2016

High-throughput mutation, selection, and phenotype screening of mutant methanogenic archaea

Mary E. Walter

Alicia Ortiz

Casey Sondgeroth

Nathan M. Sindt

Nikolas Duszenko

See next page for additional authors

Authors

Mary E. Walter, Alicia Ortiz, Casey Sondgeroth, Nathan M. Sindt, Nikolas Duszenko, Jennie L. Catlett, You Zhou, Shah R. Valloppilly, Christopher Anderson, Samodha Fernando, and Nicole R. Buan

High-throughput mutation, selection, and phenotype screening of mutant methanogenic archaea

Mary E. Walter,[1] Alicia Ortiz,[1] Casey Sondgeroth,[1] Nathan M. Sindt,[1] Nikolas Duszenko,[1] Jennie L. Catlett,[1] You Zhou,[2] Shah Valloppilly,[3] Christopher Anderson,[4] Samodha Fernando,[5] and Nicole R. Buan[1]

1 Redox Biology Center, Department of Biochemistry, University of Nebraska-Lincoln, N200 Beadle Center, Lincoln, NE 68588-0664, USA
2 Morrison Microscopy Core Research Facility, Center for Biotechnology, University of Nebraska-Lincoln, E117 Beadle Center, Lincoln, NE 68588-0664, USA
3 Nebraska Center for Materials and Nanoscience, University of Nebraska-Lincoln, 855 N 16th St, Lincoln, NE 68588, USA
4 School of Biological Sciences, University of Nebraska, Lincoln, NE 68588, USA
5 Department of Animal Science, University of Nebraska-Lincoln, C220K Animal Science, Lincoln, NE 68583, USA

Corresponding author — N. R. Buan, *email* nbuan2@unl.edu

Abstract

Bacterial and archaeal genomes can contain 30% or more hypothetical genes with no predicted function. Phylogenetically deep-branching microbes, such as methane-producing archaea (methanogens), contain up to 50% genes with unknown function. In order to formulate hypotheses about the function of hypothetical gene functions in the strict anaerobe, *Methanosarcina*

Published in *Journal of Microbiological Methods* 131 (2016), pp 113–121.
doi 10.1016/j.mimet.2016.10.010

Submitted 16 August 2016; revised 18 October 2016; accepted 18 October 2016; published 19 October 2016

acetivorans, we have developed high-throughput anaerobic techniques to UV mutagenize, screen, and select for mutant strains in 96-well plates. Using these approaches we have isolated 10 mutant strains that exhibit a variety of physiological changes including increased or decreased growth rate relative to the parent strain when cells use methanol and/or acetate as carbon and energy sources. This method provides an avenue for the first step in identifying new gene functions: associating a genetic mutation with a reproducible phenotype. Mutations in bona fide methanogenesis genes such as corrinoid methyltransferases and proton-translocating $F_{420}H_2$:methanophenazine oxidoreductase (Fpo) were also generated, opening the door to in vivo functional complementation experiments. Irradiation-based mutagenesis such as from ultraviolet (UV) light, combined with modern genome sequencing, is a useful procedure to discern systems- level gene function in prokaryote taxa that can be axenically cultured but which may be resistant to chemical mutagens.

Abbreviations: **2D**, two-dimensional; **DMS**, dimethylsulfide; **DNA**, deoxyribonucleic acid; **MeOH**, methanol; **MT**, methanethiol; **ORFs**, open reading frames; **TMA**, trimethylamine; **UV**, ultraviolet

1. Introduction

Reverse-genetics, the ability to make site-directed mutations in an organism's genome, is the gold-standard for genetic experiments to assess the role of a gene on cell physiology. However, many organisms do not have genetic systems, either because they cannot be cultured easily in the laboratory, because they have robust restriction modification systems that degrade foreign DNA, or because they have formidable structural barriers (cell walls or S-layers) that prevent entry of foreign DNA. Classical forward genetics, using ultraviolet (UV) or chemically induced random mutagenesis and selecting or screening for a desired growth phenotype, has advantages in that the experimenter does not already need a gene identity or a genetic system. Combined with modern advances in genome sequencing technology, UV mutagenesis and phenotype selection/screens may be the technical approach that unlocks the metabolic secrets of microbes which lack genetic systems (Schierenbeck et al., 2015; Farrell et al., 2014; Nguyen & Valdivia, 2013; Faivre & Baumgartner, 2015; Harper et al., 2011).

We have developed several techniques to facilitate high-throughput genetic selections and screens in the strict anaerobe, *Methanosarcina acetivorans*. *M. acetivorans* is a methane-producing archaeon (methanogen). It grows by reducing methanol, methylamines, methylsulfides, carbon monoxide, or acetate to methane gas (Sowers et al., 1984). *M. acetivorans* has the largest methanogen genome sequenced to date, and the chromosome contains a high percentage of unknown genes (Galagan et al., 2002). Because methanogens are distantly related to other model organisms like *E. coli*, up to 53% of the predicted open reading frames (ORFs) in *M. acetivorans* have unknown function (Markowitz et al., 2014).

M. acetivorans can use several carbon sources to grow. Therefore, it is possible to mutate genes needed to grow on one carbon source but not on the others. *M. acetivorans* is also fully prototrophic, although supplementation of the medium with defined vitamins increases the growth rate and biomass yield of cultures (data not shown). On the other hand, addition of yeast extract to culture medium can inhibit growth for unknown reasons. The ability to generate and screen large numbers of mutant strains would be valuable for identifying other genes that play a role in methanogenesis, and in defining novel vitamin transport, scavenging, and toxicity pathways. We investigated whether we could develop a UV mutagenesis and high-throughput method to detect previously uncharacterized genes that are involved in methanogenesis from either acetate or methanol.

Though *M. acetivorans* has a reasonably well developed genetic system, the unbiased nature of UV mutagenesis and a classical phenotype selection or screen is invaluable because it will allow us to 1) define gene functions of the many unknown and hypothetical genes, 2) identify transcriptional or post-translational regulators which may themselves be constitutively expressed, 3) isolate point mutations that can inform about enzyme mechanism and/or structure, and 4) identify genes which may be involved in a given biochemical pathway that we would not have predicted (Buan et al., 2011).

Using UV mutagenesis and classical phenotype screening requires high numbers of colonies on the order of several thousand for saturation mutagenesis. The high numbers of colonies that must be screened to find mutants with a desired phenotype represents a technical hurdle when studying relatively slow-growing, strict

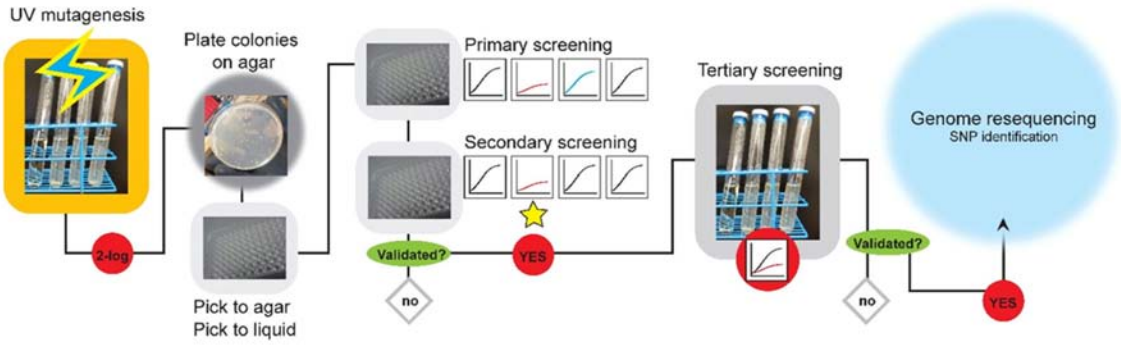

Fig. 1. Workflow for high-throughput UV mutagenesis and phenotype screening of methanogens.

anaerobes like *M. acetivorans*. For instance, classical screens in *E. coli* require several hundred petri dishes, which will take up significant bench and incubator space. However, strict anaerobes must be grown under oxygen-free conditions, and all experimental steps, including incubation of petri plates, and growth curve measurements, must be kept anaerobic (Metcalf et al., 1998). Our techniques for high-throughput genetic screening of *M. acetivorans* use a 96-well plate format (**Fig. 1**).

Although a recent paper describes the clever use of canning jars fitted with rubber gaskets to grow *Methanosarcina* in 96-well plates, specialized anaerobic equipment is necessary to grow hundreds of petri dishes and 96-well plates necessary for high-throughput genetic screening (Horne & Lessner, 2013).We have put two Wolfe incubators in our anaerobic chamber for growth of up to 306 petri dishes or 96-well plates. The Wolfe incubators are controlled by a gas manifold, allowing the growth of *Methanosarcina* under controlled atmosphere using premixed gases. Assembly of the anaerobic chamber and Wolfe incubators has been previously described (Metcalf et al., 1998). Here, we detail our current methods using this equipment for UV mutagenizing *M. acetivorans* and phenotype screening on agar and in liquid medium in 96-well plates. We also describe the growth phenotypes of 10 mutant strains generated using these procedures.

2. Materials and methods

2.1. Growth of M. acetivorans

Methanosarcina strains (**Table 1**) were grown under strictly anaerobic conditions in HS mineral salts medium (Metcalf et al., 1996; Guss et al., 2008). For growth on solid medium, cells were plated on HS medium containing 1.4% agar (w/v) with the appropriate carbon source and additions as previously described (Metcalf et al., 1998). The following anaerobic additions were added when appropriate: methanol (MeOH, 125 mM), acetate (120 mM), methanol + acetate (125 mM MeOH + 40 mM acetate), trimethylamine (TMA, 50 mM), puromycin (2 μg/ml).

2.2. Genome resequencing

Genomic DNA from 1 ml stationary-phase culture was extracted using the Qiagen Blood genomic DNA extraction kit (Qiagen, USA). Sequencing libraries were constructed using the Nextera XT DNA Library Prep Kit (Illumina, Inc.) according to the manufacturer's protocols. Bridge amplification and sequencing with the MiSeq Reagent Kit v3 (Illumina, Inc.) were performed as described by the manufacturer. The resulting sequences were demultiplexed in BaseSapce. Raw data

Table 1. Strains described in this study.

NB#	*Phenotype*	*Reference*
34	Parent (Δ*hpt*::ϕC31 *int*, *attP*)	(Guss et al., 2008)
89	Slow growth on methanol or methanol + acetate. Struvite crystals on acetate.	This study
90	Slow growth on methanol. Struvite crystals on acetate.	This study
91	Fast growth on methanol or methanol + acetate. Struvite crystals on acetate.	This study
92	Slow growth on methanol. Struvite crystals on acetate. Struvite crystals on acetate.	This study
93	Slow growth on methanol, methanol + acetate, and acetate.	This study
178	Slow growth on methanol.	This study
181	Slow growth on methanol.	This study
193	Slow growth on methanol.	This study
194	Slow growth on methanol or methanol + acetate.	This study
195	Slow growth on methanol.	This study

is available under the BioProject accession number PRJNA335646. Adapter and transposon associated sequences and the last 25 bp of the forward read, and the last 100 bp of the reverse read were removed using cutadapt (version 1.8.1) (Martin, 2011). The 'fastq_filter' command in USEARCH (v8.0.1623) was used to remove sequences containing any ambiguous bases, an expected maximum error rate N0.01, and sequences shorter than 100 bp (Edgar, 2010). Only reads where both the forward and reverse read were retained after quality control were used in downstream analysis. The Burrows-Wheeler Alignment tool (BWA) (version 0.7.12) was used to align reads against the reference genome, *Methanosarcina acetivorans* str. C2A (NC_003552.1) (Galagan et al., 2002; Li & Durbin, 2009). For each mutant strain, a combination of SAMtools, Picard Tools, and BEDTools were used to mark duplicate reads and summarize the read coverage across the reference genome (Li, 2011; Institute B, n.d.; Quinlan & Hall, 2010). SNPs and indels were identified using the 'mpileup' command (default parameters) in SAMtools (Li, 2011).

2.3. Microscopy

Crystals from acetate-grown cultures were harvested, placed on 0.2 μm filters, and split into three treatment groups; washed twice with 100 μl distilled deionized H_2O, washed twice with 70% ethanol, or washed twice with 0.4 M NaCl (iso-osmotic with the methanogen growth medium). Crystals were dried at 37 °C for 48 h. Crystals were fixed to glass slides with glutaraldehyde, sputter-coated, and observed using a Hitachi S-3000 Variable-Pressure Scanning Electron Microscope.

2.4. Powder x-ray diffraction

Crystals were harvested, washed with 400 μl 0.4MNaCl, and dried as described above. Dried crystals (0.0155 g)were ground to powder using a glass pestle in a microcentrifuge tube. The unknown powder was analyzed using a Bruker-AXS D8 Discover High-Resolution Diffractometer with a General Area Detector Diffraction System (GADDS) at the Nebraska Center for Materials and Nanoscience core facility (NCMN). A highly collimated beam of Cu K α radiation delivered by a combination of Goebel Mirror and 0.5 mm monocapillary tube was used for

the experiment. The powder sample spread on a glass slide was continuously spun to minimize preferred orientation effects in the diffraction intensities. 2D diffraction data (Debye-rings) was collected, while powder sample and the 2D detector were rotated with respected to the incident beam, in 4 frames. With step-wise increment of 25°, 2D diffraction data was obtained in the angular range of 2Θ ~ 0–95° and then the data was integrated along the diffraction cone to obtain the 2Θ vs. Intensity pattern as shown in Fig. 5E. Culture medium atom composition (Na, Mg, Ca, K, P, O, C, N, Cl) was used as a filter when matching the unknown sample data using the ICDD PDF-4+ database. A close match was observed for the entry corresponding to that of the mineral Ammonium Magnesium Phosphate Hydrate, (NH_4) Mg (PO_4) $(H_2O)_6$, also known as struvite. Furthermore, the structure was confirmed and refined by a detailed powder x-ray diffraction profile analysis with a starting model based on ICDD PDF card 04-010-2533. Struvite has orthorhombic crystal structure with a space group of *Pmn*21. Fig. 5E illustrates the experimental data, simulation, and the Bragg reflection positions of the refined structure. The refined lattice parameters are a = 6.934 Å, b = 6.129 Å and c=11.198 Å. Traces of no other chemical phases were detected in the x-ray diffraction.

3. Results

3.1. UV mutagenesis to isolate methanogenesis mutants

UV irradiation is ideal to induce mutations because it is not susceptible to hot-spot mutation bias, as can be observed with transposon or insertion element-induced mutations (Manna et al., 2001; Mullany et al., 1991; Lee et al., 1987). UV light induces thymine dimer formation in DNA, which must be repaired by site excision before chromosomal replication can commence (Witkin, 1976). Because different *Methanosarcinales* species or strains may produce varying amounts of heteropolysaccharide on the outside of the cell, and may have differences in the S-layer permeability, we wanted to avoid using a chemical mutagen (Hartmann & Konig, 1991; Sowers & Gunsalus, 1988; Steber & Schleifer, 1975; Klingl, 2014; Rohlin et al., 2012; Francoleon et al., 2009; Namboori & Graham, 2008; Karcher et al., 1993; Firtel et al., 1993; Mengele & Sumper, 1992; Zellner et al., 1989; Conway de Macario et

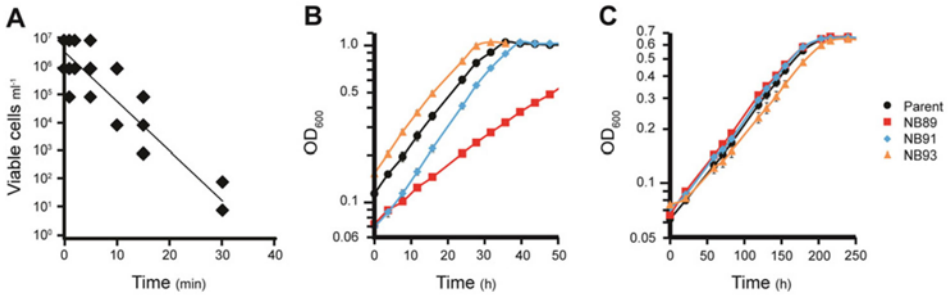

Fig. 2. Viability counts and growth curves of UV-mutagenized strains. A, UV killing of *Methanosarcina* in Balch tubes. Exponentially growing cultures were exposed to UV irradiation and diluted to extinction in 96-well plates to determine a killing curve. B and C, Cultures grown on methanol (B) or acetate (C) as energy source.

al., 1984). However, a drawback with UV irradiation is that UV lamp intensities and exposure can vary. To optimize the UV-mutagenesis we generated a UV killing standard curve method to determine the optimal UV exposure time in our experiments (**Fig. 2**A).

Before UV exposure, 0.5 ml of a recent stationary phase culture was inoculated into fresh Balch tubes containing 10 ml methanol + acetate medium and allowed to grow at 35 °C overnight to approximately mid-exponential phase. The next morning, the cultures were laid on a UV transilluminator, covered in foil, and exposed to UV irradiation for up to 30 min. Cultures were then serially 10-fold diluted to extinction in a 96-well plate in the anaerobic incubator into fresh methanol + acetate medium. The 96-well plates were placed in the Wolfe incubator at 35 °C for 10 days to grow. Plates were scored to determine the number of viable cells per ml after UV exposure. We determined that 10–15 min exposure reliably resulted in approximately 2-log killing.

3.2. Isolation of mutants with increased or decreased growth

Once the optimal UV mutagenesis conditions were determined, we mutagenized methanol+acetate cultures and plated 100 µl onto high mineral salt (HS) methanol+acetate agar plates to obtain isolated colonies. Plates were incubated for 14 days to allow visible colonies to develop. Next, 96-well agar plates were filled to 200 µl per well with molten agar dispensed from a sterile reagent trough using a serial pipettor. After the agar solidified, individual colonies were picked using

Table 2. Mutant screening results

Screen	*Format*	*Number of strains screened*	*Wild-type growth phenotype*	*Reduced growth on acetate*	*Reduced growth on methanol*	*Reduced growth on methanol + acetate*
Primary screen	96-well	960	827	14	108	11
Secondary screen	96-well	35	25	2	6	2
Tertiary screen	Balch tubes	11	1	1	9	10

sterile sticks and transferred to individual wells, sequentially on a methanol + acetate plate first, onto a methanol plate, and then onto an acetate plate, taking care to keep the same plate coordinates for each colony. Racked, sterile pipette tips were used as a frog to replicate colonies from one 96-well plate to another. Once wells were inoculated, they were placed in the Wolfe incubator to grow at 35 °C. After 2weeks, plates were compared to identify colonies that appeared to have poor or no growth on one carbon source but not on the other. Cells were then picked using sterile sticks and inoculated into methanol+acetate liquid medium in 96-well plates for secondary screening. Using this method we screened 960 colonies in two weeks. We identified 11 with decreased growth on both acetate and on methanol. 108 mutants had less growth on acetate but wild-type growth on methanol, and 14 mutants had decreased growth on methanol but wild-type growth on acetate (**Table 2**).

3.3. Growth curves in 96-well plate format

Secondary screening in liquid medium in 96-well plate format is advantageous because manual growth curves using a spec20 spectrophotometer is laborious. We identified a large number of primary hits using high-throughput screening on agar plates, and a high throughput method to screen mutant growth in liquid cultures was desirable. To do this, a 96-well plate reader and computer were set up within the anaerobic chamber. Care should be taken to keep the anaerobic chamber very dry while sensitive electronics are housed in the chamber under a CO_2 atmosphere. We have been able to keep the plate reader and computer running for at least two years with no discernible effects on performance in our anaerobic chamber.

To screen mutant strains in liquid medium, half the 96-well plate was filled with 180 µl methanol medium, and the other half was filled with 180 µl acetate medium. Mutant cultures (20 µl) were inoculated into wells of both media in the desired number of biological replicates. Growth in each well was measured automatically using the plate reader at 35 °C with high shaking intensity at 2 h intervals for 5 days (growth on methylotrophic substrates) or two weeks (growth on acetate). Using plate reader software, growth curve data was exported to Excel format for analysis.

Using this method we tested a subset of 35 mutant strains that had decreased growth on methanol or acetate. Of the 35 mutants tested, 10 were confirmed to have reproducible growth phenotypes. The mutants fell into two phenotypic classes. Class I were isolates that had decreased growth on methanol, and Class II were those that had decreased growth on acetate.

3.4. Tertiary screening in sealed glass Balch culture tubes

Methanogen medium is often supplemented with sulfide to lower the redox potential (Mountfort & Asher, 1979). Because hydrogen sulfide is volatile, there is a possibility that growth phenotypes may be different when cultures are grown in 96-well plates versus sealed glass Balch tubes. For instance, deletion of the operon encoding CoM-S-SCoB heterodisulfide reductase, *hdrABC*, causes the cell to produce methanethiol (MT) or dimethylsulfide (DMS), which diffuses into the headspace (Buan & Metcalf, 2010). In a Balch tube, the MT and DMS eventually reaches a threshold partial pressure that induces expression of dimethylsulfide methyltransferases, and the cell consumes DMS. However, the slow-growth phenotype of the Δ*hdrABC* mutant is exacerbated in a 96-well plate, presumably because the MT and DMS can easily diffuse out of the plate into the bulk gas of the anaerobic chamber (data not shown).

The 10 strains that passed secondary screening were grown in Balch tubes to determine if their growth phenotypes were specific to the 96-well format. To eliminate experimental variability, the mutants and parental control were grown in biological replicates on the same batch of culture medium. Of the 10 strains, nine had a reproducible phenotype (**Table 3**). One strain, NB93, had reduced growth rate under all

Table 3. Growth rates and maximum culture density in Balch tubes.

Carbon source	*Strain*	*Generation time (h)*	*Std dev*	*p value vs parent*[1]
Methanol	Parent	10.4	0.13	1
	NB89	16.5	0.27	0.0000016
	NB90	10.7	0.13	0.034
	NB91	9.2	0.15	0.000040
	NB92	10.2	0.05	0.045
	NB93	11.1	0.17	0.0021
	NB178	11.1	0.37	0.041
	NB181	11.2	0.11	0.00017
	NB193	12.1	0.22	0.000099
	NB194	12.0	0.17	0.000022
	NB195	11.7	0.19	0.00012
Methanol + acetate	Parent	10.5	0.12	1
	NB89	20.9	3.08	0.0099
	NB90	11.4	1.07	NS
	NB91	9.25	0.20	0.0049
	NB92	10.2	0.03	NS
	NB93	11.8	0.58	0.023
	NB178	12.4	1.53	NS
	NB181	12.0	0.94	NS
	NB193	12.9	2.03	NS
	NB194	13.5	1.52	0.040
	NB195	12.1	1.08	NS
Acetate	Parent	55.0	2.50	1
	NB89	53.4	1.57	NS
	NB90	61.9	4.38	NS
	NB91	57.1	3.89	NS
	NB92	54.9	3.14	NS
	NB93	69.1	3.11	0.0010
	NB178	60.9	5.14	NS
	NB181	56.5	5.73	NS
	NB193	58.9	4.66	NS
	NB194	57.6	3.40	NS
	NB195	53.3	3.38	NS

Calculations are based on growth phenotypes from quadruplicate biological replicates.
NS: not significant ($p > 0.05$).
1. p value determined by Student's two-tailed t-test ($p < 0.05$).

medium conditions tested (growth on methanol, methanol + acetate, and acetate alone). Eight strains had wild-type growth rates on acetate, but decreased growth rate on medium containing methanol or methanol + acetate. One of these strains, NB89, had a severe methanol- specific growth defect that resulted in a 59% slower growth rate and a 10% decrease in biomass on methanol but a 20% decrease in

biomass on methanol + acetate. Finally, one strain, NB91, had wild-type growth on acetate, but a 13% increase in growth rate on methanol and methanol + acetate (**Fig. 2**).

3.5. Identification of candidate mutations associated with growth defects

To determine the effectiveness of UV-induced mutagenesis in *M. acetivorans*, genomic DNA from the parent strain (NB34) and nine mutants was resequenced and compared to the reference genome (Table S1). (Markowitz et al., 2014) As expected, mutations were recovered from all over the chromosome (**Fig. 3**A). Two regions of the chromosome, from 2.8–3.0 Mb and 5.5–5.7 Mb accumulated mutations faster than the expected average frequency, and correlated with thymine-rich repeat sequences (Fig. 3B). The mean number of mutations per strain was 10 (8 SNPs and 2 Indels) with a median number of 5 mutations (**Table S2**). On average, 80% of the mutations were SNPs while 20% were Indels. The UV mutagenesis method described in this manuscript resulted in a frequency of 1.8×10^{-6} mutations per base (**Table S3**).

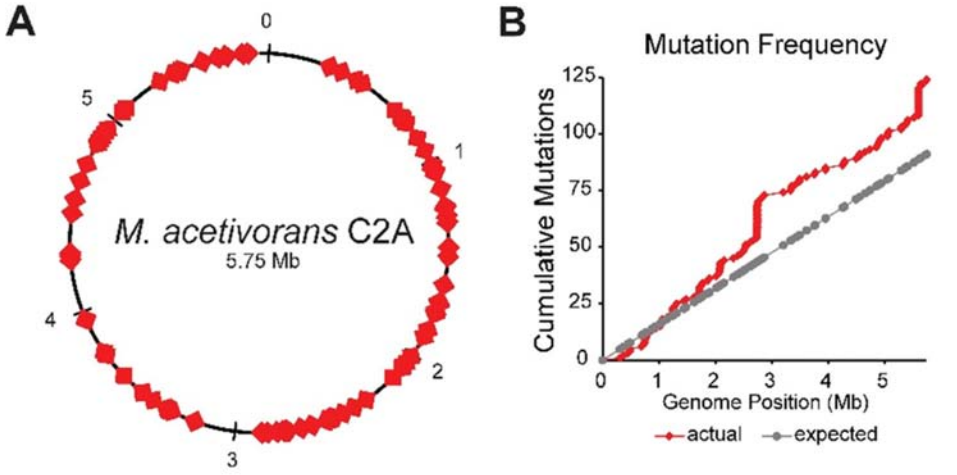

Fig. 3. Mutations generated by UV irradiation. A, Mutations recovered from the parent strain and eight mutant strains are plotted on the circular genome. B, When the cumulative mutations per site (red) is compared to the mean average cumulative mutations (gray), two mutation hotspots are identified at 2.8–3 Mb and 5.5–5.7 Mb.

To identify mutations which may be responsible for the observed phenotypes, the resequenced genomes of the NB34 parent and each mutant were compared to the reference *M. acetivorans* genome (**Fig. 4**). Sixteen mutations (of 306 total) were found to be in common between all strains. Ten mutations were found in the parent and one or more mutant strains, but seemed to have reverted back to the reference sequence in the remaining strains. Of these mutations, all but one were intergenic regions or genes with hypothetical or unknown function. One mutation, a deletion of a GC at position 5079505, was in

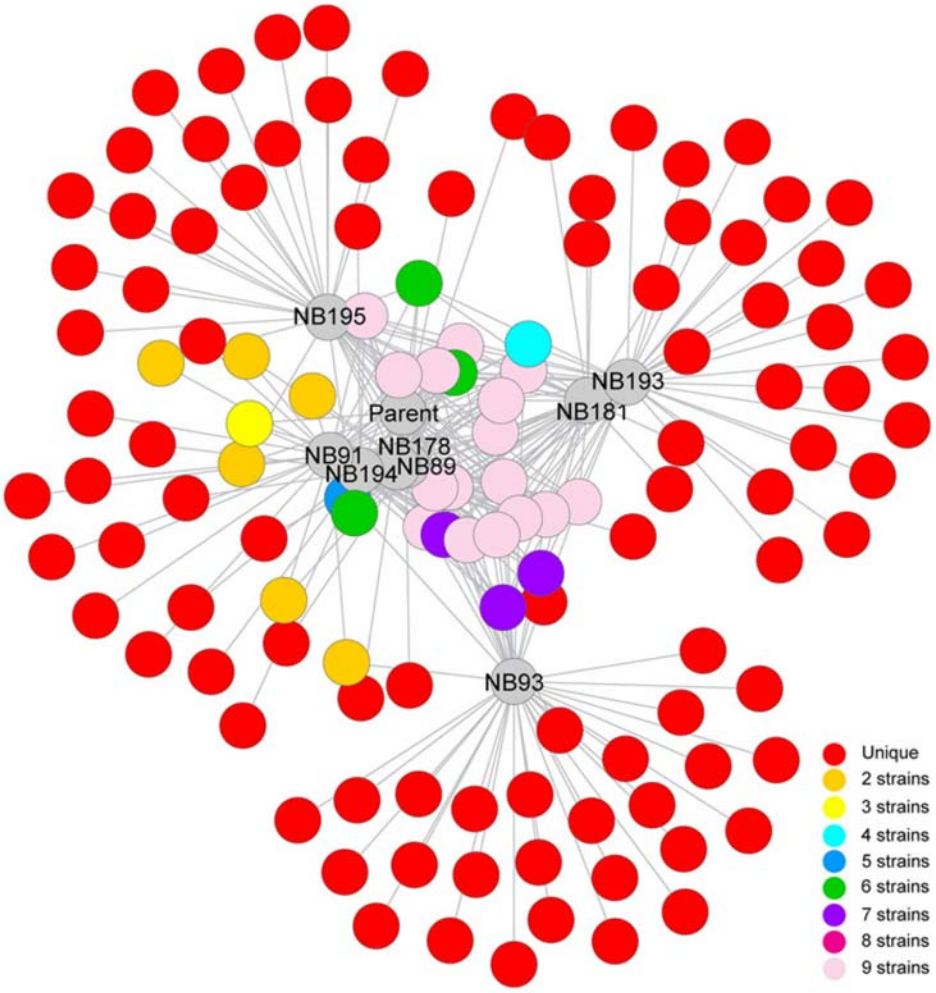

Fig. 4. Identification of unique mutations in resequenced genomes. The reference *M. acetivorans* genome was compared to the resequenced parent strain and the isolated mutant strains. Most mutations were unique (red circles), but a few were shared by two or more sequenced strains.

Table 4. Biomass measurements.

Carbon source	*Strain*	*Max* OD_{600}	*Std dev*	*p value vs parent*[1]
Methanol	Parent	1.05	0.008	1
	NB89	0.94	0.005	0.0000040
Methanol + acetate	Parent	1.03	0.030	1
	NB89	0.82	0.050	0.0018
Acetate	Parent	0.67	0.01	1
	NB89	0.68	0.01	NS

Calculations are based on growth phenotypes from quadruplicate biological replicates.
NS: not significant ($p > 0.05$).
1. p value determined by Student's two-tailed t-test ($p < 0.05$).

the coding region of *hypF*, the gene encoding the hydrogenase maturation protein. *M. acetivorans* C2A and the parent strain, NB34, do not synthesize functional hydrogenases and cannot use hydrogen as an electron donor (Sowers et al., 1984). Due to lack of functional hydrogenase expression in the C2A strain, this deletion can be reasonably expected to have no additional effect on the physiology of NB34 or the mutant strains (Guss et al., 2009).

Strain NB89 had the most severe methanol-specific growth defect. Not only was the growth rate slower than the parent strain on methylotrophic substrates, but the final culture density was also affected, suggesting an energetic or a toxicity effect (**Table 4**).

The NB89 genome contained five mutations not found in the parent or in the reference genome (**Table 5**).

Three mutations were single base changes, while two were deletions of thymine-rich sequences. Two of the five mutations were in intergenic regions, while three mutations were in predicted protein-coding regions. Based on the genome position and predicted effect, the mutations most likely to contribute to the methanol growth defect

Table 5. Mutations identified in the NB89 genome.

Genome position	*Mutation*	*Effect*	*Locus*	*Predicted function*
308383	C to T	Δ160–504	MA_RS01355	AAA+ ATPase
1146094	T to C	N199D	MA_RS24380	PGF transpeptidase
3205790	A to G	intergenic	(tnp)	none
3753590	ΔCTCTGTTGTTAACCCCCTGA	intergenic	(Na^+/H^+ antiporter)	none
5645793	ΔTGATAGATGACGTTCCCTGTATCTTCTC	in-frame ΔEKIQGTS	MA_RS23910	formyl transferase (H4PT or tMet)

are a base substitution at 1146094, an intergenic deletion at position 3753590, and a deletion at position 5645793. The mutation at position 1146094 results in a missense mutation (N199D) in a PGF transpeptidase of unknown function. The mutation at 3753590 results in a deletion 5′ upstream of the *mrpA* Na^+/H^+ antiporter. MrpA has been shown to contribute to growth on methanol under low sodium conditions (Metcalf et al., 1998). It would seem possible that the 3753590 deletion mutation may have disrupted regulation of *mrpA* gene expression, somehow affecting growth on methanol but not on acetate. However, the mutagenesis and screening was conducted in medium containing 0.4 M NaCl. Under these conditions (substrate and sodium ion concentrations), the Δ*mrpA* mutant displayed wild-type growth (Jasso-Chavez et al., 2013). Therefore the mutation at 3753590 in strain NB89 is unlikely to be the cause of the slow growth phenotype on methanol. Finally, the 5645793 mutation causes an in-frame deletion of residues 74–80 (EKIQGTS) in a predicted formyl transferase of unknown function (locus MA_RS23910). Overall the predicted polypeptide is 87 amino acid residues in length and 10.1 kDa. The deleted residues in the NB89 mutants strain lie just outside the formyltransferase core domain residues 2–68. Nearby on the genome are two methionyl tRNAs, MA_RS23925 and MA_RS23930, suggesting that MA_RS23910 may function as a methionyl-tRNA formyltransferase. A crippled methionyl-tRNA formyltransferase would seem to fit the NB89 phenotype if reduced enzyme activity was enough to satisfy the growth rate on acetate, but was not fast enough to satisfy a faster growth rate on methanol. To directly determine which of these mutations causes the growth defect of the NB89 strain it will be necessary to recreate the mutations and rescue the mutant phenotype by complementation.

Several mutations in bona fide methanogenesis genes or promoters were recovered in the mutant strains. Specifically NB93 has mutations in the methyl-coenzyme M *mcr* promoter and in the acetyl-CoA synthase/decarbonylase synthase promoter, NB181 has a frame shift in $F_{420}H_2$:methanophenazine oxidoreductase *fpoL* subunit, NB193 has a mutation in the *mtsF* methyltransferase promoter, and NB194 has a mutation that results in truncation in the *mtbC1* corrinoid protein. The mutations found in each of the sequenced strains are shown in **Figs. S1–8** and summarized in **Table S4.**

Many of the mutations detected were found in TT-rich regions such as intergenic terminators and promoters, and repeat regions within coding sequences. This is consistent with the mechanism of UV-induced mutagenesis (pyrimidine dimers) and repair by the non-homologous end-joining (NHEJ) DNA repair pathway (Blackwood et al., 2013). Therefore the amount and duration of DNA exposure affects the types of mutations recovered and must be experimentally titrated. For instance, excess UV irradiation would be expected to result in numerous severe indel mutations, such as we observed in strain NB195. Strain NB195 had 84 bases inserted at position 273884 (MA_RS11545), four insertions of the same 39 base sequence and six missense mutations in MA_RS1575, and three other deletions at positions 3958294, 4562039, and 5643060. In contrast, strain NB178 had only four point mutations and one insertion at an intergenic repeat region at position 3345530.

3.6. Struvite formation on acetate

When collecting tertiary screening data in Balch tubes, NB89 and three other strains were observed to produce crystals during growth on acetate. Crystals were large, between 2 and 6 mm in length, and rhombic (**Fig. 5**A). Small crystals are occasionally observed in parent strain culture tubes when settled, stationary-phase cultures are left on the

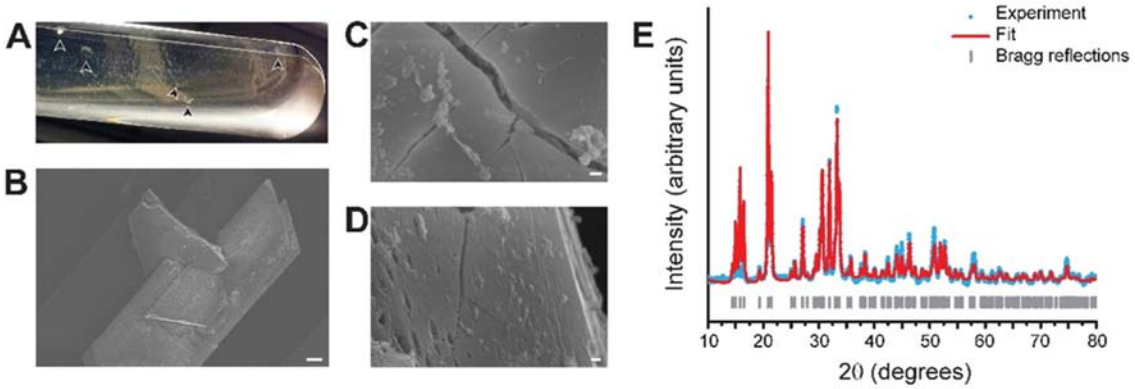

Fig. 5. Acetate-dependent struvite formation. A, Balch tube showing large rhombic struvite crystals (black arrows). B, Scanning electron microscopy of struvite crystals (bar=10 μm). C and D, Biological material suggestive of a biofilm (C) or adhered cells (D) were observed on crystal surfaces (bar = 1 μm). E, X-ray powder diffraction spectra of crystals.

benchtop for weeks (as might occur after an experiment is completed but before Balch tubes are cleaned). The crystals we observed in the mutant strains were different, however, because of the size, ordered structure, and speed of crystal formation. Crystals began to form as mutant cultures entered late exponential phase, and were reproducible. The acetate crystal phenotype appeared to be heritable when cells were passaged to fresh acetate medium, and when they were switched to methanol medium and back to acetate medium. Crystals were not observed in all mutants, nor were they observed in parent cultures.

We collected the crystals and tested their ability to dissolve in acid. The crystals were air-stable, and dissolved slowly in cold 0.1 M HCl, indicating they were not composed of carbonate. We observed crystal morphology by scanning electron microscopy (**Fig. 5**B–D). Though cells, or perhaps biofilm, were observed on the surface of the crystals, the crystal faces were generally smooth, and it did not appear as if cells were trapped within the crystals. Crystal composition was identified using powder x-ray diffraction. The XRD spectra was characteristic of the mineral $(NH_4)Mg(PO_4)(H_2O)_6$, otherwise known as struvite (**Fig. 5**E). The microscopy and XRD data therefore suggest the mutant strains catalyze struvite mineralization by increasing the pH of the medium beyond the bicarbonate/KPO_4 buffering capacity.

4. Discussion

4.1. Method evaluation

A forward genetics approach has the potential to identify every gene that is involved in expression of a phenotype. In the method described, UV mutagenesis was selected as an inexpensive, universally applicable mutagenesis method. Other forward genetics methods employ chemical mutagens or transposable elements, which have their own strengths and weaknesses (Guss et al., 2008; Sarmiento et al., 2013; Sattler et al., 2013; Rother et al., 2011; Pritchett & Metcalf, 2005; Zhang et al., 2000). A particular organism may or may not be sensitive to a particular chemical mutagen, or the cells may have extracellular structures (capsule, hydrophobic pellicle, cell wall structures, etc.) or biochemical detoxification mechanisms (P_{450}-type enzymes) that may

pose a barrier for mutagen uptake or sensitivity. Transposon mutagenesis requires a method for efficient delivery of DNA into cells, either by transformation/ transfection, conjugation, or phage/virus infection. In addition, some transposons favor insertion "hot spots", which may result in over-representation of some recovered mutants and under-representation of others (Lee et al., 1987; Gamas et al., 1987;Wang et al., 2006).

UV mutagenesis generates point mutations across the entire genome, which can result in nonsense or frame-shift mutations in either the promoter (or other cis-regulatory elements) or in the gene reading frame. The degree of phenotype change is a function of how severe the mutation is, ranging from having a subtle effect on transcription/translation, to complete gene loss due to a translational frame shift. One major difficulty with a forward genetics approach is identifying the mutation on the chromosome. Platforms such as the Illumina MySeq are ideally suited for detection of single-nucleotide polymorphisms, which are easily mapped against a genome sequence. A read depth of 20,000× coverage is generally sufficient for identifying mutations derived from the initial UV mutagenesis event that have become fixed in the culture population. Forward genetics is more experimentally involved and initially expensive than reverse genetics, however, for organisms that lack an efficient genetic system and/or have a relatively high proportion of unknown genes (such as in phylogenetically deep-branching microbes), forward genetics makes it possible to begin to link genes and transcripts to systems-level phenotypic functions.

While several methanogen species are genetically tractable, many genes remain uncharacterized, and there are likely many genes of unknown function that play a role in growth and methane production. In *M. acetivorans*, acetoclastic methanogenesis is thought to involve 303 genes (by microarray) or 257 genes (by proteomics), and methylotrophic methanogenesis involves 407 genes (by microarray) and 441 genes (by proteomics), for an overall range of 664 to 774 genes which are specifically involved in either pathway, or roughly 13.8% to 16.1% of genome (Li et al., 2007; Li et al., 2005a; Li et al., 2005b). Therefore, the probability of finding a mutation in either acetoclastic or methylotrophic pathways is 4.4×10^{-4} under the conditions we describe.

For full saturation mutagenesis of the 5.75 Mb *M. acetivorans* genome with 4807 genes, we should screen approximately 10,000

colonies (Phogat et al., 2001). Assuming a single mutation per colony (the actual number of mutations per chromosome will follow a Poisson distribution and vary depending on mutagenesis conditions), we sampled up to 20% of the genome. Theoretically, we mutated between 133 and 154 genes in either the acetoclastic or methylotrophic pathways, and we picked 133 colonies, 13.9%, that appeared to have slow or fast growth from the primary screen, in close agreement with the theoretical number predicted from transcriptomics and proteomics. After secondary and tertiary screening, the number of confirmed mutants dropped to 10, and may have resulted from human error when transferring cells from 96-well plates to Balch tubes, may reflect the change in growth conditions between 96-well plates and Balch tubes, or may be a consequence of selection for suppressor or reversion mutants upon subculturing.

Finally, we note that genome copies can affect the efficiency of forward genetic screens. Prokaryotes, and methanogens in particular, may have more than one copy of the genome. Multiple genome copies can serve as templates for DNA repair by homologous recombination. For instance, when *M. acetivorans* grows on methanol, each cell contains 16 copies of the genome (Soppa, 2011; Hildenbrand et al., 2011). In order to recover stable mutant strains, all copies of the genome must contain the mutation. If a mutation in one of sixteen gene copies causes a decrease in fitness, it is only a matter of a few generations before the genome copies segregate, homologous recombination replaces the defective lesion, and the population genotype is converted back to the most-fit allele. Therefore initial selection of single colonies under permissive conditions is critical for maximizing the chance for obtaining strains that will exhibit mutant phenotypes under an alternative (or stressed) growth condition.

4.2. Struvite formation can be stimulated by chromosomal mutations

We were intrigued that several of the mutants produced struvite crystals. Struvite can precipitate abiotically at alkaline pH when phosphate and ammonium concentrations are high. Our observation that struvite formation was reproducible, heritable, and constrained to acetoclastic growth suggests these mutants have mutations that increase nitrogen fixation, have increased CO_2 fixation, or increased proton transport into the cell. Any alteration in these metabolic activities would

be expected to tip the balance of conditions towards favoring struvite formation.

There is no current evidence directly linking methanogens with struvite mineral precipitation. However, struvite precipitation is common in anoxic environments inhabited by methanogens. For instance, struvite formation is used in municipal water treatment facilities to capture phosphate and ammonia (Doyle & Parsons, 2002; Miles & Ellis, 2001; Scott et al., 1991). However, struvite can build up in the sediment within pipes carrying anoxic, untreated waste, eventually clogging them. Clearing struvite requires removal and replacement of the obstructed sections, at considerable cost.

Struvite can also form in biological tissues, forming kidney stones, gall bladder stones, and encrusting catheters. Oxalate-producing pathogenic bacteria are known to be associated with kidney and urinary endoliths comprised of oxalate, but there is evidence that some patients may present non-oxalate struvite endoliths that are not associated with bacterial infection (Rodman, 1999). Struvite endoliths also cause significant morbidity in some breeds of canines and equines, where they are more commonly found in the colon (Hassel et al., 2001). In the human colon, struvite (non-oxalate) endoliths are associated with Crohn's disease and diverticulitis, diseases which are also linked to perturbations in the methanogen population of the gut microbiome (Hassel et al., 2001; Hassel et al., 2009; Chou et al., 2001; Lomhoff & Dubowy, 1948; Harris et al., 1997). Further research is needed to determine if and how methanogens may contribute to altering production of organic acids, CO_2 and/or ammonia by the microbiome to contribute to struvite endolith formation in vivo.

Acknowledgments — The authors would like to thank Shah Valloppilly in the Nebraska Center for Nanomaterials Research for help with x-ray diffraction.

Funding — This work was supported by National Science Foundation grant IOS-1449525 to NRB and by the Nebraska Tobacco Settlement Biomedical Research Development Funds. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the funding agencies.

Conflict of interest — NB has disclosed a significant financial interest in RollingCircle Biotech, LLC. In accordance with its Conflict of Interest policy, the University of Nebraska-Lincoln's Conflict of Interest in Research Committee has determined that this must be disclosed.

Appendix A. Supplementary data follows the **References.**

References

Blackwood, J.K., Rzechorzek, N.J., Bray, S.M., Maman, J.D., Pellegrini, L., Robinson, N.P., 2013. End-resection at DNA double-strand breaks in the three domains of life. Biochem. Soc. Trans. 41:314–320. http://dx.doi.org/10.1042/BST20120307

Buan, N.R., Metcalf, W.W., 2010. Methanogenesis by *Methanosarcina acetivorans* involves two structurally and functionally distinct classes of heterodisulfide reductase. Mol. Microbiol. 75:843–853. http://dx.doi.org/10.1111/j.1365-2958.2009.06990.xMMI6990

Buan, N., Kulkarni, G., Metcalf, W., 2011. Genetic methods for *Methanosarcina* species. Methods Enzymol. 494:23–42. http://dx.doi.org/10.1016/B978-0-12-385112-3.00002-0

Chou, Y.H., Chiou, H.J., Tiu, C.M., Chen, J.D., Hsu, C.C., Lee, C.H., Lui, W.Y., Hung, G.S., Yu, C., 2001. Sonography of acute right side colonic diverticulitis. Am. J. Surg. 181, 122–127 (S0002-9610(00)00568-7 [pii]).

Conway de Macario, E., Konig, H., Macario, A.J., Kandler, O., 1984. Six antigenic determinants in the surface layer of the archaebacterium *Methanococcus vannielii* revealed by monoclonal antibodies. J. Immunol. 132, 883–887.

Doyle, J.D., Parsons, S.A., 2002. Struvite formation, control and recovery. Water Res. 36, 3925–3940 (S0043-1354(02)00126-4).

Edgar, R.C., 2010. Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26:2460–2461. http://dx.doi.org/10.1093/bioinformatics/btq461

Faivre, D., Baumgartner, J., 2015. The combination of random mutagenesis and sequencing highlight the role of unexpected genes in an intractable organism. PLoS Genet. 11.

Farrell, A., Coleman, B.I., Benenati, B., Brown, K.M., Blader, I.J., Marth, G.T., Gubbels, M.J., 2014. Whole genome profiling of spontaneous and chemically induced mutations in *Toxoplasma gondii*. BMC Genomics 15:354. http://dx.doi.org/10.1186/1471-2164-15-354

Firtel, M., Southam, G., Harauz, G., Beveridge, T.J., 1993. Characterization of the cell wall of the sheathed methanogen *Methanospirillum hungatei* GP1 as an S layer. J. Bacteriol. 175, 7550–7560.

Francoleon, D.R., Boontheung, P., Yang, Y., Kin, U., Ytterberg, A.J., Denny, P.A., Denny, P.C., Loo, J.A., Gunsalus, R.P., Loo, R.R., 2009. S-layer, surface-accessible, and concanavalin A binding proteins of *Methanosarcina acetivorans* and *Methanosarcina mazei*. J. Proteome Res. 8:1972–1982. http://dx.doi.org/10.1021/pr800923e

Galagan, J.E., Nusbaum, C., Roy, A., Endrizzi, M.G., Macdonald, P., FitzHugh, W., Calvo, S., Engels, R., Smirnov, S., Atnoor, D., Brown, A., Allen, N., Naylor, J., Stange-Thomann, N., DeArellano, K., Johnson, R., Linton, L., McEwan, P., McKernan, K., Talamas, J., Tirrell, A., Ye, W., Zimmer, A., Barber, R.D., Cann, I., Graham, D.E., Grahame, D.A., Guss, A.M., Hedderich, R., Ingram-Smith, C.,

Kuettner, H.C., Krzycki, J.A., Leigh, J.A., Li, W., Liu, J., Mukhopadhyay, B., Reeve, J.N., Smith, K., Springer, T.A., Umayam, L.A., White, O., White, R.H., Conway de Macario, E., Ferry, J.G., Jarrell, K.F., Jing, H., Macario, A.J., Paulsen, I., Pritchett, M., Sowers, K.R., Swanson, R.V., Zinder, S.H., Lander, E., Metcalf, W.W., Birren, B., 2002. The genome of *M. acetivorans* reveals extensive metabolic and physiological diversity. Genome Res. 12:532–542. http://dx.doi.org/10.1101/gr.223902

Gamas, P., Chandler, M.G., Prentki, P., Galas, D.J., 1987. *Escherichia coli* integration host factor binds specifically to the ends of the insertion sequence IS1 and to its major insertion hot-spot in pBR322. J. Mol. Biol. 195, 261–272 (0022-2836(87)90648-6 [pii]).

Guss, A.M., Rother, M., Zhang, J.K., Kulkarni, G., Metcalf, W.W., 2008. New methods for tightly regulated gene expression and highly efficient chromosomal integration of cloned genes for *Methanosarcina* species. Archaea 2, 193–203.

Guss, A.M., Kulkarni, G., Metcalf, W.W., 2009. Differences in hydrogenase gene expression between *Methanosarcina acetivorans* and *Methanosarcina barkeri*. J. Bacteriol. 191: 2826–2833. http://dx.doi.org/10.1128/JB.00563-08 (JB.00563-08).

Harper, M.A., Chen, Z., Toy, T., Machado, I.M., Nelson, S.F., Liao, J.C., Lee, C.J., 2011. Phenotype sequencing: identifying the genes that cause a phenotype directly from pooled sequencing of independent mutants. PLoS One 6, e16517. http://dx.doi.org/10.1371/journal.pone.0016517

Harris, L.M., Volpe, C.M., Doerr, R.J., 1997. Small bowel obstruction secondary to enterolith impaction complicating jejunal diverticulitis. Am. J. Gastroenterol. 92, 1538–1540.

Hartmann, E., Konig, H., 1991. Nucleotide-activated oligosaccharides are intermediates of the cell wall polysaccharide of *Methanosarcina barkeri*. Biol. Chem. Hoppe Seyler 372, 971–974.

Hassel, D.M., Schiffman, P.S., Snyder, J.R., 2001. Petrographic and geochemic evaluation of equine enteroliths. Am. J. Vet. Res. 62, 350–358.

Hassel, D.M., Spier, S.J., Aldridge, B.M., Watnick, M., Argenzio, R.A., Snyder, J.R., 2009. Influence of diet and water supply on mineral content and pH within the large intestine of horses with enterolithiasis. Vet. J. 182:44–49. http://dx.doi.org/10.1016/j.tvjl.2008.05 016.

Hildenbrand, C., Stock, T., Lange, C., Rother, M., Soppa, J., 2011. Genome copy numbers and gene conversion in methanogenic archaea. J. Bacteriol. 193:734–743. http://dx.doi.org/10.1128/JB.01016-10

Horne, A.J., Lessner, D.J., 2013. Assessment of the oxidant tolerance of *Methanosarcina acetivorans*. FEMS Microbiol. Lett. 343:13–19. http://dx.doi.org/10.1111/1574-6968.12115

Institute B, (posting date. Picard. [Online.])

Jasso-Chavez, R., Apolinario, E.E., Sowers, K.R., Ferry, J.G., 2013. MrpA functions in energy conversion during acetate-dependent growth of *Methanosarcina acetivorans*. J. Bacteriol. 195:3987–3994. http://dx.doi.org/10.1128/JB.00581-13

Karcher, U., Schroder, H., Haslinger, E., Allmaier, G., Schreiner, R., Wieland, F., Haselbeck, A., Konig, H., 1993. Primary structure of the heterosaccharide of the surface glycoprotein of *Methanothermus fervidus*. J. Biol. Chem. 268, 26821–26826.

Klingl, A., 2014. S-layer and cytoplasmic membrane - exceptions from the typical archaeal cell wall with a focus on double membranes. Front. Microbiol. 5:624. http://dx.doi.org/10.3389/fmicb.2014.00624

Lee, S.Y., Butler, D., Kleckner, N., 1987. Efficient Tn10 transposition into a DNA insertion hot spot in vivo requires the 5-methyl groups of symmetrically disposed thymines within the hot-spot consensus sequence. Proc. Natl. Acad. Sci. U.S.A. 84, 7876–7880.

Li, H., 2011. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. Bioinformatics 27:2987–2993. http://dx.doi.org/10.1093/bioinformatics/btr509

Li, H., Durbin, R., 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25:1754–1760. http://dx.doi.org/10.1093/bioinformatics/btp324

Li, Q., Li, L., Rejtar, T., Karger, B.L., Ferry, J.G., 2005a. Proteome of *Methanosarcina acetivorans* part II: comparison of protein levels in acetate- and methanol-grown cells. J. Proteome Res. 4:129–135. http://dx.doi.org/10.1021/pr049831k

Li, Q., Li, L., Rejtar, T., Karger, B.L., Ferry, J.G., 2005b. Proteome of *Methanosarcina acetivorans* part I: an expanded view of the biology of the cell. J. Proteome Res. 4: 112–128. http://dx.doi.org/10.1021/pr049832c

Li, L., Li, Q., Rohlin, L., Kim, U., Salmon, K., Rejtar, T., Gunsalus, R.P., Karger, B.L., Ferry, J.G., 2007. Quantitative proteomic and microarray analysis of the archaeon *Methanosarcina acetivorans* grown with acetate versus methanol. J. Proteome Res. 6:759–771. http://dx.doi.org/10.1021/pr060383l

Lomhoff, I.I., Dubowy, J., 1948. Gallstone ileus associated with diverticulitis of the colon; a case report. Am. J. Roentgenol. Radium Ther. 60, 86–89.

Manna, D., Wang, X., Higgins, N.P., 2001. Mu and IS1 transpositions exhibit strong orientation bias at the *Escherichia coli* bgl locus. J. Bacteriol. 183:3328–3335. http://dx.doi.org/10.1128/JB.183.11.3328-3335.2001

Markowitz, V.M., Chen, I.M., Palaniappan, K., Chu, K., Szeto, E., Pillay, M., Ratner, A., Huang, J., Woyke, T., Huntemann, M., Anderson, I., Billis, K., Varghese, N., Mavromatis, K., Pati, A., Ivanova, N.N., Kyrpides, N.C., 2014. IMG 4 version of the integrated microbial genomes comparative analysis system. Nucleic Acids Res. 42:D560–D567. http://dx.doi.org/10.1093/nar/gkt963

Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal, 17(1), pp. 10–12. http://dx.doi.org/10.14806/ej.17.1.200

Mengele, R., Sumper, M., 1992. Drastic differences in glycosylation of related S-layer glycoproteins from moderate and extreme halophiles. J. Biol. Chem. 267, 8182–8185.

Metcalf, W.W., Zhang, J.K., Shi, X., Wolfe, R.S., 1996. Molecular, genetic, and biochemical characterization of the *serC* gene of *Methanosarcina barkeri* Fusaro. J. Bacteriol. 178, 5797–5802.

Metcalf, W.W., Zhang, J.K., Wolfe, R.S., 1998. An anaerobic, intrachamber incubator for growth of *Methanosarcina* spp. on methanol-containing solid media. Appl. Environ. Microbiol. 64, 768–770.

Miles, A., Ellis, T.G., 2001. Struvite precipitation potential for nutrient recovery from anaerobically treated wastes. Water Sci. Technol. 43, 259–266.

Mountfort, D.O., Asher, R.A., 1979. Effect of inorganic sulfide on the growth and metabolism of *Methanosarcina barkeri* strain DM. Appl. Environ. Microbiol. 37, 670–675.

Mullany, P., Wilks, M., Tabaqchali, S., 1991. Transfer of Tn916 and Tn916 ΔE into *Clostridium difficile*: demonstration of a hot-spot for these elements in the *C. difficile* genome. FEMS Microbiol. Lett. 63, 191–194.

Namboori, S.C., Graham, D.E., 2008. Acetamido sugar biosynthesis in the Euryarchaea. J. Bacteriol. 190:2987–2996. http://dx.doi.org/10.1128/JB.01970-07

Nguyen, B.D., Valdivia, R.H., 2013. Forward genetic approaches in *Chlamydia trachomatis*. J. Vis. Exp. 80, e50636. http://dx.doi.org/10.3791/50636

Phogat, S.K., Gupta, R., Burma, P.K., Sen, K., Pental, D., 2001. On the estimation of number of events required for saturation mutagenesis of large genomes. Curr. Sci. India 80, 823–824.

Pritchett, M.A., Metcalf, W.W., 2005. Genetic, physiological and biochemical characterization of multiple methanol methyltransferase isozymes in *Methanosarcina acetivorans* C2A. Mol. Microbiol. 56:1183–1194, MMI4616. http://dx.doi.org/10.1111/j.1365-2958.2005.04616.x

Quinlan, A.R., Hall, I.M., 2010. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26:841–842. http://dx.doi.org/10.1093/bioinformatics/btq033

Rodman, J.S., 1999. Struvite stones. Nephron 81 (Suppl. 1), 50–59, nef1a050

Rohlin, L., Leon, D.R., Kim, U., Loo, J.A., Ogorzalek Loo, R.R., Gunsalus, R.P., 2012. Identification of the major expressed S-layer and cell surface-layer-related proteins in the model methanogenic archaea: *Methanosarcina barkeri* Fusaro and *Methanosarcina acetivorans* C2A. Archaea 2012:873589. http://dx.doi.org/10.1155/2012/873589

Rother, M., Sattler, C., Stock, T., 2011. Studying gene regulation in methanogenic archaea. Methods Enzymol. 494:91–110. http://dx.doi.org/10.1016/B978-0-12-385112-3 00005-6.

Sarmiento, F., Ellison, C.K., Whitman, W.B., 2013. Genetic confirmation of the role of sulfopyruvate decarboxylase in coenzyme M biosynthesis in *Methanococcus maripaludis*. Archaea 2013, 185250. http://dx.doi.org/10.1155/2013/185250

Sattler, C., Wolf, S., Fersch, J., Goetz, S., Rother, M., 2013. Random mutagenesis identifies factors involved in formate-dependent growth of the methanogenic archaeon *Methanococcus maripaludis*. Mol. Gen. Genomics. 288:413–424. http://dx.doi.org/10.1007/s00438-013-0756-6

Schierenbeck, L., Ries, D., Rogge, K., Grewe, S., Weisshaar, B., Kruse, O., 2015. Fast forward genetics to identify mutations causing a high light tolerant phenotype in *Chlamydomonas reinhardtii* by whole-genome-sequencing. BMC Genomics 16:57. http://dx.doi.org/10.1186/s12864-015-1232-y

Scott, W.D., Wrigley, T.J., Webb, K.M., 1991. A computer model of struvite solution chemistry. Talanta 38, 889–895. https://doi.org/10.1016/0039-9140(91)80268-5

Soppa, J., 2011. Ploidy and gene conversion in archaea. Biochem. Soc. Trans. 39:150–154. http://dx.doi.org/10.1042/BST0390150

Sowers, K.R., Gunsalus, R.P., 1988. Adaptation for growth at various saline concentrations by the archaebacterium *Methanosarcina thermophila*. J. Bacteriol. 170, 998–1002.

Sowers, K.R., Baron, S.F., Ferry, J.G., 1984. *Methanosarcina acetivorans* sp. nov., an acetotrophic methane-producing bacterium isolated from marine sediments. Appl. Environ. Microbiol. 47, 971–978.

Steber, J., Schleifer, K.H., 1975. *Halococcus morrhuae*: a sulfated heteropolysaccharide as the structural component of the bacterial cell wall. Arch. Microbiol. 105, 173–177.

Wang, H., Smith, M.C., Mullany, P., 2006. The conjugative transposon Tn5397 has a strong preference for integration into its *Clostridium difficile* target site. J. Bacteriol. 188: 4871–4878. http://dx.doi.org/10.1128/JB.00210-06

Witkin, E.M., 1976. Ultraviolet mutagenesis and inducible DNA repair in *Escherichia coli*. Bacteriol. Rev. 40, 869–907.

Zellner, G., Stackebrandt, E., Messner, P., Tindall, B.J., Conway deMacario, E., Kneifel, H., Sleytr, U.B., Winter, J., 1989. *Methanocorpusculaceae* fam. nov., represented by *Methanocorpusculum parvum*, *Methanocorpusculum sinense* spec. nov. and *Methanocorpusculum bavaricum* spec. nov. Arch. Microbiol. 151, 381–390.

Zhang, J.K., Pritchett, M.A., Lampe, D.J., Robertson, H.M., Metcalf, W.W., 2000. In vivo transposon mutagenesis of the methanogenic archaeon *Methanosarcina acetivorans* C2A using a modified version of the insect mariner-family transposable element Himar1. Proc. Natl. Acad. Sci. U. S. A. 97:9665–9670. http://dx.doi.org/10.1073/pnas.160272597

Supplementary data follows: Tables S1–S3 & Figs. S1–S8

Supplementary Information

High-throughput mutation, selection, and phenotype screening of mutant methanogenic archaea

Mary E. Walter, Alicia Ortiz, Casey Sondgeroth, Nathan M. Sindt, Nikolas Duszenko, Jennie L. Catlett, You Zhou, Shah Valloppilly, Christopher Anderson, Samodha Fernando, Nicole R. Buan*

Table S1. Sequencing coverage

Sample	PE reads	Unmapped reads	% Reads aligned	Mean base coverage	Bases with 0 coverage	% Bases covered	Bases >5X coverage	Bases >10X coverage	Bases >20X coverage
Parent	1330642	198	99.98512	51.0114	10825	99.81179	5647527	5498355	5097164
89	3655416	1910	99.94775	141.806	646	99.98877	5749342	5746311	5732561
91	3307302	736	99.97775	125.696	32385	99.43693	5565847	5425754	5172094
93	2401006	633	99.97364	86.2128	53049	99.07765	5488416	5238767	4776651
178	1922984	214	99.98887	72.8457	1526	99.97347	5734794	5704338	5578493
181	1089214	261	99.97604	42.0348	20211	99.6486	5633275	5480214	4957415
193	442530	62	99.98599	16.7336	42397	99.26285	5225657	4282100	156684
194	806592	152	99.98116	31.3276	65295	98.86473	5341530	4964352	4057687
195	5831928	770	99.9868	222.853	777	99.98649	5746872	5738243	5707747

Table S2. Recovered Mutations (Q>20)

	All mutations			Unique Mutations		
Strain NB#	Total	Indel	SNP	Total	Indel	SNP
Parent	27	12	15	1	1	0
89	28	13	15	5	2	3
91	37	13	24	8	1	7
93	47	12	35	26	0	26
178	31	12	19	4	1	3
181	30	11	18	5	0	5
193	44	14	30	20	3	17
194	24	10	14	5	0	5
195	40	21	19	17	10	7
median	31	12	19	5	1	5
mean	34.2	13.1	21	10.1	2	8.1
std dev	7.63	3.00	6.86	8.17	2.98	7.74
%		38	61		20	80

Table S3. Frequency of Recovered Mutations

	All Mutations	Unique Mutations
Total number of mutations	308	91
Strains	9	9
Average number mutations per genome	34.2	10.1
Average frequency mutations per genome	1.68E+5	5.69E+5
Average mutation frequency per site	5.95E-6	1.76E-6

Table S4. Unique mutations in strain genomes.				
Genome position	Locus (MA_RS#)	Mutation	Effect	Predicted function
NB91				
1027125	4495	G to A		pseudogene ATPase
1481804	6455	T to A	V1305D	SIR2, P-loop NTPase
2150223	intergenic	A to G		(cell surface protein)
2322462	9830	T to C	F63I	hypothetical (permease)
2500114		T to A		repeat region
2672579	11240	T to C	I111V	SAM methyltransferase
4259860	intergenic	del ATGGAAAGAAGGA		(iron transporter)
4560343	intergenic	A to T		(cell surface protein)
1027125	4495	G to A		pseudogene ATPase
1481804	6455	T to A	V1305D	SIR2, P-loop NTPase
NB93				
453479	2015	G to A	silent	cobN cobaltochelatase
753265	intergenic	G to A		(hypothetical
1058020	intergenic	C to A		(ABC transporter, S-layer)
1243358	intergenic	T to C		(pseudo, hypothetical)
1243363	intergenic	A to C		(pseudo, hypothetical)
1243366	intergenic	T to C		(pseudo, hypothetical)
1438389	intergenic	A to G		(pyridoxamine 5-P oxidase)
1718693	7525	ins A	FS	hypothetical membrane protein
2739238	11555	T to A	R20W	hypothetical protein
2739452	11560	T to G	silent	hypothetical, repeat region
2739471	11560	G to A	P65L	hypothetical, repeat region
3330450	13985	A to G	Y38H	hypothetical
3495123	14665	A to G	silent	hypothetical
3774547	intergenic	G to A		(hypothetical)
4281256	intergenic	A to T		(hypothetical, pseudo)
4656219	intergenic	T to C		(glycosyl transferase)
5054624	intergenic	G to A		repeat region

5292664	22490	A to T	Q7L	EF hand family signaling or buffering/transport protein
5292681	22490	A to G	T13A	EF hand family signaling or buffering/transport protein
5369692	intergenic	G to A		(hypothetical)
5523929	23355	G to A	silent	hypothetical
5601773	intergenic	C to A		(mcr, SAM protein)
5601776	intergenic	G to A		(mcr, SAM protein)
5601778	intergenic	G to T		(mcr, SAM protein)
5601784	intergenic	C to T		(mcr, SAM protein)
5601788	intergenic	G to T		(mcr, SAM protein)
NB178				
2541340	intergenic	T to A		(hypothetical)
3345530	intergenic	ins		(hypothetical repeat region)
4741505	intergenic	G to T		(ACDS)
5394502	22880	C to T	S140I	putative ArsR HTH txn regulator
NB181				
693453	3085	C to T	A51T	cell surface protein
1234521	intergenic	T to C		
2453865	intergenic	T to C		(hypothetical)
2746344	intergenic	T to C		
5052335	intergenic	A to T		(repeat region)
NB193				
369001	intergenic	A to T		(hypothetical)
743829	3320	A to G	S501G	groEL
1061737	4610	G to A	D360N	PolB
1283824	intergenic	A to T		(hypothetical VanZ-like protein)
1625486	24425	G to A	G39S	acidic membrane protein
1711768	intergenic	ins TTCAG		(16S rRNA SAM methyltransferase)
1757913	intergenic	G to A		(ferredoxin)
1772584	7805	del ACAATATCCTGCAGCTCTTTGTTTC	FS	FpoL dehydrogenase
1897311	8295	G to A	silent	danorubicin DrrA family ATP transporter
2013262	intergenic	del CTTATTTT		(cation ATPase)

2073909	intergenic	T to C		(indolepyruvate ferredoxin oxidoreductase, cell surface protein)
2077910	8990	T to C	silent	hypothetical
2391801	24520	A to G	K684R	repeat protein/amidohydrolase
2513871	10525	T to C	silent	hypothetical
2759624	intergenic	A to G		(tetratricopeptide repeat protein)
2791633	intergenic	A to G		(endonuclease III)
3358138	intergenic	A to G		(amidotransferase)
4491572	intergenic	A to G		pseudo
4895405	intergenic	T to C		(aldehyde ferredoxin oxidoreductase, MutT)
5411204	intergenic	A to G		(ArsR HTH txn regulator, MtsF methylsulfide methyltransferase)
NB194				
871462	3925	T to C		pseudo
1020294	4475	C to T	del205-634	MtbC1 dimethylamine corrinoid protein
2048832	8870	A to G	D34G	peptidase U32
3953236	intergenic	T to C		(hypothetical)
5726042	intergenic	A to G		
NB195				
487686	intergenic	ins T		(pseudo)
1718473	7525	ins G	FS	membrane
2736884	11545	ins GGACCAAGACTTTCTTCCAGGACCAGGACTTCCTTCCAAGGCAACCTTTCCCAGGGAGAAATTCTTCCTTGCAGCTGCAGGCCA	FS	hypothetical
2743499	11575	C to T	A4V	hypothetical, repeat region
2743532	11575	ins GATGAGACCAGTAAAAGATATGCAGAAAATGGCAGAAAT	FS	hypothetical, repeat region
2743538	11575	ins GATGAAACCAGAAAAAGATATGCAGAAAATGGCAGAAAT	FS	hypothetical, repeat region

2743549	11575	ins GATGAAACCAGTAAAAGATATGAAGAAAATGGC AGAAAT	FS	hypothetical, repeat region
2743558	11575	ins GATGAAACCAGTAAAAGATATGCAGAAAATGTC AGAAAT	FS	hypothetical, repeat region
2743791	11575	A to G	I101M	hypothetical, repeat region
2743800	11575	A to G	I104M	hypothetical, repeat region
2743804	11575	G to A	A106T	hypothetical, repeat region
2743805	11575	C to A	A106E	hypothetical, repeat region
2743844	11575	C to A	T119K	hypothetical, repeat region
2743869	11575	A to G	I130M	hypothetical, repeat region
3958294	intergenic	del TGAGTAATTGAGTAATAGTTTGAGTAATAGCTAT		(tryptophan synthase B)
4562039	19270	del GCCGCCTGTCGCGATGACCGTAGAAAACTGCG AATCTGATCCGTTCATTCCTATCGAGCCGCCGG CGTCTCCGCCGGGTCCGCCGGCGCC	FS	cell surface protein
5643060	intergenic	del ATTATGGAATTGACACATTATGGAATTGACAC		(Zn ribbon)

Figure S1. Identification of mutations in NB89. Numbers indicate mutation site vs reference *M. acetivorans* genome. Circles represent unique mutations found on the chromosome with quality (Q) scores >20. The shade indicates Q <50 (lightest) to 228 (darkest). Cyan, parent strain; magenta, strain NB89; gray, mutations conserved between the parent strain and NB89; lavender, mutations shared by NB89 and another strain.

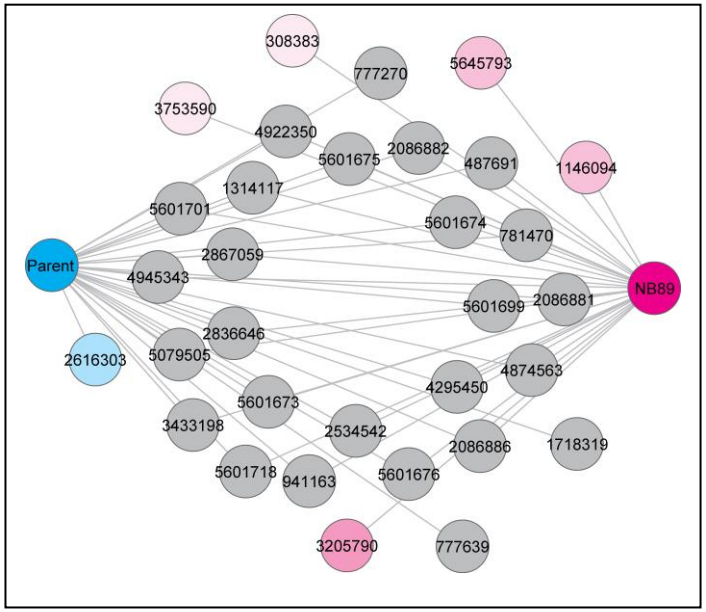

Figure S2. Identification of mutations in NB91. Numbers indicate mutation site vs reference *M. acetivorans* genome. Circles represent unique mutations found on the chromosome with quality (Q) scores >20. The shade indicates Q <50 (lightest) to 228 (darkest). Cyan, parent strain; magenta, strain NB91; gray, mutations conserved between the parent strain and NB91; lavender, mutations shared by NB91 and another strain.

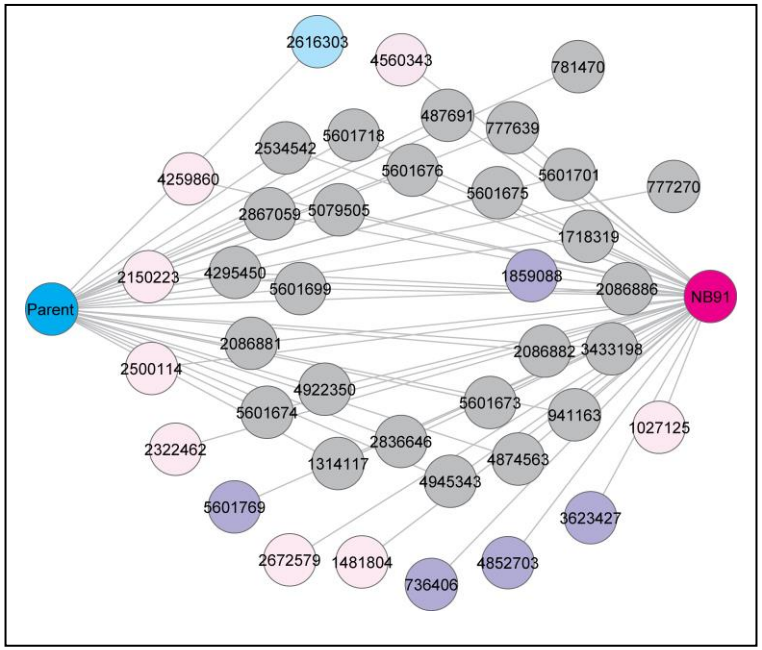

Figure S3. Identification of mutations in NB93. Numbers indicate mutation site vs reference *M. acetivorans* genome. Circles represent unique mutations found on the chromosome with quality (Q) scores >20. The shade indicates Q <50 (lightest) to 228 (darkest). Cyan, parent strain; magenta, strain NB93; gray, mutations conserved between the parent strain and NB93; lavender, mutations shared by NB93 and another strain.

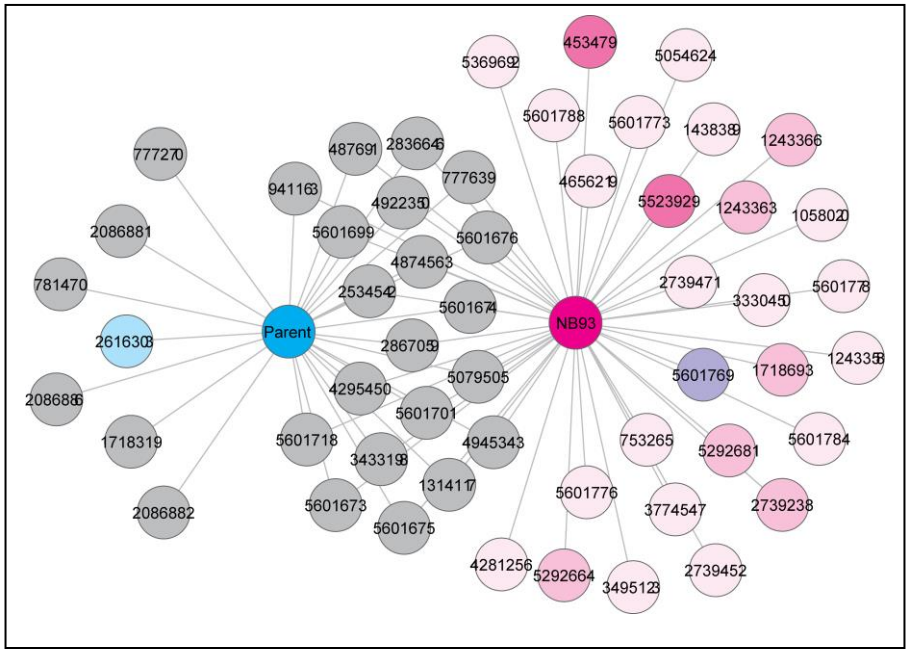

Figure S4. Identification of mutations in NB178. Numbers indicate mutation site vs reference *M. acetivorans* genome. Circles represent unique mutations found on the chromosome with quality (Q) scores >20. The shade indicates Q <50 (lightest) to 228 (darkest). Cyan, parent strain; magenta, strain NB178; gray, mutations conserved between the parent strain and NB178; lavender, mutations shared by NB178 and another strain.

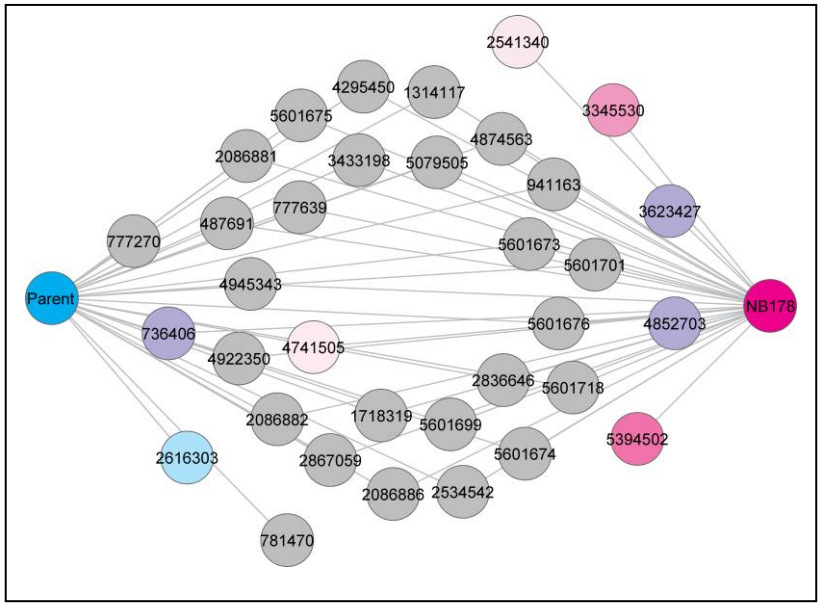

Figure S5. Identification of mutations in NB181. Numbers indicate mutation site vs reference *M. acetivorans* genome. Circles represent unique mutations found on the chromosome with quality (Q) scores >20. The shade indicates Q <50 (lightest) to 228 (darkest). Cyan, parent strain; magenta, strain NB181; gray, mutations conserved between the parent strain and NB181; lavender, mutations shared by NB181 and another strain.

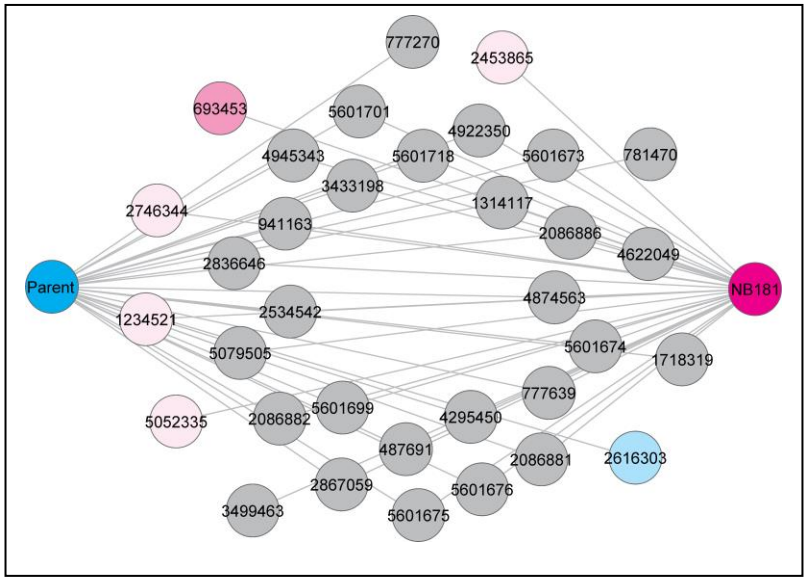

Figure S6. Identification of mutations in NB193. Numbers indicate mutation site vs reference *M. acetivorans* genome. Circles represent unique mutations found on the chromosome with quality (Q) scores >20. The shade indicates Q <50 (lightest) to 228 (darkest). Cyan, parent strain; magenta, strain NB193; gray, mutations conserved between the parent strain and NB193; lavender, mutations shared by NB193 and another strain.

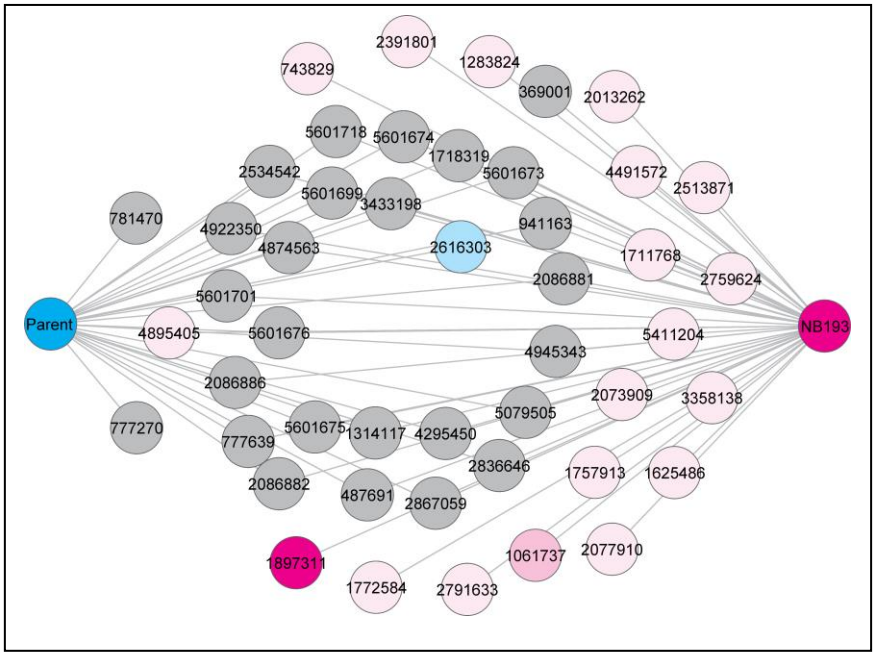

Figure S7. Identification of mutations in NB194. Numbers indicate mutation site vs reference *M. acetivorans* genome. Circles represent unique mutations found on the chromosome with quality (Q) scores >20. The shade indicates Q <50 (lightest) to 228 (darkest). Cyan, parent strain; magenta, strain NB194; gray, mutations conserved between the parent strain and NB194; lavender, mutations shared by NB194 and another strain.

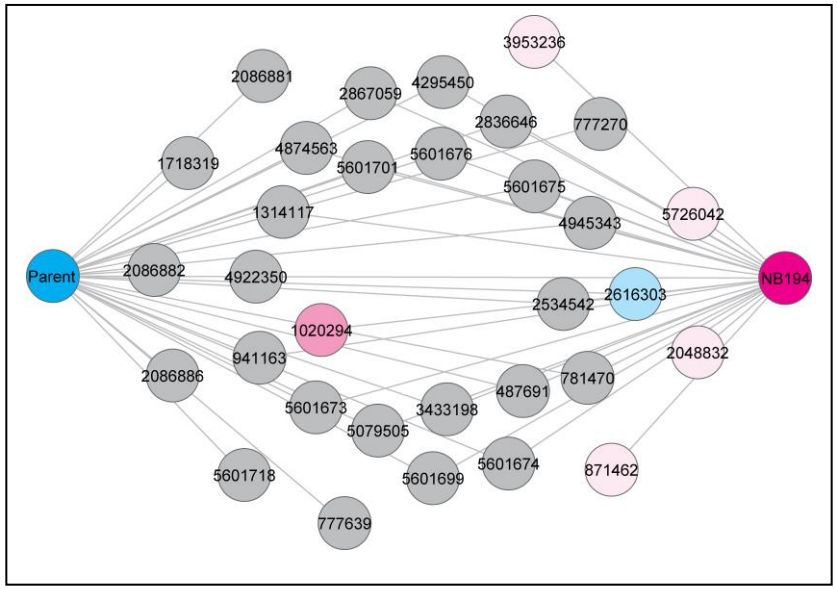

Figure S8. Identification of mutations in NB195. Numbers indicate mutation site vs reference *M. acetivorans* genome. Circles represent unique mutations found on the chromosome with quality (Q) scores >20. The shade indicates Q <50 (lightest) to 228 (darkest). Cyan, parent strain; magenta, strain NB195; gray, mutations conserved between the parent strain and NB195; lavender, mutations shared by NB195 and another strain.