Genomic reconstruction of *Shewanella oneidensis* MR-1 metabolism reveals a previously uncharacterized machinery for lactate utilization

Grigory E. Pinchuk[a,1], Dmitry A. Rodionov[b,c,1], Chen Yang[b,d,1], Xiaoqing Li[b], Andrei L. Osterman[b,e,2], Etienne Dervyn[f], Oleg V. Geydebrekht[a], Samantha B. Reed[a], Margaret F. Romine[a], Frank R. Collart[g], James H. Scott[h], Jim K. Fredrickson[a], and Alexander S. Beliaev[a,3]

[a]Biological Sciences Division, Pacific Northwest National Laboratory, Richland, WA 99352; [b]Burnham Institute for Medical Research, La Jolla, CA 92037; [c]Institute for Information Transmission Problems, Russian Academy of Sciences, 127994, Moscow, Russia; [d]Institute of Plant Physiology and Ecology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200032, China; [e]Fellowship for Interpretation of Genomes, Burr Ridge, IL 60527; [f]Genetique Microbienne, Institut National de la Recherche Agronomique, 78352 Jouy-en-Josas, France; [g]Biosciences Division, Argonne National Laboratory, Argonne, IL 60439; and [h]Department of Earth Sciences, Dartmouth College, Hanover, NH 03755

Edited by Sydney Kustu, University of California, Berkeley, CA, and approved November 14, 2008 (received for review July 14, 2008)

The ability to use lactate as a sole source of carbon and energy is one of the key metabolic signatures of Shewanellae, a diverse group of dissimilatory metal-reducing bacteria commonly found in aquatic and sedimentary environments. Nonetheless, homology searches failed to recognize orthologs of previously described bacterial D- or L-lactate oxidizing enzymes (*Escherichia coli* genes *dld* and *lldD*) in any of the 13 analyzed genomes of *Shewanella* spp. By using comparative genomic techniques, we identified a conserved chromosomal gene cluster in *Shewanella oneidensis* MR-1 (locus tag: SO_1522–SO_1518) containing lactate permease and candidate genes for both D- and L-lactate dehydrogenase enzymes. The predicted D-LDH gene (*dld-II*, SO_1521) is a distant homolog of FAD-dependent lactate dehydrogenase from yeast, whereas the predicted L-LDH is encoded by 3 genes with previously unknown functions (*lldEGF*, SO_1520–SO_1518). Through a combination of genetic and biochemical techniques, we experimentally confirmed the predicted physiological role of these novel genes in *S. oneidensis* MR-1 and carried out successful functional validation studies in *Escherichia coli* and *Bacillus subtilis*. We conclusively showed that *dld-II* and *lldEFG* encode fully functional D-and L-LDH enzymes, which catalyze the oxidation of the respective lactate stereoisomers to pyruvate. Notably, the *S. oneidensis* MR-1 LldEFG enzyme is a previously uncharacterized example of a multisubunit lactate oxidase. Comparative analysis of >400 bacterial species revealed the presence of LldEFG and Dld-II in a broad range of diverse species accentuating the potential importance of these previously unknown proteins in microbial metabolism.

central carbon metabolism | genome context analysis | lactate dehydrogenase

Many aerobic and anaerobic bacteria are able to grow by using D- and/or L-lactate as a sole source of carbon and energy (1–4). Although lactate is a common product of carbohydrate fermentation (5, 6), it is rarely detected in environmental samples (7, 8), suggesting that it is either a minor metabolic product or that its conversion rates are very high. In support of the latter possibility, Finke *et al*. (9) recently reported constant production and consumption of lactate in marine sediments, linking its high turnover rates with microbiological reduction of sulfate and metals.

Among microorganisms actively coupling lactate oxidation to the reduction of multiple electron acceptors is a diverse and ubiquitous group of dissimilatory metal-reducing bacteria, which belong to the genus *Shewanella* (10). Shewanellae are commonly found in complex microbial communities within aquatic and sedimentary systems, many of which are subject to spatial and temporal variations in the type and concentration of organic and inorganic substrates that reflect redox gradients (10). The versatile flexibility of energy-generating pathways, which enables respiration of various electron acceptors including O_2, Fe(III), Mn(IV), thiosulfate, elemental sulfur, and nitrate, contributes to the ability of *Shewanella* to compete and thrive in such environments (11). Analysis of the *Shewanella oneidensis* MR-1 genome sequence revealed an extensive electron transport system, which includes 42 putative *c*-type cytochromes that are likely involved in a broad range of energy conversion reactions (12). This anticipated diversity is matched by a variety of inferred pathways for catabolism of carbohydrates, organic acids, fatty acids, amino acids, peptides, and nucleotides (13). Although many of these genomic predictions are supported by physiological, biochemical, and genetic experimental data (14–16), significant gaps in core metabolic pathways of *S. oneidensis* MR-1 remain. Surprisingly, the genome similarity searches failed to corroborate the physiological observations for lactate utilization, because no homologs for previously characterized bacterial D- and L-lactate dehydrogenases could be identified in MR-1 or any of the other sequenced genomes of *Shewanella* spp (13).

The paucity of information on lactate metabolism in Shewanellae prompted us to address this conundrum by combining metabolic reconstruction and comparative genomic analyses with genetic and biochemical techniques for the detailed analysis of lactate utilization mechanisms. By employing the subsystems approach (17), which allows to efficiently reconstruct metabolic pathways and discover novel genes using the comparative genomic techniques (18), we report a discovery of a gene cluster encoding novel enzymes required for oxidation of D- and L-lactate to pyruvate in a large number of diverse bacteria. Function of these enzymes, named Dld-II and LldEFG, respectively, was further experimentally verified in *S. oneidensis* MR-1.

Results

Initial Physiological and Genetic Characterization of Lactate Utilization in *S. oneidensis* MR-1. Our growth studies showed that *S. oneidensis* MR-1 can use either D- or L-lactate stereoisomers as a

Author contributions: G.E.P., A.L.O., E.D., J.K.F., and A.S.B. designed research; G.E.P., D.A.R., C.Y., X.L., A.L.O., E.D., O.V.G., S.B.R., F.R.C., J.H.S., and A.S.B. performed research; M.F.R., F.R.C., and J.H.S. contributed new reagents/analytic tools; G.E.P., D.A.R., C.Y., X.L., A.L.O., E.D., O.V.G., S.B.R., M.F.R., J.K.F., and A.S.B. analyzed data; and G.E.P., D.A.R., C.Y., A.L.O., M.F.R., J.K.F., and A.S.B. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

[1]G.E.P., D.A.R., and C.Y. contributed equally to this paper.

[2]To whom correspondence may be addressed. E-mail: osterman@burnham.org.

[3]To whom correspondence may be addressed at: Biological Sciences Division, Pacific Northwest National Laboratory, P.O. Box 999, MS P7-50, Richland, WA 99352. E-mail: alex.beliaev@pnl.gov.

This article contains supporting information online at www.pnas.org/cgi/content/full/0806798106/DCSupplemental.

Table 1. Occurrence and features of genes involved in lactate utilization in representative bacterial genomes

Organism	L-LDH	Predicted L-LDH			D-LDH	Predicted D-LDH	Lactate permease	Regulators
	LldD	LldE	LldF	LldG	Dld	Dld-II	LldP	LldR
Gamma-proteobacteria (90)	44	30	30	30	29	28	60	31
Shewanella oneidensis MR-1	-	SO1520	SO1519	SO1518	-	SO1521	SO1522	SO3460 (R2)
Escherichia coli K12	LldD	YkgE	YkgF	YkgG	Dld	-	LldP	LldR (R1)
Colwellia psychrerythraea	+	+	+	+	-	+	+	R5
Pseudomonas fluorescens	-	+	+	+	-	+	+	R1
Alpha-proteobacteria (60)	37	7	7	7	5	3	9	1
Roseobacter denitrificans	+	+	+	+	+	-	-	-
Rhodospirillum rubrum	-	+	+	+	-	+	+	R1
Beta-proteobacteria (32)	21	21	21	21	11	2	24	19
Neisseria meningitidis MC58	+	+	+	+	+	-	+	R1
Dechloromonas aromatica RCB	-	+	+	+	-	+	+	R1
Delta-proteobacteria (11)	0	7	7	7	0	3	5	0
Desulfovibrio vulgariis	-	+	+	+	-	+	+	-
Epsilon-proteobacteria (9)	0	9	9	9	0	7	9	0
Helicobacter pylori	-	+	+	+	-	+	+	-
Actinobacteria (27)	18	9	9	9	3	3	7	8
Propionibacterium acnes	-	+	+	+	-	+	+	R4
Bacillus / Clostridium (50)	0	16	16	16	0	1	26	12
Bacillus subtilis	-	+	+	+	-	-	+	R3
Bacteroidetes / Flavobacteria (11)	3	3	3	3	0	1	3	0
Bacteroides fragilis	-	+	+	+	-	+	+	-
Thermus / Deinococcus (3)	0	3	3	3	0	0	1	1
Deinococcus radiodurans	-	+	+	+	-	-	+	R5
Total:	**123**	**105**	**105**	**105**	**48**	**48**	**144**	**72**

Representative species in several taxonomic groups of bacteria are shown as rows and the number of genomes analyzed within a group is given in parentheses. The presence or absence of genes for the respective functional roles (columns) is shown by + or −. For *E.coli* K12 and *S. oneidensis* MR-1, the gene names are indicated instead of +. Numbers for taxonomic group rows indicate the number of species that have a gene ortholog. Genes clustered on the chromosome (e.g., operons) are outlined by matching background colors. The genes corresponding to the lactate-specific regulators are R1 (orthologs of known LldR *E. coli* regulator), R2, R3, R4, and R5 (novel predicted regulators). Genes predicted to be regulated by one of these lactate regulators are marked in red.

sole source of carbon and energy under aerobic and anaerobic conditions. Whereas the aerobic growth rate of *S. oneidensis* MR-1 on D-lactate was significantly slower than that on L-lactate with calculated μ_{max} values of 0.135 and 0.280 h^{-1}, respectively, only negligible differences in initial growth rates on both stereoisomers (0.125 h^{-1} for D-lactate and 0.128 h^{-1} for L-lactate) were observed under anaerobic conditions with fumarate as the electron acceptor (Fig. S1 *A* and *B*).

Despite the demonstrated ability of *S. oneidensis* MR-1 to grow on D and L forms of lactate, similarity searches of 13 sequenced *Shewanella* genomes failed to identify orthologs of experimentally characterized bacterial D- or L-lactate-oxidizing enzymes. Although a gene annotated as putative lactate dehydrogenase (LDH) (SO_0968, *ldhA*) is present in the MR-1 genome, it belongs to a family of fermentative D-LDHs. Members of this family have been shown to function mainly in the opposite direction catalyzing the formation of lactate from pyruvate coupled to NADH oxidation (19). Both phenotype data for a *ldhA* knockout strain and biochemical assays (*SI Materials and Methods*) confirmed that LdhA does not contribute to the ability of *S. oneidensis* MR-1 to use D- and L-lactate, therefore leaving the identity of *Shewanella* respiratory LDH enzyme(s) in question.

Comparative Genome Analysis Predicts Novel Lactate Utilization Genes. We used genome context analysis techniques including chromosomal gene clustering, transcriptional regulons, and gene occurrence profiles (18, 20) to tentatively identify the missing components of lactate utilization machinery in *Shewanella* spp. The results of this analysis, carried out across >400 sequenced bacterial genomes in the SEED database (17), are available online (http://theseed.uchicago.edu/FIG/subsys.cgi, under the "Lactate utilization" subsystem) and illustrated in Table 1 and Table S1. Notably, the lactate permease *lldP* gene (21) appears to be the most conserved component of lactate utilization pathways. Respective genes could be readily identified in ≈150 diverse bacterial genomes, including all *Shewanella* spp. and many other species that lack orthologs of L-LDH (*lldD*) and/or D-LDH (*dld*) genes. In *Escherichia coli*, *lldP* occurs in an operon with *lldD* and *lldR* (Fig. 1), where the latter encodes L-lactate responsive transcriptional regulator (22). Whereas similarly organized chromosomal clusters are found in many bacterial genomes, a different pattern of gene conservation is observed in the genomic neighborhood of the *lldP* gene in many species that do not contain *lldD* homologs. In *S. oneidensis* MR-1, one of the two copies of *lldP* (SO_1522) is found in a conserved chromosomal cluster with 4 previously uncharacterized genes (SO_1521, SO_1520, SO_1519, and SO_1518). The entire cluster is present in all analyzed *Shewanella* genomes, with the exception of *S. denitrificans*, the only member of the group unable to grow on either D- or L-lactate (data not shown). Conservation of this cluster, with some variations, was also detected in a number of diverse bacteria such as *Pseudomonas fluorescens* and *Vibrio vulnificus* (Fig. 1 and Table 1), providing strong evidence for its involvement in lactate utilization.

Genomic organization and unique phylogenomic occurrence profiles of genes from the SO_1522–SO_1518 chromosomal locus suggest that the SO_1521 and SO_1520–SO_1518 could correspond to 2 distinct enzymatic activities. The putative SO_1521 protein revealed a long-range homology (23% amino acid sequence identity) with the yeast D-LDH (DLD1) (23). Both proteins share the FAD-binding (Pfam accession no. PF01565) and FAD-linked oxidase (PF02913) domains, whereas SO_1521 has an additional C-terminal 4Fe-4S-binding domain (Fig. S2). In contrast, no appreciable sequence similarity was observed between SO_1521 and *E. coli* D-LDH (Dld), although both of them contain an N-terminal FAD-binding domain. Orthologs of SO_1521 were detected in 48 bacterial genomes, often clustered with *lldP*. Although the stereospecificity of LldP transporters in *Shewanella* (as well as in many other species) has not been experimentally characterized, the study in *E. coli* revealed that LldP has comparable affinities toward both, L- and D-lactate, as well as D-glycolate (24). Analysis of the phylogenomic occurrence profiles (20) showed that, with the exception of 3 species (*Photobacterium profundum*, *Propionibacterium*

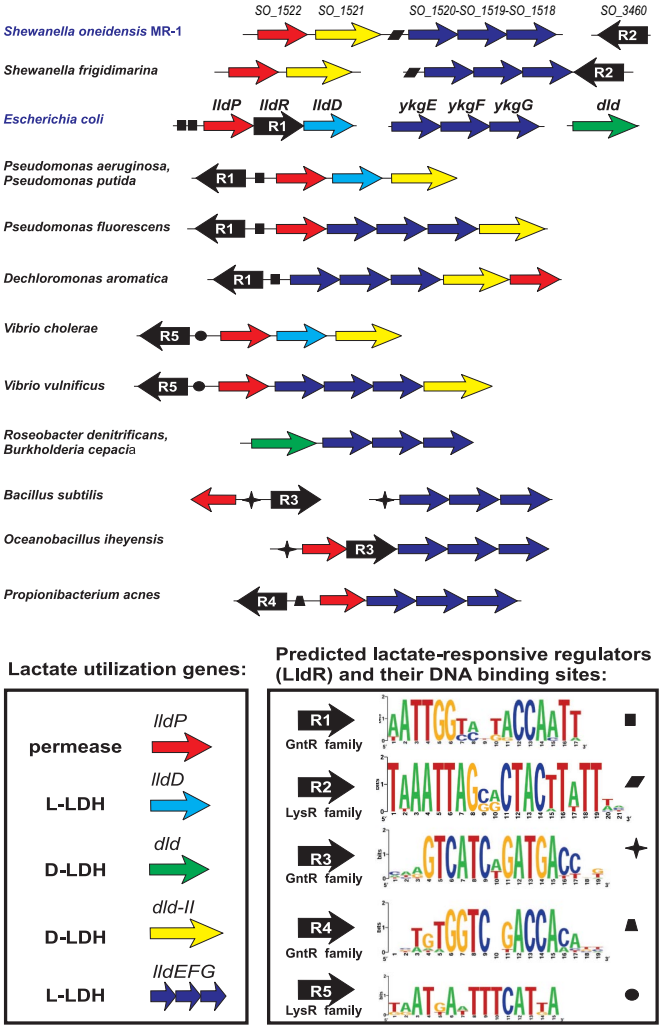

Fig. 1. Genome context analysis of lactate utilization genes across representative members of sequenced bacterial species. Orthologs in *Upper* are shown with matching colors and are explained in *Lower Left*. (*Lower Right*) Predicted binding sites of lactate-specific transcription factors are shown by black icons of different shape with consensus sequence logo depicted.

acnes, and *Brevibacterium linens*), the SO_1521 and *dld* genes were never found simultaneously in the genomes (Table 1 and Table S1). A combination of these observations allowed us to predict that SO_1521 represents a novel family of bacterial D-LDH, and therefore was tentatively designated as *dld-II* to distinguish from the nonhomologous *dld* gene of *E. coli*.

Next, our analysis revealed that orthologs of genes SO_1520, SO_1519, and SO_1518 have identical occurrence profiles and always form a putative operon conserved among 105 bacterial genomes (Table 1 and Table S1). This putative operon is often found in chromosomal clusters with *lldP* and either a newly predicted (*dld-II*) or a canonical (*dld*) form of D-LDH (Fig. 1), but never with members of the known L-LDH family (*lldD*). Moreover, a clear anticorrelation observed in the occurrence profiles of SO_1520–SO_1518 and *lldD* genes in several groups of bacteria (e.g., in Vibrionales and Pseudomonadales) additionally indicates their possible involvement in L-lactate metabolism. Long-range homology analysis revealed a distant similarity of SO_1520 protein with iron-sulfur subunits of glycerol-3-phosphate dehydrogenase (G3PDH) and glycolate oxidase (genes *glpC* and *glcF* in *E. coli*) including the presence of 2 cysteine-rich domains (PF02754). Whereas both SO_1519 and SO_1518 proteins contain a conserved domain of unknown function (PF02589), SO_1519 also contains a 4Fe-4S-binding domain (PF00037), suggesting a possible role in electron transfer (Fig. S2). Based on these observations, we concluded that SO_1520–SO_1518 genes are likely to encode a novel tripartite L-LDH in *S. oneidensis* MR-1 as well as other bacterial species. The newly identified genes were putatively designated as *lldE*, *lldF*, and *lldG* (operon *lldEFG*) by analogy with the canonical L-LDH (*lldD*) and to indicate a connection with previously uncharacterized homologous genes in *E. coli* (*ykgEFG* operon).

A comparative genomic reconstruction of lactate utilization machinery allowed us to predict a number of transcriptional factors that are likely involved in regulation of lactate utilization in several diverse groups of bacteria. Reconstruction of lactate regulons and identification of candidate transcription factor-binding sites was done using the Genome Explorer and SignalX software (22) as described in *SI Materials and Methods*. Lactate regulons inferred in several lineages of γ- and β-proteobacteria contain various combinations of known (*lldP*, *dld*, and *lldD*) and newly predicted (*dld-II* and *lldEFG*) genes (Fig. 1). They were predicted to be regulated by orthologs of the *E. coli* transcription factor LldR (R1) that are usually encoded within the respective gene clusters by similar DNA signals with a consensus sequence AATTGGnnnnnCCAATT, where "n" denotes any nucleotide (25). On the other hand, the lactate utilization genes in other lineages were predicted to be controlled by nonorthologous transcription factors from the same (GntR) or another (LysR) family recognizing quite divergent DNA motifs (Fig. 1 and Table S2). For example, in *Shewanella* spp. the putative transcriptional regulator R2 from the LysR family was predicted to control the expression of the *lldEFG* operon. Although the SO_3460 regulatory gene (*R2*) is not adjacent to the lactate catabolic genes in *S. oneidensis* MR-1, its functional assignment is supported by chromosomal clustering of *R2* and *lldEFG* genes in 3 other *Shewanella* spp. (e.g., *S. frigidimarina*) and by gene cooccurrence profile. In *Bacillus subtilis* this operon is located remotely from the lactate permease gene that forms a divergon with a putative lactate regulator of the GntR family. Nevertheless, the presence of the recognizable DNA signal suggests that together they form a lactate regulon conserved among many Gram-positive bacteria. Although the effectors of various predicted transcription factors for lactate-utilization genes have not yet been experimentally tested, the difference in their regulon content suggests that these regulators may sense different stereoisomers of lactate. In summary, the bioinformatic analysis of putative lactate regulons provided us with additional genomic evidence for the suggested functional assignments of the novel *dld-II* and *lldEFG* genes.

Mutagenesis Corroborates Predicted Lactate Utilization Genes in *S. oneidensis* MR-1. To test the role of the inferred lactate dehydrogenase genes in *S. oneidensis* MR-1, the respective chromosomal deletion mutants were constructed and tested for their ability to grow with D- or L-lactate (Fig. S1 *C* and *D*). An in-frame deletion of the gene encoding D-LDH (Δ*dld-II*) abolished the growth of the resulting strain on D-lactate but did not affect its growth on L-lactate. Conversely, Δ*lldE*, Δ*lldF*, and Δ*lldG* mutants of *S. oneidensis* MR-1 could not grow with L-lactate whereas their growth on D-lactate was not impaired. All 4 mutants (Δ*lldE*, Δ*lldF*, Δ*lldG*, and Δ*dld-II*) accumulated ≈50% less biomass compared with the wild type when grown aerobically on minimal medium supplemented with D,L-lactate (data not shown). Moreover, a Δ*dld-II*Δ*lldF* strain lost the ability to grow on and oxidize either lactate stereoisomer alone or as a mixture under aerobic or anaerobic conditions, nor did it produce any oxidation products such as pyruvate or acetate. Complementation of the Δ*dld-II*Δ*lldF* double mutant by using pBBR1MCS-5 plasmid constructs constitutively expressing *dld-II* and *lldEFG* restored the growth of MR-1 on D-lactate and L-lactate, respectively (data not shown).

To investigate potential interactions among the components of the putative LldEFG complex, affinity-tagged LldE was expressed

in the Δ*lldE* mutant. Pull-down experiments resulted in the co-purification of the recombinant LldE with 2 predominant proteins whose molecular weights corresponded to those predicted for LldF and LldG (Fig. S3). Although this evidence strongly suggests the existence of an LldEFG complex, further experiments will be necessary to determine the exact biochemical properties of this novel enzyme. Together, both co-purification and genetic evidence suggests a multiunit composition of the *S. oneidensis* MR-1 L-LDH enzyme.

Heterologous-Host Complementation Supports Functional Assignments of the Novel Lactate Utilization Genes. Plasmid constructs expressing *S. oneidensis* MR-1 *dld-II* and *lldEGF* were also introduced into *E. coli* K12 mutants from the genome-scale *Keio* collection (26), deficient in D-LDH (Δ*dld*) or L-LDH (Δ*lldD*). Expression of *dld-II* completely restored the ability of *E. coli* K12 Δ*dld* mutant to grow on D-lactate, whereas it had no appreciable effect on the impaired growth of Δ*lldD* mutant on L-lactate (Table S3 and Fig. S4). Similarly, expression of *lldEFG* from MR-1 successfully complemented the L-LDH deficiency and restored the robust growth of Δ*lldD* on L-lactate. All 3 *lldEFG* genes appeared to be required for the functionality of the L-LDH enzyme, because no appreciable growth was observed when only 1 (*lldE*) or 2 (*lldEF*) genes were used to complement the *E. coli* Δ*lldD* mutant (Fig. S4). Our results indicate that *dld-II* of *S. oneidensis* MR-1 encodes a fully functional D-LDH enzyme, whereas the L-LDH activity is linked to the expression of a 3-component enzymatic complex encoded by *lldEFG*.

Remarkably, the putative LldEFG complex of MR-1 also partially restored the ability of *E. coli* Δ*dld* mutant to grow on D-lactate, suggesting a presence of both L- and D-LDH activities. To further elucidate substrate specificity within the novel L-LDH family, we extended our studies to the previously uncharacterized *ykgEFG* operon from *E. coli*, which is orthologous to *lldEFG*. Notably, expression of *ykgEFG* from a high-copy pBAD-TOPO plasmid also restored the ability of the *E. coli* Δ*dld* mutant to grow on D-lactate (Table S3 and Fig. S4). Although the cause of such dual activity is unknown, it does not appear to be physiologically relevant in *S. oneidensis* MR-1. Complementation of *S. oneidensis* MR-1 Δ*dld-II*Δ*lldF* with *ykgEFG* carried on a low-copy broad-host range plasmid pBBR1MCS-5 restored the ability of the double mutant to use L-lactate but not D-lactate (data not shown). These observations suggest that the actual substrate stereoselectivity, e.g., preference for L-lactate over D-lactate, in the novel class of 3-component LDH enzymes (LldEFG) may be partially dictated by the genetic background and additional unknown factors expressed in the heterologous intracellular environment.

In Vitro Activity of the Novel D-LDH and L-LDH Enzymes. To extend the genetic findings and provide biochemical evidence to the proposed gene assignments, LDH activities were assayed in crude cell extracts of aerobically grown *S. oneidensis* MR-1 and *E. coli* DH10B cultures. In comparison to the *S. oneidensis* MR-1 wild-type cells, which exhibited high activities of both D- and L-LDH, Δ*dld-II* and Δ*lldF* mutants displayed only one of the corresponding activities, whereas the other decreased by >90% (Table 2). As expected, little or no D- or L-LDH activity was measured in Δ*dld-II*Δ*lldF* extracts.

The same assays were applied to crude extracts of *E. coli* DH10B strains carrying *dld-II*, *lldEGF*, or *ykgEFG* genes on a high-copy number pBAD-TOPO plasmid. Under conditions of the experiment, expression of *dld-II in trans* led to a nearly 4-fold increase of D-LDH activity as compared with the control strain carrying empty pBAD vector (Table 2). Similarly, >11-fold increase in L-LDH activity and a 4-fold increase in D-LDH activity was detected in the strain expressing the *lldEGF* genes. Finally, the plasmid-driven expression of the *E. coli ykgEFG* operon was accompanied by a comparable 3- to 4-fold increase of both activities. Analysis of reaction mixtures showed that pyruvate was the product in all Dld-II, LldEFG, and YkgEGF catalyzed reactions, confirming their function in lactate oxidation. These results are consistent with growth phenotype data suggesting that both *S. oneidensis* MR-1 lactate utilization systems, Dld-II and LldEFG, display stringent stereoselectivity when expressed in their native host. However, both representatives of the novel L-LDH class, LldEFG from *S. oneidensis* MR-1 and YkgEFG from *E. coli*, displayed an appreciable D-LDH activity when expressed in *E. coli* K12. In contrast to *S. oneidensis* MR-1, where the central role of the *lldEFG* cluster in lactate utilization is obvious, the actual physiological role of the *ykgEFG* operon in *E. coli* remains unclear.

Table 2. D- and L-LDH enzymatic activity in genetically modified strains of *S. oneidensis* MR-1 and *E. coli* DH10B

Strain	Specific activity (nmol/mg/min)	
	D-LDH	L-LDH
Wild type and mutants of *S. oneidensis* MR-1		
WT	835 ± 89	335 ± 22
Δ*lldF*	368 ± 32	23 ± 3
Δ*dld-II*	42 ± 5	367 ± 30
Δ*dld-II*/Δ*lldF*	6 ± 1	6 ± 1
E. coli DH10B containing expression plasmids		
pBAD*	38 ± 3	38 ± 4
pBAD::*dld-II*	146 ± 12	64 ± 7
pBAD::*lldEFG*	180 ± 30	430 ± 72
pBAD::*ykgEFG*	152 ± 25	136 ± 18

The activity was measured in crude cell extracts by a coupled chromogenic assay using 5 mM D- or L- lactate as an electron donor and a mixture of artificial acceptors, PMS and MTT. The activity monitored spectrophotometrically at 570 nm was normalized by the total protein concentration in crude cell extracts. Data shown are means ± standard deviations of 3 independent measurements.

**E. coli* DH10B, which encodes the wild-type copies of *dld* and *lldD*, carrying the empty pBAD vector was used as a negative control.

Mutagenesis Corroborates the Predicted Role of Lactate Oxidation Genes in *B. subtilis*. The ability to use lactate as the sole source of carbon and energy was previously reported for *B. subtilis* (27), whereas respective LDH genes remained unknown in this well-studied organism. By using genomic reconstruction of the lactate utilization machinery, we identified a single candidate L-LDH enzyme in *B. subtilis*, LldEFG, which is encoded by the *yvfV-yvfW-yvbY* (locus tags: *Bsu3402–01-00*) gene cluster. The respective gene knockout strains, *yvfV*(*lldE*)::MUTIN2, *yvfW*(*lldG*)::MUTIN2, and *yvbY*(*lldF*)::MUTIN2, were used for in vivo verification of the predicted 3-component L-LDH enzyme in *B. subtilis*. Each resulting mutant showed no growth on L-lactate as a single carbon source (Fig. S5), thus leading to a conclusion that all three components of the putative *B. subtilis* LDH enzyme are indispensable for utilization of L-lactate.

Discussion

One of the challenges in sequence homology-based functional annotation lies in the complex gene–enzyme relationship where sequence similarity does not always translate into identical activity and, moreover, cellular role. The availability of genome sequence information from related organisms provided us with a new way for applying comparative genomic approaches to develop robust predictions of biochemical and physiological functions. Here, we used genome context analysis in conjunction with physiological, genetic, and biochemical techniques to uncover novel lactate utilization machinery in *S. oneidensis* MR-1. The identified genes, *dld-II* (SO_1521) and *lldEFG* (SO_1520–SO_1518), encode fully func-

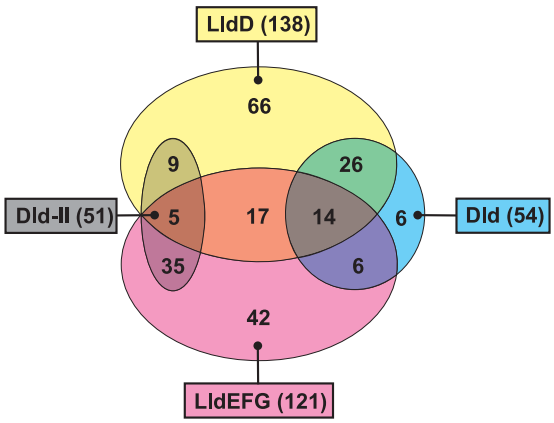

Fig. 2. Distribution of 4 different LDH types in sequenced bacterial genomes. The Venn diagram illustrates the occurrence and overlap of predicted *lldD*, *dld*, *dld-II*, and *lldEFG* genes through genome context analysis.

tional D-and L-LDHs, which catalyze the oxidation of the respective lactate stereoisomers to pyruvate. Both are novel enzymes, nonhomologous to previously characterized bacterial lactate dehydrogenases, such as those encoded by the *dld* and *lldD* genes in *E. coli*.

Although we were able to unambiguously identify the physiological role of both Dld-II and LldEFG in *S. oneidensis* MR-1, important mechanistic details including the composition and interactions of these enzymes with other components of the respiratory chain are yet to be elucidated. A distant homology with the eukaryotic D-LDH (28, 29), including the presence of FAD-binding and FAD-linked oxidase domains, suggests that the bacterial Dld-II may be mechanistically similar to flavin-dependent oxidases. In yeast, the activity of mitochondrial DLD1 enzyme is linked to the respiratory chain through ferricytochrome *c* acting as an electron acceptor. A distinctive feature of the bacterial Dld-II is the presence of an additional C-terminal cysteine-rich domain distantly related to the GlpC subunit of G3PDH enzyme. Interestingly, 2 of the 3 proteins composing the tripartite L-LDH, LldE, and LldF, also contain domains with distant homology to GlpC. Although the exact role of these 4Fe-4S cluster binding domains in the functioning of *S. oneidensis* MR-1 D- and L-lactate dehydrogenases is unclear, it is plausible to suggest that they may contribute to the interaction with electron acceptors, most likely cytochromes. These interactions may be of particular importance for Shewanellae that have an extremely rich repertoire of cytochromes (12) and could explain some of the difference in growth rates on D- and L-forms under aerobic and anaerobic conditions. Despite the apparent essential contribution of LldE, LldF, and LldG to L-LDH activity as well as the predicted multisubunit nature of the enzyme, the contribution of each component remains to be experimentally determined. It is possible that 1 or 2 of the proteins encoded by *lldEFG* may play a noncatalytic role in the functionality of the active enzyme (e.g., posttranslational modification, membrane attachment, etc.).

The identification of novel D-LDH and L-LDH enzymes, in addition to filling an important gap in the metabolic reconstruction of *S. oneidensis* MR-1, substantially expanded our knowledge of lactate utilization machinery in a broad range of bacteria. In contrast to Dld-II, which almost never occurs in bacteria containing an alternative D-LDH of the Dld family (Fig. 2), the phylogenomic distribution of LldEFG revealed 2 distinct groupings. In the first and largest group (>80 genomes), LldEFG is the only L-LDH enzyme and, as in *S. oneidensis* MR-1, its major role is likely in utilization of L-lactate. In most of these organisms the *lldEFG* operon clusters on the chromosome with *lldP* gene, and it is often predicted to be under control of a lactate transcriptional regulator (Table S1 and Table S2). In the second group (≈40 genomes), where both types of L-LDH genes (*lldD* and *lldEFG*) are present, only 1 of the 2 occurs in an operon and putative regulon with *lldP*. For example, the lactate regulon in *E. coli* includes *lldP-lldR-lldD* genes (25), whereas the *ykgEFG* operon, which is orthologous to *lldEFG*, is located remotely on the chromosome, and the mechanism of its regulation is unknown. In contrast, the *Azotobacter vinelandii* operon *lldP-lldEFG* is clustered and presumably co-regulated with divergently transcribed *lldR* gene, whereas the *lldD* gene is not co-localized or co-regulated with other lactate utilization genes. Although the genome context analysis may suggest which of the 2 L-LDH forms is primarily associated with the utilization of L-lactate, the exact interpretation of the observed functional redundancy would require a case-by-case investigation. Our data suggest that LldEFG-type enzymes from various organisms potentially have both L-LDH and D-LDH activities, and the factors contributing to their stereospecificity are yet to be elucidated.

The experimental verification and cross-genome projection of functional assignments also revealed a mosaic phylogenetic distribution of various forms of LDH genes in bacteria. Two families of D-LDH enzymes are equally represented in ≈100 bacteria from various taxonomic groups (Table S1), with only 3 cases of their simultaneous presence in the same genome. The novel L-LDH (LldEFG) is present in >80 bacteria that do not contain the canonical LldD enzyme, including *B. subtilis*, for which we confirmed the role of *yvfV-yvfW-yvbY* genes in L-lactate oxidation. Although the simultaneous presence of *lldEFG* and *lldD* genes in ≈40 bacterial species including *E. coli* and *Neisseria meningitidis* is puzzling, the existence of residual L-LDH activity in Δ*lldD*/Δ*dld* mutant of *N. meningitidis* (30) is consistent with the proposed L-LDH function of *lldEFG* operon (Table S1). Secondly, the *lldEFG* (*NMB1436–38*) operon of *N. meningitidis* was implicated in the increased resistance to H_2O_2 (31). Although no mechanistic explanation linking *lldEFG* expression with oxidative stress resistance in *N. meningitidis* is available, we believe an answer to that may lie in the co-factor composition of L-LDH enzymes.

Finally, this study provided another example of the impact of *S. oneidensis* MR-1 as a model system in genomic reconstruction of metabolism in many bacteria. Whereas the wealth of functional gene assignments and metabolic pathways was historically accumulated using *E. coli*, which remains a major source of genomic information, the comparative analysis of new genomes reveals a growing number of nonorthologous gene displacements and alternative pathways even in the most central aspects of bacterial metabolism. For example, the analysis of carbohydrate-utilization pathways in the *Shewanella* genus revealed substantial differences from *E. coli* at the level of individual enzymes, transporters, and transcriptional regulators (16). As in the case of lactate utilization machinery characterized in this study, these findings projected over a growing collection of sequenced genomes contribute to a more accurate and comprehensive understanding of metabolism in many diverse bacteria.

Materials and Methods

Strains and Growth Conditions. The list of strains and plasmids used in this study is given in Table S4. *S. oneidensis* MR-1 wild-type and mutant strains were routinely cultured at 30 °C in tryptic soy broth (TSB; pH 7.4) (32) or M1 minimal medium (15). The *E. coli* wild-type and mutant strains were routinely maintained and cultured at 37 °C on Luria-Bertani (LB) medium (pH 7.4) and M9 minimal medium (32). *B. subtilis* wild-type and mutant strains were routinely grown at 37 °C in chemically defined medium containing 20 mM glucose, 50 mg/L tryptophan, 1.5 g/L NH_4Cl, 10 g/L K_2HPO_4, 6 g/L KH_2PO_4, 0.2 g/L $MgSO_4$, 2 g/L K_2SO_4, 0.011g/L $CaCl_2$, 4 mg/L $FeCl_3$, and 0.2 mg/L $MnSO_4$. When needed, erythromycin was added to a final concentration of 0.5 mg/L. For phenotype growth assays, the glucose-grown cultures were diluted 300-fold in the defined medium and supplemented with 20 mM L-lactate or D/L-lactate mixture. Phenotype screening and growth experiments in *S. oneidensis* MR-1 and *E. coli* strains were carried out by using M1 and M9 minimal media, supplemented with 18 and 20 mM D- and/or L-lactate, respectively, or 18 mM pyruvate. For anaerobic growth of *S. oneidensis*

MR-1, fumarate was added to a final concentration of 35 mM. The growth of *S. oneidensis* MR-1, *E. coli*, and *B. subtilis* cultures was monitored spectrophotometrically at 600 nm. Organic acids were quantified by HPLC as described previously (15). The description of materials and reagents used in this study can be found in *SI Materials and Methods*.

Genetic Manipulations. In-frame deletion mutagenesis in *S. oneidensis* MR-1 was performed using a previously described method (33) with minor modifications (for details, see *SI Materials and Methods*). The *E. coli* K12 knockout strains, generated using a previously published procedure (26), were obtained from the genome-wide *Keio* collection. The *B. subtilis* knockout strains were received from the collection constructed by the joint effort of the Japanese and European *Bacillus subtilis* Functional Analysis programs. They were obtained by a standard single crossover-based protocol by using PCR-amplified fragments of target genes cloned in pMUTIN2 vector as previously described (34).

The full-length coding regions of *dld-II* and *lldEGF* from *S. oneidensis* MR-1 and *ykgEFG* from *E. coli* K-12 were PCR-amplified and initially cloned into pBAD-TOPO expression vector (Invitrogen). The resulting plasmids were transformed into *E. coli* K12 Δ*lldD* or Δ*dld* knockout mutants (26) for the complementation analysis, *E. coli* DH10B (Invitrogen) for enzyme activity assays, and *S. oneidensis* MR-1 Δ*lldE* for protein–protein interactions analysis (for details see *SI Materials and Methods*). For complementation studies in *S. oneidensis* MR-1, *dld-II* and *lldEGF* were subcloned into pBBR1MCS-5 broad-host vector (35) downstream of a lac promoter. Complementation studies were performed as described in *SI Materials and Methods*.

In Vitro Enzymatic Assays. The activities of fermentative D-LDH and lactate-oxidizing Dld-II and LldEFG enzymes were assayed in crude cell extracts of *S. oneidensis* MR-1 and *E. coli* DH10B carrying arabinose-inducible pBAD-TOPO *ldhA*, *dld-II*, and *lldEGF* constructs. In addition, the fermentative D-LDH activity was measured spectrophotometrically by following the NAD absorbance change at 340 nm using a protein partially purified by a miniscale nickel-nitrilotriacetic acid method (36). The D- and L-lactate oxidizing activities were assayed using previously published colorimetric techniques using coupling of lactate oxidation to the reduction of phenazine methosulfate (PMS) and 3-(4,5-dimethylthiazolyl-2)-2,5-diphenyltetrazolium bromide (MTT) (2, 37, 38). The experimental details describing the preparation of crude cell extracts and the assays can be found in *SI Materials and Methods*.

Genomes and Bioinformatics Tools. Analysis of the lactate utilization gene distribution, chromosomal co-localization, and co-occurrence profiles was performed using the SEED annotation environment (17). Results of the analysis are available at http://theseed.uchicago.edu/FIG/subsys.cgi under ''Lactate utilization'' subsystem. Reconstruction of lactate regulons and identification of candidate transcription factor-binding sites was performed using the Genome Explorer software (39) expressing and individual lineage-specific positional weight matrices that have been constructed by a subsystem-oriented approach reviewed in ref. 40. The Protein Families Database (Pfam) (http://pfam.sanger.ac.uk/) was used to identify conserved functional domains.

ACKNOWLEDGMENTS. We thank Drs. Kenneth H. Nealson, Anna Obraztsova, Matthew Marshall, and Liang Shi for insightful discussions and help with experimental design. We thank Dr. Hirotada Mori (Nara Institute of Science and Technology, Ikoma, Nara, Japan) for kindly providing the *E. coli* K12 lactate dehydrogenase mutants from the genome-wide *Keio* collection. We also gratefully acknowledge David W. Kennedy for help with HPLC analysis. This research was supported by the U.S. Department of Energy (DOE) Office of Biological and Environmental Research under the Genomics:GTL Program via the *Shewanella* Federation consortium and the Microbial Genome Program (MGP). Pacific Northwest National Laboratory is operated for the DOE by Battelle Memorial Institute under Contract DE-AC05–76RLO 1830.

1. Bryant MP, Campbell LL, Reddy CA, Crabill MR (1977) Growth of *Desulfovibrio* in lactate or ethanol media low in sulfate in association with H_2 utilizing methanogenic bacteria. *Appl Environ Microbiol* 33:1162–1169.
2. Erwin AL, Gotschlich EC (1993) Oxidation of D-lactate and L-lactate by *Neisseria meningitidis*: Purification and cloning of meningococcal D-lactate dehydrogenase. *J Bacteriol* 175:6382–6391.
3. Myers CR, Nealson KH (1988) Bacterial manganese reduction and growth with manganese oxide as the sole electron acceptor. *Science* 240:1319–1321.
4. Garvie EI (1980) Bacterial lactate dehydrogenases. *Microbiol Rev* 44:106–139.
5. Lynd LR, Weimer PJ, van Zyl WH, Pretorius IS (2002) Microbial cellulose utilization: Fundamentals and biotechnology. *Microbiol Mol Biol Rev* 66:506–577.
6. Shieh WY, Chen AL, Chiu HH (2000) *Vibrio aerogenes* sp. nov., a facultatively anaerobic marine bacterium that ferments glucose with gas production. *Int J Syst Evol Microbiol* 50:321–329.
7. Lovley DR, Klug MJ (1982) Intermediary metabolism of organic matter in the sediments of a eutrophic lake. *Appl Environ Microbiol* 43:552–560.
8. McMahon PB, Chapelle FH (1991) Microbial production of organic acids in aquitard sediments and its role in aquifer geochemistry. *Nature* 349:233–235.
9. Finke N, Vandieken V, Jorgensen BB (2007) Acetate, lactate, propionate, and isobutyrate as electron donors for iron and sulfate reduction in Arctic marine sediments, Svalbard. *FEMS Microbiol Lett* 59:10–22.
10. Nealson KH, Scott J (2003) in *The Prokaryotes: An Evolving Electronic Resource for the Microbiological Community*, ed Dworkin M (Springer, New York), pp 1133–1151.
11. Fredrickson JK, *et al.* (2008) Towards environmental systems biology of *Shewanella*. *Nat Rev Microbiol* 6:592–603.
12. Meyer TE, *et al.* (2004) Identification of 42 possible cytochrome *C* genes in the *Shewanella oneidensis* genome and characterization of six soluble cytochromes. *OMICS* 8:57–77.
13. Serres MH, Riley M (2006) Genomic analysis of carbon source metabolism of *Shewanella oneidensis* MR-1: Predictions versus experiments. *J Bacteriol* 188:4601–4609.
14. Driscoll ME, *et al.* (2007) Identification of diverse carbon utilization pathways in *Shewanella oneidensis* MR-1 via expression profiling. *Genome Informatics* 18:287–298.
15. Pinchuk GE, *et al.* (2008) Utilization of DNA as a sole source of phosphorus, carbon, and energy by *Shewanella* spp.: Ecological and physiological implications for dissimilatory metal reduction. *Appl Environ Microbiol* 74:1198–1208.
16. Yang C, *et al.* (2006) Comparative genomics and experimental characterization of N-acetylglucosamine utilization pathway of *Shewanella oneidensis*. *J Biol Chem* 281:29872–29885.
17. Overbeek R, *et al.* (2005) The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. *Nucleic Acids Res* 33:5691–5702.
18. Osterman A, Overbeek R (2003) Missing genes in metabolic pathways: A comparative genomics approach. *Curr Opin Chem Biol* 7:238–251.
19. Kochhar S, Hunziker PE, Leong-Morgenthaler P, Hottinger H (1992) Primary structure, physicochemical properties, and chemical modification of NAD(+)-dependent D-lactate dehydrogenase. Evidence for the presence of Arg-235, His-303, Tyr-101, and Trp-19 at or near the active site. *J Biol Chem* 267:8499–8513.
20. Morett E, *et al.* (2003) Systematic discovery of analogous enzymes in thiamin biosynthesis. *Nat Biotechnol* 21:790–795.
21. Prakash S, Cooper G, Singhi S, Saier MH (2003) The ion transporter superfamily. *Biochim Biophys Acta* 1618:79–92.
22. Dong JM, *et al.* (1993) Three overlapping *lct* genes involved in L-lactate utilization by *Escherichia coli*. *J Bacteriol* 175:6671–6678.
23. Lodi T, Alberti A, Guiard B, Ferrero I (1999) Regulation of the *Saccharomyces cerevisiae* DLD1 gene encoding the mitochondrial protein D-lactate ferricytochrome *c* oxidoreductase by HAP1 and HAP2/3/4/5. *Mol Gen Genet* 262:623–632.
24. Nunez MF, *et al.* (2002) Transport of L-Lactate, D-Lactate, and glycolate by the LldP and GlcA membrane carriers of *Escherichia coli*. *Biochem Biophys Res Commun* 290:824–829.
25. Aguilera L, *et al.* (2008) Dual role of LldR in regulation of the *lldPRD* operon, involved in L-lactate metabolism in *Escherichia coli*. *J Bacteriol* 190:2997–3005.
26. Baba, T., *et al.* (2006) Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: The *Keio* collection. *Mol Syst Biol* 2:2006.0008.
27. Schaeffer P, Millet J, Aubert JP (1965) Catabolic repression of bacterial sporulation. *Proc Natl Acad Sci USA* 54:704–711.
28. Flick MJ, Konieczny SF (2002) Identification of putative mammalian D-lactate dehydrogenase enzymes. *Biochem Biophys Res Commun* 295:910–916.
29. Lodi T, Ferrero I (1993) Isolation of the DLD gene of *Saccharomyces cerevisiae* encoding the mitochondrial enzyme D-lactate ferricytochrome *c* oxidoreductase. *Mol Gen Genet* 238:315–324.
30. Erwin AL, Gotschlich EC (1996) Cloning of a *Neisseria meningitidis* gene for L-lactate dehydrogenase (L-LDH): Evidence for a second meningococcal L-LDH with different regulation. *J Bacteriol* 178:4807–4813.
31. Grifantini R, *et al.* (2004) Characterization of a novel *Neisseria meningitidis* Fur and iron-regulated operon required for protection from oxidative stress: Utility of DNA microarray in the assignment of the biological role of hypothetical genes. *Mol Microbiol* 54:962–979.
32. Sambrook J, Russell DW (2001) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab Press, Cold Spring Harbor, New York).
33. Wan XF, *et al.* (2004) Transcriptomic and proteomic characterization of the Fur modulon in the metal-reducing bacterium *Shewanella oneidensis*. *J Bacteriol* 186:8385–8400.
34. Vagner V, Dervyn E, Ehrlich SD (1998) A vector for systematic gene inactivation in *Bacillus subtilis*. *Microbiology* 144(Pt 11):3097–3104.
35. Kovach ME, *et al.* (1994) pBBR1MCS: A broad-host-range cloning vector. *Biotechniques* 16:800–802.
36. Osterman AL, *et al.* (1995) Domain organization and a protease-sensitive loop in eukaryotic ornithine decarboxylase. *Biochemistry* 34:13431–13436.
37. Pratt EA, Fung LW, Flowers JA, Ho C (1979) Membrane-bound D-lactate dehydrogenase from *Escherichia coli*: Purification and properties. *Biochemistry* 18:312–316.
38. Friedemann T (1957) Determination of alpha-keto acids. *Methods Enzymol* 3:414–418.
39. Mironov AA, Vinokurova NP, Gel'fand MS (2000) [Software for analyzing bacterial genomes]. *Mol Biol (Mosk)* 34:253–262.
40. Rodionov DA (2007) Comparative genomic reconstruction of transcriptional regulatory networks in bacteria. *Chem Rev* 107:3467–3497.