APPLIED AND ENVIRONMENTAL MICROBIOLOGY, Jan. 2006, p. 384–391 Vol. 72, No. 1
0099-2240/06/$08.00+0 doi:10.1128/AEM.72.1.384–391.2006

Isolation of Poly-3-Hydroxybutyrate Metabolism Genes from Complex Microbial Communities by Phenotypic Complementation of Bacterial Mutants

Chunxia Wang,[1]† David J. Meek,[2] Priya Panchal,[1] Natalie Boruvka,[1] Frederick S. Archibald,[2,3] Brian T. Driscoll,[2] and Trevor C. Charles[1]*

Department of Biology, University of Waterloo, 200 University Avenue West, Waterloo, Ontario N2L 3G1, Canada[1]; Department of Natural Resource Sciences, McGill University, 21,111 Lakeshore Road, Ste-Anne-de-Bellevue, Quebec H9X 3V9, Canada[2]; and Pulp and Paper Research Institute of Canada, 570 St-John's Blvd., Pointe-Claire, Quebec H9R 3J9, Canada[3]

Received 21 June 2005/Accepted 10 October 2005

The goal of this study was to initiate investigation of the genetics of bacterial poly-3-hydroxybutyrate (PHB) metabolism at the community level. We constructed metagenome libraries from activated sludge and soil microbial communities in the broad-host-range IncP cosmid pRK7813. Several unique clones were isolated from these libraries by functional heterologous complementation of a *Sinorhizobium meliloti bdhA* mutant, which is unable to grow on the PHB cycle intermediate D-3-hydroxybutyrate due to absence of the enzyme D-3-hydroxybutyrate dehydrogenase activity. Clones that conferred D-3-hydroxybutyrate utilization on *Escherichia coli* were also isolated. Although many of the *S. meliloti bdhA* mutant complementing clones restored D-3-hydroxybutyrate dehydrogenase activity to the mutant host, for some of the clones this activity was not detectable. This was also the case for almost all of the clones isolated in the *E. coli* selection. Further analysis was carried out on clones isolated in the *S. meliloti* complementation. Transposon mutagenesis to locate the complementing genes, followed by DNA sequence analysis of three of the genes, revealed coding sequences that were broadly divergent but lay within the diversity of known short-chain dehydrogenase/reductase encoding genes. In some cases, the amino acid sequence identity between pairs of deduced BdhA proteins was <35%, a level at which detection by nucleic acid hybridization based methods would probably not be successful.

It is now well recognized that some of the more interesting and influential physiological processes in microbial communities are carried out by organisms that have not been isolated in pure culture. In recent years, many studies have involved the construction of gene libraries from DNA isolated directly from microbial communities without prior culture (40; reviewed in reference 39). Genes of interest have been identified from these libraries by phenotypic screens, usually in the *Escherichia coli* surrogate host, or by direct sequence analysis of the clones (42, 47). In some cases, the sequences of complete genomes have even been determined from complex libraries of this type (46, 47).

As part of our efforts to understand metabolic interactions in microbial communities, we are developing the use of phenotypic selection techniques as alternatives to physical, sequence-based approaches in the isolation and identification of specific metabolic genes from microbial communities in uncultivated environmental samples. The use of gene selection based on function, rather than sequence homology, has the potential for isolation of genes that are truly novel (22). The immense power of phenotypic selection methods, especially in comparison to screening, is well known in microbial genetics.

* Corresponding author. Mailing address: Department of Biology, University of Waterloo, 200 University Ave. West, Waterloo, ON N2L 3G1, Canada. Phone: (519) 888-4567, ext. 5606. Fax: (519) 746-0614. E-mail: tcharles@uwaterloo.ca.

† Present address: Virginia Bioinformatics Institute, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061.

Polyhydroxyalkanoate (PHA) deposits are accumulated by many different species of bacteria, and these deposits act as carbon stores to aid the cells in surviving nutrient carbon starvation conditions (for a comprehensive review see Madison and Huisman [33]). The most common PHA is poly-3-hydroxybutyrate (PHB). The structural diversity and range of physical properties of the known PHA polymers is very great, with >90 different constituent hydroxyalkanoic acids described (43). These compounds are of commercial interest, since they can be developed as substitutes for fossil fuel-derived polymers.

Intracellular PHA accumulation occurs when the growth of bacterial cells is inhibited by nutrient or physical limitation while carbon sufficiency is maintained. The microbial communities in pulp effluent treatment systems periodically experience occasions of such unbalanced growth conditions, as do soil bacteria in the carbon-rich rhizosphere. PHA metabolism would thus be expected to be important for the physiology of microbial communities in these systems. Although PHA synthesis has been thoroughly investigated due to the industrial interests in PHA production, a good understanding of the genetics and biochemistry of the complete PHA cycle is still lacking. Investigations in the model soil bacterium *Sinorhizobium meliloti* (2, 4, 9, 11, 12) have resulted in the mutation, isolation and investigation of most of the PHB cycle genes. As well as supporting investigation of the biological role of PHB in this organism, this genetic material has facilitated the isolation of corresponding PHB metabolism genes such as *bdhA* (encoding D-3-hydroxybutyrate dehydrogenase [EC 1.1.1.30]) and *phbC* (encoding PHB synthase) from other genomes (4, 5).

TABLE 1. Bacterial strains and plasmids

Strain, plasmid, or transposon	Relevant characteristics[a]	Reference or source
Strains		
Sinorhizobium meliloti		
Rm1021	SU47 *str-21* (Sm^r)	34
Rm11107	Rm1021 *bdhA*$_1$::Tn*5*	11
Rm11196	Rm1021 *bdhA*$_2$::Tn*5 shb-3*	3
Escherichia coli		
HB101	*supE44 hsdS20* (r_B^- m_B^-) *recA13 ara-14 proA2 lacY1 galK2 rpsL20 xyl-5 mtl-1 leuB6 thi-1*	8
LS5218	*fadR601 atoC200* (constitutive)	41
DH5α	F^- *endA1 hsdR17* (r_K^- m_K^+) *supE44 thi-1 recA1 gyrA96 relA1* Δ(*argF-lacZYA*)*U169* ϕ80d*lacZ*ΔM15, λ^-	21
DH5α(pRK600)	Mobilizing strain; Cm^r	18; this study
Plasmids		
pRK7813	IncP cosmid cloning vector; Tc^r	24
pGEM-TEasy	T-vector; Ap^r	Promega
pBSKS+	Cloning vector, ColE1 *oriV*, *bla*	Stratagene
pUC18/19	Cloning vector, ColE1 *oriV*, *bla*	50
pRCX series	pRK7813 clones, complementing *bdhA* in Rm11107	This study
pLCX series	pRK7813 clones, complementing LS5218 for growth on D-3-hydroxybutyrate	This study
Transposon		
EZ::TN <KAN-2>	Km^r	EPICENTRE Technologies

[a] Abbreviations for antibiotics: Ap, ampicillin; Cm, chloramphenicol; Km, kanamycin; Sm, streptomycin; Tc, tetracycline.

Many of the bacterial genes present in community libraries will not be expressed in a given surrogate host due to transcriptional control requirements and codon preference. To isolate genes based on their expression, it would be desirable to use vectors that are able to replicate in a broad range of bacteria. Relatively few of the large diversity of cultivated bacteria have genetic systems sufficiently developed to allow use as a surrogate host, and vectors that are able to replicate across the complete range of cultivated bacteria are not available. Nevertheless, the requirements for gene expression are sufficiently different even within the *Proteobacteria* to justify constructing libraries that can replicate and be maintained in diverse *Proteobacteria*.

IncP vectors have been extensively used within the *Proteobacteria*. We wanted to isolate genes by complementation of PHB cycle mutants in *S. meliloti*, which is well characterized (19) and genetically amenable (20). We therefore took advantage of a cosmid vector with an IncP origin of replication to construct the microbial community libraries for the present study. The described libraries should be useful for selection or screening in any strain that supports the replication of IncP plasmids and in which tetracycline resistance can be used as selectable marker.

In the present study, we describe the construction of IncP cosmid libraries from soil communities and from wastewater treatment activated sludge communities. We demonstrate the successful complementation of *S. meliloti* mutants to isolate clones from this library that contain PHB metabolism genes. Finally, we present the initial biochemical characterization of the enzymes encoded by these genes and DNA sequence characterization of some of the clones. This work is the first step in our study of the genetics of carbon metabolism at the community level.

MATERIALS AND METHODS

Strains, plasmids, and bacterial culture. Bacterial strains and plasmids are listed in Table 1. Culture methods in TY, LB, LBmc, and modified M9 minimal medium, and the concentrations of carbon sources and antibiotics, were as described previously (2, 9, 10, 11). Measurement of growth on D-3-hydroxybutyrate was carried out as previously described (3).

Molecular biology. Standard techniques were used for molecular biology. A Visible Genetics Long Read Tower system (Visible Genetics, Toronto, Ontario, Canada) was used for determination of DNA sequence. Total microbial DNA was isolated from an organic rich, sandy loam surface soil that was collected from the banks of Laurel Creek on the Campus of the University of Waterloo, Waterloo, Ontario, Canada. The Ultra Clean Soil DNA Kit (Mo Bio Laboratories, Inc., Carlsbad, CA) was used by the method suggested by the manufacturer but with the addition of a final purification step that consisted of one time phenol-chloroform, followed by two chloroform extractions. After incubation with RNase for 20 min at 37°C, the DNAs were precipitated with isopropanol, washed with 75% ethanol twice, and dissolved into $T_{10}E_1$ buffer (10 mM Tris-Cl, 1 mM Na-EDTA [pH 8.0]). DNA from wastewater communities was from a previously published study (35). Cosmid libraries were constructed by cloning fragments from BamHI partial digests into the BamHI site of the IncP Tc^r plasmid pRK7813 (24), followed by packaging with Gigapack III XL Lambda packaging extract (Stratagene) and transduction of *E. coli* HB101 to Tc^r. Representative clones from each library were analyzed by restriction digest to ensure that members of each library contained insert DNA. Colonies from single transductions were pooled and subcultured, and the resulting libraries were maintained as aliquots in LB containing 7% dimethyl sulfoxide at −70°C. For phylogenetic analysis, an ~500-bp internal segment of small subunit rRNA encoding gene was amplified. A set of primers specific for bacterial 16S rRNA gene sequences was used for amplification of a portion of the 16S gene (25). These primers were designated b341 (5′-CCTACGGGAGGCAGCAG-3′) and b758 (5′-CTACCAGGGTATCTAATCC-3′). DNA from the pooled libraries was treated with the Plasmid-Safe ATP Dependent DNase (EPICENTRE Technologies) to minimize contamination by *E. coli* host DNA prior to amplification.

TABLE 2. Characteristics of the environmental clone libraries

Library	DNA source[a]	No. of clones	No. of unique clones from LS5218 complementation[b]	No. of unique clones from Rm11107 complementation[b]
CX2	Pulp	260		
CX3	Pulp	2,544		1
CX4	Pulp	3,879	5	7
CX5	Pulp	2,049		
CX6	Municipal	3,322		
CX7	Soil	1,383		
CX8	Soil	1,315		
CX9	Soil	22,180	2	14
CX10	Soil	8,698	2	3

[a] Activated sludge DNA was from Domtar Cornwall (pulp) and Vaudreuil (municipal) treatment systems (35). Soil samples were collected from banks of Laurel Creek, University of Waterloo Campus.
[b] Unique clones were determined by comparison of restriction digests.

PCR-amplified fragments from the community genomic DNA samples or from the pooled libraries were cloned into pGEM-T Easy Vector (Promega). Individual clones were initially characterized by Sau3AI, HaeIII, CfoI, and MspI restriction enzyme digestion of reamplified fragments (amplified ribosomal DNA restriction analysis [ARDRA]), and based on comparative analysis of restriction patterns individual representative clones were sequenced by using vector primers. The RDP II database (15) was used for placement of the 16S sequences on the bacterial phylogenetic tree. The DNA sequences have been deposited in GenBank under accession numbers AY836590 to AY836610.

Genetic techniques. Triparental conjugation, using DH5α containing the mobilizing plasmid pRK600 (18), was used to transfer pRK7813-based clones between *E. coli* and *S. meliloti* strains. For isolation of clones by phenotypic complementation, selection was carried out on the selective media, followed by streak purification. Plasmids were transferred by triparental conjugation or transformation into *E. coli* DH5α for physical characterization of the cloned DNA. In each case, the complementation ability of the clone was confirmed by reintroduction from DH5α back into the mutant, followed by testing for phenotype. Mapping of the clones was carried out by using the in vitro transposon EZ::TN <KAN-2> (EPICENTRE Technologies), which also facilitated much of the DNA sequencing.

Biochemical assay and enzyme activity staining. Preparation of *S. meliloti* cell extracts, protein determination, the assay for D-3-hydroxybutyrate dehydrogenase activity (7), and nondenaturing polyacrylamide electrophoresis were carried out essentially as described previously (2, 11, 45). Crude cell extracts of the *E. coli* strains were obtained as follows. The pelleted cells were washed and suspended in the same buffers and manner as for *S. meliloti* cells above. To the cell-buffer solution, 1-mm zirconia/silica beads (Biospec Products, Inc., Bartlesville, OK) were added in a 2:1 (beads to solution) ratio. Cells were subjected to four cycles of 5 min of shaking in a bead beater at 4°C (type 2876, no. 790; Bronwill Scientific, Inc., Rochester, NY) followed by 5 min on ice. This was followed by centrifugation at 12,000 rpm (SS34 rotor) for 20 min to remove cell debris. As positive control, all extracts were assayed for malate dehydrogenase activity (17).

RESULTS

Construction of microbial community libraries. The plasmid pRK7813 (24) was chosen as the vector for library construction because its IncP replicon allows it to replicate in a broad range of *Proteobacteria*, and the *cos* sites facilitate phage lambda packaging. The libraries that were constructed in the present study are summarized in Table 2. A total of 8,732 clones were obtained from pulp effluent-activated sludge DNA, 3,322 clones were from the municipal system DNA, and 33,576 clones were from the soil DNA (Table 2). The predicted amount of DNA contained in these library types is ca. 288, 110, and 1,108 Mb, respectively, based on an average insert size of 33 kb found in our restriction analysis of randomly chosen clones (data not shown). We carried out 16S rRNA sequence analysis to ensure that the libraries were not dominated by DNA from a limited number of organisms. The internal 16S primers were used to amplify from the original DNA samples, as well as from the largest pooled soil library CX9, and 129 clones were analyzed by ARDRA (data not shown). A total of 21 clones (11 library, 10 direct) were selected as representative of the different ARDRA groups. Sequence determination was carried out on these representative clones, and their phylogenetic positions were determined by comparison with the RDPII database (15) (Table 3). In one instance (N25 and NL1), identical sequence was obtained between a library clone and a direct clone. Both of these fell within the same ARDRA group.

TABLE 3. Classification of 16S rRNA gene sequences from soil by comparison to the RDPII database[a]

Sequence[b]	Phylum	Class	Order	Family	Genus
N1	*Proteobacteria* (100)	*Alphaproteobacteria* (100)	*Rhizobiales* (100)	*Bradyrhizobiaceae* (100)	*Bradyrhizobium* (95)
N2	*Verrucomicrobia* (79)	*Verrucomicrobiae* (79)	*Verrucomicrobiales* (79)	*Verrucomicrobiaceae* (75)	*Verrucomicrobium* (75)
N5	*Actinobacteria* (92)	*Actinobacteria* (92)	*Coriobacteriales* (38)	*Coriobacteriaceae* (38)	*Slackia* (15)
N10	*Bacteroidetes* (100)	*Sphingobacteria* (70)	*Sphingobacteriales* (70)	*Flexibacteraceae* (63)	*Sporocytophaga* (29)
N12	*Proteobacteria* (100)	*Alphaproteobacteria* (100)	*Sphingomonadales* (100)	*Sphingomonadaceae* (100)	*Sphingomonas* (100)
N25 and NL1	*Bacteriodetes* (98)	*Sphingobacteria* (89)	*Sphingobacteriales* (89)	*Flexibacteraceae* (87)	*Sporocytophaga* (62)
N30	*Proteobacteria* (100)	*Alphaproteobacteria* (100)	*Rhizobiales* (100)	*Hyphomicrobiaceae* (87)	*Rhodoplanes* (80)
N34	*Proteobacteria* (100)	*Alphaproteobacteria* (100)	*Rhodospirillales* (92)	*Rhodospirillaceae* (51)	*Azospirillum* (28)
N39	*Firmicutes* (100)	*Bacilli* (100)	*Bacillales* (100)	*Bacillaceae* (99)	*Bacillus* (99)
N54	*Proteobacteria* (58)	*Deltaproteobacteria* (36)	*Desulfobacterales* (18)	*Desulfobacteraceae* (16)	*Desulfococcus* (12)
NL4	*Proteobacteria* (100)	*Alphaproteobacteria* (100)	*Rhizobiales* (100)	*Rhizobiaceae* (66)	*Sinorhizobium* (66)
NL14	*Actinobacteria* (84)	*Actinobacteria* (84)	*Actinomycetales* (55)	*Acidothermaceae* (26)	*Acidothermus* (26)
NL23	*Actinobacteria* (100)	*Actinobacteria* (100)	*Actinomycetales* (100)	*Intrasporangiaceae* (99)	*Janibacter* (73)
NL28	*Actinobacteria* (90)	*Actinobacteria* (90)	*Actinomycetales* (46)	*Nocardioidaceae* (15)	*Kribbella* (15)
NL30	*Proteobacteria* (98)	*Alphaproteobacteria* (94)	*Rhizobiales* (35)	*Hyphomicrobiaceae* (21)	*Rhodomicrobium* (21)
NL33	*Proteobacteria* (98)	*Alphaproteobacteria* (79)	*Rhizobiales* (73)	*Brucellaceae* (50)	*Mycoplana* (50)
NL45	*Proteobacteria* (100)	*Betaproteobacteria* (100)	*Burkholderiales* (100)	*Comamonadaceae* (60)	*Acidovorax* (43)
NL46	*Proteobacteria* (100)	*Alphaproteobacteria* (100)	*Rhodospirillales* (96)	*Rhodospirillaceae* (76)	*Azospirillum* (39)
NL55	*Proteobacteria* (100)	*Betaproteobacteria* (93)	*Burkholderiales* (76)	*Incertaesedis* (58)	*Schlegelella* (47)
NL67	*Verrucomicrobia* (74)	*Verrucomicrobiae* (74)	*Verrucomicrobiales* (74)	*Verrucomicrobiaceae* (67)	*Verrucomicrobium* (67)

[a] Analysis done on 27 November 2004, with Naïve Bayesian rRNA Classifier version 1.0, November 2003 (Taxonomical Hierarchy Bergey's Manual of Systematic Bacteriology, vetted sequences, release 0.9). The confidence level is indicated in parentheses.
[b] N sequences are from direct amplification of DNA extracted from soil sample; NL sequences are from amplification of the soil library CX9.

Several of the sequenced clones appear to be from organisms that are representative of soil communities, including the α, β, and δ *Proteobacteria*; high and low GC gram positive; *Verrucomicrobia*; and *Cytophaga*.

Isolation of clones complementing for D-3-hydroxybutyrate utilization. In previous studies (2, 5), we have used two different strategies to isolate genes involved in the first step in D-3-hydroxybutyrate degradation, carried out by the enzyme D-3-hydroxybutyrate dehydrogenase (EC 1.1.1.30). This enzyme is a member of the short-chain dehydrogenase/reductase (SDR) family (26). The first strategy involves the phenotypic complementation of the *S. meliloti bdhA* mutant Rm11107, which is unable to synthesize D-3-hydroxybutyrate and thus cannot form colonies on defined medium containing D-3-hydroxybutyrate as the sole carbon source. The second strategy involves selection for the gain of the ability to utilize D-3-hydroxybutyrate as the sole carbon source by the *E. coli* strain LS5218. This *E. coli* strain carries the *atoC*^{c} mutation (41), resulting in constitutive expression of the *ato* genes that are required for growth on acetoacetate as the sole carbon source. *E. coli* naturally lacks the ability to oxidize D-3-hydroxybutyrate, but synthesis of D-3-hydroxybutyrate dehydrogenase from exogenous genes results in production of acetoacetate, which is further degraded to acetyl-coenzyme A by the *ato*-encoded system. Due to different requirements for gene expression in *S. meliloti* compared to *E. coli*, it was expected that the two approaches would be complementary and that some clones that would be isolated by using one strategy would not be isolated by using the other strategy. A number of clones were isolated from each of the libraries CX3, CX4, CX9, and CX10, using both strategies, and unique clones were identified by restriction enzyme digest (Table 2).

Phenotypic characterization of complementing clones. Based on the restriction enzyme digest results, 25 representative unique clones from the Rm11107 complementation and 9 representative unique clones from the LS5218 complementation were (re)introduced into both Rm11107 and LS5218, and the resulting transconjugant strains were tested for growth on M9 D-3-hydroxybutyrate (Table 4 and Fig. 1). We were surprised to find that of the 25 unique clones isolated by phenotypic selection in *S. meliloti* Rm11107, only 1 (pRCX24) was able to complement *E. coli* LS5218 for growth on D-3-hydroxybutyrate. None of the clones isolated in LS5218 were able to complement *S. meliloti* Rm11107. Thus, despite our previous demonstration of the isolation of *bdhA*-containing clones from *Sinorhizobium* NGR234 by complementation in both *S. meliloti* and *E. coli* (5), the clones isolated from the metagenomic libraries for the most part were only able to confer D-3-hydroxybutyrate utilization ability on the host that was used for the original selection.

Cell extracts of LBmc (for *S. meliloti*) or LB (for *E. coli*) broth cultures of the representative clones in their cognate hosts were assayed for the expected D-3-hydroxybutyrate dehydrogenase activity using both the spectrophotometric assay and the native gel activity stain (Table 4). In almost all cases, the clones from the Rm11107 selection in the cognate background exhibited the enzyme activity. However, none of these clones exhibited the enzyme activity in the LS5218 background (data not shown). Furthermore, none of the clones from the LS5218 selection exhibited detectable enzyme activity in either background (data not shown). For the Rm11107 clones, there was quite a range in enzyme activity, with the most active clone pRCX32 giving activity sixfold higher than that of the *S. meliloti* wild-type Rm1021 control (Table 4).

TABLE 4. Specific D-3-hydroxybutyrate dehydrogenase activities of crude cell extracts of LBmc grown cultures, native gel activity, and growth phenotype on D-3-hydroxybutyrate minimal media

Strain[a]	Library source	Mean enzyme activity (nmol/min/mg protein)[b] ± SEM	Native gel activity	M9 + D-3-HB (5 mM)[c]	
				Generation time (h)	On plate
Rm1021	NA	83.6 ± 3.7	Yes	19.5	++
Rm11107	NA	6.4 ± 0.0	No		−
RCX2	CX3	10.7 ± 1.1	No	84.9	+
RCX6	CX4	109.9 ± 8.4	Yes	45.3	++
RCX7	CX9	117.4 ± 4.0	No	43.9	++
RCX8	CX9	21.4 ± 1.9	Yes	55.1	++
RCX9	CX10	63.8 ± 1.4	Yes	157.7	+
RCX10	CX4	10.4 ± 0.5	Yes	45.6	++
RCX11	CX4	6.4 ± 0.0	No	22.9	+
RCX12	CX4	15.0 ± 0.5	No	27.4	+
RCX13	CX4	20.9 ± 2.5	No	37.6	+
RCX14	CX4	136.7 ± 9.7	Yes	37.0	++
RCX15	CX4	3.8 ± 2.3	No	20.2	+
RCX18	CX9	83.1 ± 1.4	No	41.7	++
RCX19	CX9	79.3 ± 6.0	No	33.1	++
RCX20	CX9	293.1 ± 18.2	Yes	38.8	++
RCX21	CX9	91.1 ± 7.5	No	46.8	++
RCX23	CX9	95.4 ± 21.4	No	31.1	++
RCX24	CX9	293.1 ± 18.0	No	63.2	++
RCX25	CX9	144.7 ± 5.6	Yes	28.1	++
RCX26	CX9	20.9 ± 1.9	Yes	66.9	++
RCX27	CX9	18.2 ± 0.5	Yes	64.4	+
RCX28	CX9	95.4 ± 1.9	Yes	37.1	++
RCX30	CX9	207.9 ± 1.9	Yes	30.8	++
RCX31	CX9	102.4 ± 5.3	Yes	46.1	++
RCX32	CX10	512.3 ± 31.1	Yes	52.1	++
RCX33	CX10	33.8 ± 0.9	No	61.7	+/−

[a] The RCX strain designation indicates Rm11107 containing the corresponding pRCX plasmid.
[b] Values represent average of triplicate assays ± standard error of the mean.
[c] The optical density was measured at 600 nm after 3 days growth at 30°C. ++, large colony; +, medium-sized colony; +/−, very small colony; −, no colony. D-3-HB, D-3-hydroxybutyrate.

The growth kinetics of the transconjugant strains were examined in liquid media (Table 4). There did not appear to be a correlation between enzyme activity of LBmc grown cultures and growth rate on D-3-hydroxybutyrate as a carbon source. It is possible that several of the enzyme activities might be inducible at transcriptional or posttranscriptional level.

DNA sequence analysis and comparison to known BdhA sequences. Three clones originating from the Rm11107 complementation, pRCX32 and pRCX23 from the soil libraries CX10 and CX9, respectively, and pRCX6 from the pulp sludge library CX4 were chosen for partial DNA sequence analysis. pRCX32 is remarkable for its sixfold-higher level of D-3-hydroxybutyrate dehydrogenase activity compared to the Rm1021 control. pRCX23 had relatively high enzyme activity, but no activity was detectable in the native gel assay. Mutagenesis was carried out with the EZ::TN <KAN-2> transposon, and individual mutant clones were screened for inability to complement the *bdhA* mutants Rm11107 or Rm11196 for growth on M9 D-3-hydroxybutyrate. DNA sequence was ob-

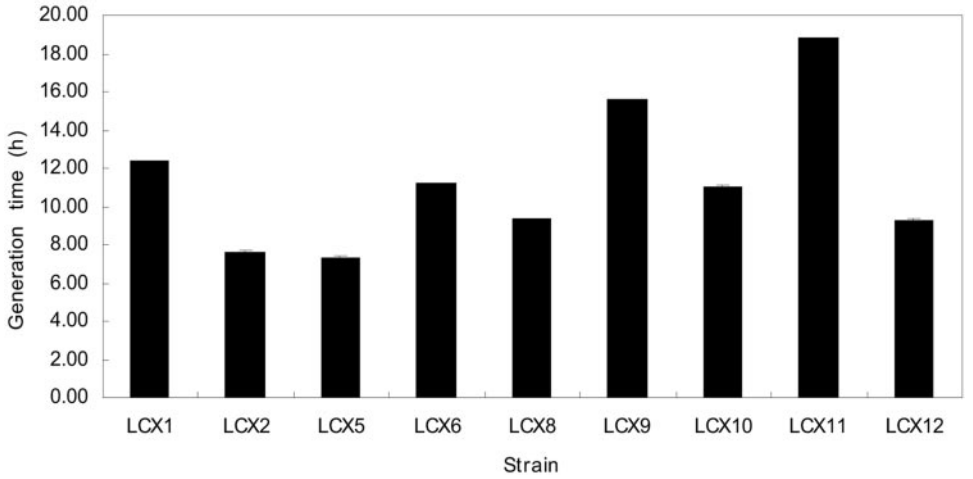

FIG. 1. Generation times of *E. coli* LS5218 carrying clones allowing growth on M9 D-3-hydroxybutyrate. No growth was observed for LS5218 alone.

tained from the sites of transposon insertion and from subclones and then assembled into contigs.

The three predicted coding sequences of $bdhA_{RCX6}$, $bdhA_{RCX23}$, and $bdhA_{RCX32}$ encode proteins of 260, 260, and 262 amino acids, corresponding to molecular weights of 273, 273, and 274, respectively. The sequences have been deposited in GenBank under accession numbers AY692352, AY692351, and AY692350, respectively. Comparison of the sequences with the NCBI database using BlastP (1) identified each of the three genes as members of the SDR family. The top hits were to sequences from *Azospirillum brasilense* (AAM00195) (16), *Legionella pneumophila* (AAU28378, locus lpg2316) (14), and *Bradyrhizobium japonicum* (BAC46753, locus blr1488) (28), respectively. Although each of these sequences is predicted to encode an SDR protein, the experimental evidence for substrate specificity has not been published.

Prior to this study, the DNA sequences of four experimentally determined D-3-hydroxybutyrate dehydrogenase-encoding genes had been published, from *S. meliloti* (O86034) (2), *Sinorhizobium* sp. strain NGR234 (AAO66470) (5), *Methylobacterium extorquens* (AAQ95632) (29), and *Rhodobacter* sp. strain DSMZ 12077 (AAD42688) (30). The deduced amino acid sequences of these four sequences and our three metagenomic sequences were analyzed with Pratt version 2.1 (23) by using the web interface at EMBL-EBI (http://www.ebi.ac.uk/pratt). The following common motif was based on that analysis: A-S-x (2)-K-x(0,1)-A-x(0,1)-Y-V-[AST]-A-K-H-G-[ILV]-x-G-[FL]-[AT]-K-[TV]-[ATV]-A-x-E-x-A-x (3)-[IV]-x-[ACV]-N-[ACSV]-I-[CS]-P-G-[FWY]-V-x-T-P-L-V-[EQ]. This motif was located in the region corresponding to positions 146 to 192 of the *S. meliloti* BdhA sequence. The motif was searched against the NR-AA database, using the MOTIF tool at GenomeNet (http://motif.genome.jp, 28 November 2004). Eight additional gene sequences were identified by this analysis, several of which had already been annotated as encoding D-3-hydroxybutyrate dehydrogenase activity, presumably based on sequence similarity to the known genes. These included one additional sequence from *B. japonicum* (BAC52294, locus blr7029) (28), one sequence from the *Agrobacterium tumefaciens* circular chromosome (AAL43297, locus Atu2308) (49), and a megaplasmid-borne sequence from *Sinorhizobium* sp. strain NGR234 (AAQ87462) (44), which is different from the one shown to be required for use of D-3-hydroxybutyrate as carbon source (5). Other sequences were from *Caulobacter crescentus* (F87668) (36), *Mesorhizobium loti* (NP_103755, locus mlr2400) (27), *Rhodopseudomonas palustris* (CAE29647, locus RPA4206) (31), *Bordetella parapertussis* (CAE38726) (37), and *Ralstonia eutropha* (Q9X642). All of these sequences, plus the *L. pneumophila* sequence (which was not identified by the motif search), were aligned by using ClustalX 1.83 (13). *E. coli* FabG (P25716) (38), a member of the SDR family, was used as outgroup. Percent identity was as low as 33 to 35 in some cases, which compares with the 25 to 35% identity commonly found between SDR paralogs of differing substrate specificity, as exemplified by the percent identity of 31 to 35 between FabG and the BdhA sequences. The resulting tree (Fig. 2) demonstrates that $BdhA_{RCX32}$ clusters with most of the experimentally determined BdhA, whereas $BdhA_{RCX6}$ groups with a clade that includes the *R. eutropha* and *A. brasilense* orthologs. $BdhA_{RCX23}$ defines a new clade with the sequences from *L. pneumophila* and *M. loti*, neither of which has been confirmed experimentally to be BdhA.

Each of the sequences identified by our motif was also identified as being a member of the TIGR01963 protein family PHB_DH. Other members of TIGR01963 were not identified by our motif, but for none of these has the enzyme activity been determined experimentally. As additional D-3-hydroxybutyrate dehydrogenase activities become known, the motif might have to be modified to accommodate them. It should be noted that the motif only identified proteins encoded by members of the *Proteobacteria*, even though it is likely that some nonproteobacterial bacteria for which genome sequence information is available exhibit D-3-hydroxybutyrate dehydrogenase activity. It will be interesting to have experimentally determined D-3-hydroxybutyrate dehydrogenase encoding genes from nonproteobacterial bacteria with which to attempt to determine a more universal motif for this enzyme.

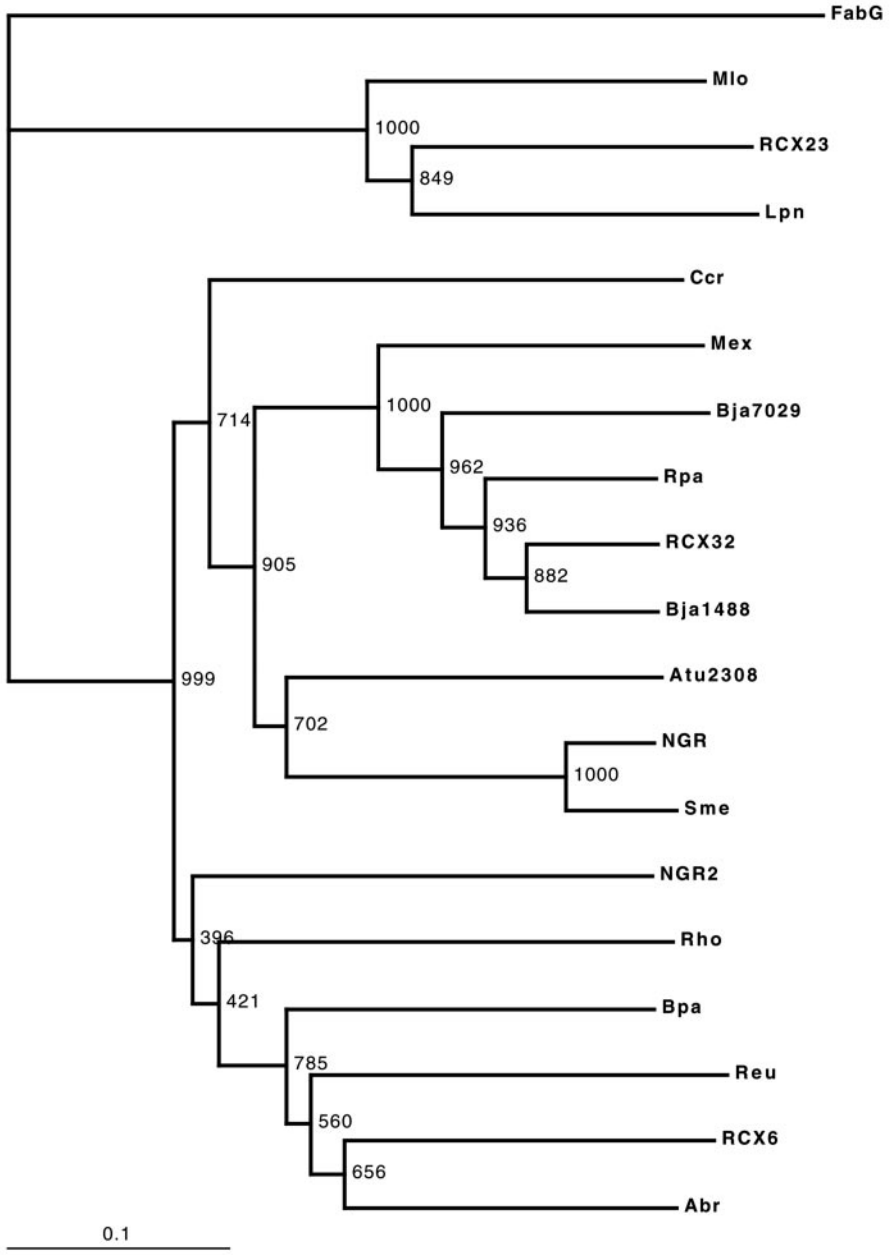

FIG. 2. N-J tree of deduced BdhA proteins based on ClustalX alignments, rooted using *E. coli* FabG (P25716) (38). Accession numbers of the BdhA sequences, described in the text, are as follows: Mlo (*M. loti*), NP_103755; RCX23, AY692351; Lpn (*L. pneumophila*), AAU28378; Ccr (*C. crescentus*), F87668; Mex (*M. extorquens*), AAQ95632; Bja7029 (*B. japonicum* blr7029), BAC52294; Rpa (*R. palustris*), CAE29647; RCX32, AY692350; Bja1488 (*B. japonicum* blr1488), BAC46753; Atu2308 (*A. tumefaciens*), AAL43297; NGR (NGR234 chromosomal locus), AA066470; Sme (*S. meliloti*), O86035; NGR2 (NGR234 megaplasmid locus), AAQ87462, Rho (*Rhodobacter* sp.), AAD42688; Bpa (*B. parapertussis*), CAE38726; Reu (*R. eutropha*), Q9X642; RCX6, AY692352; Abr (*A. brasilense*), AAM00195. Bootstrap analysis (1,000 trials) provided the indicated levels of support.

DISCUSSION

BdhA is a member of the SDR family of enzymes. Members of this family are typically proteins of 250 to 350 residues that can share as little as 15% residue identity but have strikingly high similarities in tertiary structure (26). The characterized members of the SDR family are homodimeric or homotetrameric and catalyze NAD(H)/NAD(P)(H)-dependent oxidation/reduction reactions. Surprisingly, the substrate specificities have been experimentally determined for only very few of those encoded in bacterial genomes. Members of the SDR family are likely important contributors to the catabolic capacity of bacteria in natural environments such as soil and wastewater treatment systems. The SDR-mediated oxidation of specific substrates is a critical step in the utilization of these substrates as nutrients. From available genome sequences, it is clear that a considerable amount of genomic information is

devoted to the synthesis of the SDR enzymes. For example, the *B. japonicum* genome contains 113 SDR-encoding genes, and the genome of soil- and plant-associated α-proteobacteria typically have 11 to 12 SDR genes per Mbp (Pfam database) (6). Some SDR enzymes are also potentially useful in the industrial production of chiral compounds. Nevertheless, the range and basis of substrate specificity of the bacterial SDRs remain relatively unexplored.

In the present study, we have provided additional examples of the use of functional complementation for the isolation of novel genes from microbial communities. Most advances in molecular microbial ecology research have been based on the use of sequences of known genes as probes for the detection or isolation of those genes or their transcripts in environmental samples. A significant limitation in this type of approach is that it is based on the assumption that all genes encoding enzymes involved in a particular process will exhibit sequence similarity. It is becoming apparent, however, that this assumption is incorrect (22). Considering that the majority of organisms in a given microbial community are unknown, certainly there are numerous different types of genes encoding particular enzymes that would not be identified based on DNA sequence similarity. The SDR protein family is an example of a group of proteins for which it is often not possible to determine the substrate specificity based on sequence alone.

Most function-based metagenomic studies rely on expression and screening in *E. coli* surrogate hosts. There are few examples where non-*E. coli* hosts are used (48). We have shown that even within the *Proteobacteria*, there are expression differences that influence the genes that are isolated through functional complementation. Although we had previously demonstrated the ability to isolate *bdhA* genes from single-genome libraries through functional complementation of *E. coli* LS5218, the LS5218-complementing clones isolated from our metagenomic libraries did not express this activity. The identity of the LS5218-complementing genes is currently being investigated. It is quite possible, and it will be interesting to see, whether the LS5218 complementation strategy selects for enzyme activity distinct from BdhA. The absence of detectable BdhA enzyme activity for a few of the clones isolated in the *S. meliloti* complementation could also be due to presence of genes encoding other enzymes that are able to complement the growth phenotype. This also merits further investigation.

Through the functional analysis of relatively few genes, we have considerably broadened the known diversity of bacterial BdhA proteins. It is striking that although the BdhA proteins examined in the present study are probably all encoded by proteobacterial genomes, they exhibit sequence diversity almost as great as that exhibited within the entire SDR family. It is anticipated that further functional and comparative analysis will result in better definition of motifs that will not only allow improved annotation of gene function but will also lead to greater understanding of the recognition of specific substrates by functional protein domains.

ACKNOWLEDGMENTS

We thank Saleema Saleh for collecting the soil samples.

This research was supported by a Strategic Projects research grant from Natural Sciences and Engineering Research Council of Canada and also funding from Biocap Canada Foundation. Undergraduate students (P.P. and N.B.) were partially supported by the University of Waterloo Undergraduate Research Internship Program.

REFERENCES

1. **Altschul, S. F., T. L. Madden, A. A. Schäffer, Z. Zhang, W. Miller, and D. J. Lipman.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25:**3389–3402.
2. **Aneja, P., and T. C. Charles.** 1999. Poly-3-hydroxybutyrate degradation in *Rhizobium* (*Sinorhizobium*) *meliloti*: isolation and characterization of a gene encoding 3-hydroxybutyrate dehydrogenase. J. Bacteriol. **181:**849–857.
3. **Aneja, P., R. Dziak, G.-Q. Cai, and T. C. Charles.** 2002. Identification of an acetoacetyl coenzyme A synthetase-dependent pathway for utilization of L-(+)-3-hydroxybutyrate in *Sinorhizobium meliloti*. J. Bacteriol. **184:**1571–1577.
4. **Aneja, P., M. Dai, D. A. Lacorre, B. Pillon, and T. C. Charles.** 2004. Heterologous complementation of the exopolysaccharide synthesis and carbon utilization phenotypes of *Sinorhizobium meliloti* Rm1021 polyhydroxyalkanoate synthesis mutants. FEMS Microbiol. Lett. **239:**277–283.
5. **Aneja, P., and T. C. Charles.** 2005. Characterization of *bdhA*, encoding the enzyme D-3-hydroxybutyrate dehydrogenase, from *Sinorhizobium* sp. strain NGR234. FEMS Microbiol. Lett. **242:**87–94.
6. **Bateman, A., L. Coin, R. Durbin, R. D. Finn, V. Hollich, S. Griffiths-Jones, A. Khanna, M. Marshall, S. Moxon, E. L. Sonnhammer, D. J. Studholme, C. Yeats, and S. R. Eddy.** 2004. The Pfam protein families database. Nucleic Acids Res. **32:**D138–141.
7. **Bergmeyer, H. U., K. Gawehn, H. Klotzsch, H. A. Krebs, and D. H. Williamson.** 1967. Purification and properties of crystalline 3-hydroxybutyrate dehydrogenase from *Rhodopseudomonas spheroides*. Biochem. J. **102:**423–431.
8. **Boyer, H. W., and D. Roulland-Dussoix.** 1969. A complementation analysis of the restriction and modification of DNA in *Escherichia coli*. J. Mol. Biol. **41:**459–472.
9. **Cai, G.-Q., B. T. Driscoll, and T. C. Charles.** 2000. Requirement for the enzymes acetoacetyl coenzyme A synthetase and poly-3-hydroxybutyrate (PHB) synthase for growth of *Sinorhizobium meliloti* on PHB cycle intermediates. J. Bacteriol. **182:**2113–2118.
10. **Charles, T. C., and T. M. Finan.** 1991. Analysis of a 1600-kilobase *Rhizobium meliloti* megaplasmid using defined deletions generated in vivo. Genetics **127:**5–20.
11. **Charles, T. C., G.-Q. Cai, and P. Aneja.** 1997. Megaplasmid and chromosomal loci for the PHB degradation pathway in *Rhizobium* (*Sinorhizobium*) *meliloti*. Genetics **146:**1211–1220.
12. **Charles, T. C.** 2002. *Sinorhizobium meliloti* PHB cycle genetics, p. 250–253. *In* T. M. Finan, M. R. O'Brian, D. B. Layzell, J. K. Vessey, and W. Newton (ed.), Nitrogen fixation: global perspectives. CABI Publishing, Inc., New York, N.Y.
13. **Chenna, R., H. Sugawara, T. Koike, R. Lopez, T. J. Gibson, D. G. Higgins, and J. D. Thompson.** 2003. Multiple sequence alignment with the CLUSTAL series of programs. Nucleic Acids Res. **31:**3497–3500.
14. **Chien, M., I. Morozova, S. Shi, H. Sheng, J. Chen, S. M. Gomez, G. Asamani, K. Hill, J. Nuara, M. Feder, J. Rineer, J. J. Greenberg, V. Steshenko, S. H. Park, B. Zhao, E. Teplitskaya, J. R. Edwards, S. Pampou, A. Georghiou, I. C. Chou, W. Iannuccilli, M. E. Ulz, D. H. Kim, A. Geringer-Sameth, C. Goldsberry, P. Morozov, S. G. Fischer, G. Segal, X. Qu, A. Rzhetsky, P. Zhang, E. Cayanis, P. J. De Jong, J. Ju, S. Kalachikov, H. A. Shuman, and J. J. Russo.** 2004. The genomic sequence of the accidental pathogen *Legionella pneumophila*. Science **305:**1966–1968.
15. **Cole, J. R., B. Chai, T. L. Marsh, R. J. Farris, Q. Wang, S. A. Kulam, S. Chandra, D. M. McGarrell, T. M. Schmidt, G. M. Garrity, and J. M. Tiedje, Ribosomal Database Project.** 2003. The Ribosomal Database Project (RDP-II): previewing a new autoaligner that allows regular updates and the new prokaryotic taxonomy. Nucleic Acids Res. **31:**442–443.
16. **Edelshtein, Z., D. Kadouri, E. Jurkevitch, A. Vande Broek, J. Vanderleyden, and Y. Okon, Y.** 2003. Characterization of genes involved in poly-β-hydroxybutyrate metabolism in *Azospirillum brasilense*. Symbiosis **34:**157–170.
17. **Englard, S., and L. Siegal.** 1969. Mitochondrial L-malate dehydrogenase of beef heart. Methods Enzymol. **13:**99–100.
18. **Finan, T. M., B. Kunkel, G. F. DeVos, and E. R. Signer.** 1986. Second symbiotic megaplasmid in *Rhizobium meliloti* carrying exopolysaccharide and thiamine synthesis genes. J. Bacteriol. **167:**66–72.
19. **Galibert, F., T. M. Finan, S. R. Long, A. Puhler, P. Abola, F. Ampe, F. Barloy-Hubler, M. J. Barnett, A. Becker, P. Boistard, G. Bothe, M. Boutry, L. Bowser, J. Buhrmester, E. Cadieu, D. Capela, P. Chain, A. Cowie, R. W. Davis, S. Dreano, N. A. Federspiel, R. F. Fisher, S. Gloux, T. Godrie, A. Goffeau, B. Golding, J. Gouzy, M. Gurjal, I. Hernandez-Lucas, A. Hong, L. Huizar, R. W. Hyman, T. Jones, D. Kahn, M. L. Kahn, S. Kalman, D. H. Keating, E. Kiss, C. Komp, V. Lelaure, D. Masuy, C. Palm, M. C. Peck, T. M. Pohl, D. Portetelle, B. Purnelle, U. Ramsperger, R. Surzycki, P. Thebault, M. Vandenbol, F. J. Vorholter, S. Weidner, D. H. Wells, K. Wong, K. C. Yeh, and J. Batut.** 2001. The composite genome of the legume symbiont *Sinorhizobium meliloti*. Science **293:**668–672.
20. **Glazebrook, J., and G. C. Walker.** 1991. Genetic techniques in *Rhizobium meliloti*. Methods Enzymol. **204:**398–418.

21. **Hanahan, D.** 1983. Studies on transformation of *Escherichia coli* with plasmids. J. Mol. Biol. **166:**557–580.
22. **Henne, A., R. Daniel, R. A. Schmitz, and G. Gottschalk.** 1999. Construction of environmental DNA libraries in *Escherichia coli* and screening for the presence of genes conferring utilization of 4-hydroxybutyrate. Appl. Environ. Microbiol. **65:**3901–3907.
23. **Jonassen, I., J. F. Collins, and D. Higgins.** 1995. Finding flexible patterns in unaligned protein sequences. Protein Sci. **4:**1587–1595.
24. **Jones, J. D. G., and N. Gutterson.** 1987. An efficient mobilizable cosmid vector, pRK7813, and its use in a rapid method for marker exchange in *Pseudomonas fluorescens* strain HV37a. Gene **61:**299–306.
25. **Juck, D., T. Charles, L. G. Whyte, and C. W. Greer.** 2000. Polyphasic microbial community analysis of petroleum hydrocarbon-contaminated soils from two northern Canadian communities. FEMS Microbiol. Ecol. **33:**241–249.
26. **Kallberg, Y., U. Oppermann, H. Jörnvall, and B. Persson**. 2002. Short-chain dehydrogenases/reductases (SDRs). Eur. J. Biochem. **269:**4409–4417.
27. **Kaneko, T., Y. Nakamura, S. Sato, E. Asamizu, T. Kato, S. Sasamoto, A. Watanabe, K. Idesawa, A. Ishikawa, K. Kawashima, T. Kimura, Y. Kishida, C. Kiyokawa, M. Kohara, M. Matsumoto, A. Matsuno, Y. Mochizuki, S. Nakayama, N. Nakazaki, S. Shimpo, M. Sugimoto, C. Takeuchi, M. Yamada, and S. Tabata.** 2000. Complete genome structure of the nitrogen-fixing symbiotic bacterium *Mesorhizobium loti*. DNA Res. **7:**331–338.
28. **Kaneko, T., Y. Nakamura, S. Sato, K. Minamisawa, T. Uchiumi, S. Sasamoto, A. Watanabe, K. Idesawa, M. Iriguchi, K. Kawashima, M. Kohara, M. Matsumoto, S. Shimpo, H. Tsuruoka, T. Wada, M. Yamada, and S. Tabata.** 2002. Complete genomic sequence of nitrogen-fixing symbiotic bacterium *Bradyrhizobium japonicum* USDA110. DNA Res. **9:**189–197.
29. **Korotkova, N., and M. E. Lidstrom.** 2001. Connection between poly-β-hydroxybutyrate biosynthesis and growth on C1 and C2 Compounds in the methylotroph *Methylobacterium extorquens* AM1. J. Bacteriol. **183:**1038–1046.
30. **Krüger, K., G. Lang, T. Weidner, and A. M. Engel.** 1999. Cloning and functional expression of the D-β-hydroxybutyrate dehydrogenase gene of *Rhodobacter* sp. strain DSMZ 12077. Appl. Microbiol. Biotechnol. **52:**666–669.
31. **Larimer, F. W., P. Chain, L. Hauser, J. Lamerdin, S. Malfatti, L. Do, M. L. Land, D. A. Pelletier, J. T. Beatty, A. S. Lang, F. R. Tabita, J. L. Gibson, T. E. Hanson, C. Bobst, J. L. Torres, C. Peres, F. H. Harrison, J. Gibson, and C. S. Harwood.** 2004. Complete genome sequence of the metabolically versatile bacterium *Rhodopseudomonas palustris*. Nat. Biotechnol. **22:**55–61.
32. **Law, J. H., and R. A. Slepecky.** 1961. Assay of poly-β-hydroxybutyric acid. J. Bacteriol. **82:**33–36.
33. **Madison, L. L., and G. W. Huisman.** 1999. Metabolic engineering of poly (3-hydroxyalkanoates): from DNA to plastic. Microbiol. Mol. Biol. Rev. **63:**21–53.
34. **Meade, H. M., S. R. Long, G. B. Ruvkun, S. E. Brown, and F. M. Ausubel.** 1982. Physical and genetic characterization of symbiotic and auxotrophic mutants of *Rhizobium meliloti* induced by transposon mutagenesis. J. Bacteriol. **149:**114–122.
35. **Neufeld, J. D., B. T. Driscoll, R. Knowles, and F. S. Archibald.** 2001. Quantifying functional gene populations: comparing gene abundance and corresponding enzymatic activity using denitrification and nitrogen fixation in pulp and paper mill effluent treatment systems. Can. J. Microbiol. **47:**925–934.
36. **Nierman, W. C., T. V. Feldblyum, M. T. Laub, I. T. Paulsen, K. E. Nelson, J. A. Eisen, J. F. Heidelberg, M. R. Alley, N. Ohta, J. R. Maddock, I. Potocka, W. C. Nelson, A. Newton, C. Stephens, N. D. Phadke, B. Ely, R. T. DeBoy, R. J. Dodson, A. S. Durkin, M. L. Gwinn, D. H. Haft, J. F. Kolonay, J. Smit, M. B. Craven, H. Khouri, J. Shetty, K. Berry, T. Utterback, K. Tran, A. Wolf, J. Vamathevan, M. Ermolaeva, O. White, S. L. Salzberg, J. C. Venter, L. Shapiro, and C. M. Fraser.** 2001. Complete genome sequence of *Caulobacter crescentus*. Proc. Natl. Acad. Sci. USA **98:**4136–4141.
37. **Parkhill, J., M. Sebaihia, A. Preston, L. D. Murphy, N. Thomson, D. E. Harris, M. T. Holden, C. M. Churcher, S. D. Bentley, K. L. Mungall, A. M. Cerdeno-Tarraga, L. Temple, K. James, B. Harris, M. A. Quail, M. Achtman, R. Atkin, S. Baker, D. Basham, N. Bason, I. Cherevach, T. Chillingworth, M. Collins, A. Cronin, P. Davis, J. Doggett, T. Feltwell, A. Goble, N. Hamlin, H. Hauser, S. Holroyd, K. Jagels, S. Leather, S. Moule, H. Norberczak, S. O'Neil, D. Ormond, C. Price, E. Rabbinowitsch, S. Rutter, M. Sanders, D. Saunders, K. Seeger, S. Sharp, M. Simmonds, J. Skelton, R. Squares, S. Squares, K. Stevens, L. Unwin, S. Whitehead, B. G. Barrell, and D. J. Maskell.** 2003. Comparative analysis of the genome sequences of *Bordetella pertussis*, *Bordetella parapertussis*, and *Bordetella bronchiseptica*. Nat. Genet. **35:**32–40.
38. **Rawlings, M., and J. Cronan.** 1992. The gene encoding *Escherichia coli* acyl carrier protein lies within a cluster of fatty acid biosynthetic genes. J. Biol. Chem. **267:**5751–5754.
39. **Riesenfeld, C. S., P. D. Schloss, and J. Handelsman.** 2004. Metagenomics: genomic analysis of microbial communities. Annu. Rev. Genet. **38:**525–552.
40. **Rondon, M. R., P. R. August, A. D. Bettermann, S. F. Brady, T. H. Grossman, M. R. Liles, K. A. Loiacono, B. A. Lynch, I. A. MacNeil, C. Minor, C. L. Tiong, M. Gilman, M. S. Osburne, J. Clardy, J. Handelsman, and R. M. Goodman.** 2000. Cloning the soil metagenome: a strategy for accessing the genetic and functional diversity of uncultured microorganisms. Appl. Environ. Microbiol. **66:**2541–2547.
41. **Spratt, S. K., C. L. Ginsburgh, and W. D. Nunn.** 1981. Isolation and genetic characterization of *Escherichia coli* mutants defective in propionate metabolism. J. Bacteriol. **146:**1166–1169.
42. **Stein, J. L., T. L. Marsh, K. Y. Wu, H. Shizuya, and E. F. DeLong.** 1996. Characterization of uncultivated prokaryotes: isolation and analysis of a 40-kilobase-pair genome fragment from a planktonic marine archaeon. J. Bacteriol. **178:**591–599.
43. **Steinbüchel, A., and H. E. Valentin.** 1995. Diversity of bacterial polyhydroxyalkanoate acids. FEMS Microbiol. Lett. **128:**219–228.
44. **Streit, W. R., R. A. Schmitz, X. Perret, C. Staehelin, W. J. Deakin, C. Raasch, H. Liesegang, and W. J. Broughton.** 2004. An evolutionary hot spot: the pNGR234b replicon of *Rhizobium* sp. strain NGR234. J. Bacteriol. **186:**535–542.
45. **Sukdeo, N., and T. C. Charles.** 2003. Application of crossover-PCR-mediated deletion-insertion mutagenesis to analysis of the *bdhA-xdhA2-xdhB2* mixed-function operon of *Sinorhizobium meliloti*. Arch. Microbiol. **179:**301–304.
46. **Tyson, G. W., J. Chapman, P. Hugenholtz, E. E. Allen, R. J. Ram, P. M. Richardson, V. V. Solovyev, E. M. Rubin, D. S. Rokhsar, and J. F. Banfield.** 2004. Community structure and metabolism through reconstruction of microbial genomes from the environment. Nature **428:**37–43.
47. **Venter, J. C., K. Remington, J. F. Heidelberg, A. L. Halpern, D. Rusch, J. A. Eisen, D. Wu, I. Paulsen, K. E. Nelson, W. Nelson, D. E. Fouts, S. Levy, A. H. Knap, M. W. Lomas, K. Nealson, O. White, J. Peterson, J. Hoffman, R. Parsons, H. Baden-Tillson, C. Pfannkoch, Y.-H. Rogers, and H. O. Smith.** 2004. Environmental genome shotgun sequencing of the Sargasso Sea. Science **304:**66–74.
48. **Wang, G. Y., E. Graziani, B. Waters, W. Pan, X. Li, J. McDermott, G. Meurer, G. Saxena, R. J. Andersen, and J. Davies.** 2000. Novel natural products from soil DNA libraries in a streptomycete host. Org. Lett. **2:**2401–2404.
49. **Wood, D. W., J. C. Setubal, R. Kaul, D. E. Monks, J. P. Kitajima, V. K. Okura, Y. Zhou, L. Chen, G. E. Wood, N. F. Almeida, Jr., L. Woo, Y. Chen,† I. T. Paulsen, J. A. Eisen, P. D. Karp, D. Bovee Sr., P. Chapman, J. Clendenning, G. Deatherage, W. Gillet, C. Grant, T. Kutyavin, R. Levy, M.-J. Li, E. McClelland, A. Palmieri, C. Raymond, G. Rouse, C. Saenphimmachak, Z. Wu, P. Romero, D. Gordon, S. Zhang, H. Yoo, Y. Tao, P. Biddle, M. Jung, W. Krespan, M. Perry, B. Gordon-Kamm, L. Liao, S. Kim, C. Hendrick, Z.-Y. Zhao, M. Dolan, F. Chumley, S. V. Tingey, J.-F. Tomb, M. P. Gordon, M. V. Olson, and E. W. Nester.** 2001. The genome of the natural genetic engineer *Agrobacterium tumefaciens* C58. Science **294:**2317–2323.
50. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33:**103–119.