AEM
Journals.ASM.org

Evaluating the Assignment of *alkB* Terminal Restriction Fragments and Sequence Types to Distinct Bacterial Taxa

Julia Giebler,[a] Lukas Y. Wick,[a] Michael Schloter,[b] Hauke Harms,[a] Antonis Chatzinotas[a]

Department of Environmental Microbiology, Helmholtz Centre for Environmental Research, Leipzig, Germany[a]; Research Unit for Environmental Genomics, HelmholtzZentrum München, German Research Center for Environmental Health, Neuherberg, Germany[b]

Sequence and terminal restriction fragment length polymorphism (T-RFLP) analyses revealed multiple *alkB* gene copies/cell in soil bacterial isolates and an apparently high genetic mobility among various phylogenetic groups. Identifying alkane degraders by *alkB* terminal restriction fragments (T-RFs) and sequences is strongly biased, as the phylogenetic trees based on 16S rRNA and *alkB* gene sequences were highly inconsistent.

Aliphatic hydrocarbons like *n*-alkanes occur frequently in nature, e.g., in epicuticular waxes of plant leaves which contribute to the outer defense of the plant against external stressors (1). Alkanes with a chain length between C_{10} and C_{30} (2, 3) are major components of these waxes. Hence, a large number of microbes are able to degrade alkanes due to their constant availability (4). Substrate activation is thereby accomplished by the terminal or subterminal oxidation of the alkanes followed by a complete metabolization of the alcohols formed (5). In many bacteria, the initial oxidation of alkanes with chain lengths of C_{10} to C_{20} is mediated by an alkane monooxygenase (AlkB) (4). Previous studies have indicated that AlkB is widespread among bacteria (6, 7) and that the corresponding gene occurs in multiple copies in a single organism (8, 9). It has been suggested that horizontal gene transfer plays an important role in the distribution of *alkB* among bacteria and, thus, *alkB* diversity may not follow 16S rRNA-based phylogeny (4, 9, 10). Therefore, in the present study, we validated the congruency of *alkB* and 16S rRNA gene-derived phylogenies from soil bacteria. To this end, we screened a large number of hexadecane (C_{16})-degrading soil isolates using a combined approach based on terminal restriction fragment length polymorphism (T-RFLP) fingerprinting and sequencing of the respective *alkB* genes.

Alkane-degrading bacteria were enriched from a recently described microcosm study with maize litter-amended agricultural sand soil (6). Fifty isolates out of 400 colonies growing with C_{16} as the sole carbon source on mineral medium agar and spanning a broad phylogenetic range based on partial 16S rRNA gene sequence analysis (*Alpha*-, *Beta*-, and *Gammaproteobacteria*, low-G+C Gram-positive bacteria, *Actinobacteria*, and *Verrucomicrobia*) were subjected to *alkB* gene T-RFLP analysis. *alkB* genes were

Received 4 January 2013 Accepted 27 February 2013
Published ahead of print 1 March 2013
Address correspondence to Antonis Chatzinotas, antonis.chatzinotas@ufz.de.
Supplemental material for this article may be found at http://dx.doi.org/10.1128/AEM.04028-12.

doi:10.1128/AEM.04028-12

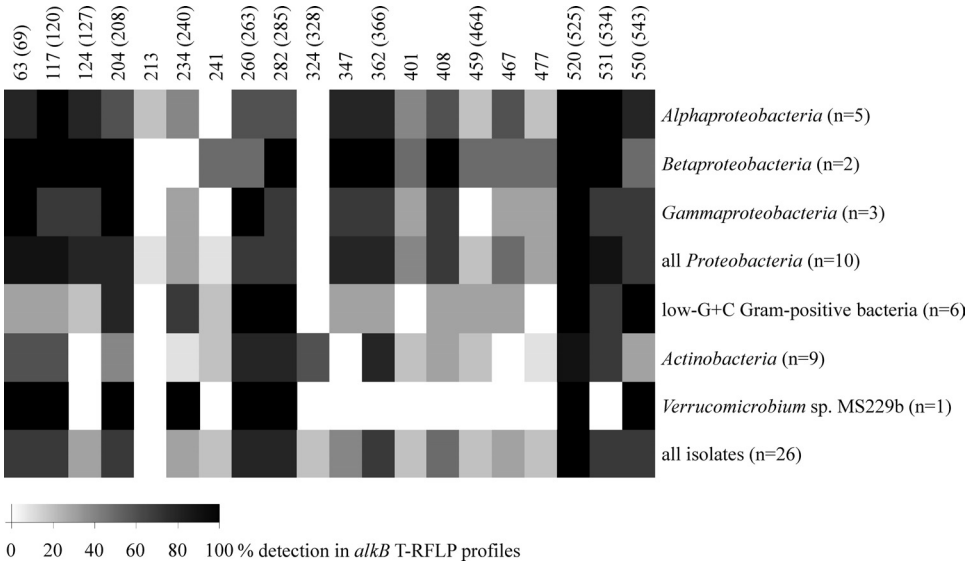

FIG 1 Heatmap displaying the occurrences of 20 T-RFs in *alkB* T-RFLP profiles of isolates. T-RF lengths (bp) are given above the map, with *in silico*-calculated sizes derived from *alkB* clone sequences in parentheses. Intensity of shading represents the percentage of occurrence in the sample group; sample size is given in parentheses. Low-G+C Gram-positive bacteria include *Firmicutes* and *Bacteroidetes* isolates.

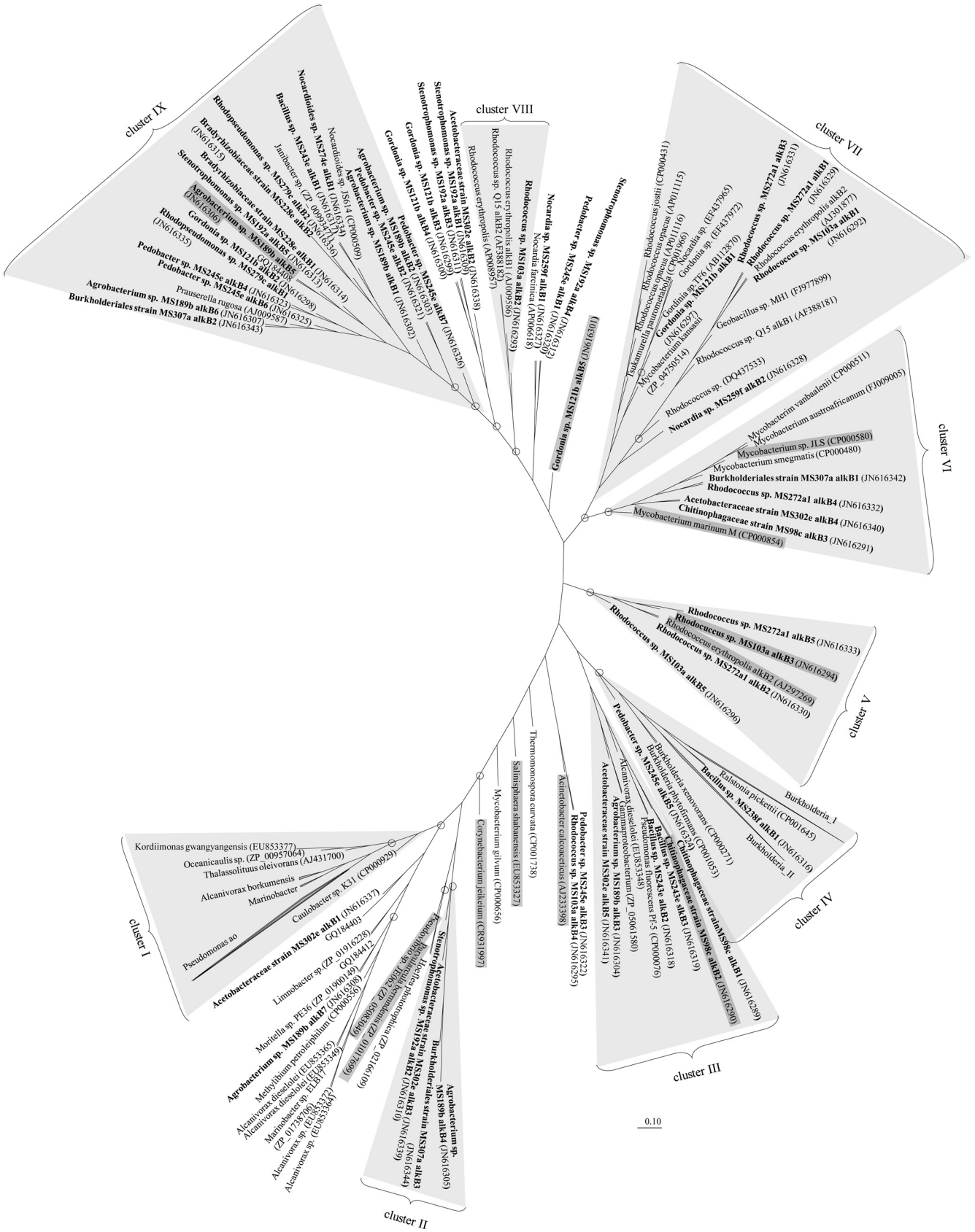

FIG 2 Phylogenetic tree of 150 partial AlkB protein sequences deduced from translated nucleotide sequences; corresponding GenBank accession numbers are given in parentheses. Tree topology was calculated with the ProML maximum-likelihood method (Jones-Taylor-Thornton substitution model, randomized input order, 500 recalculations for bootstrapping). Bootstrap values of >50% for major clusters are indicated by circles. Subclusters generally displayed high bootstrap values and are not further indicated. Gray triangles indicate the nine major AlkB clusters (defined by ≥80% protein sequence identity). AlkB sequences obtained in this study are given in boldface. Sequences yielding an exclusive T-RF are indicated by dark gray boxes. The scale bar indicates the percentage of sequence divergence.

TABLE 1 Total number of AlkB sequences per cluster in the AlkB protein dendrogram, with the relative amounts of sequences derived from *Proteobacteria*, *Actinobacteria*, or low-G+C bacteria in the nine clusters

Parameter	Value for cluster:									
	I	II	III	IV	V	VI	VII	VIII	IX	Not clustered
Total no. of sequences	23	4	9	26	5	9	17	4	21	32
% of sequences from:										
Proteobacteria	87	100	56	92	0	22	0	0	48	59
Actinobacteria	9	0	0	0	100	67	94	100	24	28
Low-G+C bacteria	4	0	44	8	0	11	6	0	24	6
% Unidentified	0	0	0	0	0	0	0	0	0	6
Dominant T-RF (bp)	550	120	—[a]	285	—	525	464	328	534	—
% of sequences displaying the dominant T-RF	48	100	—	91	—	67	29	100	57	—
No. of sequences not in the cluster displaying the same T-RF	0	1	—	9	—	11	0	0	0	—

[a] —, no dominant T-RF was observed for the respective clusters.

amplified in PCR mixtures comprising 1× *Taq* PCR master mix kit (Qiagen, Hilden, Germany), 0.12% bovine serum albumin (BSA), and 0.1 μM inosine-containing degenerate *alkB* primers (11) with modifications as described in reference 6. *alkB* gene PCR products were obtained from 34 of the 50 bacterial isolates. T-RFLP analysis was performed as described in reference 6 after overnight digestion of 10 ng of purified PCR product with 2 U HpyCH4V (New England BioLabs, Beverly, MA). Terminal restriction fragments (T-RFs) within the range of 50 to 550 bp were obtained for 27 isolates. Overall, digestion of the amplicons resulted in 20 different T-RFs, pointing to high genetic diversity of the *alkB* gene. More than one T-RF per isolate was observed in general, indicating the presence of multiple gene copies, i.e., paralogous genes as reported in former studies (7, 8, 12). Ten T-RFs occurred in more than 50% of all phylogroups differentiated (Fig. 1). Single T-RFs or T-RFLP profiles could generally not be assigned to a specific phylogenetic group, except for two rare T-RFs (213 and 324 bp).

We further cloned and sequenced *alkB* gene amplicons of 15 selected isolates from 13 different bacterial genera. The experimentally observed T-RF length generally differed from *in silico* predictions of sequenced *alkB* clones by 1 to 6 bp (Fig. 1). T-RFLP analysis of single clones further revealed that 16 clones per isolate were in most cases not sufficient to capture *alkB* sequences with low T-RF signal intensities in isolate profiles. Furthermore, T-RFs which showed low abundances in the isolate clone libraries but were not detected in the isolate profiles point to at least partial preferential amplification of certain *alkB* sequence types. Two T-RFs (87 and 95 bp) frequently detected in isolate profiles were not confirmed by *in silico* and/or experimental restriction of cloned sequences, and thus, they potentially reflect pseudo-T-RFs also known from other studies (13, 14).

Fifty-six distinct *alkB* homologous sequences were obtained. Sequence similarities to public database entries ranged from 60% to 99% on the nucleotide level and 59% to 100% on the protein level, which is consistent with published sequence analyses (10, 12, 15). All full-length amplicon sequences (550 to 560 bp) displayed histidine boxes I to IV (12). Thirty-eight distinct *alkB* gene sequences from *Proteobacteria*-, *Actinobacteria*-, *Firmicutes*-, and *Bacteroidetes*-related isolates were highly similar to sequences derived from *Actinobacteria* strains, whereas eight sequences from *Firmicutes*, *Bacteroidetes*, or *Proteobacteria* isolates were highly similar to sequences from *Proteobacteria*. The remaining sequences were most similar to *alkB* sequences from environmental clone libraries and homologous to proteins of *Pseudovibrio* sp. strain JE062 (GenBank sequence accession number ZP_05083049) or *Kordiimonas gwangyangensis* (GenBank sequence accession number ACJ22725). Comparison of a phylogenetic tree based on the AlkB amino acid alignment revealed major differences in the tree topology compared to a corresponding 16S rRNA gene tree (Fig. 2; see also Fig. S1 in the supplemental material). 16S rRNA gene sequence analysis grouped the isolates into six major taxonomic groups (phyla or classes), while the AlkB tree displayed nine major clusters. Most importantly, *alkB* sequences from taxonomically related bacteria clustered randomly and paralogous gene sequences derived from one organism were distributed all over the tree. Overall, AlkB clusters I, II, and IV were dominated by sequences derived from *Proteobacteria*, while clusters V to VIII showed a dominance of sequences from *Actinobacteria*. Cluster II comprised exclusively novel sequences that strongly diverged from previously published AlkB sequences.

Our findings on multiple *alkB* genotypes therefore differ from the findings in studies which reported on a consistency of 16S rRNA and *alkB* gene trees (16, 17). In applying partly group-specific primers to *alkB*-harboring communities with reduced diversity, *Rhodococcus*- and *Pseudomonas*-related degraders were found to host distinct types of *alkB* sequences. However, studies analyzing a broader *alkB* gene sequence diversity hypothesized a clustering according to the substrate specificity rather than to the phylogenetic origin of the bearer (4, 9). The high *alkB* sequence diversity within our isolates and the high sequence similarities (>95% nucleotide sequence identity) between sequences derived from phylogenetically distinct bacteria suggest that *alkB* sequence analyses alone do not provide unambiguous information on community composition. Our phylogenetic analysis only differentiated to some extent between *alkB* sequences derived from *Proteobacteria* and *Actinobacteria*, thus supporting the former application of group-specific primers targeting either *Pseudomonas*- or *Rhodo-*

coccus-like *alkB* sequences (7, 9). Interestingly, the *alkB* genes in *Firmicutes*- and *Bacteroidetes*-related species appear to be mainly acquired by horizontal gene transfer events from *Pseudomonas* or *Rhodococcus*, as also proposed earlier (10). These sequences clustered randomly in the tree and never built a group-specific cluster as observed for *Proteobacteria* or *Actinobacteria* (Table 1).

While many alkane degraders contain *alkB* paralogs encoded on the chromosome, alkane degradation genes have also been described as being located on large plasmids, such as the OCT plasmid in *Pseudomonas putida* GPo1 (18). The lower G+C contents of these genes compared to the G+C contents of the host strain and the plasmid (19–21) were a first clear indication of horizontal gene transfer and catabolic mobile elements (i.e., transposons) that have integrated in the OCT plasmid (22). Insertion sequences flanking *alk* gene clusters both on the OCT plasmid (*P. putida* GPo1) and the chromosome (*P. putida* P1) are likely to have played a role in mobilizing the clusters. It has been suggested that *P. putida alk* genes may in fact have originally been obtained from some closely related *Alcanivorax* strains (23). In contrast, atypical G+C contents, as well as high interspecies sequence similarities of chromosomal *alkB* genes in *Geobacillus* strains, suggest interspecific horizontal gene transfer, potentially from representatives of the phylum *Actinobacteria* (10).

In conclusion, our study revealed high *alkB* sequence diversity and the presence of a surprisingly high number of different *alkB* genes in phylogenetically diverse isolates. This indicates a genetic mobility and phylogenetic distribution of the *alkB* gene to a so-far-unknown extent (8, 12). Assigning single *alkB* gene T-RFs or T-RFLP profiles to distinct bacterial taxa is therefore strongly biased.

ACKNOWLEDGMENTS

The study was financed by the German Research Foundation (DFG) within the priority program 1315 (Biogeochemical Interfaces in Soil).

The technical support of Alexandra Marquardt and Marcus Böttcher is acknowledged.

REFERENCES

1. **Kerstiens G.** 1996. Signalling across the divide: a wider perspective of cuticular structure-function relationships. Trends Plant Sci. **1**:125–129.
2. **Eglinton G, Hamilton RJ.** 1967. Leaf epicuticular waxes. Science **156**: 1322–1335.
3. **Kolattukudy PE.** 1970. Composition of the surface lipids of pea leaves (*Pisum sativum*). Lipids **5**:398–402.
4. **van Beilen JB, Li Z, Duetz WA, Smits THM, Witholt B.** 2003. Diversity of alkane hydroxylase systems in the environment. Oil Gas Sci. Technol. **58**:427–440.
5. **Rojo F.** 2009. Degradation of alkanes by bacteria. Environ. Microbiol. **11**:2477–2490.
6. **Schulz S, Giebler J, Chatzinotas A, Wick LY, Fetzer I, Welzl G, Harms H, Schloter M.** 2012. Plant litter and soil type drive abundance, activity and community structure of alkB harbouring microbes in different soil compartments. ISME J. **6**:1763–1774.
7. **Wang W, Wang L, Shao Z.** 2010. Diversity and abundance of oil-degrading bacteria and alkane hydroxylase (*alkB*) genes in the subtropical seawater of Xiamen Island. Microb. Ecol. **60**:429–439.
8. **van Beilen JB, Funhoff EG.** 2007. Alkane hydroxylases involved in microbial alkane degradation. Appl. Microbiol. Biotechnol. **74**:13–21.
9. **Heiss-Blanquet S, Benoit Y, Marechaux C, Monot F.** 2005. Assessing the role of alkane hydroxylase genotypes in environmental samples by competitive PCR. J. Appl. Microbiol. **99**:1392–1403.
10. **Tourova TP, Nazina TN, Mikhailova EM, Rodionova TA, Ekimov AN, Mashukova AV, Poltaraus AB.** 2008. alkB homologs in thermophilic bacteria of the genus *Geobacillus.* Mol. Biol. **42**:217–226.
11. **Kloos K, Munch JC, Schloter M.** 2006. A new method for the detection of alkane-monooxygenase homologous genes (*alkB*) in soils based on PCR-hybridization. J. Microbiol. Methods **66**:486–496.
12. **Smits THM, Rothlisberger M, Witholt B, Van Beilen JB.** 1999. Molecular screening for alkane hydroxylase genes in Gram-negative and Gram-positive strains. Environ. Microbiol. **1**:307–317.
13. **Egert M, Friedrich MW.** 2003. Formation of pseudo-terminal restriction fragments, a PCR-related bias affecting terminal restriction fragment length polymorphism analysis of microbial community structure. Appl. Environ. Microbiol. **69**:2555–2562.
14. **Matsumoto M, Sakamoto M, Hayashi H, Benno Y.** 2005. Novel phylogenetic assignment database for terminal-restriction fragment length polymorphism analysis of human colonic microbiota. J. Microbiol. Methods **61**:305–319.
15. **Powell SM, Bowman JP, Ferguson SH, Snape I.** 2010. The importance of soil characteristics to the structure of alkane-degrading bacterial communities on sub-Antarctic Macquarie Island. Soil Biol. Biochem. **42**:2012–2021.
16. **Hamamura N, Fukui M, Ward DM, Inskeep WP.** 2008. Assessing soil microbial populations responding to crude-oil amendment at different temperatures using phylogenetic, functional gene (*alkB*) and physiological analyses. Environ. Sci. Technol. **42**:7580–7586.
17. **Kuhn E, Bellicanta GS, Pellizari VH.** 2009. New *alk* genes detected in Antarctic marine sediments. Environ. Microbiol. **11**:669–673.
18. **Chakrabarty AM, Chou G, Gunsalus IC.** 1973. Genetic regulation of octane dissimilation plasmid in *Pseudomonas.* Proc. Natl. Acad. Sci. U. S. A. **70**:1137–1140.
19. **Fennewald M, Prevatt W, Meyer R, Shapiro J.** 1978. Isolation of Inc P-2 plasmid DNA from *Pseudomonas aeruginosa.* Plasmid **1**:164–173.
20. **Kok M, Oldenhuis R, Van der Linden MPG, Raatjes P, Kingma J, Van Lelyveld PH, Witholt B.** 1989. The *Pseudomonas oleovorans* alkane hydroxylase gene. Sequence and expression. J. Biol. Chem. **264**:5435–5441.
21. **Panke S, Meyer A, Huber CM, Witholt B, Wubbolts MG.** 1999. An alkane-responsive expression system for the production of fine chemicals. Appl. Environ. Microbiol. **65**:2324–2332.
22. **van Beilen JB, Panke S, Lucchini S, Franchini AG, Rothlisberger M, Witholt B.** 2001. Analysis of *Pseudomonas putida* alkane-degradation gene clusters and flanking insertion sequences: evolution and regulation of the *alk* genes. Microbiology **147**:1621–1630.
23. **van Beilen JB, Marin MM, Smits THM, Rothlisberger M, Franchini AG, Witholt B, Rojo F.** 2004. Characterization of two alkane hydroxylase genes from the marine hydrocarbonoclastic bacterium *Alcanivorax borkumensis.* Environ. Microbiol. **6**:264–273.