Supplemental Methods

Transcriptional Analysis in a *Dehalococcoides*-Containing Microbial Consortium Reveals Prophage Activation

Alison S. Waller, Laura A. Hug, Kaiguo Mo, Devon R. Radford, Karen L. Maxwell, and Elizabeth A. Edwards

Supplemental Methods Table of Contents:

1. KB-1 Culture Maintenance

Experiments were set up to examine gene expression in KB-1 during vinyl chloride degradation using a robust and stable KB-1 enrichment that had been maintained on TCE and MeOH for 5 years (T3MP1). This culture (1.8 L in 2 L bottle) routinely degrades 200 μmoles of TCE to ethene every week and was grown in a medium designed to support methanogenic bacteria that had the following constituents per liter of deionized water was prepared: 10 ml of phosphate buffer (20.96 g of KH_2PO_4 per liter, 42.85 g of K_2HPO_4 per liter), 10 ml of salt solution (53.5 g of NH_4Cl per liter, 7.0 g of $CaCl_2$. $6H_20$ per liter, 2.0 g of $FeCl_2.4H_20$ per liter), 2 ml of trace mineral solution [0.3 g of H_3BO_3 per liter, 0.1 g of $ZnCl_2$ per liter, 0.1 g of $Na_2MoO_4.2H_2O$ per liter, 0.75 g of $NiCl_2.6H_20$ per liter, 1.0 g of $MnCl_2.4H_20$ per liter, 0.1 g of $CuCl_2.2H_20$ per liter, 1.5 g of $CoCl_2$ $6H_20$ per liter, 0.02 g of Na_2SeO_3 per liter, 0.1 g of $A1_2(SO_4)_3$ $16H_20$ per liter, 1 ml of H_2SO_4 per liter], 2 ml of $MgSO_4$. 7H20 solution (62.5 g/liter), 1 ml of redox indicator stock solution (1 g of resazurin per liter), 10 ml of saturated bicarbonate solution (260 g of $NaHCO_3$ per liter), 10 ml of filter-sterilized vitamin stock solution (0.02 g of biotin per liter, 0.02 g of folic acid per liter, 0.1 g of pyridoxine hydrochloride per liter, 0.05 g of riboflavin per liter, 0.05 g of thiamine per liter, 0.05 g of nicotinic acid per liter, 0.05 g of

pantothenic acid per liter, 0.05 g of p-aminobenzoic acid per liter, 0.05 g of cyanocobalamin per liter, 0.05 g of thioctic acid per liter, 1 g of mercaptoethanesulfonic acid [coenzyme M] per liter), and 10 ml of an amorphous ferrous sulfide solution [39.2 g of $(NH_4)_2Fe(SO_4)_2$ 6H20 per liter, 24.0 g of Na_2S $9H_20$ per liter] that had been washed three times with deionized water to remove free sulfide. The vitamins, bicarbonate, and ferrous sulfide were added from sterile stock solutions after the medium had been autoclaved and cooled while being gassed with N_2-CO_2 (80:20 [vol/vol]).

2. Microarray Fabrication

Clone Library Construction

UofT library. Genomic DNA was extracted from the mixed culture KB-1 using the Mo Bio Ultraclean Soil DNA Isolation Kit (Mo Bio Laboratories Inc., Solana Beach, CA) according to the manufacturer's instructions. Subsets of the genomic DNA were fragmented by sonication, bead beating, enzyme digestion or a nebulizer. The fragments underwent electrophoresis on a 1 % agarose gel and the 1 to 3 kilobase fragments were excised and eluted by QIAquick gel extraction kit (QIAGEN, Valencia, California). The fragments were then blunt-end repaired and dephosphorylated before cloning. Before blunt end cloning the concentration of the DNA fragments was adjusted to 200 ng/μl. The following 50 μl blunt-end repair reaction was first set up on ice (35 μl Sheared DNA, 5 μl of 10X Blunting Buffer, 1 μl of BSA, 5 μl of dNTPs, 2 μl of T4 polymerase and 2 μl of Klenow DNA polymerase), and incubated at room temperature for 30 min followed by 75°C for 20 min to inactivate the enzymes. Subsequently, the DNA was dephosphorylated and 35 μl of deionized water, 10 μl of 10X dephosphorylation buffer and 5 μl of calf intestinal phosphatase (CIP) were added to the above reaction mixture and incubated at 37°C for 60 min. After purification of the reaction sample the sample was concentrated to 100 ng/μl using a SpeedVac Concentrator Savant DNA 120 (Thermo Scientific, Waltham, MA) TOPO cloning technology (Invitrogen, Carlsbad, California) was used to ligate sheared blunt-end, dephosphorylated DNA into pCR 4Blunt-TOPO vector, and transform the recombinant vector into chemically competent TOP10 *E. coli* cells. Two different ligation mixtures were made each with 1 μL of vector (10 ng/μL), 1 μL of salt and 4 μL of DNA (29 ng/μL - 60 ng/μL). After ligation for 20 min at room temperature, transformation reactions were set up with 2 and 4 μl of ligation mixture. After incubation at 37°C for 1 hr with shaking at 250 rpm, the sample was plated onto medium containing ampicillin and X-Gal, and incubated overnight at 37°C. The white colonies containing plasmids with inserts were selected and grown overnight in 96-well plates. These plates were then used for PCR amplification of the inserts, and comprise 4,608 spots of the 19,200 final spots on the array.

TAGC Library. Purified genomic DNA was sheared by passing through a Millipore solvent delivery system Model 510 set at 3.5 ml/minute for 30 minutes. Fragments were blunt ended using mung bean nuclease, then 1-2.5 kb and 2.5-4 kb fragments were excised from a 1% agarose gel. A second blunt ending reaction was carried out using Klenow fragment T4 polymerase and kinase, followed by purification on a second agarose gel. Isolated fragments were cloned into the SmaI site of dephosphorylated

pUC19. After transformation into *E. coli*, individual bacterial colonies were picked into 384 well plates containing LB/glycerol using an automated colony picking robot (QPix2 Genetix) . These plates were incubated overnight at 37°C, then stored at -80°C until used for microarray preparation or sequencing. These 38 plates were labelled Kbgs001-Kbgs0038, and comprise 14,592 spots of the 19,200 spots on the final array.

PCR Amplification of Inserts

UofT library. PCR amplification was performed using 0.5 μl of culture in a 50 μl PCR reaction containing 2 U Taq DNA polymerase, 10 mM Tris, pH 8.4, 50 mM KCl, 1.5 mM $MgCl_2$, 800 μM dNTPs, and 0.2 pM each primer (T7f/M13r). The PCR mixture was held at 94ºC for 2 min and then cycled 40 times at 94ºC for 30 s, 53.5ºC for 30 s, and 72ºC for 1 min, followed by holding at 4ºC after the final cycle. From each reaction 2.5 μl were electrophoresed on a 2% agarose gel to check the yield and quality of PCR products. The remaining reaction volume was purified with QIAquick PCR purification kit (QIAGEN) and quantified spectrophotometrically with a plate reader. The PCR products were transferred to 384-well plates and adjusted to 200 ng/μl in 1XSSC buffer and sent for printing.

TAGC library. Inserts from the TAGC clone library were amplified at the University Health Network Microarray Center (UHN, Toronto, Canada) using a 96-well format. The cultures were stabbed and transferred into a total volume of 100 μl containing reaction reagent concentrations as above. The PCR products were filtered through Unifilter 800 filter plates (Whatman, 7700-2803) and the purified PCR products were quantified by spectrophotometric measurements at 260 and 280 nm. Then approximately 2 μg of DNA from each clone was transferred into 384-well polyfiltronic plates (Whatman, 7701-5101) using Evolution P3 liquid-handling robotics (Perkin Elmer). The PCR products were subsequently dried by overnight centrifugation in a vacuum dessicator and resuspended in 10 μl of 3XSSC for array printing.

Microarray Printing

The microarrays were printed from 384-well plates containing around 0.2 μg/μl of DNA in 3X SSC in the VersArray ChipWriter Pro System (Bio-Rad, 169-0006) using a 48-pin (SMP-3 pins, ArrayIT, product #SMP3) configuration. Each clone was spotted once on GAPS II–coated slides (Corning, 40003). The printed arrays were processed following manufacturer's protocol (Corning) and kept in a dessicator for later use. Printing was carried out at the UHN microarray centre (www.microarrays.ca).

3. RNA Extraction

For RNA extraction, 50 ml samples were withdrawn from the culture bottles inside an anaerobic chamber (Coy Laboratory Products). The culture samples were dispensed into anaerobic centrifuge tubes on ice. Cells were collected by centrifugation at 9900 rpm for 25 minutes at 4°C. After discarding the supernatant, the pellets were resuspended in 300 µL of ice-cold lysis solution (1.4 M NaCl, 22 mM EDTA, 35 mM SDS) and transferred into 1.5 mL screw cap microcentrifuge tubes containing 100 µL of zirconium beads (0.5

mm) (BioSpec Products Inc., Bartlesville, OK). Subsequently, 900 µL of ice-cold acid-phenol:chloroform:isoamyl alcohol (125:24:1, pH 4.5) (Ambion, Austin, TX) was added to the microfuge tubes. The tubes were then agitated horizontally on a vortexer (VELP Scientifica, Plainview, NY) for 4 minutes at maximum speed followed by centrifugation at 14000 x g for 3 minutes at 4°C. The aqueous supernatants were removed and transferred to new 1.5 mL microcentrifuge tubes. Ammonium acetate (0.1 volumes of a 5 M solution) and isopropanol (1.1 volumes) were added and the RNA was precipitated overnight at –20°C. The RNA solutions were then purified using an RNeasy spin column (Qiagen,Valencia, California). Contaminating DNA was removed using two successive treatments of the DNA-free kit (Ambion, Austin, TX). The quantity and quality of the RNA was then assessed using a ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE) as well as electrophoresis on a 1% agarose gel or using the 2100 Bioanalyzer (Agilent Technologies).

4. Reverse Transcription, Labelling and Hybridization

The RNA was labelled using an indirect method in which amino allyl-dUTPs were incorporated during reverse transcription (RT), and then the monoreactive Cyanine dyes were coupled to the aminoallyl cDNA in a separate reaction. First strand cDNA synthesis was carried out using 10-20 µg of total RNA in a 40µL reaction volume. The RNA was added to a master mix containing 10 mM dithiothreitol (0.4 µmoles), 42.5µg/mL random hexamers (5.1 µg) (Invitrogen, Carlsbad, CA), 0.5 mM dNTP-dTTP (20 nmoles), 0.15 mM dTTP (6 nmoles), 0.15mM amino allyl-dUTP (6 nmoles) and 10 ng control RNA (*Arabidopsis* chlorophyll synthetase gene). After heating this mixture to 65°C for 5 minutes to denature the RNA in a PTC-200 DNA engine thermocycler (MJ research), it was cooled at 42°C for 2 minutes and then at room temperature for 5 minutes to promote primer annealing. After this, 2µL of SuperScript II was added (Invitrogen). The RT reaction was then run for 2 hours at 42°C. The RNA was hydrolyzed by addition of 8 µmoles of NaOH to each RT reaction tube and heating to 65 for 5 minutes, and then neutralized with 8 µmoles of HCl and 4 µmoles of Tris-HCl (pH7.5). The remaining cDNA was then purified using the CyScribe kit (Amersham) following the kit protocol. The cDNA was eluted in 60 µL of 0.1667M sodium bicarbonate (pH 9.0). The cDNA was then quantified using the NanoDrop ND-1000 (Nanodrop Technologies). For labelling, the cDNA was first concentrated in a SpeedVac Concentrator Savant DNA 120 (Thermo Scientific). The dried pellet was then resuspended in 7µL of DNAse free water and 3µL of either Cy3 or Cy5 dye was added (Amersham). The labelling reaction was then incubated for 1 hour in the dark at room temperature. The un-reacted dyes were subsequently removed by addition of 18 µmoles of hydroxylamine and incubation in the dark for 15 minutes at room temperature. The labelled cDNA was again purified using the CyScribe columns and eluted in 60µL elution buffer. The purified labelled cDNA was then quantified using the nanodrop to ensure that an appropriate amount of dye was incorporated (>100 pmoles) and at an appropriate frequency (25-50 nucleotides per dye molecules). It was then concentrated and resuspended in 5uL DNAase free water. The cDNA was then combined with 80 µL of hybridization solution (80 µL DIG Easy Hyb (Roche, Indianapolis, IN), 4 uL calf thymus DNA (SIGMA, 10 mg/mL) and 4 uL yeast tRNA (Invitrogen, 10 mg/mL). After mixing the hybridization mixture was heated up to

65°C for 3 minutes and then allowed to cool to room temperature before being applied to the slide. The slide was then placed in a hybridization chamber and placed in a 37°C incubator for 16-18 hours.

5. Array Washing and Scanning

After removing the coverslip in 1X SSC, two 15 minute washes were carried out at 55°C in 1XSSC/0.1%SDS each with intermittent agitation. Then an additional wash was performed in 0.1XSSC/0.1%SDS. The slides were then rinsed twice at room temperature in 0.1XSSC. The slides were subsequently dried by centrifugation at 600 rpm for 5 minutes in a slide box lined with Whatman paper and then immediately scanned with an Axon scanner using GenePix Pro software.

6. Array Data Analysis

Data analysis was performed in the R programming language using programs available from Bioconductor such as Linear Analysis of Microarrays (Limma), gplots and other graphing programs (8). Raw intensities were imported from GenePix and visualized using various plots to assess the need for and effect of different background correction (6) and normalization techniques (9). The intensities were then corrected for background, by fitting a convolution of normal and exponential distributions to the foreground intensities using the background intensities as a covariate, resulting in a smooth transformation of the background subtracted intensities such that all the corrected intensities are positive (backgroundCorrect, method=normexp). The log ratios (M-values) were then normalized to correct for dye biases within each array using the Loess method which corrects for positional and intensity-dependent dye biases (NormalizeWithinArrays, method=Loess). Subsequently the log ratios were scaled to have the same median-absolute-deviation across the arrays (NormalizeBetweenArrays, method=scale). To determine which spots contained genes that had differential transcript levels, Limma was used to fit a linear model between the data from the six different arrays and Bayesian analysis was performed to estimate the statistical significance (7). This resulted in an average M value ($\log_2$[VC+MeOH/MeOH]) for each spot across the set of six arrays and a corresponding statistical significance (B and P value). To generate the list of top differential intensity spots for sequencing, a cutoff of an M value > 2 or < -1.5, and a B value of >4 was chosen. In addition to differential intensity spots, spots that had high intensities in both treatments (VC+MeOH and MeOH-only) were also identified. The intensities for each channel were read into Limma and were corrected for background intensity. Using R commands, the intensities for each array channels (*i.e.*, treatment) were ranked, and then a consistent set was created by determining which genes were ranked in the top 5% on all 6 arrays.

7. Sequencing

From the statistical analysis of differential transcript levels and the list of spots with high intensity levels, 780 spots were chosen to be sequenced. Sequencing was performed at The Atlantic Genome Center (TAGC) (Halifax). Clones corresponding to the selected

spotted fragments were selected from the genomic libraries (at the UHN microarray centre, Toronto) and sequencing was performed directly from the glycerol stocks. The bacterial suspensions were inoculated into lysis buffer using a 384-pronged tool and denatured at 95°C for 5 minutes. DNA from each clone was amplified using TempliPhi™ DNA polymerase according to manufacturer's instructions. DNA sequencing was performed using ET terminator chemistry, and reactions processed using magnetic beads to remove excess fluorescent terminators before loading onto GE Healthcare MegaBace capillary DNA sequencers. Traces were base-called using phred, and were trimmed for vector and poor quality sequence.

8. Bioinformatic Analyses of Sequences

A series of basic local alignment searches (BLAST) (1) were performed to ascribe putative functions and phylogenies to the sequences. Databases that were used include the NCBI non-redundant protein and nucleotide databases, Greengenes 16S rRNA database (3), and the protein database from the eggNOG website and the corresponding mappings to COG groups (5). BLAST hit scores of > 60 bits were considered significant. Sequences with significant hits to 16S rRNA genes were further classified using the Greengenes alignment and classification tools. In addition, sequences with BLASTX hit scores <100 bits were also subject to BLASTN against NCBI/nr-nt database and further identified 23S rRNA genes. Stand-alone BLASTS were performed and the results were parsed and manipulated using a series of Perl scripts. The information was then imported into a Microsoft Access database. Array data was also uploaded to the database and linked to the corresponding sequence information for each spot. The resulting Access database was used to query the data based on sequence information of the spots, or intensity data from the array experiments, and to produce the data in the supplementary tables.

9. Quantitative PCR for verification of array data

Nine spots encoding differentially transcribed genes were selected for verification using qPCR. Primers for SYBR green detection were designed using Primer3 software (http://frodo.wi.mit.edu/primer3/input.htm). Each 30-μl reaction contained 15 μl of SYBR Green JumpStart *Taq* ReadyMix (Sigma-Aldrich Co.), 11.8 μl of sterile water, 2 μl of DNA template, and each forward and reverse primer at 0.5 μM. The qPCRs were conducted using an Opticon 2 (MJ Research) thermocycler with the following program: initial denaturation for 10 min at 94°C; 45 cycles of denaturation at 94°C for 30 s, annealing at 60°C for 30 s, and extension at 72°C for 30 s; and a final melting curve analysis from 72 to 95°C, measuring fluorescence every 0.5°C. Quantification standards were created with dilutions of the plasmids containing the relevant spot sequence. Results are presented in supplementary figure S3.

10. Representation of Clusters of Orthologous Groups (COGs)

Sequences were ascribed to certain Clusters of Orthologous Groups (COGs) based on a BLASTX search against proteins from the eggNOG database (5). To determine if the

Dehalococcoides genes with differential transcript levels were enriched in sequences belonging to certain COGs, an enrichment ratio was calculated for each COG category in the differentially transcribed spots and compared to the proportion of that COG in an average *Dehalococcoides* genome. Enrichment ratio=[(proportion of COG X in *Dehalococoides* sequences with higher transcript levels)/(proportion of COG X in *Dehalococcoides* genomes)]. A ratio of 1 indicates that there is the same proportion of that COG in the sequenced spots as in an average *Dehalococcoides* genome. Proportions of COGs in the *Dehalococcoides* genomes was determined by averaging numbers from the NCBI COG tables for *Dehalococcoides* CBDB1 and *Dehalococcoides* str. 195. In addition, statistical analysis was performed to evaluate if the over representation was statistically significant. Using a hypergeometric distribution, a P-value was calculated using the formula:

$$\text{P value} = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$$

where n= the total number of sequences in the sample (*e.g.,* set with higher transcript levels); x=the number of sequences in a specific COG in the sample (*e.g.,* set with higher transcript levels); N=the total number of proteins in an average *Dehalococcoides* genome; M=the average number of proteins in a specific COG from the *Dehalococcoides* genomes (4, 10).

11. Phage DNA extraction

All glassware and centrifuge bottles were acid washed, washed with soap and water, rinsed with DI water, and autoclaved prior to use. A 500 mL aliquot of T3 MP1 KB-1 culture was taken after 5 days of starvation following a completed dechlorination cycle. The 500 mL of culture was sonicated anaerobically, on ice, for 5 min in 1s pulses of 40 W amplitude to maximize viral particle yield. The sonicated culture was then centrifuged at 8000xg for 30 min at 4°C. The supernatant was filtered through a 0.2 μM sterivex filter and treated with DNase I (2 U/mL) and RNase A (5 μg/mL) at room temperature for 1 hour. For PEG precipitation of viral particles, NaCl was added to a 1M concentration (29.2 g per 500 mL). After dissolution of the NaCl, the bottle was incubated on ice for 1 hour. An optional centrifugation spin was omitted, as no further precipitate was observed following this incubation. Solid PEG 8000 was added to 10% w/v (50 g into 500 mL), and dissolved by slow stirring at RT. Once the PEG was dissolved, the bottle was incubated overnight at 4°C. The PEG-precipitated viral particles were recovered by centrifugation at 11,000xg for 10 min at 4°C. The supernatant was removed, and the bacteriophage pellet resuspended in 6 mL of SM buffer (50mM Tris-HCL, pH 7.5, 100mM NaCl, 8mM magnesium sulfate, 0.01% gelatin). PEG and cell debris were extracted by adding an equal volume of chloroform and vortexing for 30s. The aqueous phase, containing viral particles, was recovered by centrifugation at 5000xg for 15 min at 4°C. The phage particles were disrupted by addition of EDTA (pH 8) to 20 mM, Proteinase K to 50 μg/mL, and SDS to 0.5% followed by incubation at 56°C for 1 hour. The digestion mixture was cooled to room temperature, and the phage DNA extracted through addition of an equal volume of phenol (equilibrated with 50 mM Tris). The

phenol mixture was vortexed and then centrifuged at 3,000xg for 5 min at RT. The DNA fraction was further purified through phenol-chloroform and then chloroform extractions, with centrifugation of each extraction as for the phenol step. The final 5 mL of aqueous phase contained the KB-1 viral metagenomic DNA. This DNA was precipitated through addition of 10% vol. 3M sodium acetate (0.5 mL) and two volumes of ice-cold 99.9% ethanol (11 mL). The DNA was precipitated for 1 hour on ice, following which it was harvested into a single pellet through repeated centrifugation of 2 mL aliquots at 12,000xg for 15 min at 8°C. The DNA pellet was washed with 70% ethanol (diluted with DNase-free water), and centrifuged. The supernatant was removed, and the pellet allowed to dry at RT. Once dry, the pellet was resuspended in 50 µL of DNase/RNase-free ddH_2O.

The presence of phage genomic DNA (capsid, tail, and resolvase proteins), the *tceA* gene, two hypothetical proteins, and bacterial contamination was tested with PCR. The reactions contained 0.25 mM dNTPs, 0.5 mM primer (F and R separate, see supplemental table S7 for primer sequences), 1x NEB *Taq* polymerase reaction buffer, and 1 U *Taq* polymerase (NEB). All reactions were amended with 1 µL of the template of interest. The PCR was conducted with an initial denaturation step for 5 min at 95°C, followed by 27 cycles of denaturation for 30s at 95°C, annealing of primers for 30s at 55 or 59°C (see table S7), and extension of product for 1min 45s at 72°C, and then a final extension for 5 min at 72°C. PCR products were visualized on a 1% agarose gel stained post-electrophoresis in an ethidium bromide bath.

12. Cloning and heterologous expression

In total, 37 genes of interest identified from the microarray data were nominated for cloning and protein characterization by our collaborators in Toronto (Dr. Alexei Savchenko & Dr. Alexander Yakunin) who are affiliated with the Midwest Center for Structural Genomics (MCSG). Of these, 16 were expressed well and were soluble (Table S1). The genes were cloned, expressed and purified using a protocol described previously (11), which was developed at the MCSG. In brief the genes were cloned into a derivative of pET15b (Novagen), using a ligation-independent cloning protocol. The vector encodes for a 6×His-tag followed by a spacer and a tobacco etch virus (TEV) protease cleavage site on the N-terminus of the expressed protein.

References

1. **Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman.** 1990. Basic local alignment search tool. J Mol Biol **215:**403-10.

2. **Cardarelli, L., R. Lam, A. Tuite, L. A. Baker, P. D. Sadowski, D. R. Radford, J. L. Rubinstein, K. P. Battaile, N. Chirgadze, K. L. Maxwell, and A. R.**

Davidson. 2010. The crystal structure of bacteriophage HK97 gp6: defining a large family of head-tail connector proteins. J Mol Biol **395:**754-68.

3. **DeSantis, T. Z., Jr., P. Hugenholtz, K. Keller, E. L. Brodie, N. Larsen, Y. M. Piceno, R. Phan, and G. L. Andersen.** 2006. NAST: a multiple sequence alignment server for comparative analysis of 16S rRNA genes. Nucleic Acids Res **34:**W394-9.

4. **Mahadevan, R., B. Yan, B. Postier, K. P. Nevin, T. L. Woodard, R. O'Neil, M. V. Coppi, B. A. Methe, and J. Krushkal.** 2008. Characterizing regulation of metabolism in *Geobacter sulfurreducens* through genome-wide expression data and sequence analysis. OMICS **12:**33-59.

5. **Muller, J., D. Szklarczyk, P. Julien, I. Letunic, A. Roth, M. Kuhn, S. Powell, C. von Mering, T. Doerks, L. J. Jensen, and P. Bork.** 2010. eggNOG v2.0: extending the evolutionary genealogy of genes with enhanced non-supervised orthologous groups, species and functional annotations. Nucleic Acids Res **38:**D190-5.

6. **Ritchie, M. E., J. Silver, A. Oshlack, M. Holmes, D. Diyagama, A. Holloway, and G. K. Smyth.** 2007. A comparison of background correction methods for two-colour microarrays. Bioinformatics **23:**2700-7.

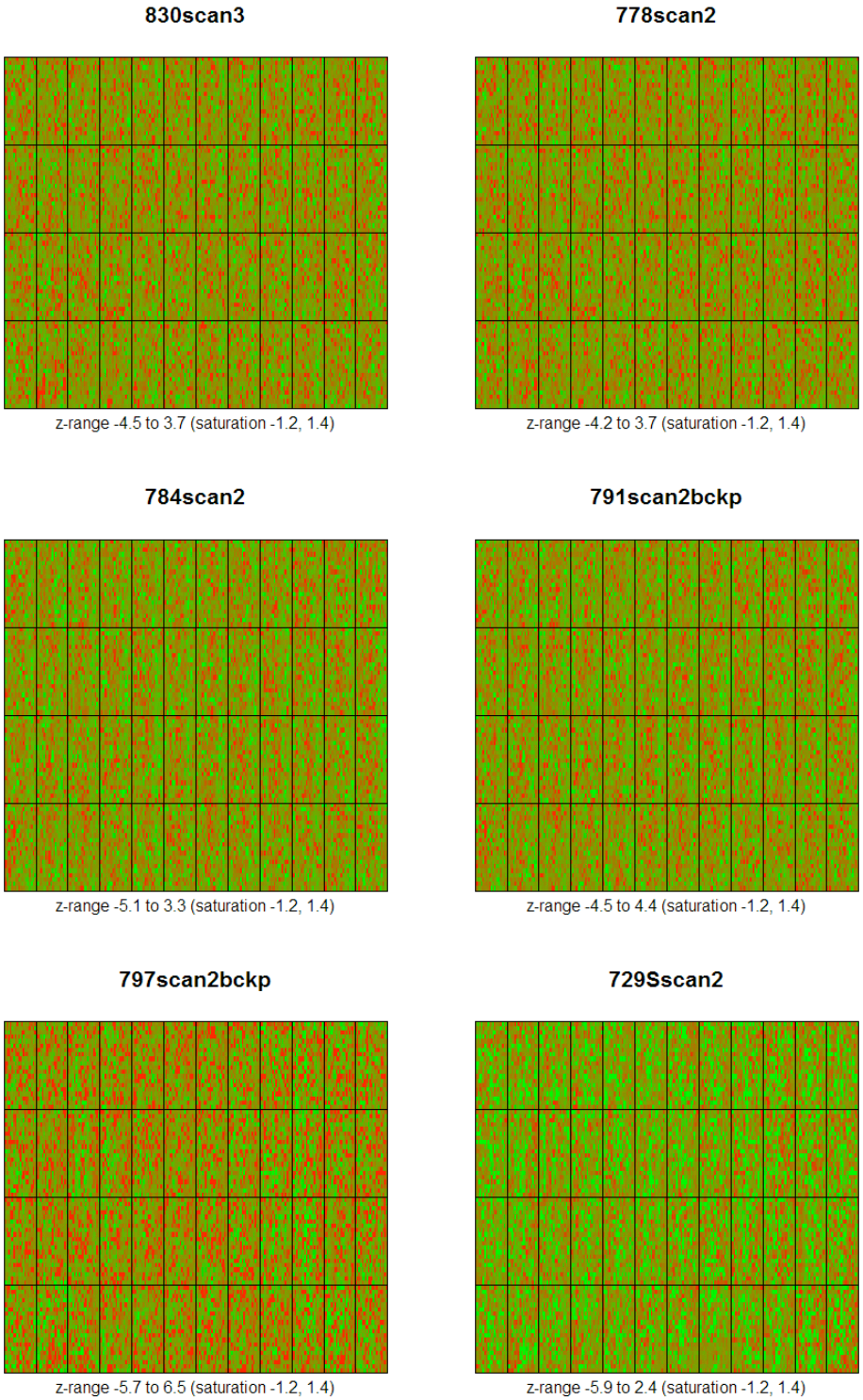
7. **Smyth, G. K.** 2004. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. Stat Appl Genet Mol Biol **3:**Article3.

8. **Smyth, G. K., J. Michaud, and H. S. Scott.** 2005. Use of within-array replicate spots for assessing differential expression in microarray experiments. Bioinformatics **21:**2067-75.

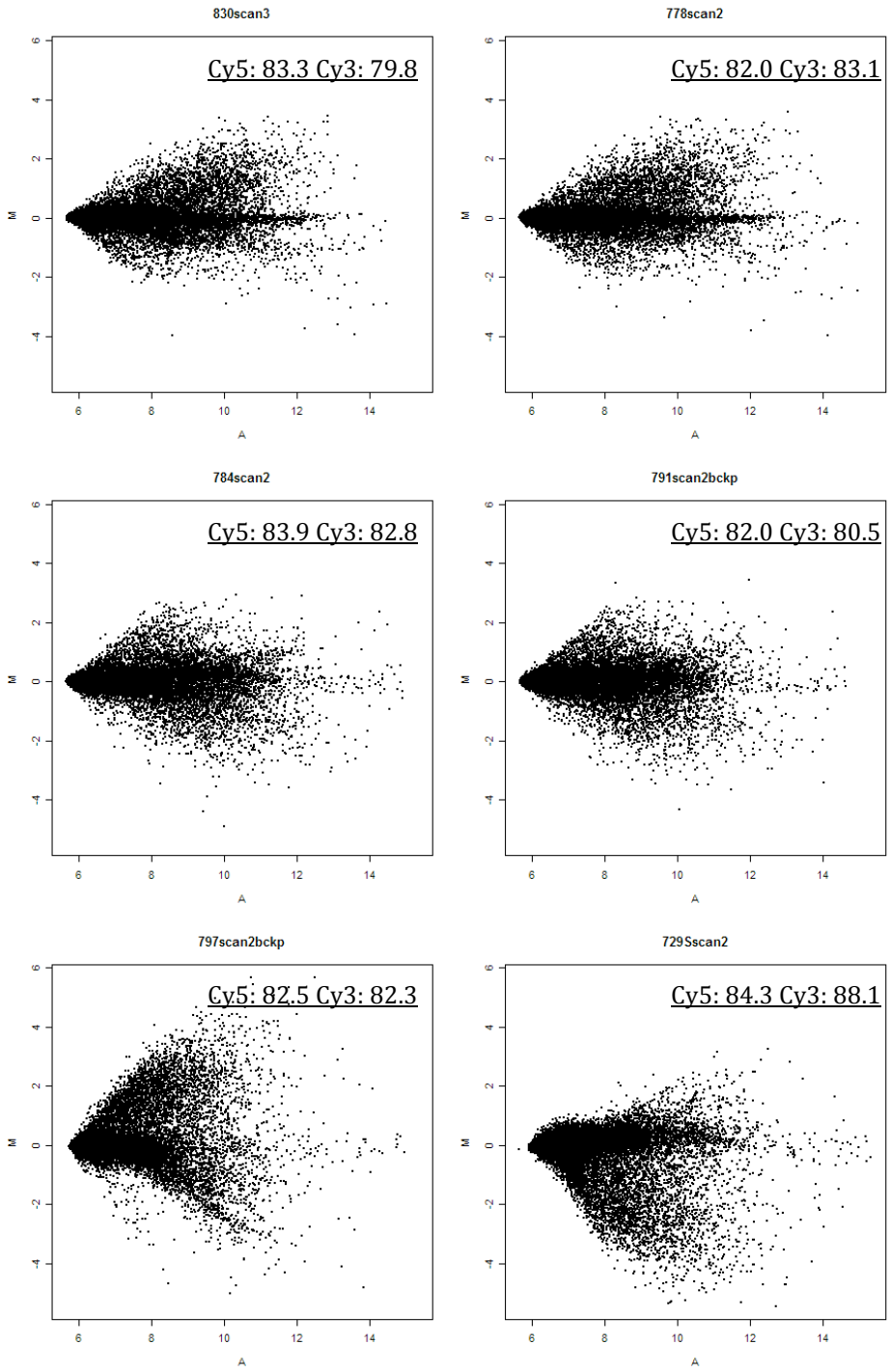
9. **Smyth, G. K., and T. Speed.** 2003. Normalization of cDNA microarray data. Methods **31:**265-73.

10. **Tavazoie, S., J. D. Hughes, M. J. Campbell, R. J. Cho, and G. M. Church.** 1999. Systematic determination of genetic network architecture. Nat Genet **22:**281-5.

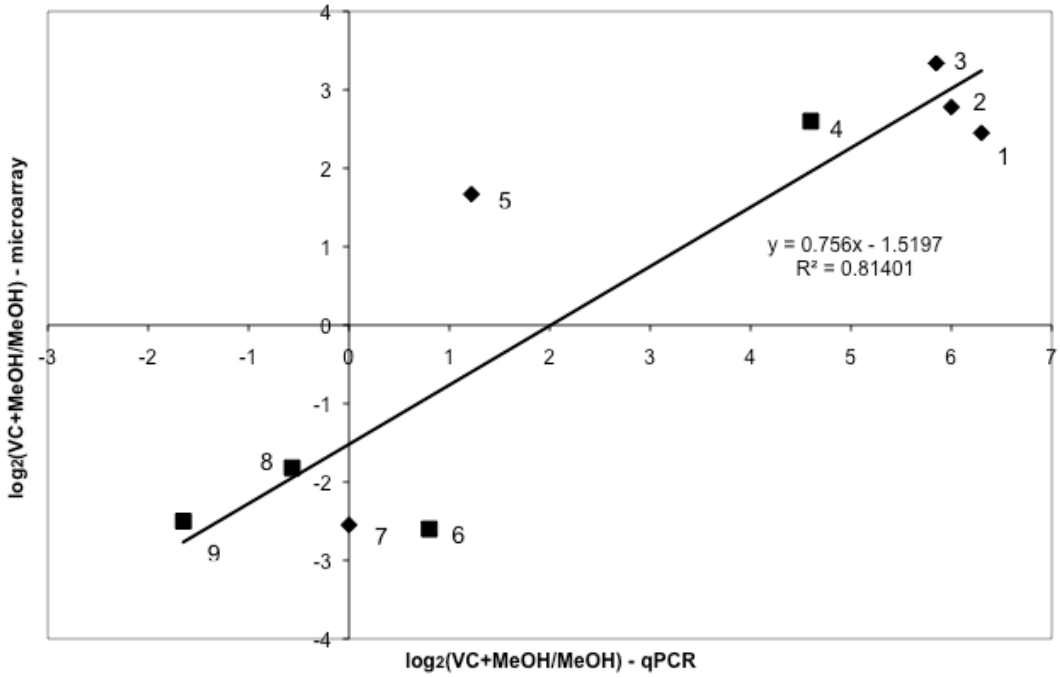
11. **Zhang RG, Skarina T, Katz JE, Beasley S, Khachatryan A, Vyas S, Arrowsmith CH, Clarke S, Edwards A, Joachimiak A, Savchenko A** (2001) Structure of *Thermotoga maritima* stationary phase survival protein SurE: a novel acid phosphatase. Structure 9:1095–1106



Supplementary Figure 1 – Image plot of the normalized log ratios (M values = [log2(Cy5/Cy3)]), where red represents high M values and green low M values. There are no spatial trends indicating minor background and hybridization variation was accounted for. Dyes were assigned as follows: for arrays 830, 778, and 797, cy5=VC+MeOH, cy3 = MeOH, for arrays 784, 791, and 729, cy5 = MeOH, cy3 = VC+MeOH.

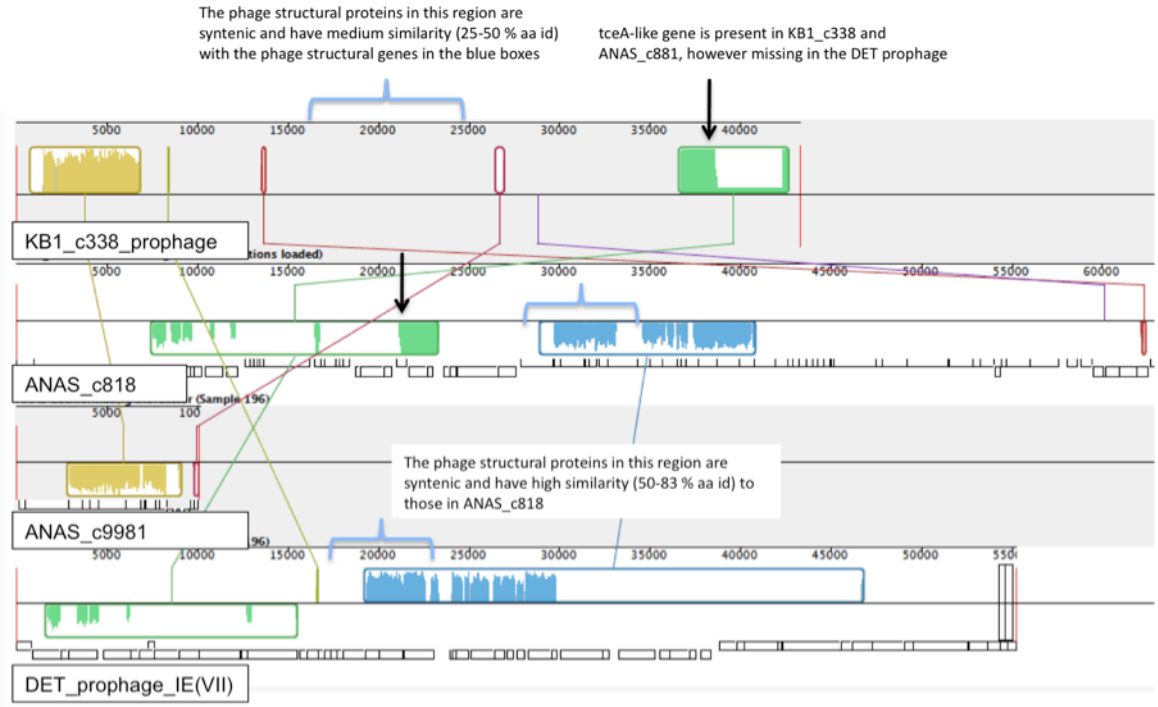


Supplementary Figure 2 – MA-plots for all six arrays after normalization
On the x-axis is plotted the A – the average intensity [1/2(Cy5+Cy3)] on the y-axis is plotted M – the log2 fold change [log2(Cy5/Cy3)]. Dyes were assigned as follows: for arrays 830, 778, and 797, cy5=VC+MeOH, cy3 = MeOH (therefore stats are based on VC+MeOH/MeOH), for arrays 784, 791, and 729, cy5 = MeOH, cy3 = VC+MeOH (therefore stats are based on MeOH/VC+MeOH). The percentage of the spots with signal-to-noise ratio (SNR) greater than 2 is indicated in the upper right corner. SNR = [(Intensity-background)/standard deviation of the background].



Supplemental Figure 3 - Comparison of qPCR Log Fold Change to Microarray Log Fold Change.

qPCR primers were designed for 9 genes detected in the spot sequences. Transcript copies were then quantified from VC+MeOH RNA-generated cDNA and compared to transcript numbers obtained from MeOH only RNA-generated cDNA. The log fold change as determined by qPCR was then plotted against the log fold change of the spot with the corresponding gene. RDH gene spots are plotted with squares. Top BLASTX hits for the spots are as follows: 1 putative magnesium and cobalt transport protein [Dehalococcoides sp. CBDB1], 2 SAM-dependent methyltransferase UbiE/COQ5 family [Dehalococcoides sp. CBDB1], 3 MaoC domain protein dehydratase [Desulfatibacillum alkenivorans AK-01], 4 - KB1 RdhA14, 5 - thioredoxin-related protein [Methanococcoides burtonii DSM 6242], 6 – KB1 RdhA13, 7- protein of unknown function DUF192 [Dehalococcoides sp. BAV1], 8- KB1 RdhA1, 9 – KB1RdhA5



Supplementary Figure 4 – An alignment of genomic fragments of putative *Dehalococcoides* prophage regions: KB1_c338_prophage, ANAS_c818, ANAS_c9981, DET_prophage_Insertion Element VII, produced by Mauve .
Boxes of the same colour indicate Locally Collinear Blocks that are shared between two or more genome fragments. The height of the lines within the boxes indicates the degree of similarity, and areas that are white correspond to areas that were not aligned, and thus specific to one sequence. Boxes that appear below the center-line are reversed in orientation than the reference (KB1_c338).

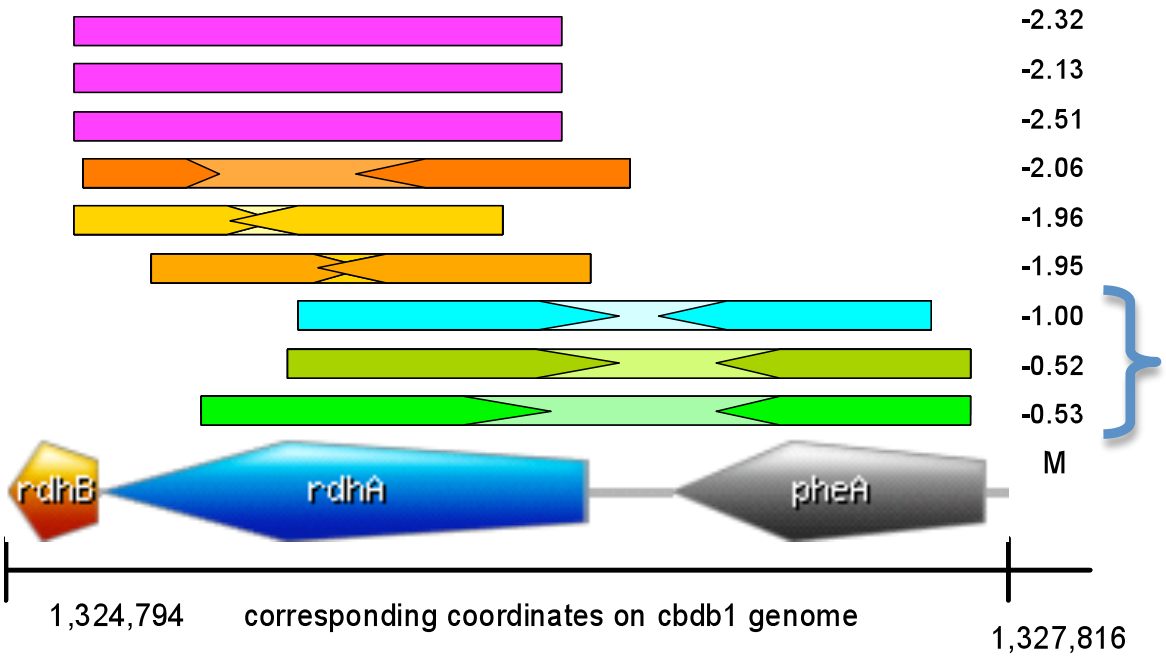


Figure S5. Alignment of *rdh*A5 Spot Sequences to the CBDB1 genome with log-fold change intensity data for each spot. The rectangles represent different spots containing rdhA5 fragments. The arrows indicate sequence information while the lighter color in between is inferred sequence. The top three pink rectangles represent cloned rdhA5 spots, the bottom three (blue and green rectangles) represent random gene-fragment spots containing multiple genes. To the right of each spot the corresponding log fold change is indicated (M value). The 3 bottom spots which contained fragments from 2 different genes did not show significant differential expression (ie. log2<1) whereas all of the spots with log2 >1.5 only contained one gene in the spotted DNA.