

2007

## The complete genome sequence and analysis of the Epsilonproteobacterium *Arcobacter butzleri*

William G. Miller

Craig T. Parker

Mark Rubenfield

George L. Mendz

*University of Notre Dame Australia*, gmendz@nd.edu.au

Marc MSM Wo

*See next page for additional authors*

Follow this and additional works at: [https://researchonline.nd.edu.au/med\\_article](https://researchonline.nd.edu.au/med_article)



Part of the [Medicine and Health Sciences Commons](#)

This article was originally published as:

Miller, W. G., Parker, C. T., Rubenfield, M., Mendz, G. L., Wo, M. M., Ussery, D. W., Stolz, J. F., Binneweis, T. T., Hallin, P. F., Wang, G., Malek, J. A., Rogosin, A., Stanker, L. H., & Mandrell, R. E. (2007). The complete genome sequence and analysis of the Epsilonproteobacterium *Arcobacter butzleri*. *PLoS ONE*, 2 (12), e1358.

<http://doi.org/10.1371/journal.pone.0001358>

This article is posted on ResearchOnline@ND at  
[https://researchonline.nd.edu.au/med\\_article/1](https://researchonline.nd.edu.au/med_article/1). For more  
information, please contact [researchonline@nd.edu.au](mailto:researchonline@nd.edu.au).



---

## Authors

William G. Miller, Craig T. Parker, Mark Rubenfield, George L. Mendz, Marc MSM Wo, David W. Ussery, John F. Stolz, Tim T. Binneweis, Peter F. Hallin, Guilin Wang, Joel A. Malek, Andrea Rogosin, Larry H. Stanker, and Robert E. Mandrell

# The Complete Genome Sequence and Analysis of the Epsilonproteobacterium *Arcobacter butzleri*

William G. Miller<sup>1\*</sup>, Craig T. Parker<sup>1</sup>, Marc Rubenfield<sup>2</sup>, George L. Mendz<sup>3</sup>, Marc M. S. M. Wösten<sup>4</sup>, David W. Ussery<sup>5</sup>, John F. Stolz<sup>6</sup>, Tim T. Binnewies<sup>5</sup>, Peter F. Hallin<sup>5</sup>, Guilin Wang<sup>1</sup>, Joel A. Malek<sup>2</sup>, Andrea Rogosin<sup>2</sup>, Larry H. Stanker<sup>7</sup>, Robert E. Mandrell<sup>1</sup>

**1** Produce Safety and Microbiology Research Unit, Agricultural Research Service, U.S. Department of Agriculture, Albany, California, United States of America, **2** Agencourt Bioscience Corporation, Beverly, Massachusetts, United States of America, **3** School of Medicine, Sydney, The University of Notre Dame Australia, Broadway, New South Wales, Australia, **4** Department of Infectious Diseases and Immunology, Utrecht University, Utrecht, The Netherlands, **5** Center for Biological Sequence Analysis, Technical University of Denmark, Lyngby, Denmark, **6** Department of Biological Sciences, Duquesne University, Pittsburgh, Pennsylvania, United States of America, **7** Foodborne Contaminants Research Unit, Agricultural Research Service, U.S. Department of Agriculture, Albany, California, United States of America

**Background.** *Arcobacter butzleri* is a member of the epsilon subdivision of the Proteobacteria and a close taxonomic relative of established pathogens, such as *Campylobacter jejuni* and *Helicobacter pylori*. Here we present the complete genome sequence of the human clinical isolate, *A. butzleri* strain RM4018. **Methodology/Principal Findings.** *Arcobacter butzleri* is a member of the *Campylobacteraceae*, but the majority of its proteome is most similar to those of *Sulfuromonas denitrificans* and *Wolinella succinogenes*, both members of the *Helicobacteraceae*, and those of the deep-sea vent Epsilonproteobacteria *Sulfurovum* and *Nitratiruptor*. In addition, many of the genes and pathways described here, e.g. those involved in signal transduction and sulfur metabolism, have been identified previously within the epsilon subdivision only in *S. denitrificans*, *W. succinogenes*, *Sulfurovum*, and/or *Nitratiruptor*, or are unique to the subdivision. In addition, the analyses indicated also that a substantial proportion of the *A. butzleri* genome is devoted to growth and survival under diverse environmental conditions, with a large number of respiration-associated proteins, signal transduction and chemotaxis proteins and proteins involved in DNA repair and adaptation. To investigate the genomic diversity of *A. butzleri* strains, we constructed an *A. butzleri* DNA microarray comprising 2238 genes from strain RM4018. Comparative genomic indexing analysis of 12 additional *A. butzleri* strains identified both the core genes of *A. butzleri* and intraspecies hypervariable regions, where <70% of the genes were present in at least two strains. **Conclusion/Significance.** The presence of pathways and loci associated often with non-host-associated organisms, as well as genes associated with virulence, suggests that *A. butzleri* is a free-living, water-borne organism that might be classified rightfully as an emerging pathogen. The genome sequence and analyses presented in this study are an important first step in understanding the physiology and genetics of this organism, which constitutes a bridge between the environment and mammalian hosts.

Citation: Miller WG, Parker CT, Rubenfield M, Mendz GL, Wösten MMSM, et al (2007) The Complete Genome Sequence and Analysis of the Epsilonproteobacterium *Arcobacter butzleri*. PLoS ONE 2(12): e1358. doi:10.1371/journal.pone.0001358

## INTRODUCTION

The epsilon subdivision of the Gram-negative Proteobacteria comprises multiple genera contained within three major families: *Campylobacteraceae*, *Helicobacteraceae* and *Nautiliaceae*. The majority of well-characterized species in this subdivision are members of genera within the first two families, including *Campylobacter*, *Arcobacter* and *Sulfurospirillum* in the *Campylobacteraceae*, and *Helicobacter* and *Wolinella* in the *Helicobacteraceae*. Many of these species are pathogenic, e.g. *Campylobacter jejuni* [1] and *Helicobacter pylori* [2], and/or are associated with a particular host or hosts, e.g. *Campylobacter upsaliensis* [3] and *Helicobacter mustelae* [4]; however, several species are free-living, e.g. *Sulfurospirillum* spp. [5], and they are not considered to be pathogenic.

The genus *Arcobacter* is an unusual taxon within the epsilon subdivision in that it contains both pathogenic and free-living species found in a wide range of environments. Currently, *Arcobacter* contains four recognized species: *A. butzleri* [6], *A. cryaerophilus* [7], *A. skirrowii* [8] and *A. nitrofigilis* [9]. *A. butzleri*, *A. cryaerophilus* and *A. skirrowii* have been isolated from animals and humans [10], while *A. nitrofigilis* is a nitrogen-fixing bacterium isolated originally from *Spartina atrovirens* roots in an estuarine marsh [9]. In addition to these established *Arcobacter* species, three new species have been described recently: 1) the obligate halophile *A. halophilus* sp. nov., isolated from a Hawaiian hypersaline lagoon [11], 2) *A. cibarius* sp. nov., isolated from broiler carcasses [12] and 3) *Candidatus A. sulfidicus*, a sulfide-oxidizing marine organism that

produces filamentous sulfur [13]. In addition, several potential, novel *Arcobacter* species, based so far on only 16S rDNA sequence data, have been identified in: the flora of deep-sea hydrothermal vents [14], hydrocarbon-contaminated seawater [15], a low-salinity petroleum reservoir [16], infected or dead coral surfaces [17], deep-sea sediments [18], tube worms [19], anaerobic sludge [20], and a circulated dairy wastewater lagoon [21]. These studies demonstrate clearly that the genus is associated strongly with freshwater and marine environments. In fact, although *A. butzleri*, *A. cryaerophilus* and *A. skirrowii* have been isolated often from animals

.....  
**Academic Editor:** Cecile Fairhead, Pasteur Institute, France

**Received** October 26, 2007; **Accepted** November 19, 2007; **Published** December 26, 2007

This is an open-access article distributed under the terms of the Creative Commons Public Domain declaration which stipulates that, once placed in the public domain, this work may be freely reproduced, distributed, transmitted, modified, built upon, or otherwise used by anyone for any lawful purpose.

**Funding:** This work was funded by the United States Department of Agriculture, Agricultural Research Service, CRIS projects 5325-42000-230-043 and 5325-42000-230-045.

**Competing Interests:** The authors have declared that no competing interests exist.

\* **To whom correspondence should be addressed.** E-mail: bmillier@pw.usda.gov

or food sources, they have been isolated frequently also from water or water systems [22–30].

*Arcobacter butzleri* is the best characterized of all Arcobacters. *A. butzleri* cells are small, spiral and motile [10], similar morphologically to the taxonomically-related *Campylobacter*. Nonetheless, notable differences exist between *A. butzleri* and *Campylobacter* spp. Classified initially as an “aerotolerant *Campylobacter*”, along with *A. cryaerophilus* [10], *A. butzleri* is able to grow aerobically, at variance with most *Campylobacters* which are microaerophilic. However, *A. butzleri* grows also under microaerobic and anaerobic conditions [10]; thus, this bacterium can grow at all oxygen concentrations. Additionally, *Campylobacter* spp. grow generally between 37°C and 42°C [31], whereas *A. butzleri* is more psychrophilic with a temperature range between 15°C and 37°C, although some strains can grow at 42°C [10]. Furthermore, *A. butzleri* is more halotolerant than most *Campylobacter* spp., with some strains able to grow at 3.5% NaCl [10].

While *A. butzleri* is isolated often from aqueous environments, it is isolated also from multiple animals and food sources. It has been found in pigs [32] and ground pork [33,34], chicken carcasses [35,36] and other poultry [37], as well as in beef [38,39], lamb [34] and the feces of other animals [40]. *A. butzleri* has also been isolated increasingly from human diarrheal stool samples [41–47]. The clinical symptomatology described for *A. butzleri* typically includes diarrhea and recurrent abdominal cramps [10], although *A. butzleri*-related bacteremia has also been reported [48,49]. Prouzet-Mauléon et al. [46] reported an isolation frequency of 1% from human clinical stool samples. Additionally, Vandenberg et al. [47] reported an isolation frequency of 3.5% from diarrheic stool samples. Although co-infection with other enteric pathogens was reported by Prouzet-Mauléon and Vandenberg, in the majority of samples (14/15 clinical stool samples and 55/67 patients, respectively) no other enteric pathogen was detected. Houf et al. reported that Arcobacters were isolated from 1.4% (7/500) of asymptomatic human stool samples [50]. However, all seven isolates were typed as *A. cryaerophilus*; *A. butzleri* was not isolated. Similarly, Vandenberg et al. [47] reported also that *A. butzleri* was isolated more frequently from diarrheic stool samples than from non-diarrheic stool samples. Thus, the isolation of *A. butzleri* from diarrheic stool samples is likely to be relevant clinically and is probably not due to the organism being merely a human commensal. Therefore, these data suggest strongly that *A. butzleri* is an emerging pathogen [10], where transmission, as with *C. jejuni*, occurs probably through consumption of contaminated food or water. The low level of incidence reported in human clinical samples is most likely an underestimate, due to sub-optimal isolation and/or detection methods [46].

Relatively little is known about *A. butzleri*, compared to other members of the epsilon subdivision, but the wealth of genomic information from other epsilonproteobacterial taxa provides a solid foundation to compare and contrast *A. butzleri* to its taxonomic relatives. The genomes of multiple species of Epsilonproteobacteria have been sequenced; these include: four strains of *C. jejuni* subsp. *jejuni* ([51–54]: strains NCTC 11168, RM1221, 81-176 and 81116, respectively); *C. jejuni* subsp. *doylei* strain 269.97 (CP000768.1); *Campylobacter coli* strain RM2228, *Campylobacter lari* strain RM2100 and *C. upsaliensis* strain RM3195 [52]; *Campylobacter fetus* subsp. *fetus* strain 82-40 (CP000487.1); *Campylobacter curvus* strain 525.92 (CP000767.1); *Campylobacter concisus* strain 13826 (CP000792.1); *Campylobacter hominis* strain ATCC BAA-381 (CP000776.1); *Sulfuromonas denitrificans* strain ATCC 33889 (formerly *Thiomicrospira denitrificans* [55]: CP000153); *Wolinella succinogenes* strain DSM 1740 [56]; *Helicobacter hepaticus* strain ATCC 51449 [57]; three strains of *H. pylori* ([58-60]: 26695,

J99 and HPAG1, respectively); *Helicobacter acinonychis* strain Sheeba [61]; and the deep-sea vent taxa *Nitratiruptor* sp. and *Sulfurovum* sp. (strains SB155-2 and NBC37-1, respectively [62]). This study presents the genomic sequence of a human clinical isolate, *A. butzleri* strain RM4018, a derivative of the type strain ATCC 49616. The genomic data revealed multiple differences between *A. butzleri* and other members of the *Campylobacteraceae*, as well as pathways and systems vital for its survival in diverse environments.

## RESULTS AND DISCUSSION

### General features

The genome of *Arcobacter butzleri* strain RM4018 contains 2,341,251 bp; as such it is the second largest characterized epsilonproteobacterial genome to date, smaller than the genome of *Sulfurovum* strain NBC37-1 (2,562,277 bp) but larger than both the genomes of *S. denitrificans* strain ATCC 33889 (2,201,561 bp) and *W. succinogenes* strain DSM 1740 (2,110,355 bp). The G+C content of the RM4018 genome (27%) is remarkably low. A summary of the features of the strain RM4018 genome is provided in Table 1. A diagrammatic representation of the RM4018 genome is presented in Figure S1.

Consistent with its size, the RM4018 genome is predicted to encode 2259 coding sequences (CDSs). Based on pairwise BLASTP comparisons of proteins predicted to be encoded by these CDSs against proteins in the NCBI non-redundant (nr) database (release 10/13/2007), and on the presence of various Pfam and PROSITE motifs, 1011 (45%) of the predicted proteins were assigned a specific function, 505 (22%) were attributed only a general function, and 743 (33%) were considered proteins of unknown function (Table 1). A complete list of the CDSs predicted to be present within the genome of strain RM4018 and their annotation is presented in the supplementary Table S1. A breakdown of the CDSs by function is presented in supplementary Table S2.

### Relationship of *A. butzleri* to other taxa

*Arcobacter butzleri* is a member of the family *Campylobacteraceae* which includes also the genera *Campylobacter* and *Sulfurospirillum*. Given the close taxonomic relationship between *Arcobacter* and *Campylobacter*, it is noteworthy that 17.2% and 12.4% of the RM4018 proteins have their best match in proteins encoded by *S. denitrificans* and *W. succinogenes*, respectively, both members of the *Helicobacteraceae* (Table 2). Moreover, approximately 25% of the RM4018 proteins have their best match in proteins encoded by *Sulfurovum* or *Nitratiruptor*, deep-sea vent Epsilonproteobacteria isolated from a sulfide mound off the coast of Japan [62]. The percentage of *S. denitrificans* homologs with best matches is greater than that of the eight sequenced *Campylobacteraceae* species combined (13.1%). These differences are reduced somewhat by comparing the top five matches and not just the best match; however, even using these parameters, *S. denitrificans* (9.6%) remains the closest related organism (Table 2). Among the *Campylobacters*, *C. fetus* (3.1% best matches) is the closest to *A. butzleri*, although it is possible that other *Campylobacters*, whose genomes are as yet unsequenced, may have higher degrees of similarity to *A. butzleri* than *C. fetus*.

In the entire *A. butzleri* strain RM4018 proteome (unique and non-unique), 61.5% (1390/2259) of the proteins have their best matches in proteins encoded by the epsilonproteobacterial taxa, and 79.0% (1785/2259) are similar to Proteobacterial proteins. Most of the matches to non-Epsilonproteobacteria are found within the gamma subdivision (Table 2), in genera such as *Marinobacter* (*Alteromonadales*), *Oceanospirillum* (*Oceanospirillales*) and *Pseudomonas* (*Pseudomonadales*). Other phyla with a moderate number

**Table 1.** Features of the *Arcobacter butzleri* RM4018 genome

General features	Number or % of total
Chromosome size (bp)	2,341,251
G+C content	27.05%
CDS numbers <sup>a</sup>	2259
Assigned function	1011 (45%)
Pseudogenes	5
General function	505 (22%)
Conserved hypothetical/hypothetical	743 (33%)
Prophage	1
Genetic islands	3
Ribosomal RNA operons	5
Plasmids	0
IS elements	0
Poly GC tracts	0
<b>Gene classes</b>	
Chemotaxis proteins	46
Che/Mot proteins	11
Methyl-accepting chemotaxis proteins	29
Redox-sensing PAS domain proteins	3
Cyclic diguanylic acid proteins	25
Restriction/modification systems	
Type I	0 <sup>b</sup>
Type II/III	0
Type III	0
Transcriptional regulators	
Regulatory proteins	36
Non-ECF family $\sigma$ factors	1
ECF family $\sigma$ /anti- $\sigma$ factor pairs	7
Two-component systems	
Response regulator	42
Sensor histidine kinase	37
<b>Taxon-specific genes</b>	
<i>A. butzleri</i> proteins found:	
Within $\epsilon$ -Proteobacterial taxa	1754 (77.6%)
<i>Campylobacter</i> only	98 (4.3%)
<i>Helicobacter</i> only	10 (0.4%)
<i>S. denitrificans</i> or <i>W. succinogenes</i> only	133 (5.9%)
<i>Sulfurovum</i> or <i>Nitratiruptor</i> only	56 (24.8%)
Only in non $\epsilon$ -Proteobacterial taxa	190 (8.4%)
Unique <i>A. butzleri</i> proteins	315 (13.9%)

<sup>a</sup>Total does not include pseudogenes.

<sup>b</sup>One *hsdM* pseudogene and an *hds/hsdM* pair are present but no complete Type I R/M system (including *hdsR*) is present.

doi:10.1371/journal.pone.0001358.t001

of protein matches include Firmicutes (2.5%), especially *Clostridiales* (1.5%), and Bacteroidetes/Chlorobi (2.8%). Among the 2259 predicted CDSs of strain RM4018, 315 (13.9%) had no homolog within the nr database using a minimum expect (*E*) value of  $1 \times 10^{-5}$ , a minimum identity of 25% and a minimum alignment length of 75% (Figure 1, Table 2, Table S3).

Although approximately 30% of the *A. butzleri* RM4018 proteins with homologs within the nr database have their best matches to non-epsilonproteobacterial proteins, the approximately 550 genes

**Table 2.** Similarity of predicted *A. butzleri* proteins to proteins from other taxa.

Taxon	Best match <sup>a</sup>		Within best 5 matches <sup>b</sup>
	#	%	%
$\epsilon$ -Proteobacteria	1390	71.50	63.43
<i>Campylobacteraceae</i>	258	13.27	22.54
<i>Campylobacter coli</i>	19	0.98	1.48
<i>Campylobacter concisus</i>	47	2.42	3.88
<i>Campylobacter curvus</i>	50	2.57	4.73
<i>Campylobacter fetus</i>	60	3.09	4.24
<i>Campylobacter hominis</i>	19	0.98	1.90
<i>Campylobacter jejuni</i>	35	1.80	3.55
<i>Campylobacter lari</i>	15	0.77	1.59
<i>Campylobacter upsaliensis</i>	10	0.51	1.00
Other	3	0.15	0.17
<i>Helicobacteraceae</i>	603	31.02	21.21
<i>Helicobacter acinonychis</i>	4	0.21	0.59
<i>Helicobacter hepaticus</i>	21	1.08	1.98
<i>Helicobacter pylori</i>	3	0.15	0.75
<i>Sulfuromonas denitrificans</i>	335	17.23	9.56
<i>Wolinella succinogenes</i>	240	12.35	8.31
Other	0	0.00	0.02
<i>Nautiliaceae</i>	47	2.42	3.33
Unclassified	482	24.79	16.34
<i>Nitratiruptor</i> sp.	240	12.35	7.98
<i>Sulfurovum</i> sp.	235	12.09	7.96
Other	7	0.36	0.41
Other Proteobacteria	395	20.32	24.53
$\alpha$ -Proteobacteria	30	1.54	2.47
$\beta$ -Proteobacteria	47	2.42	3.11
<i>Burkholderiales</i>	34	1.75	2.15
Other	13	0.67	0.96
$\gamma$ -Proteobacteria	242	12.45	14.30
<i>Alteromonadales</i>	60	3.09	4.06
<i>Enterobacteriales</i>	19	0.98	1.60
<i>Oceanospirillales</i>	36	1.85	1.58
<i>Pseudomonadales</i>	34	1.75	2.27
<i>Thiotrichales</i>	22	1.13	0.75
<i>Vibrionales</i>	26	1.34	1.68
Other	45	2.31	2.35
$\delta$ -Proteobacteria	68	3.50	4.23
<i>Desulfuromonadales</i>	46	2.37	2.64
Other	22	1.13	1.59
Actinobacteria	3	0.15	0.35
Aquificae	5	0.26	0.38
Bacteroidetes/Chlorobi	55	2.83	3.82
<i>Chlorobiales</i>	14	0.72	1.22
Other	41	2.11	2.60
<i>Chlamydiae/Verrucomicrobia</i>	1	0.05	0.25
<i>Chloroflexi</i>	3	0.15	0.25
<i>Cyanobacteria</i>	11	0.57	0.94
<i>Deinococcus-Thermus</i>	2	0.10	0.06

Table 2. cont.

Taxon	Best match <sup>a</sup>		Within best 5 matches <sup>b</sup>
	#	%	%
<i>Fibrobacteres/Acidobacteria</i>	2	0.10	0.17
<i>Firmicutes</i>	48	2.47	3.19
<i>Bacillales</i>	14	0.72	1.28
<i>Clostridiales</i>	30	1.54	1.61
Other	4	0.21	0.30
<i>Fusobacterium</i>	4	0.21	0.29
<i>Nitrospirae</i>	0	0.00	0.05
<i>Planctomycetes</i>	3	0.15	0.35
<i>Spirochaetes</i>	7	0.36	0.48
<i>Thermotogae</i>	1	0.05	0.13
Other Bacteria	0	0.00	0.04
<i>Archaea</i>	7	0.36	0.68
<i>Crenarchaeota</i>	1	0.05	0.06
<i>Euryarchaeota</i>	5	0.26	0.55
Other	1	0.05	0.07
Phage/Plasmid	1	0.05	0.29
<i>Eukaryota</i>	6	0.31	0.33

Top matches are derived from Table S3. Predicted *Arcobacter butzleri* proteins were compared to proteins in the NCBI non-redundant (nr) database by BLASTP. Matches with an Expect (E) value of  $>1 \times 10^{-5}$ , an identity of  $<25\%$ , and an alignment length across either the query or match sequences of  $<75\%$  were excluded. The top five matches (where applicable) were identified at each locus (Table S3) and classified by taxon. Individual orders are listed if they constitute  $>1\%$  of the total.

<sup>a</sup>A total of 1944 loci contained matches to proteins in the nr database, according to the criteria listed above. Percentage values for the defined taxa do not include the 315 loci (13.9% of the 2259 identified loci) for which no homologs were identified.

<sup>b</sup>Multiple strains of the same species (e.g. *C. jejuni* RM1221 and *C. jejuni* NCTC 11168) at a given locus were counted once. Percentage values for the defined taxa do not include the 315 loci for which no homologs were identified. doi:10.1371/journal.pone.0001358.t002

encoding these proteins are not distributed randomly through the RM4018 genome, but rather clustered with respect to position and protein function (Table S2, Table S3). Interestingly, many of these clusters contain proteins involved in transport, and several contain discrete loci associated with a single function, such as the urease and quinohemoprotein loci described below. The genes contained in these loci are found often in the same order as in other taxa, suggesting lateral transfer. However, the G+C content of these gene clusters is not significantly different from the genome as a whole, suggesting that if these clusters were acquired from other non-*Arcobacter* taxa the acquisition was not recent.

Generally, strain RM4018 proteins involved in “housekeeping” functions (e.g. amino acid biosynthesis, fatty acid biosynthesis and protein synthesis) have homologs well-conserved among the other Epsilonproteobacteria (Table 3). Nevertheless, a number of other major functional categories are more divergent. For example, 19% (13/70) of the strain RM4018 proteins involved in DNA replication and repair have either no homologs or homologs of low similarity ( $\leq 35\%$  identity) among the other epsilonproteobacterial taxa; this proportion increases to 37% (26/70) if proteins with homologs only within the *Campylobacteraceae* are included (Table 3). Other divergent functional categories include: sulfur/nitrogen metabolism (34%), transcriptional regulators/ $\sigma$ -factors (47%), signal transduction (49%), cell envelope (37%), chemotaxis (70%) and antibiotic resistance (36%) (Table 3). These divergent functional categories will be discussed in further detail.

## Methyl-directed mismatch repair and the absence of polynucleotide G:C tracts

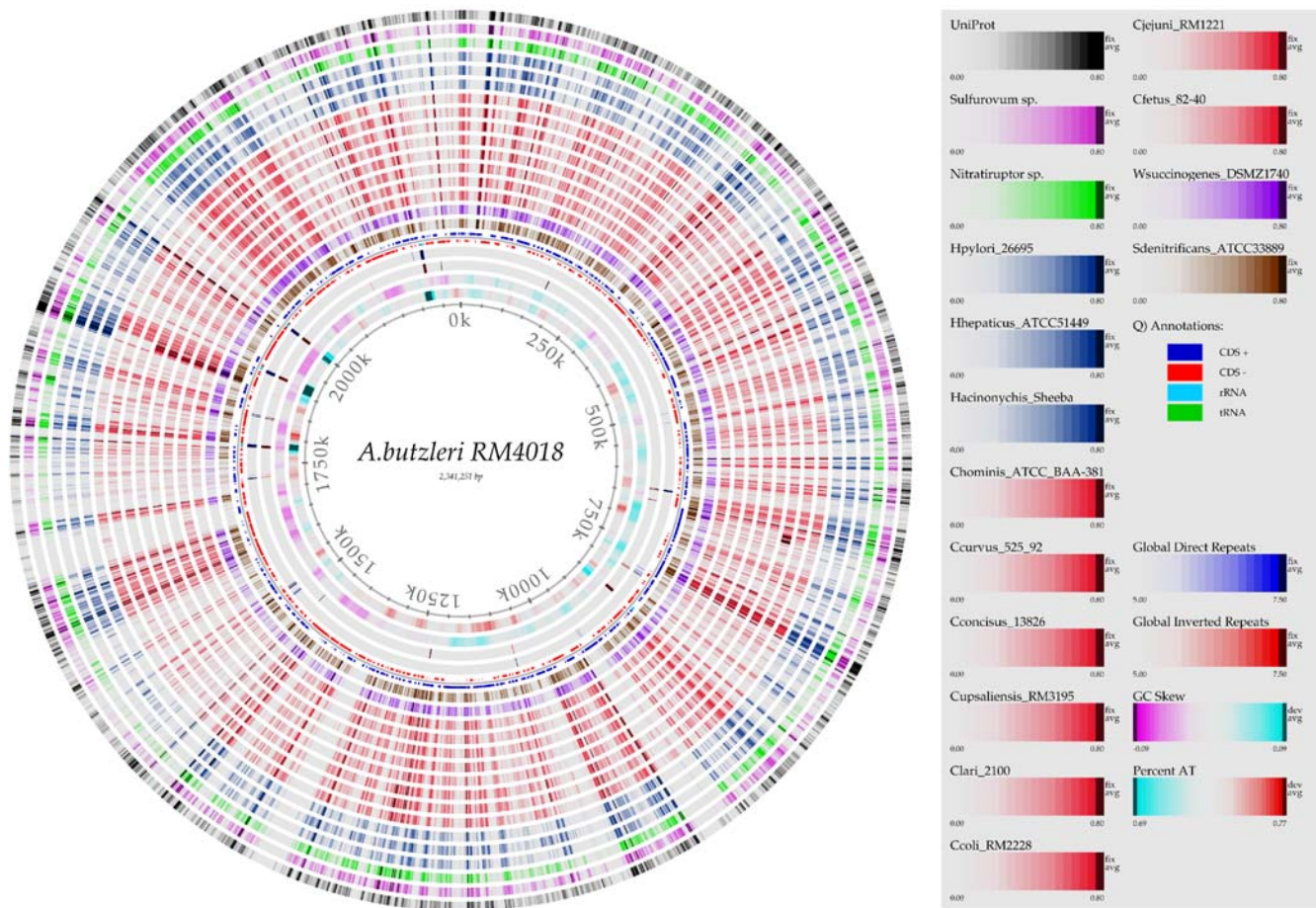
Members of the Epsilonproteobacteria lack multiple genes in the methyl-directed mismatch repair system (MMR). The MMR system depends upon the presence of three main functions: 1) the MutSLH endonuclease complex, 2) a methylation system to identify parental vs. daughter strands and 3) multiple 5'→3' and 3'→5' single-strand DNA exonucleases, e.g. RecJ, ExoI and ExoVII [63]. The *A. butzleri* RM4018 genome is not predicted to encode MutL or MutH. However, it is predicted to encode a MutS2 family MutS protein, which is distinguished from MutS1 by the presence of a C-terminal Smr domain. It has been proposed that the Smr domain has a MutH-like nicking endonuclease function [64]. Thus, the *A. butzleri* MutS2 protein may contain both MutS and MutH domains, obviating the need for a MutL scaffold protein. *A. butzleri* strain RM4018 is not predicted to encode a Dam DNA-adenine methyltransferase (data not shown) used commonly in MMR systems, but the presence of a Dcm DNA-cytosine methyltransferase was demonstrated experimentally in strain RM4018 (data not shown); this Dcm function may be used by the *A. butzleri* MMR system to distinguish between parental and daughter strands. Although the annotated epsilonproteobacterial genomes contain the 5'→3' exonuclease RecJ, they do not contain the 3'→5' exonucleases ExoI or ExoVII. A 32-fold increase in +1 frameshifts and an 11-fold increase in -1 frameshifts were observed in an *E. coli* ExoI<sup>-</sup> ExoVII<sup>-</sup> mutant [65]; therefore, one possible outcome of such a defect in MMR systems would be the formation and extension of hypervariable G:C tracts, such as those identified in all characterized *Campylobacter* and *Helicobacter* genomes. A major distinguishing characteristic of the *A. butzleri* RM4018 genome is the lack of such G:C tracts, a feature shared by the *S. denitrificans* and *W. succinogenes* genomes. A comparison of these three genomes revealed the presence of two to four members of the DnaQ superfamily, to which the 3'→5' exonuclease ExoX belongs [66]. Thus, one or more of these DnaQ homologs may provide the missing 3'→5' exonuclease function, and it is conceivable that the absence of hypervariable G:C tracts in *A. butzleri* strain RM4018 and the presence of these tracts in *Campylobacter* and *Helicobacter* may be due to the presence or absence of a functional MMR system, respectively.

## Sulfur assimilation, oxidation and the biosynthesis of sulfur-containing amino acids

*Arcobacter butzleri* strain RM4018 contains a number of genes required for sulfur uptake and assimilation (Figure 2). These genes include those encoding the sulfate ABC transporter CysATW, the sulfate binding protein Sbp, the ATP sulfhydrylase CysDN, the adenosine phosphosulfate (APS) reductase CysH, the sulfite reductase proteins CysI and CysJ and the siroheme synthase CysG. CysD and CysN have been identified also in *C. coli*, but the position of these genes, along with the APS kinase-encoding gene *cysC* and the 3'(2'),5'-bisphosphate nucleotidase-encoding gene *cysQ*, is in the capsular locus, suggesting that these genes in *C. coli* are involved in the formation or modification of the capsule and not sulfur assimilation per se.

Intracellular sulfate is reduced to sulfite in bacteria by one of two pathways. Both pathways first convert sulfate to APS via the ATP sulfhydrylase CysDN [67]. The first pathway converts APS to PAPS (phosphoadenosine phosphosulfate) using the kinase CysC, and then reduces PAPS to sulfite using the PAPS reductase CysH [67]; PAPS toxicity in some taxa is decreased through the conversion of PAPS to APS via CysQ [68,69]. The second





**Figure 1. Genome BLAST atlas of the *A. butzleri* strain RM4018.** *Arcobacter butzleri* strain RM4018 is the reference genome and is compared to a set of 15 other epsilonproteobacterial genomes, including different *Campylobacter* (rings 7–14 from center) and *Helicobacter* (rings 15–17 from center) strains, as well as the UniProt database (outermost ring in black). A web-based “zoomable atlas” can be found at [147]. doi:10.1371/journal.pone.0001358.g001

pathway, identified originally in plants and subsequently in taxa such as *P. aeruginosa* [70] and *M. tuberculosis* [71], reduces APS to sulfite directly using the APS reductase CysH. The presence of conserved two-cysteine motifs in strain RM4018 CysH, characteristic of bacterial APS reductases [70,71], and the absence of both CysC and CysQ homologs, suggests that sulfite reduction to sulfite in this strain does not use a PAPS intermediate.

The sulfite generated by strain RM4018 CysH has two potential fates: reduction to sulfide and assimilation into the sulfur-containing amino acids L-cysteine and L-methionine (Figure 2), or oxidation to sulfate, facilitated by a cytochrome *c* multienzyme complex encoded by the *sox* genes (*ab0563-ab0570*). Initial data indicated that sulfite is oxidized by strain RM4018 (data not shown), and also that it can grow in minimal media without added cysteine or methionine (data not shown). Hence, both metabolic fates are possible, and the genetic switch modulating oxidation or reduction of sulfite remains to be identified. Homologs of the *sox* genes have been identified in multiple taxa, including *Paracoccus pantotrophus* [72,73], *Chlorobium tepidum* [74] and *Rhodovulum sulfidophilum* [75]. The Sox clusters of some organisms can contain as many as 15 genes, but seven genes (*soxXYZABCD*) are essential for sulfur oxidation [76]. These seven genes encode four proteins: the heterodimeric *c*-type cytochrome SoxXA, the heterodimeric

sulfur-binding protein SoxYZ, the heterotetrameric SoxCD sulfur dehydrogenase and the thiol sulfate esterase SoxB [76]. Sulfite, sulfide, sulfur and thiosulfate are possible substrates for the Sox complex; thiosulfate and sulfide oxidation requires the entire complex, while sulfite oxidation does not require SoxCD [76]. The presence of SoxCD in strain RM4018 would suggest that thiosulfate is oxidized by this organism. However, other *sox* genes important for thiosulfate oxidation, e.g. *soxV* [77,78], are absent in this strain, as well as the Sbp-related thiosulfate-binding protein CysP. Thus, it is possible that only sulfite and sulfide are oxidized in strain RM4018. Sox proteins are present also in the related epsilonproteobacterial taxa *Sulfurovum*, *Nitratiruptor* and *S. denitrificans* (Table S2). However, there are multiple differences between the sulfur oxidation systems of these related organisms and those of strain RM4018: 1) the *sox* genes in strain RM4018 form a single cluster instead of two clusters, 2) the strain RM4018 *sox* cluster contains only one copy of *soxY* and *soxZ* and 3) no significant similarity exists between the SoxXA protein of strain RM4018 and the SoxXA proteins of *Sulfurovum*, *Nitratiruptor* and *S. denitrificans*. The *sox* cluster of strain RM4018 was not detected in 12 additional *A. butzleri* strains (see below). However, a cluster of *sox* genes similar to that of strain RM4018 has been identified in *A. halophilus* (data not shown). Thus, it appears that strain RM4018 may have acquired the

**Table 3.** Divergence within the major functional categories

Functional category	Divergent genes <sup>a</sup>		
	$\epsilon$ -proteobacteria	Campylobacteraceae	Helicobacteraceae
Glycolysis/TCA cycle/gluconeogenesis	5/34 (15%)	5/34 (15%)	6/34 (18%)
Respiration/electron transport	10/90 (11%)	24/90 (27%)	20/90 (22%)
Sulfur/nitrogen metabolism <sup>b</sup>	10/29 (34%)	22/29 (76%)	12/29 (41%)
Amino acid biosynthesis	3/85 (4%)	11/85 (13%)	3/85 (4%)
Purine/pyrimidine biosynthesis	2/29 (7%)	5/29 (17%)	2/29 (7%)
Biosynthesis of cofactors <sup>c</sup>	6/86 (7%)	16/86 (19%)	10/86 (12%)
Fatty acid biosynthesis	0/17 (0%)	1/17 (6%)	0/17 (0%)
Transcriptional regulators/ $\sigma$ factors	24/51 (47%)	41/51 (80%)	29/51 (57%)
Signal transduction	51/104 (49%)	77/104 (74%)	58/104 (56%)
Protein translation/modification	2/143 (1%)	9/143 (6%)	5/143 (3%)
DNA replication/repair	13/70 (19%)	26/70 (37%)	16/70 (23%)
RNA synthesis	1/17 (6%)	5/17 (29%)	1/17 (6%)
Macromolecule degradation	4/45 (9%)	13/45 (29%)	7/45 (16%)
Cell envelope <sup>d</sup>	70/190 (37%)	91/190 (48%)	85/190 (45%)
Transport/secretion	47/191 (25%)	88/191 (46%)	71/191 (37%)
Chemotaxis	32/46 (70%)	38/46 (83%)	33/46 (72%)
Antibiotic resistance	10/28 (36%)	20/28 (71%)	10/28 (36%)
General function	458/782 (59%)	597/782 (76%)	527/782 (67%)

<sup>a</sup>*Arcobacter butzleri* strain RM4018 proteins that have either no homologs or homologs of low similarity ( $\leq 35\%$  identity) within the given taxa/taxon. Categories and values are derived from Table S2.

<sup>b</sup>Includes [Fe-S] centers.

<sup>c</sup>Includes biosynthesis of prosthetic groups and carriers.

<sup>d</sup>LOS, flagella and membrane proteins.

<sup>e</sup>Includes conserved hypothetical and hypothetical proteins. Does not include general function proteins assigned to the other functional categories.

doi:10.1371/journal.pone.0001358.t003

*sox* cluster through lateral gene transfer from another *Arcobacter* species and that the sulfur oxidation machinery of *Arcobacter* and, e.g., *Sulfurovum* are likely to be distinct evolutionarily.

The genome of strain RM4018 contains all of the genes necessary for the biosynthesis of L-cysteine, L-methionine and iron-sulfur clusters (Figure 2). Consistent with the proposed absence of thiosulfate metabolism, the bacterium has only a gene encoding the CysK cysteine synthase A and not the CysM cysteine synthase B that utilizes thiosulfate instead of sulfide to produce S-sulfocysteine [67]. Strain RM4018 also contains both the cobalamin-dependent and cobalamin-independent homocysteine transmethyases MetH and MetE.

### The central metabolism of *A. butzleri*

The general function of the citric acid/tricarboxylic acid (TCA) cycle is to oxidize organic tri- and di-carboxylic acids to provide energy and biosynthetic precursors for metabolism. *Arcobacter butzleri* strain RM4018 encodes several proteins homologous to other epsilonproteobacterial TCA cycle enzymes (Figure 3), e.g. isocitrate dehydrogenase (AB1321), 2-oxoglutarate dehydrogenase (AB0852-AB0855), malate dehydrogenase (AB1322) and citrate synthase (AB0307). However, *A. butzleri* strain RM4018 is predicted putatively to encode two aconitate hydratases and two fumarate dehydratases. Additionally, enzymes that catalyze two TCA cycle steps are absent apparently in strain RM4018.

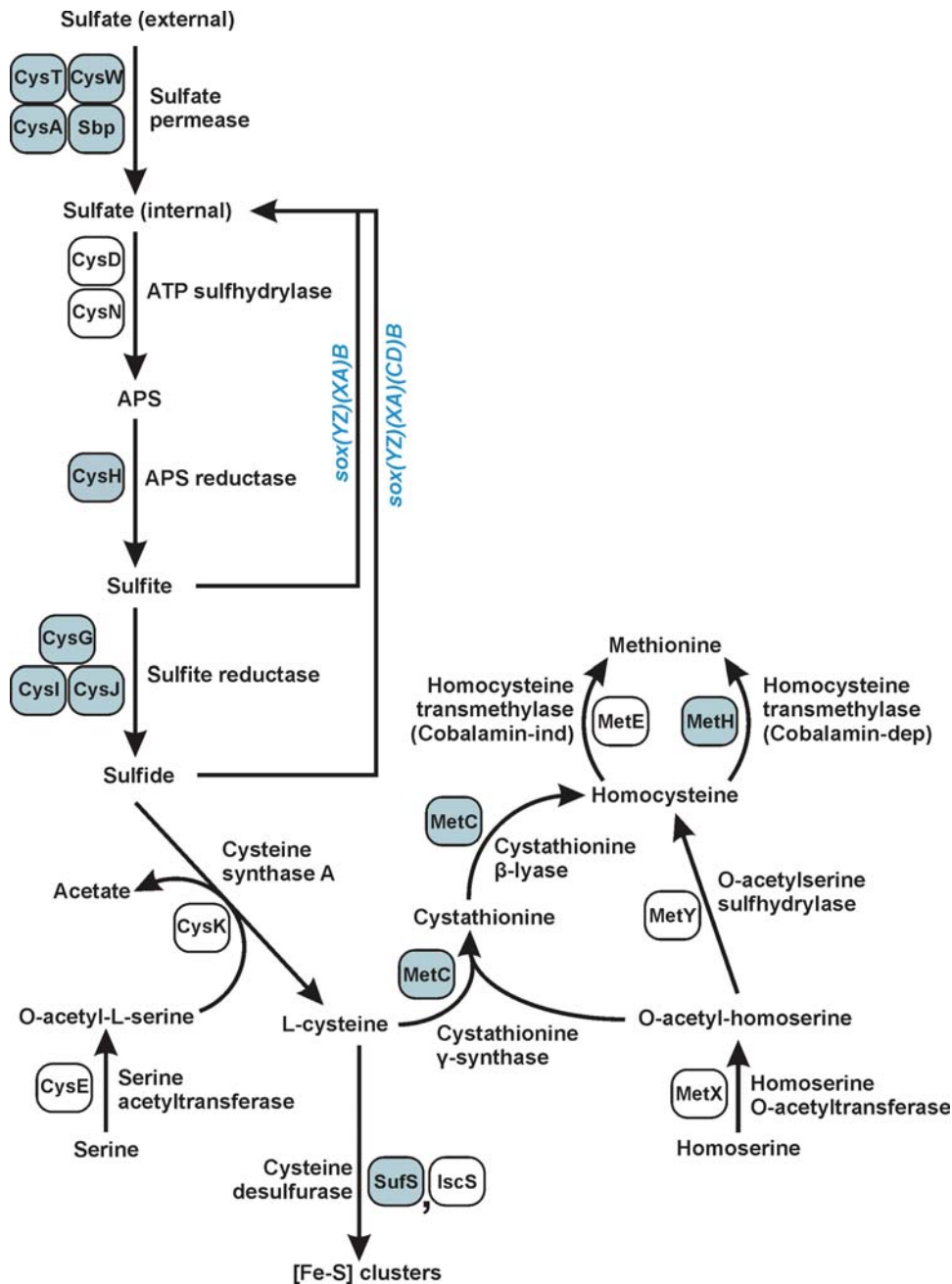
The genes *ab0275* and *ab1447* encode proteins homologous to the *E. coli* aconitate hydratases AcnA and AcnB, respectively. *Arcobacter butzleri* AcnB shares between 62% and 76% identity with other epsilonproteobacterial aconitases. *Arcobacter butzleri* AcnD,

encoded by *ab0275*, does not have significant similarity with any of the known epsilonproteobacterial proteins, but has an 82% similarity with the PrpD 2-methylisocitrate dehydratase of *Alkalilimnicola ehrlichei* as well as 68% similarity with *E. coli* AcnA. In addition, the proximity of *ab0275* to genes that encode the methylcitrate pathway enzymes PrpB and PrpC suggests that in strain RM4018 AcnD is the aconitase of the methylcitrate pathway and AcnB is the aconitase of the TCA cycle. Also, the absence of the 2-methylcitrate dehydratase PrpD and the presence of *acnD* next to *ab0276*, which encodes the AcnD-associated protein PrpF, suggests that *A. butzleri* strain RM4018 contains the alternate AcnD/PrpF methylcitrate pathway [79].

*Escherichia coli* can express three fumarases: A, B and C. The first two contain iron-sulfur clusters and are unstable aerobically; fumarase C is stable in the presence of oxygen. The genome of strain RM4018 is predicted to encode two fumarases: AB0722 and AB1921. AB0722 has 74% similarity to *E. coli* fumarase C and would be predicted to be active under aerobic conditions. AB1921 has 48% similarity to *E. coli* fumarases A and B and would require microaerobic or anaerobic conditions. AB1921 also has 55–64% similarities with two proteins encoded by the obligate microaerophile *W. succinogenes* and two encoded by *H. hepaticus*. Interestingly, the sequence similarities with the *W. succinogenes* proteins WS1766 and WS1767, and the *H. hepaticus* proteins HH1702 and HH1793 correspond to the first 282 and the last 185 residues of AB1921, respectively, suggesting that the two polypeptides in *W. succinogenes* or *H. hepaticus* are fused into one protein in *A. butzleri*.

No genes encoding the SucCD succinyl-CoA synthetase or the SdhABCD succinate dehydrogenase have been identified in the *A. butzleri* genome, but the gene cluster *ab0296-ab0298* encodes





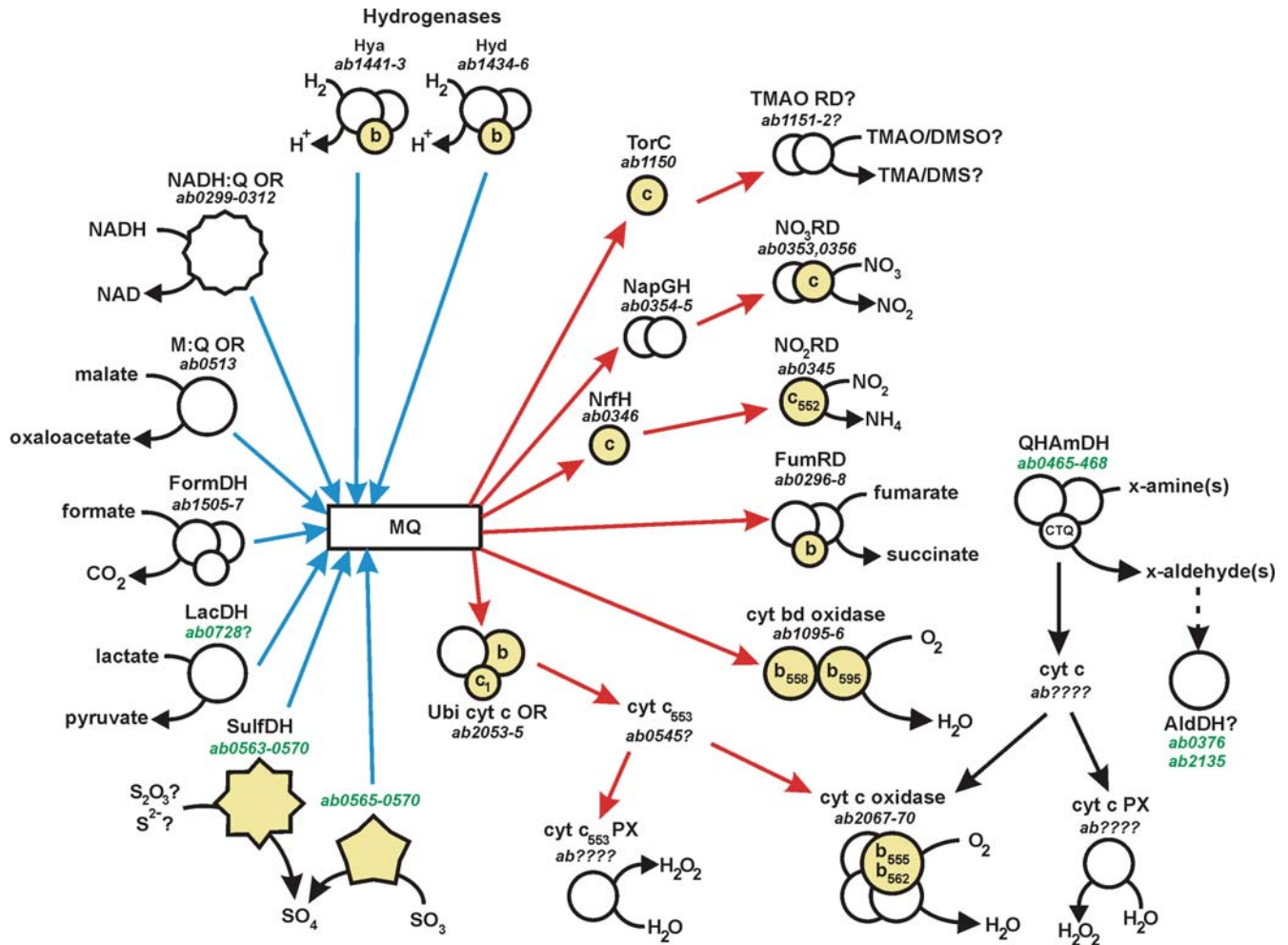
**Figure 2. Sulfur assimilation and biosynthesis of the sulfur-containing amino acids.** Genes/proteins in strain RM4018 unique within *Campylobacteraceae* or proteins with *Campylobacteraceae* orthologs of low similarity are labeled/shaded in blue. doi:10.1371/journal.pone.0001358.g002

proteins with high similarities (71–85%) to the fumarate reductase FrdABC of *C. jejuni*, *H. hepaticus*, *H. pylori*, and *W. succinogenes*, which catalyzes the reaction in the reductive direction converting fumarate to succinate. Some bacteria such as *E. coli* and *C. jejuni* have genes encoding both Sdh and Frd, and others, e.g. *H. pylori* and *W. succinogenes*, encode only Frd. The similar structures of Sdh and Frd preclude predicting, solely from sequence analyses, whether an enzyme is one or the other. Regulation of transcriptional levels by the oxygen content in the atmosphere permits differentiation between both enzymes; obligate aerobes encode Sdh and anaerobes encode Frd. Initial experiments with *A. butzleri* indicated the presence of fumarate reduction and no succinate oxidation, and this activity increased several-fold in

bacteria grown under anaerobic conditions relative to bacteria grown under aerobic conditions (data not shown). These results provided evidence supporting the identification of *A. butzleri* FrdABC (*ab0296-ab0298*) as a fumarate reductase (Figure 3).

The genome of strain RM4018 may contain also a pathway, encoded by the gene cluster *ab1917-ab1921*, which resembles segments of the 3-hydroxypropionate cycle [80,81] and the citramalate cycle [82,83], and interconverts glyoxylate and propionyl-CoA with pyruvate and acetyl-CoA (Figure 3). AB1917 and AB1920 are similar to the CitE citrate lyase. In addition, AB1918 is an acetyl-CoA synthetase with a CaiC domain in residues 50-550, and a CitE domain in residues 550-825. CitE domains have strong similarity to malyl-CoA lyases and have lesser





**Figure 4. Respiratory pathways in *A. butzleri* strain RM4018.** Transfer of electrons to the menaquinone (MQ) pool is represented by blue arrows; transfer of electrons from the menaquinone pool is represented by red arrows. b- and c-type cytochromes are shaded yellow. Genes unique within the *Campylobacteraceae* are labeled in green. NADH:Q OR: NADH:quinone oxidoreductase; M:Q OR: malate:quinone oxidoreductase; FormDH: formate dehydrogenase; LacDH: L-lactate dehydrogenase; SulfDH: sulfur dehydrogenase; Ubi cyt c OR: Ubiquinol cytochrome c oxidoreductase; TMAO RD: TMAO reductase; NO<sub>3</sub>RD: nitrate reductase; NO<sub>2</sub>RD: nitrite reductase; FumRD: fumarate reductase; QHAmDH: quinohemoprotein amine dehydrogenase; AldDH: aldehyde dehydrogenase; PX: peroxidase; CTQ: cysteine tryptophylquinone. doi:10.1371/journal.pone.0001358.g004

Potential electron donors, in addition to NADH, are hydrogen, malate and formate. One large gene cluster encodes three FeNi hydrogenases with one uptake hydrogenase (*hupSL*) and two membrane associated proteins encoded by *hydABCDF* and *hyaABCD*, the latter two predicted to be anchored to the membrane by the b-type cytochromes HydC and HyaC, respectively. There are two formate dehydrogenases, one seleno-cysteine homolog and one cysteine homolog, suggesting that selenium may be important in their regulation. Also present is a malate:quinone oxidoreductase (Mqo) and a putative lactate dehydrogenase (AB0728), suggesting that lactate may be a potential electron donor. Electrons may be transferred also to the menaquinone pool through the Sox system described above, the number of electrons depending on the substrate oxidized, 2 e<sup>-</sup> for sulfite and 8 e<sup>-</sup> for sulfide.

Fumarate, nitrate and nitrite are used by strain RM4018 as electron acceptors. Fumarate can be reduced to succinate by the fumarate reductase FrdABC. Nitrate can be reduced to ammonia via a periplasmic nitrate reductase NapAB and a pentaheme nitrite reductase NrfAH. The *nap* operon has the same gene

number and order as that seen in *Campylobacter* species. It is possible also that trimethylamine-oxide (TMAO) and/or dimethylsulfoxide (DMSO) may serve as alternative electron acceptors. Analysis of the strain RM4018 genome indicates the presence of the pentaheme c-type cytochrome TorC encoded by *ab1150*. The CDS *ab1151* has been annotated as *bisC*, but the BisC family includes also other anaerobic dehydrogenases, such as the TMAO and DMSO reductases TorA and DmsA, respectively. The genes *torC* and *torA* are co-transcribed usually with *torD*, but no TorD homolog was detected in the strain RM4018 genome. Other electron acceptors may be present: CDS *ab1360* encodes a putative cytochrome b-type nitric oxide reductase, and *ab1987* encodes a putative nitrite/nitric oxide reductase. Additional investigations will be necessary to determine the function of these putative electron acceptors.

In addition to the respiratory proteins described, *A. butzleri* strain RM4018 is predicted to encode a quinohemoprotein amine dehydrogenase (QHAmDH), the presence of which is novel in the epsilon subdivision. QHAmDH is a heterotrimeric (αβγ) protein that deaminates oxidatively a variety of aliphatic and aromatic

amines, e.g. n-butylamine and benzylamine [84], and is unusual in that the small ( $\gamma$ ) subunit contains an intrinsic quinone cofactor [in *A. butzleri*, cysteine tryptophylquinone (CTQ)], formed by the covalent linkage of a cysteine residue to an oxidized tryptophan residue [85]. The presence of this intrinsic quinone cofactor permits the transfer of electrons directly to the *cbb3* oxidase through either a cytochrome *c* [86] or blue-copper protein (e.g. azurin [87]) intermediate (Figure 4). The QHAmDH locus of strain RM4018 encodes also a radical SAM (S-Ado-Met) family protein (AB0466). A radical SAM protein (ORF2) is found also in the QHAmDH locus of *Paracoccus denitrificans* and has been proposed to play an important role in the formation of the CTQ cofactor [88]. Amino acid motifs conserved in ORF2 are 100% identical to those in AB0466, as well as to the relevant amino acids in the  $\gamma$  subunit AB0467, suggesting that the strain RM4018 QHAmDH locus encodes a functional dehydrogenase. M9 minimal media amended with 0.5% n-butylamine-HCl (v/v), benzylamine-HCl (w/v) or methylamine-HCl (w/v) as a sole carbon source does not support the growth of strain RM4018 (data not shown). However, substrate specificity has been demonstrated among the amine-utilizing taxa [89]; therefore, it is likely that RM4018 utilizes an as yet unidentified aliphatic or aromatic amine. Finally, two aldehyde dehydrogenases, encoded putatively by *ab0376* and *ab2135*, also novel in the Epsilonproteobacteria, were identified in strain RM4018 (Figure 4). Thus, it is possible that the aldehydes generated from the QHAmDH are oxidized further.

## Urease

*Arcobacter butzleri* strain RM4018 contains six genes (*ab0808-ab0813*) involved in the degradation of urea. In bacteria this catabolism involves generally three sets of genes: a nickel-containing urease (composed of  $\alpha$ ,  $\beta$  and  $\gamma$  subunits), urease accessory proteins which deliver the nickel to the urease and a nickel uptake system. The urease of strain RM4018 is functional (data not shown), as determined by a phenol red/urea assay [90], although the level of activity is not as high as that found in other urease-producing taxa (e.g. UPTC *C. lari*). As in *Helicobacter*, the urease  $\alpha$  and  $\beta$  subunits are fused, and the strain RM4018 urease subunits show high homology to *Helicobacter* urease subunits. However, differences exist with the urease loci of both *Helicobacter* and UPTC *C. lari*. First, the gene order of the locus itself, *ureD(AB)CEFG*, is similar to the gene order of the urease loci in *Klebsiella*, *Proteus*, *E. coli* O157:H7 and *Vibrio*, but not the *Helicobacter* locus *ure(AB)IEFGH*. Second, although the urease enzyme itself is similar to the *Helicobacter* urease, the accessory proteins UreD, UreE and UreF, had greater similarity to those identified in *Bacillus*, *Lactobacillus* and *Psychromonas* (Table S3); additionally, the nickel-binding protein UreE has a histidine-rich C terminus, found in multiple UreE proteins, but not in those from *Helicobacter* or *C. lari*. Also, unlike *Helicobacter*, no obvious nickel uptake system, such as the *H. hepaticus nixABDE* or *H. pylori nixA*, was found in strain RM4018. A putative nickel transporter, AB1752, was identified, but it is unclear whether it is specific for nickel, or is a heavy-metal ion transporter. Finally, although the *A. butzleri* urease may serve to degrade exogenous urea, it may also degrade endogenous urea, formed during putrescine biosynthesis, specifically during the conversion of agmatine to putrescine by SpeB (AB1578).

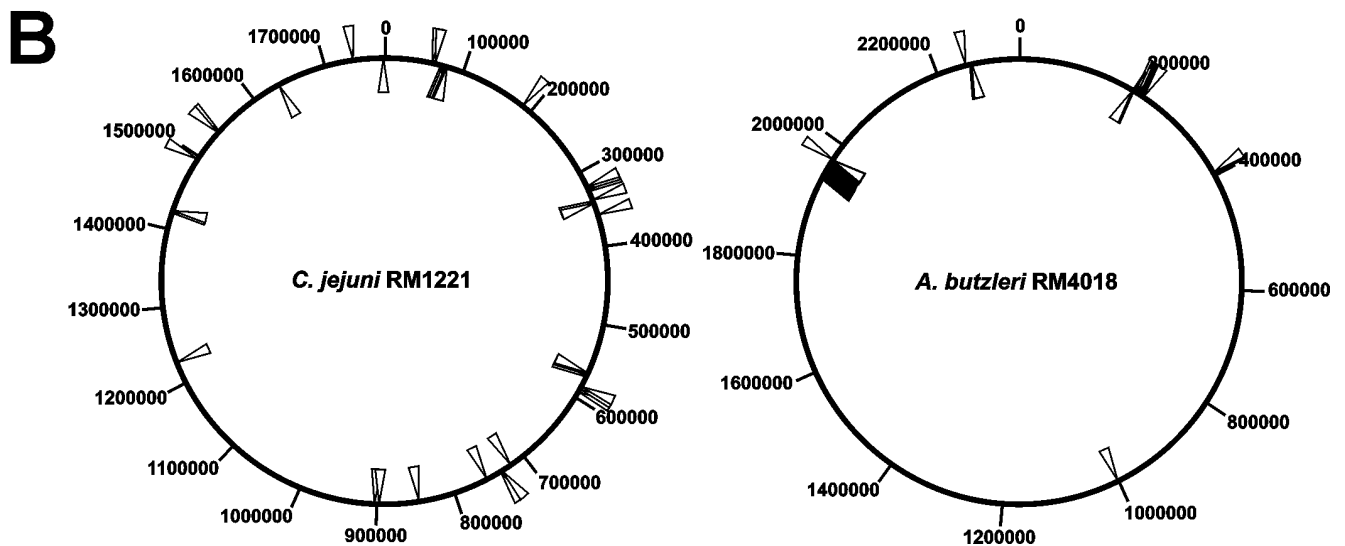
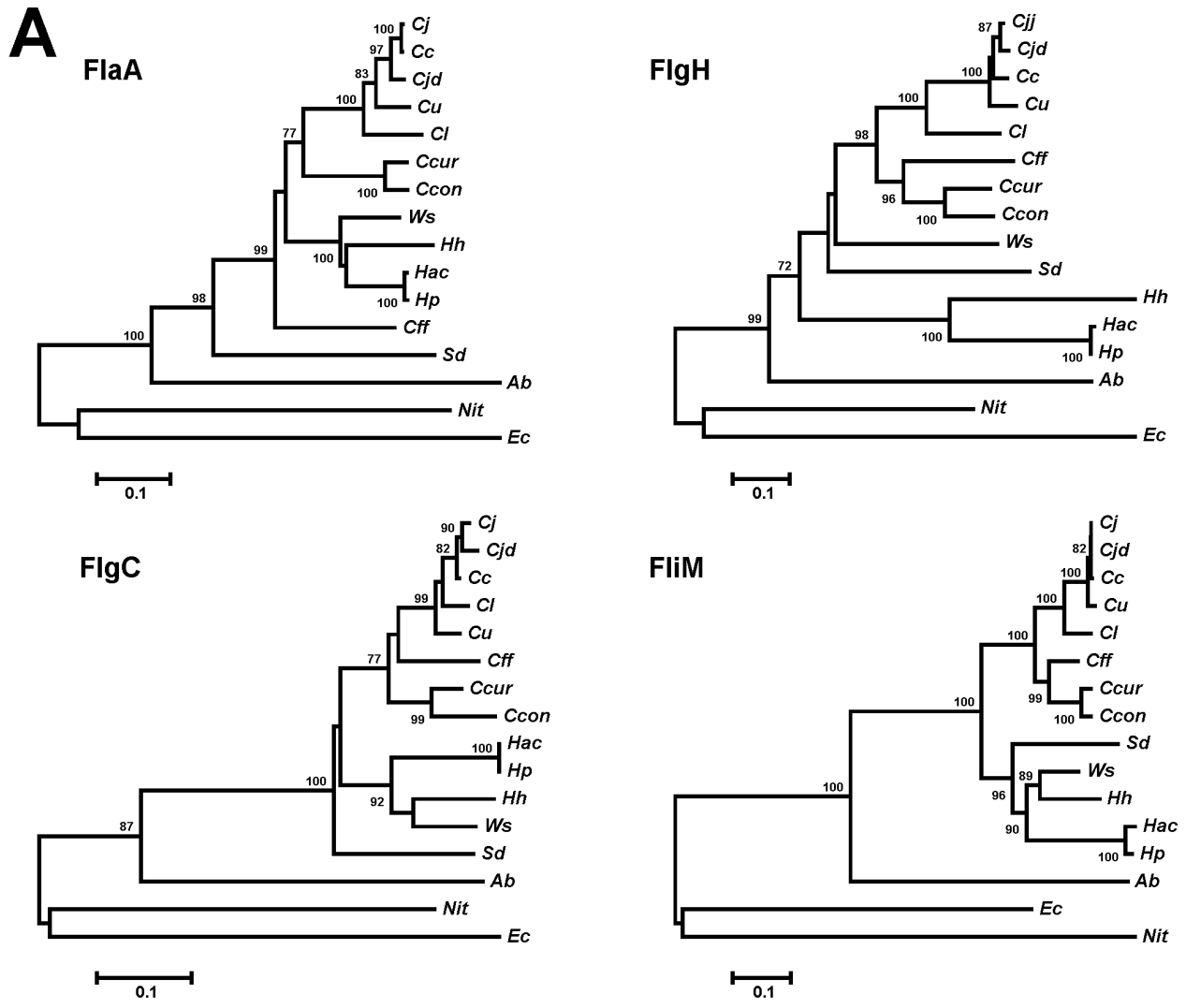
## Surface structures

SDS-PAGE analysis suggested that *A. butzleri* strain RM4018 can express lipooligosaccharide (LOS). To date, little is known about the roles of these molecules in *A. butzleri*. They generate great

attention in bacterial pathogens since LOS/lipopolysaccharide (LPS) are major inducers of proinflammatory responses, are immunodominant antigens, and play a role in host cell interactions. The LOS biosynthesis locus of strain RM4018 (*ab1805-ab1833*) showed a similar organization to those of *Campylobacter* [51,52,91,92], and is thus dissimilar from the loci of *Helicobacter* and *Wolinella*. At both ends of the locus are genes involved in the addition of heptose to the oligosaccharides and the encoded proteins are similar to those of other Epsilonproteobacteria (Table S2). Within the locus, genes are found whose products have functions related to LOS/LPS biosynthesis, including several glycosyltransferases, but the proteins encoded have greater similarity to proteins outside of the Epsilonproteobacteria (Table S3). This LOS/LPS biosynthesis region is conserved among 12 unrelated *A. butzleri* strains, based on comparative genomic indexing (described in detail later), which likely distinguishes *A. butzleri* from *C. jejuni*, where the LOS biosynthesis region is an intraspecies hypervariability region [93]. The gene conservation of this region in *A. butzleri* resembles more closely the conservation of the LPS core biosynthesis region occurring among many of the *Salmonella enterica* serovars [94].

Many bacteria synthesize structurally diverse polysaccharide polymers, O-antigen and capsule that are major antigenic determinants. It is possible that *A. butzleri* strain RM4018 produces O-antigen, since there is a locus (*ab0661-ab0697*) that encodes several additional glycosyltransferases. This region has two copies of *wbpG*, *hisH*, and *hisF*, found also in the *Pseudomonas aeruginosa* B-band O-antigen locus [95], and many of the other encoded proteins have greater similarity to proteins from bacteria outside of the Epsilonproteobacteria (Table S2). Although this region could be involved putatively in capsular formation, the absence of conserved *kps* capsular genes in this region, combined with the presence of O-antigen-related genes, suggests that *A. butzleri* strain RM4018 produces O-antigen and not capsule; however, further investigations will be necessary to determine the nature of the *A. butzleri* cell-surface structures. Like many O-antigen biosynthesis regions, the *A. butzleri* region appears to represent an intraspecies hypervariability region with the RM4018 region, present in only 1 of the 12 *A. butzleri* strains examined using comparative genomic indexing.

*Arcobacter butzleri* strain RM4018 is a motile bacterium that synthesizes a polar flagellum. Many of the flagellar apparatus proteins encoded by strain RM4018 have homologs in other epsilonproteobacterial taxa. However, phylogenetic analysis of selected flagellar proteins suggests that the flagellar apparatus of strain RM4018 has an evolutionary history distinct from those of *Campylobacter* and *Helicobacter* (Figure 5A). This distinct history is supported also by predicted differences between strain RM4018 and *Campylobacter/Helicobacter* in flagellar gene regulation (see below). Additionally, the flagellar genes of strain RM4018 are highly clustered, compared to the flagellar genes of *Campylobacter jejuni* (Figure 5B). The primary cluster in strain RM4018 contains 20 flagellar genes (*ab1931-ab1961*) with the other two flagellar clusters containing eight (*ab0197-ab0208*) and three (*ab2238-ab2244*) genes. The flagellar genes of the related organism *Nitratiruptor* are also highly clustered with the primary cluster containing 36 flagellar and chemotaxis genes. Significantly, the *Nitratiruptor* flagellar proteins also appear to be distinct phylogenetically from those of both *A. butzleri* and *Campylobacter/ Helicobacter*. Moreover, the primary flagellar cluster of *Nitratiruptor* has an atypical G+C content, suggesting acquisition through horizontal gene transfer [62]. Although the G+C content of the strain RM4018 flagellar genes is not atypical, it is possible that the flagellar genes of this organism were acquired via a similar mechanism.





←

**Figure 5. Phylogenetic analysis of four representative flagellar proteins and flagellar clustering in strain RM4018. A.** Each dendrogram was constructed using the neighbor-joining algorithm and the Kimura two-parameter distance estimation method. Bootstrap values of >75%, generated from 500 replicates, are shown at the nodes. The scale bar represents substitutions per site. *Cjj*: *Campylobacter jejuni* subsp. *jejuni*; *Cjd*: *C. jejuni* subsp. *doylei*; *Cc*: *C. coli*; *Cu*: *C. upsaliensis*; *Cl*: *C. lari*; *Cff*: *C. fetus* subsp. *fetus*; *Ccur*: *C. curvus*; *Ccon*: *C. concisus*; *Ab*: *A. butzleri*; *Ws*: *W. succinogenes*; *Sd*: *S. denitrificans*; *Hh*: *H. hepaticus*; *Hac*: *H. acinonychis*; *Hp*: *H. pylori*; *Nit*: *Nitratiruptor sp.*; *Ec*: *Escherichia coli*. **B.** Location of the flagellar genes of *C. jejuni* strain RM1221 and *A. butzleri* strain RM4018. doi:10.1371/journal.pone.0001358.g005

## Prophage and genomic islands

The genome of the *A. butzleri* strain RM4018 is predicted to contain a prophage. The size of this prophage is approximately 38 kb and spans genes *ab1655-ab1706*. BLASTP comparison of the predicted phage proteins to proteins from other bacteriophage indicates that this prophage is a member of the mutator (Mu) bacteriophage family. The size of the prophage is similar to other Mu-like bacteriophage, and it contains proteins similar to the Mu transposition proteins A and B in addition to coat, baseplate, and tail proteins. Mu-like bacteriophage have been found in multiple bacterial taxa, including *E. coli*, *Neisseria meningitidis*, *Deinococcus radiodurans*, *Haemophilus influenzae*, *Burkholderia cenocepacia* [96,97] and notably *C. jejuni* (CMLP1: strain RM1221 [52,93]). Indeed, 26 of the 50 predicted RM4018 Mu-like phage proteins are similar to those encoded by CMLP1 (Table S2). Hence, it is proposed to name this bacteriophage AMLP1 (Arcobacter Mu-like phage). Mu-like bacteriophage have been identified in other epsilonproteobacterial taxa, including *Campylobacter* (11 species), *Arcobacter* (3 species), and *Helicobacter bilis* ([93]; Miller and Mendoza, unpublished data). Some of these Mu-like phage are similar to CMLP1, but many show marked variation in gene content and gene sequence, indicating that CMLP1 and AMLP1 are members of a diverse Mu bacteriophage family common to Proteobacteria.

The genome of strain RM4018 also contains three small genomic islands, termed ABG11, ABG12 and ABG13 (for *Arcobacter butzleri* genomic island). ABG11 is 26,918 bp (bp 1,324,568-1,351,485) and contains 29 genes (*ab1330-ab1358*). ABG12 is 15,973 bp (bp 1,703,320-1,719,292) and contains 13 genes (*ab1721-ab1733*). ABG13 is 4,907 bp (bp 2,103,976-2,108,882) and contains eight genes (*ab2090-ab2097*). Additionally, ABG11-3 are bordered by direct repeats of 21, 21 and 25 bp, respectively. The presence of genomic islands within the epsilon subdivision is not unusual. They have been identified in *C. jejuni* (CJIE2 and CJIE3; [52]), *H. hepaticus* (HHGI; [57]) and *H. pylori* (*cagPAI*; [58]). CJIE2 and CJIE3 are bounded by direct repeats, contain an integrase gene at one end, and have integrated into the 3' end of a tRNA [52,93]. Consistent with the *C. jejuni* islands, all three RM4018 genomic islands contain terminal integrase genes and ABG11 has inserted into the 3' end of a leucyl-tRNA; however, ABG12 and ABG13 did not integrate into tRNAs, although ABG13 is in close proximity to a cysteinyl-tRNA. CJIE3 has been proposed to be an integrated plasmid, based on the similarity of this element to the *C. coli* strain RM2228 megaplasmid [52,93]; in contrast, none of the protein functions encoded by elements ABG11-3 suggested a plasmid origin. Interestingly, ABG12 does encode proteins similar to the Type I restriction enzymes HsdS and HsdM. Thus, the only Type I, II, or III restriction/modification enzymes present in strain RM4018 are encoded by a genomic island. The role of ABG13 is unknown.

## Antibiotic resistance

*Arcobacter butzleri* strain RM4018 was resistant to 42 of the 65 antibiotics tested (see supplementary Table S4). This level of

resistance is remarkably high, considering that the multi-drug-resistant *C. coli* strain RM2228 was resistant to only 33 of the same 65 antibiotics ([52] and Table S4). Strain RM4018 was resistant to all macrolides and sulfonamides tested and to all of the  $\beta$ -lactam antibiotics, with the exception of the  $\beta$ -lactam cephalosporin ceftazidime. Strain RM4018 had resistance to some quinolones, i.e. nalidixic acid and oxolinic acid, and also to chloramphenicol and 5-fluorouracil (5FU). The pattern of antibiotic resistance in strain RM4018 is consistent, in part, with the resistances of 39 *A. butzleri* strains tested against a smaller set of 23 antibiotics by Atabay and Aydin [98], with the exception that strain RM4018 was resistant to chloramphenicol and nalidixic acid. No plasmids were detected in strain RM4018; therefore all resistance mechanisms would be chromosomal in nature.

In many cases, the antibiotic resistance observed for strain RM4018 was due to the presence or absence of genes characterized previously in terms of antibiotic resistance. For example, chloramphenicol resistance is due most likely to the presence of a *cat* gene (*ab0785*), which encodes a chloramphenicol O-acetyltransferase.  $\beta$ -lactam resistance is due probably to the three putative  $\beta$ -lactamases (AB0578, AB1306 and AB1486) present in the RM4018 genome;  $\beta$ -lactam resistance may be enhanced also by the presence of the *lgAB* operon (*ab0179*, *ab0180*) which modulates penicillin tolerance in *Staphylococcus* [99,100]. Uracil phosphoribosyltransferase, encoded by the *upp* gene, catalyzes the first step in the pathway that leads to the production of the toxic analog 5-fluorodeoxyuridine monophosphate; mutations in the *upp* gene have been shown to lead to increased 5FU resistance [101,102]. Thus, the absence of *upp* in *A. butzleri* strain RM4018 results presumably in high 5FU resistance.

Although strain RM4018 is resistant to some quinolones, mutations implicated previously in *Campylobacter* [103] and *Arcobacter* [104] quinolone resistance at Thr-86, Asp-90 and Ala-70 in the DNA gyrase subunit GyrA are not present in strain RM4018. It is probably not a coincidence that strain RM4018 is susceptible to hydrophilic quinolones (e.g. ciprofloxacin and norfloxacin) and resistant to hydrophobic ones (e.g. nalidixic acid and oxolinic acid). These data suggest that the mechanism of quinolone resistance in strain RM4018 is not at the level of the gyrase, but rather of uptake. Decreased quinolone uptake is associated with either increased impermeability or the activity of efflux pumps [105]. Hydrophobic quinolones alone can transit across the phospholipid bilayer but all quinolones can enter the cell through porins [105]. It is possible that the phospholipid bilayer of strain RM4018 has reduced permeability towards hydrophobic quinolones in conjunction with modifications in the porins to permit passage of only hydrophilic quinolones. A more likely scenario is the presence of a hydrophobic quinolone-specific efflux pump. Examples of pumps with specificity towards one class of quinolone are known; for example, the NorA protein of *Staphylococcus aureus* has been shown to be involved in the specific efflux of hydrophilic quinolones [106].

## Virulence determinants

Putative virulence determinants have been identified in *Campylobacter*, but little is known about potential virulence factors in *A. butzleri*. Since campylobacterioses and reported *A. butzleri*-related illnesses have similar clinical outcomes [10], it might be expected that some *C. jejuni* virulence factors would be found in *Arcobacter*. In fact, some virulence determinants identified in *C. jejuni*, have homologs within *A. butzleri*. For example, the fibronectin binding proteins CadF and Cj1349 have homologs in strain RM4018 (AB0483 and AB0070, respectively). Moreover, homologs of the invasin protein CiaB, the virulence factor MviN, the phospholipase PldA and the TlyA hemolysin are present in strain RM4018 (AB1555, AB0876, AB0859 and AB1846, respectively). However, it has not been determined if these putative virulence determinants are functional or if their function and role in *A. butzleri* biology is similar to the function of their *Campylobacter* homologs.

On the other hand, several *Campylobacter* virulence-associated genes are not present in the RM4018 genome. Most notably, the genes encoding the cytolethal distending toxin CDT (*cdtABC*) are absent from strain RM4018. CDT is an exotoxin which irreversibly blocks eukaryotic cells in the G<sub>1</sub> or G<sub>2</sub> phase [107]; *cdtABC* genes have been identified in *Helicobacter hepaticus* and several characterized *Campylobacter* genomes. The absence of *cdtABC* in strain RM4018 correlated well with a study by Johnson and Murano [108] which was unable to detect *cdt* genes in *Arcobacter* by PCR. *A. butzleri* strain RM4018 also contains no PEB1 or JlpA adhesin homologs [109,110].

Analysis of the strain RM4018 genome identified two additional putative virulence determinants: *irgA* (*ab0729*) and *hecAB* (*ab0941-ab0940*). The *irgA* gene in *V. cholerae* encodes an iron-regulated outer membrane protein [111], and an IrgA homolog has been demonstrated to play a role in the pathogenesis of urinary tract infections by uropathogenic *E. coli* [112]. Adjacent to *irgA* is *ab0730*, which encodes a putative IroE homolog. IroE is a siderophore esterase found also in uropathogenic *E. coli* [113]. The other novel virulence determinant, HecA, is a member of the filamentous hemagglutinin (FHA) family, and *hecB* encodes a related hemolysin activation protein. FHA proteins are distributed widely among both plant and animal pathogens, e.g. HecA in *Erwinia crysthanthemi* contributes to both attachment and aggregation and is involved in epidermal cell killing [114]. Consistent with this distribution, RM4018 HecA homologs occur in both plant- (*Pseudomonas syringae*, *Ralstonia solanacearum*) and animal-pathogens (*Burkholderia cepacia*, *Acinetobacter spp.* and uropathogenic *E. coli*).

## Interacting with the environment: chemotaxis and signal transduction

Microbial life is characterized by continuous interactions between bacteria and their environment. The ability of microorganisms to monitor environmental parameters is a prerequisite for survival. Hence, bacteria have evolved different mechanisms such as sensory histidine kinases, methyl-accepting chemotaxis proteins, sigma( $\sigma$ )/anti-sigma factor pairs, and adenylate and diguanylate cyclases to monitor and rapidly adapt to changes in their environment.

*Arcobacter butzleri* is well equipped with a large number of these systems, allowing it to survive in diverse ecological niches. Prime bacterial mechanisms of environmental adaptation and gene regulation are the two-component systems (2CS), consisting of a membrane-bound sensor histidine protein kinase (HPK) and a cytoplasmic response regulator (RR). *A. butzleri* has one of the highest densities of 2CS genes per megabase, with a total of 78 genes of this type (Table S2). Analysis of the genome sequence

revealed 36 HPK genes, 41 RR genes, and one hybrid HPK/RR gene. On the genome, 31 pairs of HPK and RR genes are present that likely form functional 2CS involved in responses to environmental changes. In three instances, a more complex regulation may exist, where an additional sensor or response regulator gene is adjacent to the pair, namely AB0416, AB0417, AB0418; AB0453, AB0454, AB0455; and AB0795, AB0796, AB0797.

A fully functional chemotaxis system modulated by the 2CS CheA-CheY is present in *A. butzleri* strain RM4018. Its genome encodes a large number of putative chemotaxis proteins, many of which are novel in the epsilon subdivision. *A. butzleri* is predicted to have the CheA, CheY, CheV, CheW, CheR and CheB core enzymes, although most of these proteins have low similarity to their epsilonproteobacterial orthologs, and some have key differences in terms of domain composition. For instance, CheA in *Campylobacter* spp. is fused to a CheY-like response regulator (RR) domain [115] which is absent from *A. butzleri* CheA. Also, the methyltransferase CheB in *A. butzleri* has both the C-terminal methyltransferase and N-terminal RR domains whereas the *C. jejuni* CheB protein contains only the methyltransferase domain [115]. The presence of this CheB RR domain may support adaptation in *A. butzleri*, allowing the cell to “sense” increasing concentrations of attractants. In addition to these core chemotaxis proteins, strain RM4018 is predicted to encode the CheC phosphatase, the CheD glutamine amidase, involved in maturation of the methyl-accepting chemotaxis proteins (MCPs), and two additional CheY proteins. CheC hydrolyzes phosphorylated CheY and is employed by multiple taxa as an alternative to CheZ in signal removal [116]. The additional CheY proteins also could modulate signal removal by serving as a “phosphate sink” [116]. In the chemotaxis network, the receptor and kinase functions are separated in order to allow the cell to execute an integrated response to multiple stimuli [117]. The elaborate sensing capabilities of *A. butzleri* are reflected also in its high number (N = 25) of chemotaxis-specific receptors (MCPs). Some MCPs, e.g. AB1764, contain Cache (Ca<sup>2+</sup> channels and chemotaxis) domains which have been identified in multiple bacterial taxa [118], and are similar to those found in animal voltage-gated Ca<sup>2+</sup> channels. In addition, the periplasmic sensor region of AB1496 contains a putative nitrate/nitrite sensory domain, suggesting that nitrate and/or nitrite serves as an attractant for *A. butzleri*.

The limited similarity between the remaining non-chemotaxis-associated 2CS of *A. butzleri* and those of other bacteria do not allow functional prediction. However, six histidine kinases possess PAS domains, which are involved often in sensing changes in cellular energy levels, oxygen levels, or redox potential. Remarkable is the apparent lack of genes encoding homologs of the NtrC/NtrB family of two-component proteins that act together with the  $\sigma$  factor RpoN. Genes encoding RpoN as well as the flagellar sigma factor FliA are also missing in the *A. butzleri* genome. This indicates that the transcriptional regulation of the *A. butzleri* polar flagella machinery is completely different from that of other Epsilonproteobacteria where the 2CS FlgS/FlgR together with RpoN and FliA regulate the formation of the basal-body hook and filament complex [119,120]. Regulation of *A. butzleri* flagella synthesis may involve a completely novel system, since also absent from the genome are genes encoding homologs of the lateral master flagella transcription factors, such as FlhC, FlhD of *E. coli* [121] or members of the LuxR-type of transcription regulators like *Sinorhizobium meliloti* VisNR [122].

There are no predicted alternate sigma factors in the *A. butzleri* genome. Instead, and unlike the genomes of *Campylobacter* spp. or *Helicobacter* spp., it contains seven ECF-family (extracytoplasmic

**Table 4.** Strains used in the comparative genomic indexing analysis.

Strain	Description	Location	Source
RM4018	ATCC 49616, type strain	USA (Calif.)	Human diarrheal stool
RM1588	NADC 5276	USA	Chicken
RM1591	NADC 5377	USA	Turkey carcass
RM4128	29.97	South Africa	Human stool
RM4462	NADC 3553	USA (Tex.)	Human stool
RM4467	NADC 3566	USA	Primate rectal swab
RM4596	NADC 5262	USA (Iowa)	Turkey
RM4843	NADC 5278	USA (Iowa)	Chicken carcass
RM4850	NADC 6830	USA (Iowa)	Horse
RM5516	NADC 3156	USA (Iowa)	Pig
RM5538	CDC D2725	USA (Mass.)	Human stool
RM5541	CDC D2901	USA (Colo.)	Human stool
RM5544	CDC D2778	Thailand	Human stool

All test strains selected had unique MLST sequence types (data not shown).  
doi:10.1371/journal.pone.0001358.t004

function)  $\sigma$ /anti- $\sigma$  factor pairs. Genes encoding these pairs are in tandem. Anti- $\sigma$  factors are usually membrane-bound and prevent their cognate  $\sigma$  factor from associating with core RNA polymerase [123]. To date, only the *W. succinogenes* proteins WS1422 and WS1423 have been identified as an ECF-family  $\sigma$  factor pair within the epsilon subdivision. It is not known what genes are regulated in *A. butzleri* by these ECF-family  $\sigma$  factors. In other bacterial taxa, these ECF-family  $\sigma$  factors regulate functions such as alginate biosynthesis, iron-citrate transport, pyoverdine biosynthesis and cytochrome expression [123], and have been associated with virulence in *S. enterica*, *P. aeruginosa* and *M. tuberculosis* [124].

Finally, *A. butzleri* strain RM4018 encodes a class of proteins involved in formation and turnover of the “second messenger” bis-(3'-5')-cyclic dimeric guanosine monophosphate (cyclic-di-GMP). In bacteria cyclic-di-GMP is synthesized by diguanylate cyclases containing a conserved GGDEF motif (GGDEF(DUF1) domain proteins) and degraded into diguanylate (pGpG) by phosphodiesterase proteins containing domains enriched in glutamic acid, alanine and leucine (EAL(DUF2) domain proteins) [125,126]. The genome of strain RM4018 encodes 25 proteins, either membrane-associated or cytoplasmic/periplasmic, containing GGDEF and/or EAL domains (Figure S2b). In other bacteria, GGDEF/EAL domain proteins have been implicated in biofilm formation [127,128], motility [129,130] and virulence [131].

It is likely that many of the signal transduction proteins described above interact across group boundaries, e.g. GGDEF proteins with chemotaxis proteins. Thus, taken together, the large repertoire of 2C systems, chemotaxis proteins, ECF sigma factors and cyclic-di-GMP factors suggest the presence of a truly complex signal transduction network.

### Genomic comparisons with other *A. butzleri* strains

The genomic diversity of 13 *A. butzleri* strains, of both human and animal origin, was examined by microarray-based comparative genomic indexing (CGI) analysis. A list of the strains used in the CGI analysis is presented in Table 4. The CGI analysis allowed the assessment of CDS content for each *A. butzleri* strain relative to the *A. butzleri* DNA microarray, which comprises 2238 CDS from strain RM4018. Genomic DNA from strain RM4018 was used as

a reference DNA and competitively hybridized with genomic DNA from each of the other *A. butzleri* strains. The GENCOM software described previously [132,133] was used to assign the CDS as present, absent or multicopy for each *A. butzleri* strain. The complete CGI data sets as trinary scores (present = 1; absent = 0; and multicopy = 2) are available in supplementary Table S5. It was observed that 74.9% (1676 of 2238) of the CDS represented on the microarray were present in all *A. butzleri* strains, and served to define approximately the core genes of *A. butzleri* (supplementary Table S6).

To gain more information concerning how genomic diversity may affect the physiology of these *A. butzleri* strains, the *A. butzleri* CDS were grouped into defined functional categories. Genes in each functional category for the *A. butzleri* RM4018 genome were analyzed against the CGI data set with GeneSpring software. The results show that all *A. butzleri* strains in this study possessed every gene in our data set assigned to the functional categories: polyamine biosynthesis; purines, pyrimidines, nucleosides and nucleotides biosynthesis; aminoacyl tRNA synthetases and tRNA modification; protein translation and modification; and protein and peptide secretion (Table 5). The results demonstrated also that more than 95% of the genes in our data set assigned to the following functional categories were present for all strains: energy metabolism; amino acid biosynthesis; and ribosomal protein synthesis, modification and maturation (Table 5). In contrast, functional categories with less than 70% of the genes present in at least two strains were: central intermediary metabolism; transcriptional regulation; surface polysaccharides, lipopolysaccharides, and antigens; surface structures; phage-related functions and prophage; and hypothetical proteins (Table 5). There were also 42 CDS that were found only in strain RM4018 (Table S7), including the Sox cluster of seven genes (*soxXYZABCD*), suggesting that this system was acquired by RM4018. Some genes within the bacteriophage AMLP1 were present in the other 12 *A. butzleri* strains, but this prophage was not present in its entirety in any of the other 12 *A. butzleri* strains. The same applies to the small genomic islands ABG11 and ABG12. Finally, urea degradation does not appear to be a core function of *A. butzleri*, considering that only six of the 12 strains possessed the gene clusters *ab0802-ab0806* and *ab0808-ab0813*.

### Conclusions

The human pathogen *A. butzleri* is a member of the family *Campylobacteraceae* in the epsilon subdivision of the Proteobacteria, and it was found, based on 16S rDNA sequence similarity, to be most closely related to the *Campylobacter*s. However, based on complete genome analysis, *A. butzleri* appeared more closely related to non-*Campylobacteraceae* taxa, specifically *S. denitrificans*, *W. succinogenes*, *Sulfurovum* and *Nitratiruptor*. This unexpected finding may be less surprising considering the limited number of completed epsilonproteobacterial genomes. Considering that to date the Epsilonproteobacteria genome sequences annotated belong in their majority to pathogens, further refinements of the structure of the epsilonproteobacterial taxon are likely to emerge, as additional genomes of this subdivision are sequenced, especially those of other *Arcobacter* species, and of related genera such as *Sulfurospirillum*.

Even though *A. butzleri* has a large number of *S. denitrificans*, *W. succinogenes*, *Sulfurovum* and *Nitratiruptor* orthologs, strain RM4018 has several features in common with *Campylobacter*, such as the inability to utilize sugars as carbon sources and the presence of defined LOS/LPS loci. Interesting differences between *A. butzleri* and the *Campylobacter*s are the absence of Type I, II or III restriction enzymes and the absence of polynucleotide G:C tracts.

**Table 5.** Assignment of absent CDS to functional categories of *A. butzleri*.

	Genes in data set	Absent genes (%) <sup>a</sup>
<b>I. Small molecule metabolism</b>		
I.A Degradation	12	8.3
I.B Energy metabolism	129	3.8
I.C Central intermediary metabolism <sup>b</sup>	37	40.5
I.D Amino acid biosynthesis	85	3.5
I.E Polyamine biosynthesis	4	0
I.F Purines, pyrimidines, nucleosides and nucleotides	41	0
I.G Biosynthesis of cofactors, prosthetic groups and carriers	96	7.3
I.H Fatty acid biosynthesis	17	11.8
<b>II. Broad regulatory functions</b>	12	8.3
II.A Transcriptional regulation	51	33.3
II.B Signal transduction	104	23.1
<b>III. Macromolecule metabolism</b>		
III.A Synthesis and modification of macromolecules	240	7.1
III.A.2-3 Ribosome and ribosomal protein synthesis, modification and maturation	63	4.8
III.A.5 Aminoacyl tRNA synthetases and tRNA modification	53	0
III.A.7 DNA replication, restriction/modification, repair, and recombination	69	13
III.A.8 Protein translation and modification	27	0
III.A.9 RNA synthesis, and RNA modification, and DNA transcription	18	5.6
III.A.11 Phospholipids	11	27.2
III.B Degradation of macromolecules	45	8.9
III.C Cell envelope		
III.C.1 Membrane proteins, lipoproteins, and porins	93	16.1
III.C.2 Surface polysaccharides, lipopolysaccharides, and antigens	57	36.8
III.C.3 Surface structures	41	34.1
III.C.4 Murein sacculus and peptidoglycan	21	9.5
<b>IV. Cell processes</b>		
IV.A Transport/binding proteins	169	24.9
IV.B Chaperones	16	12.5
IV.C Cell division	14	7.1
IV.D Chemotaxis and mobility	46	21.7
IV.E Protein and peptide secretion	21	0
IV.G Detoxification	9	11.1
IV.I Pathogenicity	12	25
IV.J DNA uptake/competence	8	12.5
<b>V. Other</b>		
V.A Phage-related functions and prophage	54	96.3
V.D Drug/analog sensitivity and antibiotic resistance	28	17.9
V.F Adaptations and atypical conditions	9	11.1
<b>VI. Miscellaneous proteins, general function proteins, and hypothetical proteins</b>		
VI.A Miscellaneous/General function	143	16.1
VI.B Domain of unknown function (DUF) proteins	61	9.8
VI.C Conserved hypothetical proteins-no conserved domains	303	19.9
VI.D Hypothetical proteins	262	39.7

<sup>a</sup>Genes absent in at least 2 of 12 additional strains of *A. butzleri*.

<sup>b</sup>Primarily *sox* and *ure* genes.

doi:10.1371/journal.pone.0001358.t005

These tracts in *Campylobacter* are found often in LOS and capsular loci, and changes in the lengths of these tracts could modify the surface structure of the organism, perhaps altering the serotype.

Therefore, the absence of these tracts and the resulting absence of contingency genes may disadvantage *A. butzleri* as a pathogen. Alternatively, the absence of restriction enzymes may facilitate the

uptake and incorporation of phage, genomic islands, and novel genes/loci, as seen in strain RM4018, thus providing the genomic plasticity that is associated often with virulence.

Notwithstanding its relationships to *Campylobacter* and *Helicobacter*, *A. butzleri* is probably an environmental organism that may cause disease through either water-mediated food contamination or ingestion of *A. butzleri*-contaminated water, and whose biology is not associated tightly with any particular host or hosts. *A. butzleri* is more psychrophilic than *Campylobacter* spp. and may not proliferate in some hosts, especially avian hosts, to the same extent as, e.g. *C. jejuni*, resulting in *A. butzleri*-related food-borne illnesses occurring less frequently. The enhanced ability to survive in the environment, compared to such taxa as *Campylobacter*, is due most likely to the genes and systems described in this study, such as the increased number of respiratory enzymes, chemotaxis proteins and two-component systems. *A. butzleri* has a number of genes that confer increased environmental survival and the ability to grow under a wider range of atmospheric conditions and lower temperatures. An example of the first would be the photoreactivating lyase *phr*, which reduces DNA damage induced by UV radiation, a condition encountered frequently by environmental organisms. Examples of the second would include: 1) more oxygen-stable enzymes (e.g. pyruvate dehydrogenase), 2) a larger repertoire of TCA cycle enzymes active over a greater range of oxygen concentrations, 3) novel cold shock and stress proteins and 4) proteins such as the fatty acid cis/trans isomerase, Cti, which modulate membrane fluidity.

Although *A. butzleri* has been isolated from clinical diarrheal stool samples, the association of this organism with human illness is not as defined as the association of *C. jejuni* and human gastroenteritis. Obviously, the presence of *Campylobacter* virulence determinant homologs in strain RM4018 is not predictive of the pathogenicity of this strain; however, it is consistent with such a designation. It is likely, if *A. butzleri* is indeed pathogenic, that additional uncharacterized virulence determinant genes are present within the strain RM4018 genome. Therefore, additional investigations, using in part the genomic information presented here, will be required to elucidate further the mechanisms supporting the pathogenicity of *A. butzleri*. Indeed, the genome sequence and analyses presented in this study are an important first step in understanding the physiology and genetics of this organism, which constitutes a bridge between the environment and mammalian hosts.

## MATERIALS AND METHODS

### Growth conditions and chemicals

All *Arcobacter* strains were cultured routinely at 28°C on Brain Heart Infusion agar (Becton Dickinson, Sparks, MD) amended with 5% (v/v) laked horse blood (Hema Resource & Supply, Aurora, OR). The incubation atmosphere for all strains was 5% H<sub>2</sub>, 10% CO<sub>2</sub>, and 85% N<sub>2</sub>. PCR enzymes and reagents were purchased from New England Biolabs (Beverly, MA) or Epicentre (Madison, WI). All chemicals were purchased from Sigma-Aldrich Chemicals (St. Louis, MO) or Fisher Scientific (Pittsburgh, PA). DNA sequencing chemicals and capillaries were purchased from Applied Biosystems (Foster City, CA).

### Genomic DNA preparation and library construction

A sample of the *Arcobacter butzleri* type strain ATCC 49616 was analyzed by two methods: 1) an *Arcobacter* multiplex speciation PCR [134] and 2) amplification and sequencing of the 16S rDNA locus, using the primers 16ARCO-1 (5' ACA ATG GAC GAA

AGT CTG AT 3') and 16ARCO-2 (5' CGC AAT CGG TAT TCC TTC 3'). Results of these tests indicated that the type strain isolate was a mixed culture containing an uncharacterized *A. cryaerophilus* strain. Therefore, a phosphate-buffered saline suspension of the type strain isolate was sonicated and dilution plated. Eight well-isolated colonies were picked and tested using the above two methods. Both tests confirmed that the DNA from each of the eight colonies was from a pure culture of *A. butzleri* and one purified isolate was given the new designation of RM4018. Moreover, the multilocus sequence type of strain RM4018 was identical to the sequence types of multiple isolates of strain ATCC 49616, obtained from independent laboratories (W. Miller, manuscript in preparation).

*Arcobacter* genomic DNA was prepared as described previously [135]. The genomic DNA was checked for high molecular weight using pulsed field gel electrophoresis, quantified using a PicoGreen assay (Invitrogen, Carlsbad, CA), and then sheared physically to the desired average sizes of approximately 4.0 kb and 40.0 kb, for the small- and large-fragment libraries, respectively. The 3–4 kb plasmid insert library was constructed in the pAGEN proprietary vector system (Agencourt). A large-insert fosmid library was constructed in the CopyControl pCC1FOS vector system (Epicentre, Madison, WI). Libraries were not amplified and were kept as ligation mixtures.

Initial quality control of each library included quantification of: the titer of each one, the frequency of recombinant/insert-less clones and the average insert size (using a subset of 20 recombinant clones). Prior to high throughput sequencing, 384 clones were picked and sequenced bi-directionally. A BLAST analysis was performed on the 768 sequencing reads to verify the integrity of the library. Sequence data was further reviewed to provide more precise information on the percentage of recombinant clones.

### Template preparation

Solid Phase Reversible Immobilization (SPRI) [136,137] technology was used to purify the templates. SPRI technology uses carboxylate-coated, iron-core, paramagnetic particles to capture DNA of a desired fragment length based on tuned buffering conditions. Once the desired DNA is captured on the particles, they can be concentrated magnetically and separated so that contaminants can be washed away.

High-copy plasmid templates (insert size range 800 bp–4 kb) were purified using a streamlined SprintPrep SPRI protocol. Plasmid DNA was harvested directly from lysed bacterial cultures by trapping both plasmid and genomic DNA with the functionalized bead particles and selectively eluting only the plasmid DNA. Briefly, the purification procedure involves addition of alkaline lysis buffer (containing RNase A) to the bacterial culture, addition of an alcohol-based precipitation reagent including paramagnetic particles, separation of the magnetic particles using custom ring-based magnetic separator plates, 5× washing of the beads with 70% EtOH and elution of the plasmid DNA with water. The fosmid templates were purified using a SPRI protocol optimized to accommodate larger growth volumes and increased sample recovery needed for the low copy nature of these vectors.

### Polymerase chain reactions

Standard amplifications were performed on a Tetrad thermocycler (Bio-Rad, Hercules, CA) with the following settings: 30 s at 94°C; 30 s at 53°C; 2 min at 72°C (30 cycles). Each amplification mixture contained 50 ng genomic DNA, 1× PCR buffer (Epicentre), 1× PCR enhancer (Epicentre), 2.5 mM MgCl<sub>2</sub>,



250  $\mu$ M each dNTP, 50 pmol each primer, and 1 U polymerase (New England Biolabs). Amplicons were purified on a BioRobot 8000 Workstation (Qiagen, Santa Clarita, CA). Sequencing and PCR oligonucleotides were purchased from Qiagen or MWG-Biotech (High Point, NC).

### DNA sequencing

DNA templates from shotgun clones were sequenced in a 384-well format. Cycle sequencing reactions were performed on a 96-well or 384-well Tetrad thermocycler (Bio-Rad, Hercules, CA) using the ABI PRISM BigDye terminator cycle sequencing kit (version 3.1) and standard protocols. Shotgun template extension products were purified using the CleanSeq dye-terminator removal kit (Agencourt) and amplicon extension products were purified using DyeEx 96 well plates (Qiagen) according to the manufacturer's protocols. DNA sequencing was performed on an ABI PRISM 3130XL or ABI PRISM 3730 Genetic Analyzer (Applied Biosystems) using the POP-7 polymer and ABI PRISM Genetic Analyzer Data Collection and ABI PRISM Genetic Analyzer Sequencing Analysis software. Shotgun reads were processed using Phred base calling software [138,139] and monitored constantly against quality metrics using the Phred Q20. The quality scores for each run were monitored through a proprietary Galaxy LIMS system (Agencourt). Any substantial deviation from the normal range was investigated immediately.

### Contig assembly

Shotgun reads were assembled using the Paracel GenomeAssembler (Paracel, Pasadena, CA) with default program parameters and quality scores. To assist in the final assembly and to monitor read quality, Paracel-derived contig sequences and their associated reads were entered into SeqMan Pro (v7.0, DNASTar, Madison, WI). Within each contig, all reads were then trimmed for quality at the 5' and 3' ends. Those reads that did not meet minimum quality requirements were discarded. A large number of sequence and contig gaps were present after assembly of the initial shotgun reads. Arrangement of these contigs into a final single contig used both standard PCR and combinatorial long and standard PCR to bridge the contig gaps; due to low sequence similarity and minimal synteny, existing epsilonproteobacterial genome sequences could not be used as a scaffold for the *A. butzleri* genome. All gaps, both sequence and contig, present in the assembly were closed by PCR using primers designed with Primer Premier (v 5.0, Premier Biosoft, Palo Alto, CA). The *A. butzleri* strain RM4018 genome contains five nearly identical ribosomal RNA (rRNA) loci of ~6 kb. Each locus was amplified independently using unique flanking primer sets and long PCR. rRNA region amplicons were sequenced, using rRNA primers common to all five loci, and assembled. The assembled rRNA loci were then integrated into the larger genome assembly manually. Other large repeat regions were analyzed similarly.

The final assembly contained 27,377 reads: 12,990 small insert reads (average insert size  $\cong$  4200 bp), 7,364 large insert reads (average insert size  $\cong$  31,650 bp) and 7,023 PCR-based reads. The final assembly contained also contiguous sequences ( $>2\times$  coverage/nt) on both strands for an average coverage of  $5.5\times$ ; ambiguous bases were re-sequenced at least twice. Validation of the final assembly was achieved by comparison of mock in silico digestion patterns with recently published PFGE patterns using the type strain of *A. butzleri* [140]. Although the PFGE band sizes were on average  $\sim 10\%$  larger than those predicted in silico, there was an almost 100% correlation between the two banding patterns.

### Genome analysis

Open reading frames were predicted using ORFscan (Rational Genomics, South San Francisco, CA) with the parameters set to include all three potential start codons (ATG, GTG and TTG) and a coding sequence (CDS) cut-off of 70 amino acids. Spurious CDSs (e.g. CDSs contained within larger CDSs) were deleted manually. Initial annotation of the putative CDSs was accomplished by comparing the predicted proteins to the nr database using BLASTP. The list of putative CDSs was then used to create a preliminary GenBank-formatted (.gbk) file that was entered into Artemis [141,142]. Annotation within Artemis included the fusion of split CDSs into pseudogenes and the identification of small (usually ribosomal protein) genes overlooked in the initial ORFscan analysis. The start codon of each putative CDS was curated manually, either through visual inspection within Artemis of the ribosome binding site, or through BLAST comparison of each CDS to its epsilonproteobacterial homologs, where present; CDSs with dubious ribosome binding sites, lengths shorter than 200 bp and a low BLASTP score ( $E > 10^{-10}$ ) were removed from the final list of CDSs. tRNAs were annotated using tRNAscan-SE [143,144]. 16S and 23S rRNA loci were identified through BLASTN comparison to *Campylobacter* 16S and 23S loci; 5S rRNA loci were identified through BLASTN comparison to *Helicobacter* 5S loci.

Final annotation was performed using 1) BLASTP comparison to proteins in the NCBI nr database, 2) CDS identification, COG function prediction, cellular location prediction, and identification of Pfam domains and PROSITE motifs using the BASys Bacterial Annotation System [145] and 3) additional Pfam domain identification using the Sanger Centre Pfam search engine [146]. Four levels of gene assignment were used for each CDS: 1) CDSs predicted to encode proteins with high similarity across multiple taxa to proteins of known function (e.g. TrpF, RpsL, PurE) were assigned that function, 2) CDSs predicted to encode proteins with only a general function, but similar to members of various Pfam families or superfamilies, or containing various PROSITE motifs, were annotated to reflect the tentative general function (e.g. topoisomerase, ABC transporter, glycosyl transferase), 3) CDSs predicted to encode proteins with no defined function or motifs but similar to other non-defined proteins ( $E < 10^{-10}$ ) across multiple bacterial taxa were designated as 'conserved hypothetical', and 4) CDSs predicted to encode proteins with no defined function and either unique to this genome or with a low similarity ( $E > 10^{-10}$ ) to proteins from other taxa were designated as 'hypothetical'. The complete nucleotide sequence and annotation of the *A. butzleri* strain RM4018 has been deposited at GenBank under the accession number CP000361.

### Construction of the *A. butzleri* DNA microarray

DNA fragments of individual ORFs within *A. butzleri* strain RM4018 were amplified using primers designed with ArrayDesigner 3.0 (Premier Biosoft, Palo Alto, CA) and purchased from MWG Biotech Inc. (High Point, NC). Each PCR reaction (total reaction volume, 100  $\mu$ l) consisted of  $1\times$  MasterAmp *Taq* PCR buffer,  $1\times$  MasterAmp *Taq* Enhancer, 2.5 mM  $MgCl_2$ , 200  $\mu$ M each dNTP, forward and reverse primers at 0.2  $\mu$ M each, 0.5 U of MasterAmp *Taq* DNA polymerase (Epicentre), and approximately 50 ng of genomic DNA from *A. butzleri* strain RM4018. Thermal cycling was performed using a Tetrad thermal cycler (Bio-Rad, Hercules, CA) with the following amplification parameters: 30 cycles of 25 s at 94°C, 25 s at 52°C, and 2 min at 72°C and a final extension at 72°C for 5 min. PCR products were analyzed by gel electrophoresis in a 1% (w/v) agarose gel in  $1\times$  Tris-borate-EDTA buffer and stained with ethidium bromide. The DNA bands were

examined under UV illumination. We amplified successfully a total of 2238 PCR products from *A. butzleri* strain RM4018. These PCR products were purified on a Qiagen 8000 robot using a Qiaquick 96-well Biorobot kit (Qiagen, Valencia, CA), dried, and resuspended to an average concentration of 0.1–0.2  $\mu\text{g } \mu\text{l}^{-1}$  in 20  $\mu\text{l}$  of 50% dimethyl sulfoxide (DMSO) containing 0.3 $\times$  Saline-Sodium Citrate (SSC). All of the PCR probes were then spotted in duplicate on UltraGAPS slides (Corning) using an OmniGrid Accent (GeneMachines, Ann Arbor, MI) producing a final array that contained a total of 4476 features. The DNA was UV cross-linked to the microarray slides with a Stratalink at 300 mJ (Stratagene, La Jolla, CA).

## Preparation and fluorescent labeling of genomic DNA

Genomic DNA from *A. butzleri* was isolated as described previously [92]. For each microarray hybridization reaction, genomic DNA from the *A. butzleri* reference strain RM4018 and a test strain of *A. butzleri* were labeled fluorescently with indodicarbocyanine (Cy5)-dUTP and indodicarbocyanine (Cy3)-dUTP, respectively. An aliquot (2  $\mu\text{g}$ ) of DNA was mixed with 5  $\mu\text{l}$  10 $\times$  NEBlot labeling buffer containing random sequence octamer oligonucleotides (NEB, Beverly, MA.) and water to a final volume of 41  $\mu\text{l}$ . This mixture was heated to 95°C for 5 min and then cooled for 5 min at 4°C. After this time, the remainder of the labeling reaction components were added: 5  $\mu\text{l}$  of 10 $\times$  dNTP labeling mix (1.2 mM each dATP, dGTP, dCTP; 0.5mM dTTP in 10 mM Tris pH 8.0; 1 mM EDTA), 3  $\mu\text{l}$  of Cy3 dUTP or Cy5 dUTP (GE Biosciences, Piscataway, NJ) and 1  $\mu\text{l}$  of Klenow fragment. The labeling reactions were incubated overnight at 37°C. Labeled DNA was purified from unincorporated label using Qiaquick PCR Cleanup kits. Labeling efficiencies were determined using a NanoDrop ND-1000 (NanoDrop, Wilmington, DE) and efficiently-labeled samples were dried by vacuum.

## Microarray hybridization

Labeled reference and test DNAs were combined in 45  $\mu\text{l}$  Pronto! cDNA hybridization solution (Corning, Corning, NY) and heated to 95°C for 5 min. Then, 15  $\mu\text{l}$  of the hybridization mixture was put onto a microarray slide and sealed with a cover slip. The microarray slide was placed in a hybridization chamber (Corning) and incubated at 42°C for 18 h. Following hybridization, the slides were washed twice in 2 $\times$  SSC, 0.1% sodium dodecyl sulfate at 42°C for 10 min, followed by twice in 1 $\times$  SSC at room temperature for 10 min, and finally twice in 0.2 $\times$  SSC at room temperature for 5 min. The microarray slides were dried by centrifugation at 300 $\times g$  for 10 min before scanning. At least two hybridization reactions were performed for each test strain.

## Microarray data analysis

Microarrays were scanned using an Axon GenePix 4000B microarray laser scanner (Axon Instruments, Inc. Union City, CA) as described previously [93]. Features (spots) and the local background intensities were detected and quantified with GenePix 4.0 software (Axon Instruments, Inc.). The data were filtered so that spots with a reference signal lower than the background plus 2 standard deviations of the background were discarded. Signal intensities were corrected by subtracting the local background, and then the Cy5/Cy3 ratios were calculated. The method for comparative genomic indexing (CGI) analysis utilized the GENCOM software described previously [132,133] to assign present and absent CDS of *A. butzleri* strains when compared to reference strain RM4018. The Cy5 and Cy3 signal intensities were

corrected by subtracting the local background before submission into the GENCOM program. GENCOM analysis was performed for each hybridization and the present/absent assignments for each CDS as determined by GENCOM were converted into scores (present = 1; absent = 0). If the CDS was assigned absent for the reference strain, this indicates that the test strain hybridized better than the reference. Considering all the CDS are from the reference strains, this may indicate that the CDS is in multicopy in the test strain. The CDS assigned absent for the reference strain, were given a score of 2 to indicate that the CDS may be in multicopy. The trinary CDS scores for all strains were analyzed further with GeneSpring microarray analysis software version 7.3 (Agilent Technologies, Redwood City, CA) and subjected to average-linkage hierarchical clustering with the standard correlation and bootstrapping. Microarray data have been deposited in the NCBI GEO repository (<http://www.ncbi.nlm.nih.gov/geo/>) with the accession number GSE8725.

## SUPPORTING INFORMATION

**Figure S1** Diagram of the *A. butzleri* strain RM4018 genome. Genes and features are drawn to scale. Genes are colored according to role category. tRNA and rRNA loci are included. Prophage and genetic islands are shaded in grey.  
Found at: doi:10.1371/journal.pone.0001358.s001 (0.17 MB PDF)

**Figure S2** Selected *A. butzleri* RM4018 signal transduction proteins. Diagrammatic representation of the Non-Che chemotaxis proteins and GGDEF domain diguanylate cyclases/EAL domain c-di-GMP phosphodiesterases predicted to be encoded by strain RM4018. Relevant motifs and transmembrane helices are indicated and drawn to scale.  
Found at: doi:10.1371/journal.pone.0001358.s002 (0.03 MB PDF)

**Table S1** *Arcobacter butzleri* predicted open reading frames and their functional annotation. Complete list of *A. butzleri* strain RM4018 genes, tRNA loci and rRNA loci. Start and end coordinates are provided for each feature along with its functional annotation.  
Found at: doi:10.1371/journal.pone.0001358.s003 (1.43 MB PDF)

**Table S2** Functional prediction of *Arcobacter butzleri* genes: similarity to proteins from other epsilonproteobacteria. *A. butzleri* strain RM4018 genes are sorted by functional category. The % amino acid identity of each protein to homologs in other epsilonproteobacteria is listed.  
Found at: doi:10.1371/journal.pone.0001358.s004 (3.33 MB PDF)

**Table S3** Similarity of *Arcobacter butzleri* proteins to proteins from other taxa. The five best homologs (where applicable) for each protein in strain RM4018 are listed, as well as the Expect (E) value, % identity and % similarity.  
Found at: doi:10.1371/journal.pone.0001358.s005 (4.71 MB PDF)

**Table S4** Antibiotic resistance profiles of strain RM4018 and selected Campylobacters. Comparison of the antibiotic resistance profiles of *A. butzleri* strain RM4018, *C. jejuni* strain NCTC 11168, *C. coli* strain RM2228, *C. lari* strain RM2100 and *C. upsaliensis* strain RM3195.  
Found at: doi:10.1371/journal.pone.0001358.s006 (0.02 MB PDF)

**Table S5** Trinary CGI data as determined with GENCOM. Trinary *A. butzleri* DNA microarray data for the reference strain RM4018 and the 12 additional *A. butzleri* tester strains. Found at: doi:10.1371/journal.pone.0001358.s007 (0.91 MB XLS)

**Table S6** CDS contained in all *A. butzleri* strains examined. Core set of 1676 genes identified in RM4018 and all 12 additional *A. butzleri* strains by comparative genomic indexing. The functional annotation of each gene is provided. Found at: doi:10.1371/journal.pone.0001358.s008 (0.51 MB PDF)

**Table S7** CDS present only in *A. butzleri* strain RM4018. Set of 42 genes identified by comparative genomic indexing of 13 *A. butzleri* strains to be present only in strain RM4018. The functional annotation of each gene is provided.

## REFERENCES

- Skirrow MB, Blaser MJ (2000) Clinical aspects of *Campylobacter* infection. In: Nachamkin I, Blaser MJ, eds (2000) *Campylobacter*. 2nd ed. ed. Washington, DC: ASM Press. pp 69–88.
- Marshall B (2001) One hundred years of discovery and rediscovery of *Helicobacter pylori* and its association with peptic ulcer disease. In: Mobley HL, Mendz GL, Hazell SL, eds (2001) *Helicobacter pylori*: Physiology and genetics. Washington, DC: ASM Press. pp 19–24.
- Sandstedt K, Ursing J, Walder M (1983) Thermotolerant *Campylobacter* with no or weak catalase activity isolated from dogs. *Curr Microbiol* 8: 209–213.
- Fox JG, Cabot EB, Taylor NS, Laraway R (1988) Gastric colonization by *Campylobacter pylori* subsp. *mustelae* in ferrets. *Infect Immun* 56: 2994–2996.
- Vandamme P (2000) Taxonomy of the Family *Campylobacteraceae*. In: Nachamkin I, Blaser MJ, eds (2000) *Campylobacter*. Washington, DC: ASM Press. pp 3–26.
- Kiehlbauch JA, Brenner DJ, Nicholson MA, Baker CN, Patton CM, et al. (1991) *Campylobacter butzleri* sp. nov. isolated from humans and animals with diarrheal illness. *J Clin Microbiol* 29: 376–385.
- Neill SD, Campbell JN, O'Brien JJ, Weatherup STC, Ellis WA (1985) Taxonomic position of *Campylobacter cryaeophila* sp. nov. *Int J Syst Bacteriol* 35: 342–356.
- Vandamme P, Vancanneyt M, Pot B, Mels L, Hoste B, et al. (1992) Polyphasic taxonomic study of the emended genus *Arcobacter* with *Arcobacter butzleri* comb. nov. and *Arcobacter skirrowii* sp. nov., an aerotolerant bacterium isolated from veterinary specimens. *Int J Syst Bacteriol* 42: 344–356.
- McClung CR, Patriquin DG, Davis RE (1983) *Campylobacter nitrofigidis* sp. nov., a nitrogen fixing bacterium associated with roots of *Spartina alterniflora* Loisel. *Int J Syst Bacteriol* 33: 605–612.
- Forsythe SJ (2006) *Arcobacter*. In: Motarjemi Y, Adams M, eds (2006) *Emerging foodborne pathogens*. Cambridge, England: Woodhead Publishing Ltd. pp 181–221.
- Donachie SP, Bowman JP, On SL, Alam M (2005) *Arcobacter halophilus* sp. nov., the first obligate halophile in the genus *Arcobacter*. *Int J Syst Evol Microbiol* 55: 1271–1277.
- Houf K, On SL, Coenye T, Mast J, Van Hoof J, et al. (2005) *Arcobacter cibarius* sp. nov., isolated from broiler carcasses. *Int J Syst Evol Microbiol* 55: 713–717.
- Wirsen CO, Sievert SM, Cavanaugh CM, Molyneux SJ, Ahmad A, et al. (2002) Characterization of an autotrophic sulfide-oxidizing marine *Arcobacter* sp. that produces filamentous sulfur. *Appl Environ Microbiol* 68: 316–325.
- Moussard H, Corre E, Cambon-Bonavita MA, Fouquet Y, Jeanthon C (2006) Novel uncultured Epsilonproteobacteria dominate a filamentous sulphur mat from the 13 degrees N hydrothermal vent field, East Pacific Rise.
- Prabakaran SR, Manorama R, Delille D, Shivaji S (2006) Predominance of *Roseobacter*, *Sulfotobacter*, *Glaciicola* and *Psychrobacter* in seawater collected off Ushuaia, Argentina, Sub-Antarctica. *FEMS Microbiol Ecol*.
- Grabowski A, Nercessian O, Fayolle F, Blanchet D, Jeanthon C (2005) Microbial diversity in production waters of a low-temperature biodegraded oil reservoir. *FEMS Microbiol Ecol* 54: 427–443.
- Frias-Lopez J, Zerkle AL, Bonheyo GT, Fouke BW (2002) Partitioning of bacterial communities between seawater and healthy, black band diseased, and dead coral surfaces. *Appl Environ Microbiol* 68: 2214–2228.
- Thamdrup B, Rossello-Mora R, Amann R (2000) Microbial manganese and sulfate reduction in Black Sea shelf sediments. *Appl Environ Microbiol* 66: 2888–2897.
- Naganuma T, Kato C, Hirayama H, Moriyama N, Hashimoto J, et al. (1996) Intracellular occurrence of epsilonproteobacterial 16S rDNA sequences in the vestimentiferan trophosome. *J Oceanography* 53: 193–197.
- Snaird J, Amann R, Huber I, Ludwig W, Schleifer KH (1997) Phylogenetic analysis and in situ identification of bacteria in activated sludge. *Appl Environ Microbiol* 63: 2884–2896.
- McGarvey JA, Miller WG, Sanchez S, Silva CJ, Whitehand LC (2005) Comparison of bacterial populations and chemical composition of dairy wastewater held in circulated and stagnant lagoons. *J Appl Microbiol* 99: 867–877.
- Assanta MA, Roy D, Lemay MJ, Montpetit D (2002) Attachment of *Arcobacter butzleri*, a new waterborne pathogen, to water distribution pipe surfaces. *J Food Prot* 65: 1240–1247.
- Dhamabutra N, Kamol-Rathanakul P, Pienthaweechai K (1992) Isolation of *Campylobacters* from the canals of Bangkok metropolitan area. *J Med Assoc Thai* 75: 350–364.
- Fera MT, Maugeri TL, Gugliandolo C, Beninati C, Giannone M, et al. (2004) Detection of *Arcobacter* spp. in the coastal environment of the Mediterranean Sea. *Appl Environ Microbiol* 70: 1271–1276.
- Jacob J, Lior H, Feuerpfeil I (1993) Isolation of *Arcobacter butzleri* from a drinking water reservoir in eastern Germany. *Zentralbl Hyg Umweltmed* 193: 557–562.
- Jacob J, Woodward D, Feuerpfeil I, Johnson WM (1998) Isolation of *Arcobacter butzleri* in raw water and drinking water treatment plants in Germany. *Zentralbl Hyg Umweltmed* 201: 189–198.
- Maugeri TL, Gugliandolo C, Carbone M, Caccamo D, Fera MT (2000) Isolation of *Arcobacter* spp. from a brackish environment. *New Microbiol* 23: 143–149.
- Maugeri TL, Irrera GP, Lentini V, Carbone M, Fera MT, et al. (2005) Detection and enumeration of *Arcobacter* spp. in the coastal environment of the Straits of Messina (Italy). *New Microbiol* 28: 177–182.
- Rice EW, Rodgers MR, Wesley IV, Johnson CH, Tanner SA (1999) Isolation of *Arcobacter butzleri* from ground water. *Lett Appl Microbiol* 28: 31–35.
- Stampi S, De Luca G, Varoli O, Zanetti F (1999) Occurrence, removal and seasonal variation of thermophilic campylobacters and *Arcobacter* in sewage sludge. *Zentralbl Hyg Umweltmed* 202: 19–27.
- Mandrell RE, Miller WG (2006) *Campylobacter*. In: Motarjemi Y, Adams M, eds (2006) *Emerging foodborne pathogens*. Cambridge, England: Woodhead Publishing Ltd. pp 476–521.
- Hume ME, Harvey RB, Stanker LH, Droleskey RE, Poole TL, et al. (2001) Genotypic variation among *Arcobacter* isolates from a farrow-to-finish swine facility. *J Food Prot* 64: 645–651.
- Collins CI, Wesley IV, Murano EA (1996) Detection of *Arcobacter* spp. in ground pork by modified plating methods. *J Food Protect* 59: 448–452.
- Rivas L, Fegan N, Vanderlinde P (2004) Isolation and characterisation of *Arcobacter butzleri* from meat. *Int J Food Microbiol* 91: 31–41.
- Atabay HI, Aydin F, Houf K, Sahin M, Vandamme P (2003) The prevalence of *Arcobacter* spp. on chicken carcasses sold in retail markets in Turkey, and identification of the isolates using SDS-PAGE. *Int J Food Microbiol* 81: 21–28.
- Houf K, Devriese LA, De Zutter L, Van Hoof J, Vandamme P (2001) Development of a new protocol for the isolation and quantification of *Arcobacter* species from poultry products. *Int J Food Microbiol* 71: 189–196.
- Manke TR, Wesley IV, Dickson JS, Harmon KM (1998) Prevalence and genetic variability of *Arcobacter* species in mechanically separated turkey. *J Food Prot* 61: 1623–1628.
- Golla SC, Murano EA, Johnson LG, Tipton NC, Cureington EA, et al. (2002) Determination of the occurrence of *Arcobacter butzleri* in beef and dairy cattle from Texas by various isolation methods. *J Food Prot* 65: 1849–1853.
- Wesley IV, Wells SJ, Harmon KM, Green A, Schroeder-Tucker L, et al. (2000) Fecal shedding of *Campylobacter* and *Arcobacter* spp. in dairy cattle. *Appl Environ Microbiol* 66: 1994–2000.
- van Driessche E, Houf K, van Hoof J, De Zutter L, Vandamme P (2003) Isolation of *Arcobacter* species from animal feces. *FEMS Microbiol Lett* 229: 243–248.
- Engberg J, On SL, Harrington CS, Gerner-Smidt P (2000) Prevalence of *Campylobacter*, *Arcobacter*, *Helicobacter*, and *Sutterella* spp. in human fecal

- samples as estimated by a reevaluation of isolation methods for Campylobacters. *J Clin Microbiol* 38: 286–291.
42. Lastovica AJ, Skirrow MB (2000) Clinical significance of *Campylobacter* and related species other than *Campylobacter jejuni* and *C. coli*. In: Nachamkin I, Blaser MJ, eds (2000) *Campylobacter*. Washington, DC: ASM Press. pp 89–120.
  43. Lerner J, Brumberger V, Preac-Mursic V (1994) Severe diarrhea associated with *Arcobacter butzleri*. *Eur J Clin Microbiol Infect Dis* 13: 660–662.
  44. Taylor DN, Kichlbauch JA, Tee W, Pitarangsi C, Echeverria P (1991) Isolation of group 2 aerotolerant *Campylobacter* species from Thai children with diarrhea. *J Infect Dis* 163: 1062–1067.
  45. Vandamme P, Pugina P, Benzi G, Van Etterijck R, Vlaes L, et al. (1992) Outbreak of recurrent abdominal cramps associated with *Arcobacter butzleri* in an Italian school. *J Clin Microbiol* 30: 2335–2337.
  46. Prouzet-Mauleon V, Labadi L, Bouges N, Menard A, Megraud F (2006) *Arcobacter butzleri*: underestimated enteropathogen. *Emerg Infect Dis* 12: 307–309.
  47. Vandenberg O, Dediste A, Houf K, Ibeqwem S, Souayah H, et al. (2004) *Arcobacter* species in humans. *Emerg Infect Dis* 10: 1863–1867.
  48. Lau SK, Woo PC, Teng JL, Leung KW, Yuen KY (2002) Identification by 16S ribosomal RNA gene sequencing of *Arcobacter butzleri* bacteraemia in a patient with acute gangrenous appendicitis. *Mol Pathol* 55: 182–185.
  49. Yan JJ, Ko WC, Huang AH, Chen HM, Jin YT, et al. (2000) *Arcobacter butzleri* bacteremia in a patient with liver cirrhosis. *J Formos Med Assoc* 99: 166–169.
  50. Houf K, Stephan R (2007) Isolation and characterization of the emerging foodborn pathogen *Arcobacter* from human stool. *J Microbiol Methods* 68: 408–413.
  51. Parkhill J, Wren BW, Mungall K, Ketley JM, Churcher C, et al. (2000) The genome sequence of the food-borne pathogen *Campylobacter jejuni* reveals hypervariable sequences. *Nature* 403: 665–668.
  52. Fouts DE, Mongodin EF, Mandrell RE, Miller WG, Rasko DA, et al. (2005) Major structural differences and novel potential virulence mechanisms from the genomes of multiple campylobacter species. *PLoS Biol* 3: e15.
  53. Hofreuter D, Tsai J, Watson RO, Novik V, Altman B, et al. (2006) Unique features of a highly pathogenic *Campylobacter jejuni* strain. *Infect Immun* 74: 4694–4707.
  54. Pearson BM, Gaskin DJ, Segers RP, Wells JM, Nuijten PJ, et al. (2007) The complete genome sequence of *Campylobacter jejuni* strain 81116 (NCTC11828). *J Bacteriol* 189: 8402–8403.
  55. Takai K, Suzuki M, Nakagawa S, Miyazaki M, Suzuki Y, et al. (2006) *Sulfurimonas parvalinellae* sp. nov., a novel mesophilic, hydrogen- and sulfur-oxidizing chemolithoautotroph within the Epsilonproteobacteria isolated from a deep-sea hydrothermal vent polychaete nest, reclassification of *Thiomicrospira denitrificans* as *Sulfurimonas denitrificans* comb. nov. and emended description of the genus *Sulfurimonas*. *Int J Syst Evol Microbiol* 56: 1725–1733.
  56. Baar C, Eppinger M, Raddatz G, Simon J, Lanz C, et al. (2003) Complete genome sequence and analysis of *Wolnella succinogenes*. *Proc Natl Acad Sci U S A* 100: 11690–11695.
  57. Suerbaum S, Josenhans C, Sterzenbach T, Drescher B, Brandt P, et al. (2003) The complete genome sequence of the carcinogenic bacterium *Helicobacter hepaticus*. *Proc Natl Acad Sci U S A* 100: 7901–7906.
  58. Tomb JF, White O, Kerlavage AR, Clayton RA, Sutton GG, et al. (1997) The complete genome sequence of the gastric pathogen *Helicobacter pylori*. *Nature* 388: 539–547.
  59. Alm RA, Ling LS, Moir DT, King BL, Brown ED, et al. (1999) Genomic-sequence comparison of two unrelated isolates of the human gastric pathogen *Helicobacter pylori*. *Nature* 397: 176–180.
  60. Oh JD, Kling-Backhed H, Giannakis M, Xu J, Fulton RS, et al. (2006) The complete genome sequence of a chronic atrophic gastritis *Helicobacter pylori* strain: evolution during disease progression. *Proc Natl Acad Sci U S A* 103: 9999–10004.
  61. Eppinger M, Baar C, Linz B, Raddatz G, Lanz C, et al. (2006) Who ate whom? Adaptive *Helicobacter* genomic changes that accompanied a host jump from early humans to large felines. *PLoS Genet* 2: e120.
  62. Nakagawa S, Takaki Y, Shimamura S, Reysenbach AL, Takai K, et al. (2007) Deep-sea vent epsilon-proteobacterial genomes provide insights into emergence of pathogens. *Proc Natl Acad Sci U S A* 104: 12146–12150.
  63. Iyer RR, Pluciennik A, Burdett V, Modrich PL (2006) DNA mismatch repair: functions and mechanisms. *Chem Rev* 106: 302–323.
  64. Malik HS, Henikoff S (2000) Dual recognition-incision enzymes might be involved in mismatch repair and meiosis. *Trends Biochem Sci* 25: 414–418.
  65. Viswanathan M, Lovett ST (1998) Single-strand DNA-specific exonucleases in *Escherichia coli*. Roles in repair and mutation avoidance. *Genetics* 149: 7–16.
  66. Viswanathan M, Lovett ST (1999) Exonuclease X of *Escherichia coli*. A novel 3'-5' DNase and DnaQ superfamily member involved in DNA repair. *J Biol Chem* 274: 30094–30100.
  67. Kredich NM (1996) Biosynthesis of cysteine. In: Neidhardt FC, ed (1996) *Escherichia coli* and *Salmonella*: Cellular and molecular biology. 2nd ed. ed. Washington DC: ASM Press. pp 514–527.
  68. Neuwald AF, Krishnan BR, Brikun I, Kulakauskas S, Suziedelis K, et al. (1992) *cysQ* a gene needed for cysteine synthesis in *Escherichia coli* K-12 only during aerobic growth. *J Bacteriol* 174: 415–425.
  69. Valdes J, Veloso F, Jedlicki E, Holmes D (2003) Metabolic reconstruction of sulfur assimilation in the extremophile *Acidithiobacillus ferrooxidans* based on genome analysis. *BMC Genomics* 4: 51.
  70. Bick JA, Dennis JJ, Zylstra GJ, Nowack J, Leustek T (2000) Identification of a new class of 5'-adenylylsulfate (APS) reductases from sulfate-assimilating bacteria. *J Bacteriol* 182: 135–142.
  71. Williams SJ, Senaratne RH, Mougous JD, Riley LW, Bertozzi CR (2002) 5'-adenosinephosphosulfate lies at a metabolic branch point in mycobacteria. *J Biol Chem* 277: 32606–32615.
  72. Friedrich CG, Quantmeier A, Bardischewsky F, Rother D, Kraft R, et al. (2000) Novel genes coding for lithotrophic sulfur oxidation of *Paracoccus pantotrophus* GB17. *J Bacteriol* 182: 4677–4687.
  73. Wodara C, Bardischewsky F, Friedrich CG (1997) Cloning and characterization of sulfite dehydrogenase, two c-type cytochromes, and a flavoprotein of *Paracoccus denitrificans* GB17: essential role of sulfite dehydrogenase in lithotrophic sulfur oxidation. *J Bacteriol* 179: 5014–5023.
  74. Eisen JA, Nelson KE, Paulsen IT, Heidelberg JF, Wu M, et al. (2002) The complete genome sequence of *Chlorobium tepidum* TLS, a photosynthetic, anaerobic, green-sulfur bacterium. *Proc Natl Acad Sci U S A* 99: 9509–9514.
  75. Appia-Ayme C, Little PJ, Matsumoto Y, Leech AP, Berks BC (2001) Cytochrome complex essential for photosynthetic oxidation of both thiosulfate and sulfide in *Rhodoculum sulfidophilum*. *J Bacteriol* 183: 6107–6118.
  76. Friedrich CG, Rother D, Bardischewsky F, Quantmeier A, Fischer J (2001) Oxidation of reduced inorganic sulfur compounds by bacteria: emergence of a common mechanism? *Appl Environ Microbiol* 67: 2873–2882.
  77. Appia-Ayme C, Berks BC (2002) SoxV, an orthologue of the CcdA disulfide transporter, is involved in thiosulfate oxidation in *Rhodoculum sulfidophilum* and reduces the periplasmic thioredoxin SoxW. *Biochem Biophys Res Commun* 296: 737–741.
  78. Bagchi A, Ghosh TC (2006) Structural insight into the interactions of SoxV, SoxW and SoxS in the process of transport of reductants during sulfur oxidation by the novel global sulfur oxidation reaction cycle. *Biophys Chem* 119: 7–13.
  79. Grimek TL, Escalante-Semerena JC (2004) The *acdD* genes of *Shewanella oneidensis* and *Vibrio cholerae* encode a new Fe/S-dependent 2-methylcitrate dehydratase enzyme that requires *prpF* function in vivo. *J Bacteriol* 186: 454–462.
  80. Friedmann S, Steindorf A, Alber BE, Fuchs G (2006) Properties of succinyl-coenzyme A:L-malate coenzyme A transferase and its role in the autotrophic 3-hydroxypropionate cycle of *Chloroflexus aurantiacus*. *J Bacteriol* 188: 2646–2655.
  81. Hallam SJ, Mincer TJ, Schleper C, Preston CM, Roberts K, et al. (2006) Pathways of carbon assimilation and ammonia oxidation suggested by environmental genomic analyses of marine Crenarchaeota. *PLoS Biol* 4: e95.
  82. Berg IA, Filatova LV, Ivanovsky RN (2002) Inhibition of acetate and propionate assimilation by itaconate via propionyl-CoA carboxylase in isocitrate lyase-negative purple bacterium *Rhodospirillum rubrum*. *FEMS Microbiol Lett* 216: 49–54.
  83. Ivanovsky RN, Krasilnikova EN, Berg IA (1997) A proposed citramalate cycle for acetate assimilation in the purple non-sulfur bacterium *Rhodospirillum rubrum*. *FEMS Microbiol Lett* 153: 399–404.
  84. Takagi K, Torimura M, Kawaguchi K, Kano K, Ikeda T (1999) Biochemical and electrochemical characterization of quinohemoprotein amine dehydrogenase from *Paracoccus denitrificans*. *Biochemistry* 38: 6935–6942.
  85. Vandenberghe I, Kim JK, Devreese B, Hacısalihoglu A, Iwabuki H, et al. (2001) The covalent structure of the small subunit from *Pseudomonas putida* amine dehydrogenase reveals the presence of three novel types of internal cross-linkages, all involving cysteine in a thioether bond. *J Biol Chem* 276: 42923–42931.
  86. Takagi K, Yamamoto K, Kano K, Ikeda T (2001) New pathway of amine oxidation respiratory chain of *Paracoccus denitrificans* IFO 12442. *Eur J Biochem* 268: 470–476.
  87. Adachi O, Yoshihara N, Tanasupawat S, Toyama H, Matsushita K (2003) Purification and characterization of membrane-bound quinoprotein quinate dehydrogenase. *Biosci Biotechnol Biochem* 67: 2115–2123.
  88. Ono K, Okajima T, Tani M, Kuroda S, Sun D, et al. (2006) Involvement of a putative [Fe-S]-cluster-binding protein in the biogenesis of quinohemoprotein amine dehydrogenase. *J Biol Chem* 281: 13672–13684.
  89. Hacısalihoglu A, Jongejans JA, Duine JA (1997) Distribution of amine oxidases and amine dehydrogenases in bacteria grown on primary amines and characterization of the amine oxidase from *Klebsiella oxytoca*. *Microbiology* 143 ( Pt 2): 505–512.
  90. Goldie J, Veldhuyzen van Zanten SJ, Jalali S, Hollingsworth J, Riddell RH, et al. (1989) Optimization of a medium for the rapid urease test for detection of *Campylobacter pylori* in gastric antral biopsies. *J Clin Microbiol* 27: 2080–2082.
  91. Gilbert M, Godschalk P, Parker CT, Endtz H, Wakarchuk W (2005) Genetic basis for the variation in the lipooligosaccharide outer core of *Campylobacter jejuni* and possible association of glycosyltransferase genes with post-infectious neuropathies. In: Ketley JM, Konkel ME, eds (2005) *Campylobacter*: Molecular and Cellular Biology. Norfolk, UK: Horizon Bioscience. pp 219–248.
  92. Parker CT, Horn ST, Gilbert M, Miller WG, Woodward DL, et al. (2005) Comparison of *Campylobacter jejuni* lipooligosaccharide biosynthesis loci from a variety of sources. *J Clin Microbiol* 43: 2771–2781.
  93. Parker CT, Quinones B, Miller WG, Horn ST, Mandrell RE (2006) Comparative genomic analysis of *Campylobacter jejuni* strains reveals diversity due to genomic elements similar to those present in *C. jejuni* strain RM1221. *J Clin Microbiol* 44: 4125–4135.

94. Kaniuk NA, Monteiro MA, Parker CT, Whitfield C (2002) Molecular diversity of the genetic loci responsible for lipopolysaccharide core oligosaccharide assembly within the genus *Salmonella*. *Mol Microbiol* 46: 1305–1318.
95. Burrows LL, Charter DF, Lam JS (1996) Molecular characterization of the *Pseudomonas aeruginosa* serotype O5 (PAO1) B-band lipopolysaccharide gene cluster. *Mol Microbiol* 22: 481–495.
96. Morgan GJ, Hatfull GF, Casjens S, Hendrix RW (2002) Bacteriophage Mu genome sequence: analysis and comparison with Mu-like prophages in *Haemophilus*, *Neisseria* and *Deinococcus*. *J Mol Biol* 317: 337–359.
97. Summer EJ, Gonzalez CF, Carlisle T, Mebane LM, Cass AM, et al. (2004) *Burkholderia cenocepacia* phage BcepMu and a family of Mu-like phages encoding potential pathogenesis factors. *J Mol Biol* 340: 49–65.
98. Atabay HI, Aydin F (2001) Susceptibility of *Arcobacter butzleri* isolates to 23 antimicrobial agents. *Lett Appl Microbiol* 33: 430–433.
99. Bayles KW (2000) The bactericidal action of penicillin: new clues to an unsolved mystery. *Trends Microbiol* 8: 274–278.
100. Groicher KH, Firek BA, Fujimoto DF, Bayles KW (2000) The *Staphylococcus aureus* *hlgAB* operon modulates murein hydrolase activity and penicillin tolerance. *J Bacteriol* 182: 1794–1801.
101. Glaab WE, Mitchell LS, Miller JE, Vlasakova K, Skopek TR (2005) 5-fluorouracil forward mutation assay in *Salmonella*: determination of mutational target and spontaneous mutational spectra. *Mutat Res* 578: 238–246.
102. Martinussen J, Hammer K (1994) Cloning and characterization of *uppB*, a gene encoding uracil phosphoribosyltransferase from *Lactococcus lactis*. *J Bacteriol* 176: 6457–6463.
103. Taylor DE, Tracz DM (2005) Mechanisms of antimicrobial resistance in *Campylobacter*. In: Ketley JM, Konkel ME, eds (2005) *Campylobacter*: Molecular and Cellular Biology. Norfolk, UK: Horizon Bioscience. pp 193–204.
104. Abdelbaqi K, Menard A, Prouzet-Mauleon V, Bringaud F, Lehours P, et al. (2007) Nucleotide sequence of the *gyrA* gene of *Arcobacter* species and characterization of human ciprofloxacin-resistant clinical isolates. *FEMS Immunol Med Microbiol* 49: 337–345.
105. Ruiz J (2003) Mechanisms of resistance to quinolones: target alterations, decreased accumulation and DNA gyrase protection. *J Antimicrob Chemother* 51: 1109–1117.
106. Yoshida H, Bogaki M, Nakamura S, Ubukata K, Konno M (1990) Nucleotide sequence and characterization of the *Staphylococcus aureus* *norA* gene, which confers resistance to quinolones. *J Bacteriol* 172: 6942–6949.
107. Pickett CL, Lee RB (2005) Cytolethal distending toxin. In: Ketley JM, Konkel ME, eds (2005) *Campylobacter*: Molecular and Cellular Biology. Norfolk, UK: Horizon Bioscience. pp 385–396.
108. Johnson LG, Murano EA (2002) Lack of a cytolethal distending toxin among *Arcobacter* isolates from various sources. *J Food Prot* 65: 1789–1795.
109. Jin S, Joe A, Lynett J, Hani EK, Sherman P, et al. (2001) JlpA, a novel surface-exposed lipoprotein specific to *Campylobacter jejuni*, mediates adherence to host epithelial cells. *Mol Microbiol* 39: 1225–1236.
110. Pei Z, Blaser MJ (1993) PEB1, the major cell-binding factor of *Campylobacter jejuni*, is a homolog of the binding component in gram-negative nutrient transport systems. *J Biol Chem* 268: 18717–18725.
111. Goldberg MB, DiRita VJ, Calderwood SB (1990) Identification of an iron-regulated virulence determinant in *Vibrio cholerae*, using *TnphoA* mutagenesis. *Infect Immun* 58: 55–60.
112. Johnson JR, Jelacic S, Schoening LM, Clabots C, Shaikh N, et al. (2005) The IrgA homologue adhesin Iha is an *Escherichia coli* virulence factor in murine urinary tract infection. *Infect Immun* 73: 965–971.
113. Larsen NA, Lin H, Wei R, Fischbach MA, Walsh CT (2006) Structural characterization of enterobactin hydrolase IroE. *Biochemistry* 45: 10184–10190.
114. Rojas CM, Ham JH, Deng WL, Doyle JJ, Collmer A (2002) HecA, a member of a class of adhesins produced by diverse pathogenic bacteria, contributes to the attachment, aggregation, epidermal cell killing, and virulence phenotypes of *Ercinia chrysanthemi* EC16 on *Nicotiana glauca* seedlings. *Proc Natl Acad Sci U S A* 99: 13142–13147.
115. Korolik V, Ketley JM (2006) *Campylobacter* chemotaxis. In: Ketley JM, Konkel ME, eds (2006) *Campylobacter*: Molecular and Cellular Biology. Norfolk, UK: Horizon Bioscience. pp 349–367.
116. Szurmant H, Ordal GW (2004) Diversity in chemotaxis mechanisms among the bacteria and archaea. *Microbiol Mol Biol Rev* 68: 301–319.
117. Tsang N, Macnab R, Koshland DE Jr (1973) Common mechanism for repellents and attractants in bacterial chemotaxis. *Science* 181: 60–63.
118. Anantharaman V, Aravind L (2000) Cache - a signaling domain common to animal Ca(2+)-channel subunits and a class of prokaryotic chemotaxis receptors. *Trends Biochem Sci* 25: 535–537.
119. Niehus E, Gressmann H, Ye F, Schlapbach R, Dehio M, et al. (2004) Genome-wide analysis of transcriptional hierarchy and feedback regulation in the flagellar system of *Helicobacter pylori*. *Mol Microbiol* 52: 947–961.
120. Wosten MM, Wagenaar JA, van Putten JP (2004) The FlgS/FlgR two-component signal transduction system regulates the *fla* regulon in *Campylobacter jejuni*. *J Biol Chem* 279: 16214–16222.
121. Chilcott GS, Hughes KT (2000) Coupling of flagellar gene expression to flagellar assembly in *Salmonella enterica* serovar typhimurium and *Escherichia coli*. *Microbiol Mol Biol Rev* 64: 694–708.
122. Sourjik V, Muschler P, Scharf B, Schmitt R (2000) VisN and VisR are global regulators of chemotaxis, flagellar, and motility genes in *Sinorhizobium (Rhizobium) meliloti*. *J Bacteriol* 182: 782–788.
123. Hughes KT, Mathee K (1998) The anti-sigma factors. *Annu Rev Microbiol* 52: 231–286.
124. Kazmierczak MJ, Wiedmann M, Boor KJ (2005) Alternative sigma factors and their roles in bacterial virulence. *Microbiol Mol Biol Rev* 69: 527–543.
125. Romling U, Amikam D (2006) Cyclic di-GMP as a second messenger. *Curr Opin Microbiol* 9: 218–228.
126. Romling U, Gomelsky M, Galperin MY (2005) C-di-GMP: the dawning of a novel bacterial signalling system. *Mol Microbiol* 57: 629–639.
127. Gjermansen M, Ragas P, Tolker-Nielsen T (2006) Proteins with GGDEF and EAL domains regulate *Pseudomonas putida* biofilm formation and dispersal. *FEMS Microbiol Lett*.
128. Da Re S, Ghigo JM (2006) A CsgD-independent pathway for cellulose production and biofilm formation in *Escherichia coli*. *J Bacteriol* 188: 3073–3087.
129. Kazmierczak BI, Lebron MB, Murray TS (2006) Analysis of FimX, a phosphodiesterase that governs twitching motility in *Pseudomonas aeruginosa*. *Mol Microbiol* 60: 1026–1043.
130. O'Shea T M, Klein AH, Geszvain K, Wolfe AJ, Visick KL (2006) Diguanylate cyclases control magnesium-dependent motility of *Vibrio fischeri*. *J Bacteriol*.
131. Tischler AD, Camilli A (2005) Cyclic diguanylate regulates *Vibrio cholerae* virulence gene expression. *Infect Immun* 73: 5873–5882.
132. Pearson BM, Pin C, Wright J, l'Anson K, Humphrey T, et al. (2003) Comparative genome analysis of *Campylobacter jejuni* using whole genome DNA microarrays. *FEBS Lett* 554: 224–230.
133. Pin C, Reuter M, Pearson B, Friis L, Overweg K, et al. (2006) Comparison of different approaches for comparative genetic analysis using microarray hybridization. *Appl Microbiol Biotechnol* 72: 852–859.
134. Houf K, Tutenel A, De Zutter L, Van Hoof J, Vandamme P (2000) Development of a multiplex PCR assay for the simultaneous detection and identification of *Arcobacter butzleri*, *Arcobacter cryaerophilus* and *Arcobacter skirrovii*. *FEMS Microbiol Lett* 193: 89–94.
135. Miller WG, Pearson BM, Wells JM, Parker CT, Kapitonov VV, et al. (2005) Diversity within the *Campylobacter jejuni* type I restriction-modification loci. *Microbiology* 151: 337–351.
136. DeAngelis MM, Wang DG, Hawkins TL (1995) Solid-phase reversible immobilization for the isolation of PCR products. *Nucleic Acids Res* 23: 4742–4743.
137. Hawkins TL, McKernan KJ, Jacotot LB, MacKenzie JB, Richardson PM, et al. (1997) A magnetic attraction to high-throughput genomics. *Science* 276: 1887–1889.
138. Ewing B, Green P (1998) Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res* 8: 186–194.
139. Ewing B, Hillier L, Wendl MC, Green P (1998) Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res* 8: 175–185.
140. Stoeva K, Bruce Ward F (2006) Genome mapping of *Arcobacter butzleri*. *FEMS Microbiol Lett* 256: 290–297.
141. Artemis (release 8.0). [<http://www.sanger.ac.uk/Software/Artemis/>].
142. Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, et al. (2000) Artemis: sequence visualization and annotation. *Bioinformatics* 16: 944–945.
143. tRNAscan-SE (v. 1.23). [<http://bioweb.pasteur.fr/seqanal/interfaces/trnscan.html>].
144. Lowe TM, Eddy SR (1997) tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res* 25: 955–964.
145. BASys Bacterial Annotation System. [<http://wishart.biology.ualberta.ca/basys/cgi/submit.pl>].
146. Sanger Centre Pfam search engine. [<http://www.sanger.ac.uk/Software/Pfam/>].
147. Zoomable *Arcobacter butzleri* strain RM4018 BLAST atlas. [[http://www.cbs.dtu.dk/services/GenomeAtlas/suppl/zoomatlas/?zpid=Arcobacter\\_butzleri](http://www.cbs.dtu.dk/services/GenomeAtlas/suppl/zoomatlas/?zpid=Arcobacter_butzleri)].