. loti* and *S. meliloti* sheds light on forces behind these events. The autonomous inheritance of plasmid genomes from the cell enables them to transfer to broad host groups. Previous studies suggested that the evolution of antibiotics resistance and nitrogen fixation in bacteria involves plasmid-mediated lateral gene transfers (Tauxe et al. 1989; Chen et al. 2003). In both *M. loti* and *S. meliloti*, the plasmid genomes contain genes for replication and conjugation transfer. Although their affinity cannot be definitely determined in our study, the plasmid phn gene clusters in *M. loti* have an apparently different evolutionary history from its chromosomal counterpart, indicating its past involvement in lateral gene transfer. The phn gene cluster in *S. meliloti* has been identified only in plasmid and is closely related to its chromosomal homologues in *A. tumefaciens*, *M. loti*, and *Bradyrhizobium japonicum*. These species all are members of the α-proteobacteria and common inhabitants of rhizosphere, where interspecific interactions are active due to abundant moisture and nutrients (Elsas et al. 2003), rendering the scenario of lateral gene transfer among them via plasmid more likely.

Another interesting finding from our analyses is the neighboring position between the phn gene cluster and mobile elements on the chromosome (or plasmid). This phenomenon has been observed in chromosomes of *T. erythraeum*, *R. gelatinosus*, and *P. stutzeri* (htx operon) and plasmids of *M. loti*, and *S. meliloti* (Fig. 1). In all cases, they are likely associated with lateral gene transfers. Particularly, the *P. stutzeri* htx operon and its homologues in *R. gelatinosus* and *T. erythraeum* are members of a branch that is involved in extensive lateral gene transfers (Fig. 2). Although mobile elements usually do not directly transfer genes between cells, their abilities to mobilize genes onto transmissible plasmids could indirectly contribute to lateral gene transfers. Even in the absence of a transmissible vector, conjugative transposons could also mobilize genes across species boundaries (Scott 1992). Whether these mobile elements have indeed contributed to the spread of phn gene clusters across a wide spectrum of bacteria remains to be further investigated.

The evolution of the capability among microbes to utilize phosphonates as potential nutrients is not only important for their survival and growth, but also vital for the ecosystem to sustain effective element recycling. Our data point to the importance of lateral gene transfer in the evolution of phosphonate utilization. The finding of this study should add further insights into the role of gene transfer in the evolution of other biochemical activities in general and the ecology of the natural system.

Acknowledgments. We thank Dr. Anna Karls for discussions and two anonymous reviewers for helpful comments and suggestions. This research was supported in part by the U.S. Department of Energy's Genomes to Life program (http://doegenomestolife.org), under the project "Carbon Sequestration in *Synechococcus* sp.: From Molecular Machines to Hierarchical Modeling" (www.genomes2life.org), and also by the National Science Foundation (Grants NSF/DBI-0354771 and NSF/ITR-IIS-0407204).

References

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410

Chen WM, Moulin L, Bontemps C, Vandamme P, Bena G, Boivin-Masson C (2003) Legume symbiotic nitrogen fixation by β-proteobacteria is widespread in nature. J Bacteriol 185:7266–7272

Cook AM, Daughton CG, Alexander M (1978) Phosphonate utilization by bacteria. J Bacteriol 133:85–90

Elsas JD, Turner S, Bailey MJ (2003) Horizontal gene transfer in the phytosphere. New Phytol 157:525–537

Felsenstein J (2004) PHYLIP (Phylogenetic Inference Package) version 3.6. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle

Hilderbrand RL, Henderson TO (1983) Phosphonic acids in nature. In: Hilderbrand RL (ed) The role of phosphonates in living systems. CRC Press, Boca Raton, FL, pp 5–29

Itoh T, Takemoto K, Mori H, Gojobori T (1999) Evolutionary instability of operon structures disclosed by sequence comparisons of complete microbial genomes. Mol Biol Evol 16:332–346

Jones DT, Taylor WR, Thornton JM (1992) The rapid generation of mutation data matrices from protein sequences. Comput Appl Biosci 8:275–282

Kononova SV, Nesmeyanova MA (2002) Phosphonates and their degradation by microorganisms. Biochemistry (Moscow) 67:184–195

Koski LB, Morton RA, Golding GB (2001) Codon bias and base composition are poor indicators of horizontally transferred genes. Mol Biol Evol 18:404–412

Krzysko-Lupicka T, Strof W, Kubs K, Skorupa M, Wieczorek P, Lejczak B, Kafarski P (1997) The ability of soil-borne fungi to degrade organophosphonate carbon-to-phosphorus bonds. Appl Microbiol Biotechnol 48:549–552

Kulakova AN, Kulakov LA, Quinn JP (1997) Cloning of the phosphonoacetate hydrolase gene from *Pseudomonas fluorescens* 23F encoding a new type of carbon–phosphorus bond cleaving enzyme and its expression in *Escherichia coli* and *Pseudomonas putida*. Gene 195:49–53

Kulakova AN, Kulakov LA, Akulenko NV, Ksenzenko VN, Hamilton JTG, Quinn JP (2001) Structural and functional analysis of the phosphonoacetate hydrolase (phnA) gene region in *Pseudomonas fluorescens* 2F. J Bacteriol 183:3268–3275

Lawrence JG, Roth JR (1996) Selfish operons: horizontal transfer may drive the evolution of gene clusters. Genetics 143:1843–1860

Lee KS, Metcalf WW, Wanner BL (1992) Evidence for two phosphonate degradative pathways in *Enterobacter aerogenes*. J Bacteriol 174:2501–2510

Liu Y, Harrison PM, Kunin V, Gerstein M (2004) Comprehensive analysis of pseudogenes in prokaryotes: widespread gene decay and failure of putative horizontally transferred genes. Genome Biol 5:R64

Logsdon JM, Faguy DM (1999) *Thermotoga* heats up lateral gene transfer. Curr Biol 9:R747–R751

McGrath JW, Kulakova AN, Quinn JP (1999) A comparison of three bacterial phosphonoacetate hydrolases from different environmental sources. J Appl Microbiol 86:834–840

Metcalf WW, Wanner BL (1993) Mutational analysis of an *Escherichia coli* fourteen-gene operon for phosphonate degradation, using TnphoA' elements. J Bacteriol 175:3430–3442

Obojska A, Lejczak B (2003) Utilisation of structurally diverse organophosphonates by Streptomycetes. Appl Microbiol Biotechnol 62:557–563

Ochman H, Jones IB (2000) Evolutionary dynamics of full genome content in *Escherichia coli*. EMBO J 19:6637–6643

Ochman H, Moran NA (2001) Genes lost and genes found: evolution of bacterial pathogenesis and symbiosis. Science 292:1096–1099

Ochman H, Lawence JG, Groisman EA (2000) Lateral gene transfer and the nature of bacterial innovation. Nature 405:299–304

Omelchenko MV, Makarova KS, Wolf YI, Rogozin IB, Koonin EV (2003) Evolution of mosaic operons by horizontal gene transfer and gene displacement *in situ*. Genome Biol 4:R55

Parker GF, Higgins TP, Hawkes T, Robson RL (1999) *Rhizobium* (*Sinorhizobium*) *meliloti* phn genes: characterization and identification of their protein products. J Bacteriol 181:389–395

Ragan MA (2001) On surrogate methods for detecting lateral gene transfer. FEMS Microbiol Lett 201:187–191

Schmidt HA, Strimmer K, Vingron M, Haeseler A (2002) TREE–PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. Bioinformatics 18:502–504

Schowanek D, Verstraete W (1990) Phosphonate utilization by bacterial cultures and enrichments from environmental samples. Appl Environ Microbiol 56:895–903

Schwartz D, Recktenwald J, Pelzer S, Wohlleben W (1998) Isolation and characterization of the PEP-phosphomutase and the phosphonopyruvate decarboxylase genes from the phosphinothricin tripeptide producer *Streptomyces viridochromogenes* Tu494. FEMS Microbiol Lett 163:149–157

Scott JR (1992) Sex and the single circle: conjugative transposition. J. Bacteriol 174:6005–6010

Sundareshwar PV, Morris JT, Koepfler EK, Fornwalt B (2003) Phosphorus limitation of costal ecosystem processes. Science 299:563–565

Tauxe RV, Cavanagh TR, Cohen ML (1989) Interspecies gene transfer *in vivo* producing an outbreak of multiply resistant shigellosis. J Infect Dis 160:1067–1070

Ternan NG, McGrath JW, McMullan G, Quinn JP (1998) Organophosphonates: occurrence, synthesis and biodegradation by microorganisms. World J Microbiol Biotech 14:635–647

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 25:4876–4882

Wanner BL, Boline JA (1990) Mapping and molecular cloning of the phn (psiD) locus for phosphonate utilization in *Escherichia coli*. J Bacteriol 172:1186–1196

Wanner BL, Metcalf WW (1992) Molecular genetic studies of a 10.9-kb operon in *Escherichia coli* for phosphonate uptake and biodegradation. FEMS Microbiol Lett 15:133–139

White AK, Metcalf WW (2004) Two C–P lyase operons in *Pseudomonas stutzeri* and their roles in the oxidation of phosphonates, phosphite, and hypophosphite. J Bacteriol 186:4730–4739

Wilson MM, Metcalf WW (2005) Genetic diversity and horizontal transfer of genes involved in oxidation of reduced phosphorus compounds by *Alcaligenes faecalis* WM2072. Appl Environ Microbiol 71:290–296