Evauluation of Oligonucleotide Microarrays and Expression Analysis of *Desulfovibrio vulgaris* Cells in Metal-Reducing Conditions

M.E. Clark[1*], Q. He[2*], Z. He[2], L. Wu[2], J.D. Wall[3], T.C. Hazen[4], K. Huang[4], E. Alm[4], A.P. Arkin[4], J. Zhou[2], and M.W. Fields[1]

[1]Miami University, Oxford, OH; [2]Oak Ridge National Laboratory, Oak Ridge, TN; [3]University of Missouri, Columbia, MO; [4]Lawrence Berkeley National Laboratory, Berkeley, CA

Abstract

Objectives: (i) Construct whole-genome microarrays for *Desulfovibrio vulgaris*; (ii) Determine optimal conditions for hybridization conditions; and (iii) Identify whole-genome expression patterns of *D. vulgaris* cells grown under metal-reducing conditions. **Results**: Recent studies have shown that the design characteristics of probes can significantly affect hybridization behavior, and therefore, possibly affect the dataset of genes with altered expression profiles. In order to experimentally establish the criteria for the design of gene-specific probes, an oligonucleotide array was constructed that contained perfect-match and mismatch probes. The effects of probe-target identity, continuous stretch, mismatch position, and hybridization free energy on specificity were examined. Second, in order to determine the effect of probe length on signal intensities, microarrays with different length probes were used to monitor gene expression. Based on the experimental results, a set of criteria were suggested for the design of gene-specific and group-specific probes, and these criteria should provide valuable information for the development of new software and algorithms for microarray-based studies.

The whole-genome microarrays of *D. vulgaris* were constructed using 70mer probes. Currently, growth conditions and medium are being evaluated for transcriptomic characterization with respect to growth phase and heavy metal responsive genes. The elucidation of growth-phase dependent gene expression is essential for a general understanding of growth physiology that is also crucial for data interpretation of stress-responsive genes (e.g., chromium). We recently cultivated *D. vulgaris* in a defined medium and sampled biomass over time for approximately 70 h to characterize the shifts in gene expression as cells transitioned from logarithmic phase growth to stationary phase. Each respective time point during exponential phase had a similar number of ORFs (7 to 9% of the genome) that were up-expressed (>3-fold) and approximately 5.5 to 6.0% of the ORFs were down-expressed (<3-fold). Also, a majority of the predicted ORFs did not display altered expression patterns when early- and late-exponential phases were compared. As the cells entered into early stationary phase, approximately 7% of the ORFs were up-expressed, whereas 3.5% to 4% of the genome was up-expressed as the cells experienced prolonged stationary-phase. Within the first hours of stationary-phase, up-expressed ORFs included proteases, chaperonins, permeases, and hypothetical proteins. Interestingly, preliminary comparisons suggest that many of the ORFs that were up-expressed in early stationary phase also had altered expression levels in response to other cellular stresses. Preliminary data suggests that 0.05 mM Cr(VI) but not 0.01 mM Cr(VI) decreased growth in the defined medium with lactate and sulfate. In addition, *D. vulgaris* grew well in a defined, minimal medium designed to be analogous to FRC groundwater. We are presently analyzing microarray data to identify sub-sets of genes that respond to growth phase changes and/or stresses, as well as the determination of the optimal growth conditions in the presence of Cr(VI).

ORFs with Predicted PAS-Domains that have Altered Expression Levels Between Exponential- and Stationary-Growth Phases

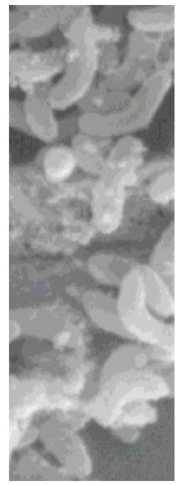
ORF ID	ORF Description	Time-Point Change	Additional Responses
206995*	sensory box protein	↑ T6, T9	↑ heat-shock
207582^	σ^{54}-dependent transcript. regulator	↓ T5, T6, T7	↓ O_2
209009^	sensory box/ HK / RR	↑ T6, T7	
209192^	sensory box/ HK / RR	↑ T6, T7	
209265*	sensory box/ HK	↑ T6 – T9	
206146^	sensory box/ HK	↑ T6	
206316	RR	↑ T6, T8	↑ NO_3
207350^	MAC protein	↑ T5, T6, T9	↑ heat-shock
207606^	sensory box/ HK / RR	↑ T6, T7, T8	
208701^	GGDEF domain protein	↑ T5, T7, T8	↓ NaCl
208703^	dcrA MAC-protein	↑ T9	↑ heat-shock
208671^	sensory box/ HK	↑ T5, T8, T9	↑ NO_2

Up-Expression T4 versus T1

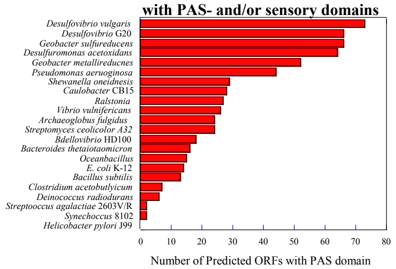
Gene ID	Predicted Function	Change Expression	Additional Responses
208068 208069	feoB feoA	4.83 (z: 6.02) 3.96 (z: 3.55)	↑ heat ↑ NO_2 ↑ NaCl
208071	transport component of feoA	1.89 (z: 1.97)	↑ NO_2 ↑ heat
208180	hypothetical protein	3.36 (z: 2.75)	↑ NO_2 ↑ O_2 ↑ heat ↑ NaCl
209207	conserved hypothetical protein	2.99 (z: 2.25)	↑ NO_2 Putative FUR-box
208070	hypothetical protein	2.70 (z: 2.63)	↑ NO_2 ↑ heat
209551	ahcY, adenosylhomocysteinase	2.53 (z: 4.63)	
207938	hypothetical protein	2.47 (z: 2.95)	↓ heat ↓ NaCl ↓ co-culture-CH_4
209629	katA, catalase	2.41 (z: 3.14)	↓ heat
208964	hypothetical protein	2.36 (z: 3.08)	↑ NO_2 ↓ heat
209542 209543	carbon starvation protein A	1.81 (z: 2.62) 2.34 (z: 3.76)	↑ heat
207937	flaB3, flagellin	2.30 (z: 4.44)	
209462 209463	flgM hypothetical protein	2.04 (z: 3.61) 2.24 (z: 4.17)	
206409	hypothetical protein	2.12 (z: 2.91)	
209052	hypothetical protein	2.07 (z: 2.61)	↑ NO_2 ↑ NaCl
208179	fld; flavodoxin (Fe-repressed)	2.07 (z: 2.48)	
207942	metK, S-adenosylmethionine synthetase	2.04 (z: 2.80)	
206429	metF, 5,10-mth reductase	1.97 (z: 2.94)	↑ NO_3 ↑ cold
208898	metE, 5-methyltetrahydropteroyltri-E-hc-S-methyltransferase	1.94 (z: 1.61)	↑ NO_3 ↑ cold ↑ cold

Down-Expression T4 versus T1

Gene ID	Predicted Function	Change Expression	Additional Responses
208063	lysA2 diaminopimelate dc	-2.39 (z: -4.06)	↑ O_2 ↑ heat
206150	Na/A symporter	-2.19 (z: -3.95)	↓ heat
206743-206772	ribosomal proteins	-0.54 to -2.0	↓ heat
206094	conserved hypothetical protein	-1.97 (z: -3.43)	↑ heat

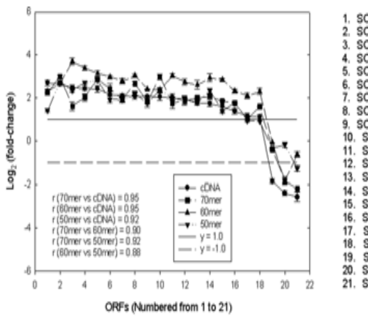


D. vulgaris has a high number of predicted ORFs with PAS- and/or sensory domains

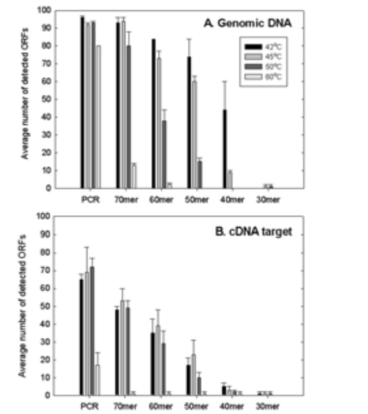


Up-Expression T5 versus T1

Gene ID	Predicted Function	Change Expression	Additional Responses
209625	major facilitator	5.72 (z: 5.71)	↑ NO_3 ↑ heat
208626	nhaC-2, Na/H antiporter	5.66 (z: 6.00)	↑ heat
207324	cold shock domain protein	5.45 (z: 4.24)	↑ NO_3 ↑ heat
209679	polysaccharide deacetylase	5.02 (z: 7.00)	
208597	membrane protein	4.69 (z: 5.25)	↑ NO_3 ↑ heat
207275	hypothetical protein	4.55 (z: 3.36)	
208628	oxidoreductase	4.13 (z: 3.96)	
208332	dctM, dicarboxylate transporter	4.11 (z: 2.35)	↑ NO_3
208069 208068	feoA feoB	4.05 (z: 3.49) 2.47(z: 1.86)	↑ NO_2 ↑ heat ↑ NaCl
208807	ADP-ribosylglycohydrolase	3.98 (z: 4.04)	
209685	sugar transferase domain	3.95 (z: 6.65)	
208486	integration host factor	3.89 (z: 7.29)	
208364	tail/DNA circulation protein	3.81 (z: 2.90)	↑ heat
206564	lysozyme	3.76 (z: 2.50	
208791	hypothetical protein	3.70 (z: 2.86)	
208851	hypothetical protein	3.65 (z: 2.45)	↑ heat
208370-208380	presumptive phage proteins	1.20 to 3.51	
208282	dsrC	3.48 (z: 5.70)	↑ O_2
206856	aa transport, putative	3.42 (z: 3.03)	
209131-209137	presumptive phage proteins	1.80 to 3.32	↓ co-culture ↓ NaCl
207823	carboxyl-terminal protease	3.25 (z: 2.12)	
208912	zraP,Zn-resistance	3.22 (z: 2.86	↑ NO_3 ↑ NaCl ↑ heat ↑ co-culture
206982 206983	membrane proteins	3.22(z: 5.25) 2.31 (z: 3.30	
206613	mviN, integral membrane protein	3.21 (z: 2.90)	↑ heat
206369	isoamylase	2.97 (z: 4.32)	
206808	putative, lipoprotein	2.50 (z: 4.28)	
207638-207641	hypothetical proteins	2.11 to 2.50	
207942	metK, S-adenosylmethionine syn	2.03 (z: 3.54)	
209182	conserved hypothetical (PO_4/phosphonate)	2.19 (z: 2.02)	
207901	sodB	2.27 (z: 3.15)	
206465	universal stress protein	2.30 (z: 2.82	↑ NO_3 ↑ heat
207934	hspC, heat shock	2.30 (z: 2.84)	↑ NO_2 ↑ heat ↑ NaCl ↑ co-culture



Comparison of fold changes of 21 ORFs consistently detected by PCR amplicon (circles), 70mer (squares), 60mer (up triangles) and 50mer (down triangles) arrays when *S. oneidensis* MR-1 cells were exposed to pH 4 for 30 min. ORFs were up-regulated above the line y = 1.0 (solid), ORFs were down-regulated below the line y = -1.0 (middle dash), and ORFs were not changed between the two lines.



Average number of genes detected among 96 ORFs selected from *Shewanella oneidensis* MR-1 at an SNR >3.0. 500 ng of genomic DNA (A) or 10 µg of cDNA target RNA (B) was labeled with Cy-dyes and hybridized with this array. The data are represented as average with standard deviation of three slides (6 spots).

Up-Expression T6 versus T1

Gene ID	Predicted Function	Change Expression	Additional Responses
207324	cold shock domain protein	5.76 (z: 6.87)	↑ NO_2 ↑ heat
208597	membrane protein	4.98 (z: 5.10)	↑ NO_3 ↑ heat
209014	hypothetical protein	4.91 (z:2.96)	
209625	major facilitator	4.89(z: 7.23)	↑ NO_3 ↑ heat
208626	nhaC-2, Na/H antiporter	4.74 (z: 3.65	↑ heat
206552-206556	presumptive phage proteins	2.04 to 4.45	
208370-208380	presumptive phage proteins	1.60 to 3.83	
209131-209137	presumptive phage proteins	2.34 to 4.54	
209138-209141	hypothetical proteins	3.27 to 4.25	
209679	polysaccharide deacetylase	5.02 (z: 7.00)	
209192	sensory box/ HK / RR	4.12 (z: 3.51)	
207275	hypothetical protein	3.99 (z: 2.55)	
208282	dsrC	3.85 (z: 6.29)	↑ O_2
206898	ispD	3.80 (z: 2.16)	↓ heat
208077	TPR domain	3.66 (z: 3.66)	
208316	TonB domain	3.56 (z: 2.66)	
208069 208068	feoA feoB	3.44(z: 4.69) 1.71(z: 2.09)	↑ NO_2 ↑ heat ↑ NaCl ↓ co-culture
208912	zraP,Zn-resistance	3.19(z: 2.66)	↑ NO_3 ↑ NaCl ↑ heat ↑ co-culture
206469	glucokinase	3.17 (z: 2.17)	↑ heat
207275	hypothetical protein	4.55 (z: 3.36)	
209300	5-formyltetrahydrofolate cyclo-ligase	3.03 (z: 4.34)	
209384	GDP-mannose-4,6-dehydratase	2.94 (z: 1.69)	
206564	lysozyme	3.12 (z: 4.68)	
208628	oxidoreductase	4.22 (z: 4.37)	
208486	integration host factor	4.17 (z: 3.02)	
208851	hypothetical protein	3.65 (z: 2.45)	↑ heat
207823	carboxyl-terminal protease	3.25 (z: 2.12)	
207934	hspC, heat shock	2.50 (z: 2.04)	↑ NO_2 ↑ heat ↑ NaCl ↑ co-culture
209665	glycosyl hydrolase	2.48 (z: 3.28)	
206613	mviN, integral membrane protein	2.44(z: 3.88)	↑ heat
209709	glnB-3 regulatory P-II	2.42 (z: 2.47)	
209738	putative DNA repair	2.30 (z: 2.33)	
207604	von Willebrand factor; Ankyrin	2.30 (z: 2.33)	

Up-Expression T5 to T9

Gene ID	Predicted Function	Change Expression	Time-Points
207324	cold shock domain protein	6.36 (z: 7.09)	T5, T6, T7, T8, T9
208626	nhaC-2, Na/H antiporter	5.53 (z: 3.92	T5, T6, T7, T8, T9
209625	major facilitator	5.45 (z: 4.13)	T7, T8, T9
209678	aminotransferase	5.35 (z: 10.42)	T7, T8
209679	polysaccharide deacetylase	5.31 (z: 9.37)	T5, T6, T7, T8
209685	sugar transferase domain	4.68(z: 5.47)	T5, T7, T8, T9
206469	glucokinase	3.17 (z: 2.17)	T6, T7, T8, T9
207948	epimerase/dehydratase	3.93 (z: 4.43)	T7
208077	TPR domain	3.93(z: 4.89)	T6, T7, T8
206564	lysozyme	3.93 (z: 3.97)	T5, T6, T7, T8, T9
208370-208380	presumptive phage proteins	3.68 (z: 4.35)	T5, T6, T7, T8, T9
209191	acetyltransferase, GNAT family	3.81(z: 4.00)	T7
209384	GDP-mannose 4,6-dehydratase	3.45 (z: 1.79)	T7, T8, T9
208069 208068	feoA feoB	1.83(z: 2.28) 3.28(z: 3.69)	T4 – T9
209738	putative DNA repair	3.23 (z: 2.75)	T6, T7, T8, T9
209285	polysaccharide biosynthesis protein	3.20 (z: 3.13)	T7, T9
208912	zraP,Zn-resistance	2.95 (z: 2.61)	T5, T6, T7, T9
206613	mviN, integral membrane protein	2.94 (z: 2.50)	T5, T6, T7, T8, T9
208072 208073	peptidase oligopeptide ABC transporter	2.67 (z: 2.92) 3.01 (z: 2.16)	T7 T7
2 08200	transglycosylase SLT domain protein	2.57 (z: 2.82)	T7, T9
209009	sensory box/ HK/ RR	2.47 (z: 3.22)	T6, T7
209012	aIF-2BI family	2.98(z: 2.50)	T7
209604	type III secretion system protein	2.41 (z: 4.34)	T7, T8, T9
209666	σ-54 dependent transcriptional regulator/ RR	2.32 (z: 2.85)	T7, T9

Major Conclusions:

Down-Expression T5 to T9

Gene ID	Predicted Function	Representative Values	Time-Points
206956	Plasmid maintenance toxin	-3.52 (z: -5.41)	T5, T6, T7, T8, T9
208347 208348	conserved hypothetical protein	-3.30 (z: -4.56)	T5, T6, T7, T8, T9
206699	outer membrane protein P1	-3.54 (z: -3.31)	T4, T5, T6
207771	L-lactate permease	-2.94 (z: -3.28)	T6
209502	glycosyl transferase (LPS or EPS)	-2.89 (z: -3.95)	T5, T6, T7, T8, T9
206371	GGDEF (putative adenyl cyclase)	-2.84 (z: -3.18)	T5, T6, T8
207124	3-dehydroquinate dehydratase (aromatic aa biosynthesis)	-2.67 (z: -2.73)	T5, T6
206915	peptidyl-prolyl cis-trans isomerase C	-2.55 (z: -2.16)	T5, T6, T7
209213	sulfate permease	-2.29 (z: -3.31)	T5, T6, T7, T9
208427	UDP-3-0-acyl N-acetylglucosamine deacetylase	-2.27 (z: -2.97)	T5, T6, T7
206681	vacJ lipoprotein, putative	-2.21 (z: -3.12)	T6, T9
207307	fructose-1,6-bisphosphatase	-2.29 (z: -2.19)	T6, T7, T8, T9
207508	plasmid stabilization system family	-2.04 (z:-3.40)	T7, T8, T8

Survey of Down-Expressed ORFs at T5 to T9

Ribosomal Proteins and Translation: 6 SSU and 10 LSU proteins were down-expressed greater than 2-fold. These ribosomal proteins were also down-expressed in response to nitrate, nitrite, and heat. Other proteins associated with translation were also down-expressed: nusG, secE, EF-Ts, IF-1, and SecF.

Biosynthesis: ribose-P-pyrophosphokinase, RNA-binding protein, ACP-reductase, ACP-synthase, ribonuclease P, gycerol-P acyl transferase, and histone-acetylation.

Hypothetical and Conserved Hypotheticals: Many hypothetical or conserved hypothetical proteins were down-expressed and some were also down-expressed in response to heat, nitrite, and nitrate while others appeared to be unique to stationary-phase growth. Between 40 and 50% of the significantly down-expressed ORFs (T5-T9) were hypothetical or conserved hypothetical proteins.

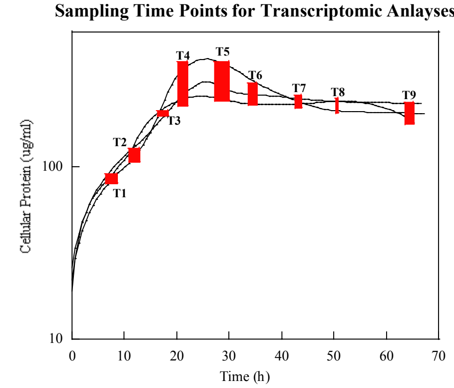
Sulfate Permease: Three mono-cistronic permeases. Each displayed a different pattern of expression at T5: 208981 Up-expressed; 209213 Down-expressed; 207471 unchanged.

Ferritin: *D. vulgaris* has two predicted ferritin-like proteins: 206840 expression was unchanged at T5 but 207018 was down-expressed.

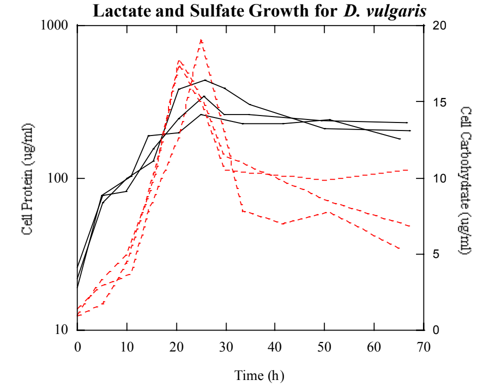
Occurrence of Predicted σ-Factors from a Range of Different Genomes The *D. vulgairs* Genome Does Not Appear to Contain a Recognized σ^{38}

	40	32	54	70	140	160		24	38	28
Desulfovibrio vulgaris	**A**	**H**	**N**	**D**	**B**	**C**	**Z**		?	**FliA**
Desulfovibrio desulfuricans G20	**A**	**H**	**N**	**D**	**B**	**C**	**Z**		?	**FliA**
Geobacter sulfureducens PCA	**A**	**H**	**N**	**D1, D2**	**B**	**C**	**Z**		?	**FliA**
Bdellovibrio bacteriovorus HD100	**A**		**N**	**D**	**B**	**C**		**E1-2**	?	**FliA**
Desulfotalea psychrophila		**H**	**N**	**D**					?	
Desulfuromonas acetoxidans		**H**	**N**	**D**				**E**	?	
Escherichia coli K-12	A	H	N	D	B	C	Z	E	S	FliA
Shewanella oneidensis MR-1	A	H	N	D	B	C	Z	E	S	FliA
Deinococcus radiodurans				D						
Bacteroides thetaiotaomicron				D					S	
Caulobacter crescentus CB15		H		D						
Chlorobium tepidum TLS	A		N		B	C				
Fusobacterium nucleatum				D						
Sinorhizobium meliloti	A	H1, 2	N		B	C	Z	E1-6	S	
Ralstonia solanacearum	A	H	N1,2	D	B	C		E1,2	S	FliA
Pseudomonas aeruginosa PAO1	A	H	N	D	B	C	Z		S	FliA
Vibrio cholerae		H	N	D				E	S	FliA
Clostridium acetobutylicum	A		N	D	B	C	Z			

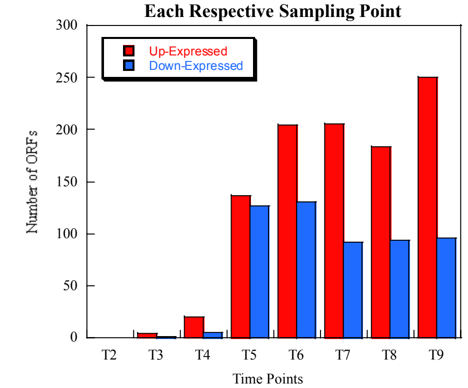
Sampling Time Points for Transcriptomic Anlayses



Cell Protein and Carbohydrate During Lactate and Sulfate Growth for *D. vulgaris*



Number of ORFs with Altered Expression at Each Respective Sampling Point



OAK RIDGE NATIONAL LABORATORY
U. S. DEPARTMENT OF ENERGY

MIAMI UNIVERSITY