Structural conservation of chemotaxis machinery across Archaea and Bacteria

Ariane Briegel[1], Davi R. Ortega[1], Audrey Huang[1], Catherine M. Oikonomou[1], Robert P. Gunsalus[2] and Grant J. Jensen[1,3*]

[1] California Institute of Technology, 1200 E. California Blvd., Pasadena, CA 91125

[2]University of California Los Angeles, 609 Charles E. Young Dr. S., Los Angeles, CA 90095

[3] Howard Hughes Medical Institute, 1200 E. California Blvd., Pasadena, CA 91125

[*]Correspondence: Tel (626) 395-8827, Fax (626) 395-5730, Email jensen@caltech.edu

Running title: Structure of archaeal chemoreceptor arrays

Summary

Chemotaxis allows cells to sense and respond to their environment. In bacteria, stimuli are detected by arrays of chemoreceptors that relay the signal to a two-component regulatory system. These arrays take the form of highly stereotyped super-lattices comprising hexagonally packed trimers-of-receptor-dimers networked by rings of histidine kinase and coupling proteins. This structure is conserved across chemotactic bacteria, and between membrane-bound and cytoplasmic arrays, and gives rise to the highly cooperative, dynamic nature of the signaling system. The chemotaxis system, absent in eukaryotes, is also found in archaea, where its structural details remain uncharacterized. Here we provide evidence that the chemotaxis machinery was not present in the last archaeal common ancestor, but rather was introduced in one of the waves of lateral gene transfer that occurred after the branching of Eukaryota but before the diversification of Euryarchaeota. Unlike in Bacteria, the chemotaxis system then evolved largely vertically in Archaea, with very few subsequent successful lateral gene transfer events. By electron cryotomography (ECT), we find that the structure of both membrane-bound and cytoplasmic chemoreceptor arrays is conserved between Bacteria and Archaea, suggesting the fundamental importance of this signaling architecture across diverse prokaryotic lifestyles.

Introduction

Single-celled organisms rely on signal transduction pathways to sense and respond to their environments. In Bacteria, one such pathway, the chemotaxis system, relays information on the chemical environment to the flagellar motor to bias swimming

direction. Roughly half of all bacteria are chemotactic (Wuichet and Zhulin, 2010). The chemotaxis system, typified by *Escherichia coli*, consists of chemoreceptors (methyl-accepting chemotaxis proteins, or MCPs) networked into cooperative arrays by coupling protein (CheW) and a two-component signaling kinase (CheA). Adaptation to stimulus is mediated by methylation and de-methylation of the receptors, performed by a methyltransferase (CheR) and methylesterase (CheB). CheB is one of the two response regulators controlled by CheA; the other is CheY, which, when phosphorylated, can bind to the flagellar motor, inducing a shift in its direction of rotation, and therefore the swimming behavior of the cell, inducing a "tumble" rather than a "run" (Hazelbauer et al., 2008).

In Bacteria, the structure of the chemosensory array is universal across all species imaged to date (Briegel et al., 2009). It is even conserved between membrane-bound and cytoplasmic arrays (Briegel et al., 2014).

Many archaea also contain chemotaxis genes. Decades of work have shown that the halophilic archaeon *Halobacterium salinarum* senses chemical attractants such as oxygen, as well as light, through a chemosensory system that, as in bacteria, translates into a change in direction of flagellar rotation. Interestingly, the archaeal flagellum, or archaellum, is not homologous to that of bacteria (Jarrell and Albers, 2012). Archaeal chemotaxis systems use an additional protein, CheF, to translate signal from CheY to the motor (Schlesner et al., 2009).

Evolutionary genomics has revealed that the majority of archaeal chemotaxis genes are found in Euryarchaeota and exhibit high sequence similarity to those found in Firmicutes and Thermotogales, suggesting that the system was acquired via lateral gene

66 transfer (LGT) (Wuichet and Zhulin, 2010). However, recently sequenced genomes
67 (Blainey et al., 2011; Spang et al., 2012) from the deeply branching Thaumarchaeota
68 (Brochier-Armanet et al., 2008) also contain chemotaxis systems, raising the possibility
69 of a chemotactic Last Archaeal Common Ancestor (LACA).
70 Here, we apply evolutionary genomics to distinguish between these hypotheses,
71 and electron cryotomography (ECT) to investigate the structure of archaeal chemotaxis
72 systems.
73
74 **Results and Discussion**
75 To identify chemotactic archaea, we selected 240 genomes spanning all major
76 archaeal branches: Euryarchaeota (178), Crenarchaeota (51), Thaumarchaeota (9),
77 Korarchaeota (1), and Nanoarchaeota (1). Using the MiST2 database (Ulrich and Zhulin,
78 2010), we determined that half of these genomes contained chemotaxis systems, defined
79 as the presence of at least one CheA and one MCP. Chemotactic species all belonged to
80 one of two archaeal phyla – Euryarchaeota (118) or Thaumarchaeota (2).
81 Chemotaxis systems are commonly classified on the basis of their signaling
82 kinase, CheA (Wuichet and Zhulin, 2010). In bacteria, there are more than a dozen such
83 classes: multiple systems that signal to the flagellum (denoted F), another that signals to
84 the Type IV pilus (Tfp), and one involved in alternate cellular functions (ACF). To
85 classify the archaeal chemotaxis systems identified above, we examined the protein
86 sequences of their associated CheA. There are 139 such CheA genes in the pangenome,
87 representing only two classes: F1 (134) and ACF (5), consistent with previous work
88 (Wuichet and Zhulin, 2010). ACF systems are found only in the Methanomicrobiales

order of Archaea, and, based on sequence similarity (Supplemental Table 1), likely arose from LGT from Deltaproteobacteria.

In contrast to ACF systems, F1 systems are widespread across Euryarchaeota, supporting the idea that this system was present in their common ancestor. F1 systems are also found in Thaumarchaeota, a deeply branched phylum, raising the question of whether the LACA was chemotactic. To investigate this question, we used a concatenated alignment of CheA, CheB, and CheR to construct a phylogenetic tree of F1 systems in Bacteria and Archaea. In many genomes, there are multiple copies of these genes, making it difficult to unambiguously assign them to shared classes. Therefore, we chose only species that either contained all three genes in the same cluster or contained only a single copy of each gene of the F1 system. The final dataset contains 203 CheA:CheB:CheR concatenated sequences from 193 organisms: Euryarchaeota (82), Thaumarcheota (2), Firmicutes (93), Thermotogae (3), Synergistetes (3), Cyanobacteria (3), Chloroflexi (1), Planctomycetes (1), Actinobacteria (1), Chlorobi (2) and Nitrospirae (2). The resulting tree is shown in Supplemental Figure 1. We find that the F1 chemotaxis system in Archaea is monophyletic. In general, the topology of our F1 chemotaxis tree recapitulates that of the organismal tree, with two notable exceptions, highlighted in Figure 1.

First, our results suggest that the F1 chemotaxis system was laterally transferred from the Euryarchaeota to the Thaumarchaeota sometime between the branching of the Methanococcales and the Archaeoglobales (Supplemental Figure 2), which rejects the idea of a chemotactic LACA. Rather, our data suggest that the ancestral archaeal chemotaxis system was laterally transferred from an ancestor of Bacteria before the

112 divergence of the earliest branching bacterial lineages of Firmicutes, Thermotogales and
113 Synergistetes

114 Our data also suggest more recent LGT events. We find that the chemotaxis
115 system of Methanosarcinales appears to be misplaced in the CheA:CheB:CheR tree
116 (Supplemental Figure 2). Further inspection shows that most Methanosarcinales have a
117 version of the F1 chemotaxis system more similar to that of Clostridia than to any other
118 archaeal clade, consistent with a relatively recent LGT, as has been suggested
119 (Deppenmeier et al., 2002). Some Methanosarcinales (e.g. *Methanosarcina mazei*,
120 *Methanosarcina acetivorans*) appear to have the vertically inherited system, while others
121 (e.g. *Methanococcoides burtonii*) have the laterally transferred system.

122 Given the bacterial source of archaeal chemotaxis systems, we wanted to see
123 whether their structure was conserved. Electron microscopy of Archaea has lagged
124 behind that of Bacteria. Many species require complex media, extremes of temperature,
125 or anaerobic conditions for growth. Halobacteria are relatively easy to culture but require
126 high salt concentrations that interfere with ECT, and cells often lose structural integrity
127 upon removal of salt (Trachtenberg et al., 2000). Additionally, many archaeal cells are
128 too large to allow adequate transmission of electrons. With these constraints in mind, we
129 selected four archaeal species, representing three diverse orders of Euryarchaeota:
130 Thermococcales (*Thermococcus kodakarensis*), Methanomicrobiales (*Methanoregula*
131 *formicica, Methanosprillum hungatei*), and Halobacteriales (*Halobacterium salinarum*).

132 By ECT, we observed membrane-bound chemoreceptor arrays in *T. kodakarensis*,
133 *M. hungatei*, and *H. salinarum* (Figure 2). The order and packing of the chemoreceptors
134 was identical to that of bacteria, with 12 nm between the centers of adjacent hexamers of

trimers-of-MCP-dimers. Interestingly, chemoreceptor arrays may have been observed previously in EM images of Archaea, but not identified as such. For instance, in *M. hungatei*, what are identified as cytoskeletal structures are likely chemoreceptor arrays (Toso et al., 2011). Similarly, work in *M. hungatei* and *H. salinarum* identified a "polar organelle," which is likely the chemoreceptor array (Cruden et al., 1989; Metlina, 2004). The "polar organelle" was first described in *(Aqua)Spirillum serpens* in 1962 (Murray and Birch-Andersen, 1963). That study preceded Julius Adler's seminal 1966 paper on bacterial chemotaxis, which stimulated much work in the field (Adler, 1966). However, the "polar organelle" is described as a structure distinct from the chemoreceptor array as recently as 2001 (Lybarger and Maddock, 2001). Our results, however, lead us to believe that they are, in fact, the same structure (Supplemental Figure 3).

We also observe cytoplasmic chemoreceptor arrays in *M. formicicum* (Figure 3). Again, their structure is strikingly similar to that of bacterial cytoplasmic arrays: two hexagonally-packed lattices of trimers-of-MCP-dimers, presumably interacting at their ligand-binding tips, sandwiched between two CheA/CheW baseplates. As observed in the bacterium *Rhodobacter sphaeroides*, these arrays frequently curve, exhibiting both positive and negative curvature in opposing halves of the array (Briegel et al., 2014). The data from all four species are summarized in Supplemental Table 2.

Both cytoplasmic and membrane-bound arrays exhibited more variation in subcellular localization than their bacterial counterparts. Rather than clustering tightly at the cell pole, as in most chemotactic bacteria, archaeal arrays were frequently observed nearer to mid-cell.

The F1 chemotaxis system was likely transferred from the last bacterial common ancestor in one of the waves of LGT that gave rise to the diversification of Euryarchaeota more than 3.5 billion years ago (Battistuzzi et al., 2004; Nelson-Sathi et al., 2014). It has been suggested that these influxes of genes allowed metabolic divergence, and it may be that chemotaxis helped these diversifying organisms find and colonize new niches. That the transfer occurred after the branching of Eukaryotes from Archaea may explain why chemotaxis systems are not found in Eukaryota (Koretke et al., 2000). It is currently believed that there has been much less LGT to eukaryotes than between prokaryotes, and most of these events happened relatively recently(Keeling and Palmer, 2008). In fact, of the dozen or so subfamilies of two-component signaling systems, only one is found in Eukaryota(Wolanin et al., 2002). Bacteria evolved several classes of chemotaxis systems that were widely exchanged via LGT. In Archaea, by contrast, evolution was largely vertical, with successful LGT of the F1 chemotaxis system occurring maybe as few as three times. It is striking that this signaling architecture, developed in Bacteria, should remain unchanged over billions of years in Archaea adapting to new lifestyles.

Acknowledgements We thank Dr. Christine Moissl-Eichinger for sharing *Halobacterium salinarum* (DSM 3754). We thank Drs. Matthias Koch and Kristin Wuichet for discussions. This work was funded by the Department of Energy Biosciences Division award DE-FG02-O8ER64689 and the UCLA-DOE Institute of Genomics and Proteomics award DE-FC03-02ER6342 to R.P.G. and NIGMS award GM101425 to G.J.J.

Figure Legends

Figure 1. Mapping evolutionary events of the archaeal chemotaxis system onto the tree of life. Tree of life scheme is as in (Brochier-Armanet et al., 2011). Presence (thick lines) and absence (thin lines) of the F1 chemotaxis system are marked by black and white nodes, respectively. Loss events are marked by a bar across the branch. Dashed arrows denote LGT events of the indicated systems. The Euryarchaeota are shaded in grey. LUCA: Last Universal Common Ancestor.

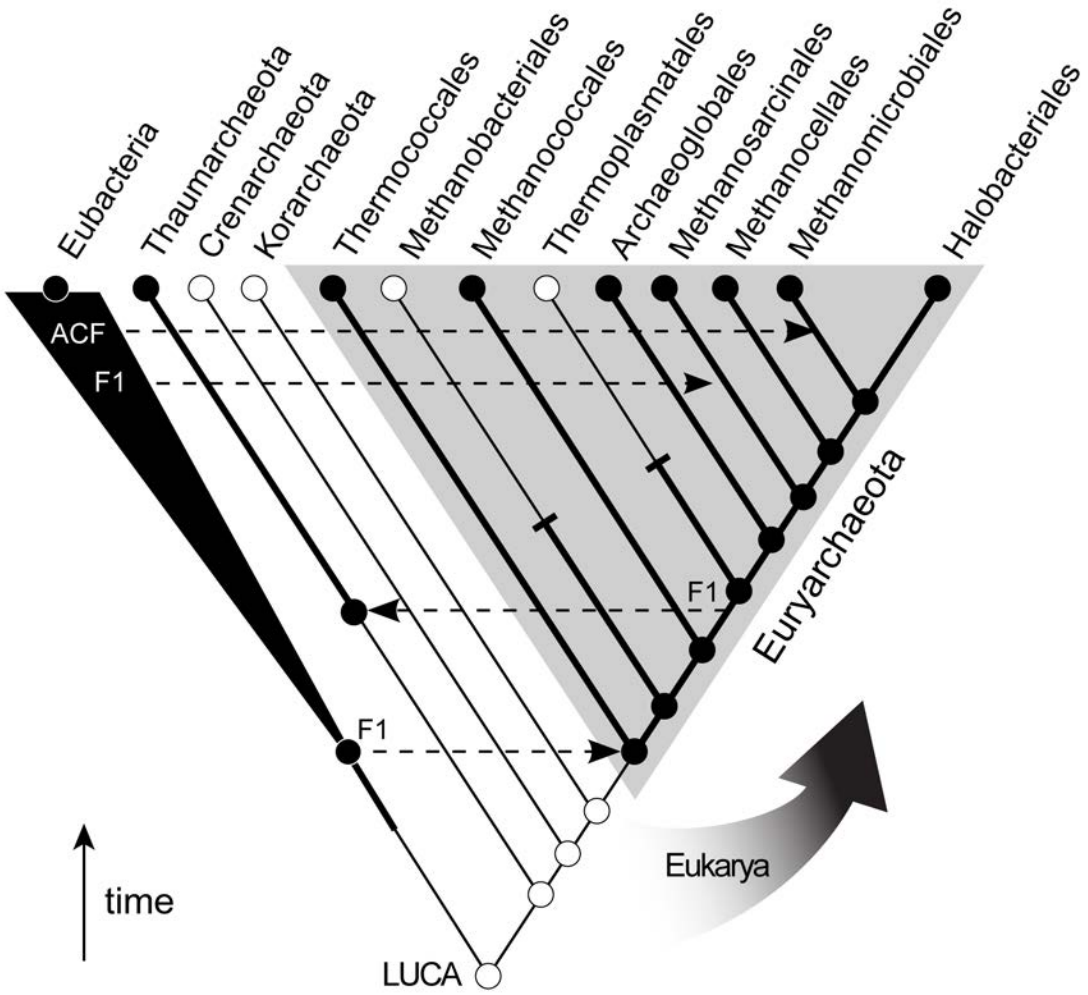
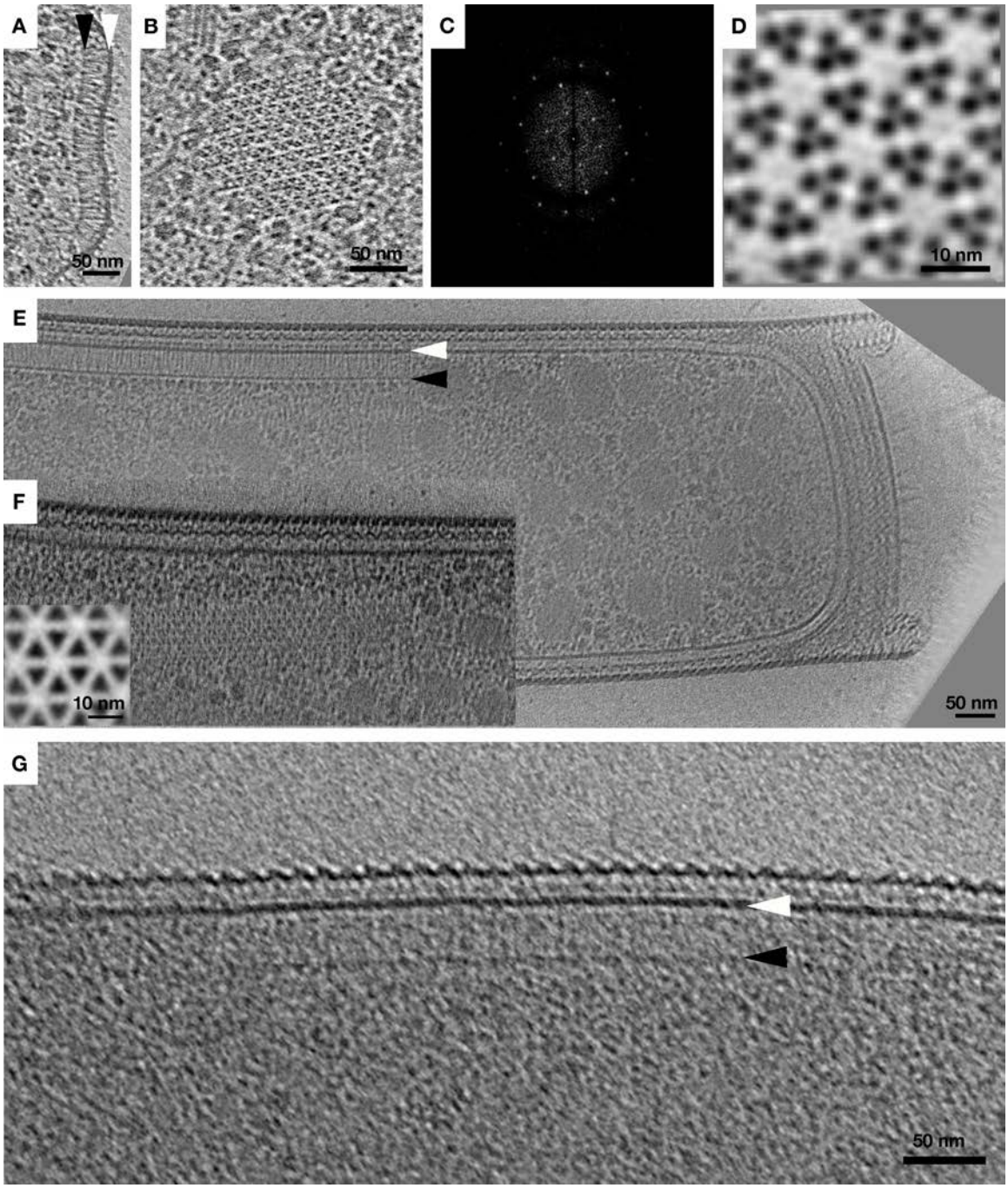


Figure 2. Membrane-bound chemoreceptor arrays in Archaea. *T. kodakarensis* (**A-D**) was grown anaerobically on elemental sulfur as previously described (Atomi et al.,

2004). *M. hungatei* (**E-F**) was grown on hydrogen and carbon dioxide (Toso et al., 2011). *H. salinarum* (**G**) was grown aerobically in complex medium (Oesterhelt and Krippahl, 1983) at 37°C. *H. salinarum* cells were fixed with 2.5% glutaraldehyde then washed into low-salt buffer with 7% PEG-6000. Cells of all strains were mixed with fiducial markers (Iancu et al., 2007) and plunge-frozen on EM grids (Tivol et al., 2008) and imaged by ECT (Gan and Jensen, 2012). Images were aligned, CTF corrected, and reconstructed with IMOD (Kremer et al., 1996). SIRT reconstructions were calculated using TOMO3D (Agulleiro and Fernandez, 2011), and subvolume averages generated using PEET (Nicastro et al., 2006). (**A**, **E**, and **G**) show representative side views of membrane-bound arrays in *T. kodakarensis, M. hungatei*, and *H. salinarum*, respectively. (**B** and **F**) show top views, demonstrating the hexagonal packing of the trimers-of-MCP-dimers (power spectrum shown in **C** and subtomogram averages in **D** and inset). Scale bars 50 nm (A, B, E, F, G) or 10 nm (D, F inset); power spectrum (C) not to scale.



203

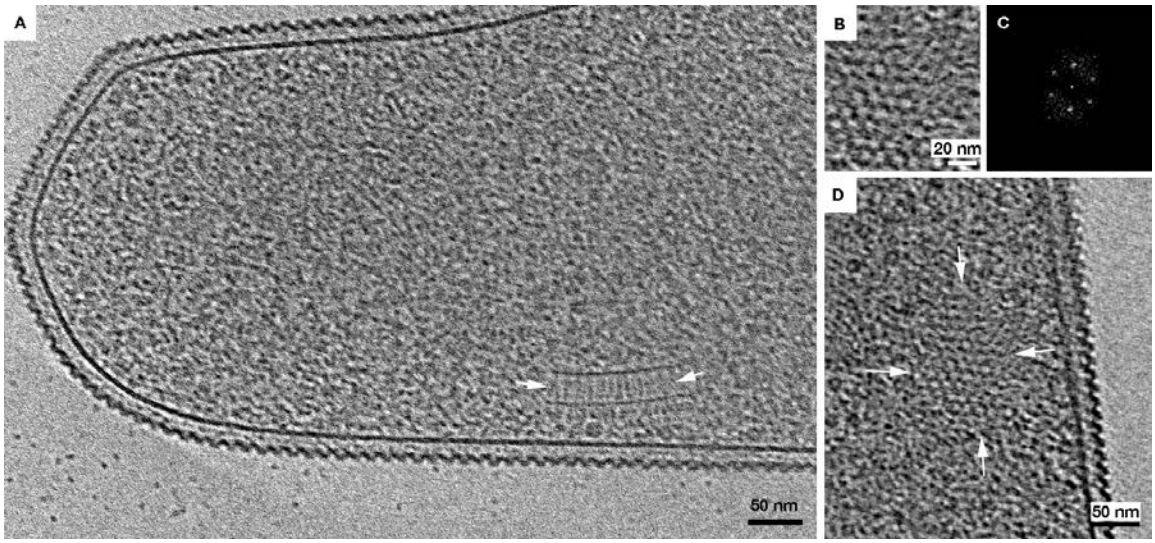
204 **Figure 3. Cytoplasmic chemoreceptor arrays in *M. formicicum*.** Live cultures of *M.*

205 *formicicum* (DSM 22288) were purchased from the DSMZ and frozen on EM grids upon

206 arrival. ECT data collection and image processing was performed as described in Figure

207 2. (**A**) shows a side view of a cytoplasmic array (white arrows). (**B**, associated power

spectrum in **C**) and (**D**) show cross-sections through a top view of an array revealing hexagonal packing of two layers of trimers-of-MCP-dimers, sandwiched between CheA/CheW baseplates. Scale bars 50 nm (A, D) or 20 nm (B); power spectrum (C) not to scale.



References

Adler, J. (1966) Chemotaxis in Bacteria. *Science* **153**: 708-716.

Agulleiro, J.I., and Fernandez, J.J. (2011) Fast tomographic reconstruction on multicore computers. *Bioinformatics* **27**: 582-583.

Atomi, H., Fukui, T., Kanai, T., Morikawa, M., and Imanaka, T. (2004) Description of Thermococcus kodakaraensis sp. nov., a well studied hyperthermophilic archaeon previously reported as Pyrococcus sp. KOD1. *Archaea* **1**: 263-267.

Battistuzzi, F.U., Feijao, A., and Hedges, S.B. (2004) A genomic timescale of prokaryote evolution: insights into the origin of methanogenesis, phototrophy, and the colonization of land. *BMC Evol Biol* **4**: 44.

223 Blainey, P.C., Mosier, A.C., Potanina, A., Francis, C.A., and Quake, S.R. (2011) Genome
224 of a low-salinity ammonia-oxidizing archaeon determined by single-cell and
225 metagenomic analysis. *PLoS One* **6**: e16626.
226 Briegel, A., Ladinsky, M.S., Oikonomou, C., Jones, C.W., Harris, M.J., Fowler, D.J. et
227 al. (2014) Structure of bacterial cytoplasmic chemoreceptor arrays and implications for
228 chemotactic signaling. *Elife* **3**: e02151.
229 Briegel, A., Ortega, D.R., Tocheva, E.I., Wuichet, K., Li, Z., Chen, S. et al. (2009)
230 Universal architecture of bacterial chemoreceptor arrays. *Proc Natl Acad Sci U S A* **106**:
231 17181-17186.
232 Brochier-Armanet, C., Forterre, P., and Gribaldo, S. (2011) Phylogeny and evolution of
233 the Archaea: one hundred genomes later. *Curr Opin Microbiol* **14**: 274-281.
234 Brochier-Armanet, C., Boussau, B., Gribaldo, S., and Forterre, P. (2008) Mesophilic
235 Crenarchaeota: proposal for a third archaeal phylum, the Thaumarchaeota. *Nat Rev*
236 *Microbiol* **6**: 245-252.
237 Cruden, D., Sparling, R., and Markovetz, A.J. (1989) Isolation and Ultrastructure of the
238 Flagella of Methanococcus thermolithotrophicus and Methanospirillum hungatei. *Appl*
239 *Environ Microb* **55**: 1414-1419.
240 Deppenmeier, U., Johann, A., Hartsch, T., Merkl, R., Schmitz, R.A., Martinez-Arias, R.
241 et al. (2002) The genome of Methanosarcina mazei: evidence for lateral gene transfer
242 between bacteria and archaea. *J Mol Microbiol Biotechnol* **4**: 453-461.
243 Gan, L., and Jensen, G.J. (2012) Electron tomography of cells. *Quart Rev Biophys* **45**:
244 27-56.

245 Hazelbauer, G.L., Falke, J.J., and Parkinson, J.S. (2008) Bacterial chemoreceptors: high-
246 performance signaling in networked arrays. *Trends Biochem Sci* **33**: 9-19.
247 Iancu, C.V., Tivol, W.F., Schooler, J.B., Dias, D.P., Henderson, G.P., Murphy, G.E. et al.
248 (2007) Electron cryotomography sample preparation using the Vitrobot. *Nat Protoc* **1**:
249 2813-2819.
250 Jarrell, K.F., and Albers, S.V. (2012) The archaellum: an old motility structure with a
251 new name. *Trends Microbiol* **20**: 307-312.
252 Keeling, P.J., and Palmer, J.D. (2008) Horizontal gene transfer in eukaryotic evolution.
253 *Nat Rev Genet* **9**: 605-618.
254 Koretke, K.K., Lupas, A.N., Warren, P.V., Rosenberg, M., and Brown, J.R. (2000)
255 Evolution of two-component signal transduction. *Mol Biol Evol* **17**: 1956-1970.
256 Kremer, J.R., Mastronarde, D.N., and McIntosh, J.R. (1996) Computer visualization of
257 three-dimensional data using Imod. *J Struct Biol* **116**: 71-76.
258 Lybarger, S.R., and Maddock, J. (2001) Polarity in Action: Asymmetric Protein
259 Localization in Bacteria. *J Bacteriol* **183**: 3261-3267.
260 Metlina, A.L. (2004) Bacterial and archaeal flagella as prokaryotic motility organelles.
261 *Biochemistry (Mosc)* **69**: 1203-1212.
262 Murray, R.G.E., and Birch-Andersen, A. (1963) Specialized Structure in the region of the
263 flagella tuft in *Spirillum serpens*. *Can J Microbiol* **9**.
264 Nelson-Sathi, S., Sousa, F.L., Roettger, M., Lozada-Chavez, N., Thiergart, T., Janssen,
265 A. et al. (2014) Origins of major archaeal clades correspond to gene acquisitions from
266 bacteria. *Nature*.

267 Nicastro, D., Schwartz, C.L., Pierson, J., Gaudette, R., Porter, M.E., and McIntosh, J.R.
268 (2006) The molecular architecture of axonemes revealed by cryoelectron tomography.
269 *Science* **313**: 944-948.
270 Oesterhelt, D., and Krippahl, G. (1983) Phototrophic growth of halobacteria and its use
271 for isolation of photosynthetically-deficient mutants. *Ann Microbiol (Paris)* **134B**: 137-
272 150.
273 Schlesner, M., Miller, A., Streif, S., Staudinger, W.F., Muller, J., Scheffer, B. et al.
274 (2009) Identification of Archaea-specific chemotaxis proteins which interact with the
275 flagellar apparatus. *BMC Microbiol* **9**: 56.
276 Spang, A., Poehlein, A., Offre, P., Zumbragel, S., Haider, S., Rychlik, N. et al. (2012)
277 The genome of the ammonia-oxidizing Candidatus Nitrososphaera gargensis: insights
278 into metabolic versatility and environmental adaptations. *Environ Microbiol* **14**: 3122-
279 3145.
280 Tivol, W., Briegel, A., and Jensen, G.J. (2008) An Improved Cryogen for Plunge
281 Freezing. *Microsc Micoanal* **14**: 375-379.
282 Toso, D.B., Henstra, A.M., Gunsalus, R.P., and Zhou, Z.H. (2011) Structural, mass and
283 elemental analyses of storage granules in methanogenic archaeal cells. *Environ Microbiol*
284 **13**: 2587-2599.
285 Trachtenberg, S., Pinnick, B., and Kessel, M. (2000) The cell surface glycoprotein layer
286 of the extreme halophile Halobacterium salinarum and its relation to Haloferax volcanii:
287 cryo-electron tomography of freeze-substituted cells and projection studies of negatively
288 stained envelopes. *J Struct Biol* **130**: 10-26.

Ulrich, L.E., and Zhulin, I.B. (2010) The MiST2 database: a comprehensive genomics resource on microbial signal transduction. *Nucleic Acids Res* **38**: D401-407.

Wolanin, P.M., Thomason, P.A., and Stock, J.B. (2002) Histidine protein kinases: key signal transducers outside the animal kingdom. *Genome Biol* **3**: REVIEWS3013.

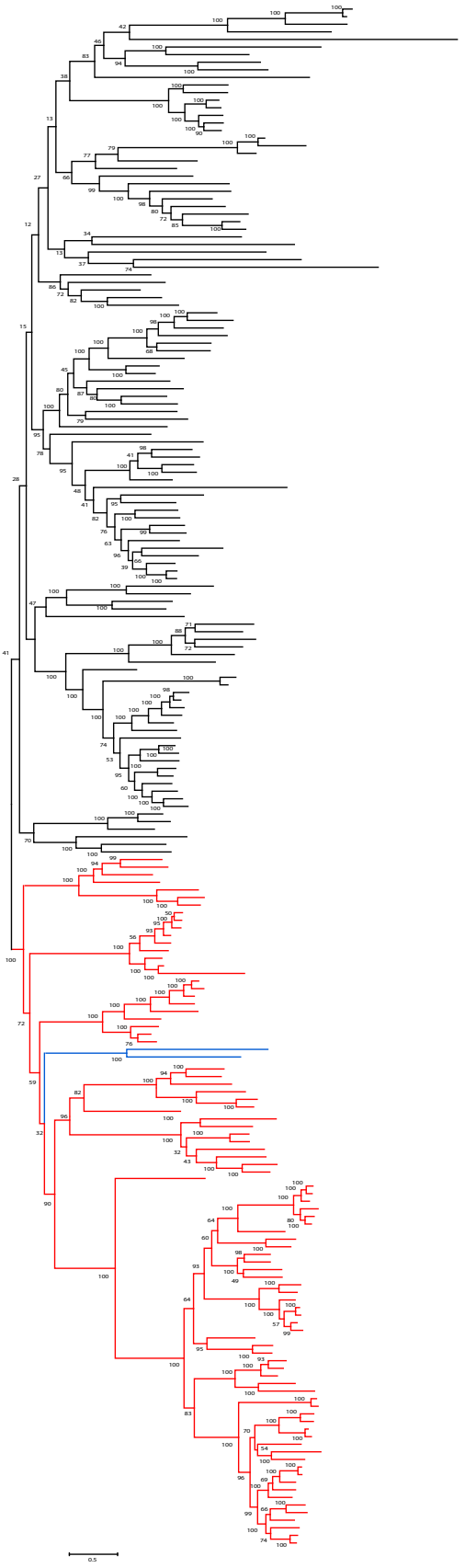
Wuichet, K., and Zhulin, I.B. (2010) Origins and diversification of a complex signal transduction system in prokaryotes. *Sci Signal* **3**: ra. 50.

Supplemental Information

Structural conservation of chemotaxis machinery across Archaea and Bacteria

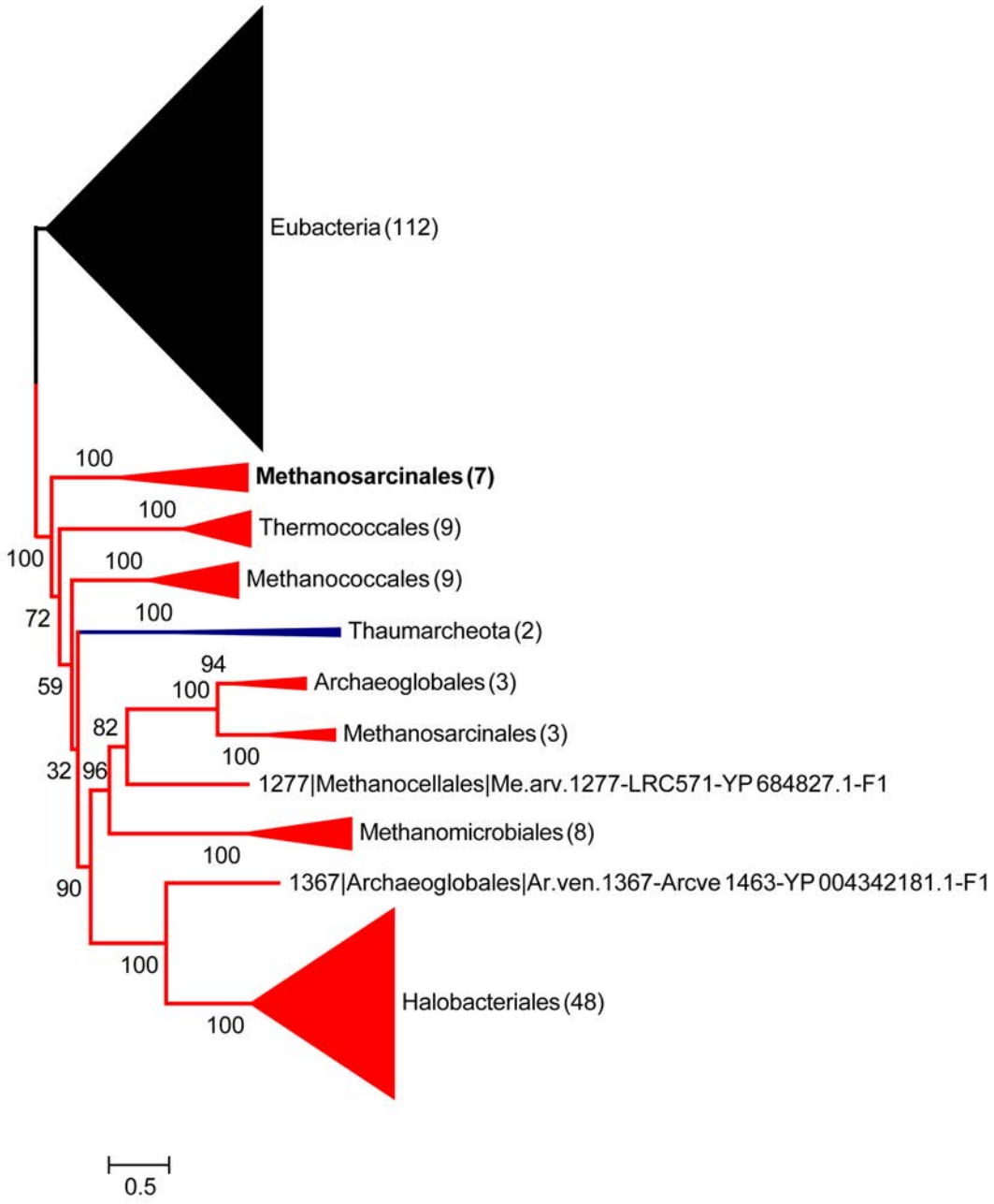
Briegel, A., Ortega, D.R., Huang, A., Oikonomou, C.M., Gunsalus, R.P., Jensen, G.J.

Supplemental Figure 1. CheA:CheB:CheR phylogenetic tree of the F1 chemotaxis system in Bacteria (black) and Archaea (color). Thaumarchaeota are in blue and Euryarchaeota are in red. Bold font indicates the F1 system in Methanosarcinales resulting from secondary LGT. All CheA, CheB, and CheR genes from the MiST2 database (Ulrich and Zhulin, 2010) were downloaded and classified into chemotaxis classes using HMMER3 (Eddy, 2011) and HMM provided by the authors (Wuichet and Zhulin, 2010). Of those, 193 CheA, CheB, and CheR homologs from genomes containing unambiguously interacting CheA, CheB, and CheR genes from F1 systems were independently aligned using the E-INS-I algorithm from the MAFFT v7.182 package (Marucci et al., 2014). The resultant alignments were then concatenated. RAxML (Stamatakis, 2014) was used to produce 1,000 rapid bootstrapped trees using an LG + Γ_4 + I evolutionary model, which were then searched for the best-scoring maximum likelihood trees. Bootstrap values were calculated using the 100 best scoring trees and are shown above nodes (values indicate the number of trees, out of 100, that placed the node as shown). Leaf names represent a description of the system: A | B | C – D – E, where A is the MiST ID of the genome, B is the order, C is the sequence tag (two letter genus and species codes plus MiST ID), D is the locus of the respective CheA, and E is the chemotaxis system (F1). Scale bar indicates average number of substitutions per site.

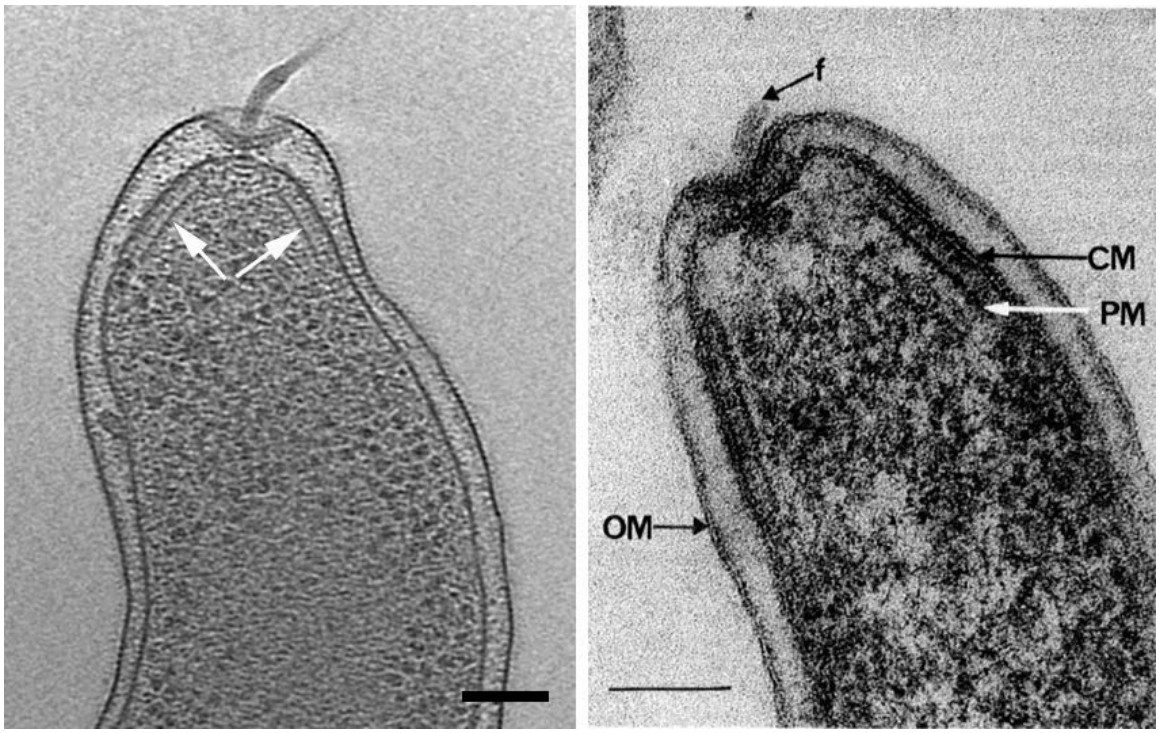


3492|Oscillatoriales|Ar.pla.3492-APPUASWS_20327-ZP_11276632.1-F1
2649|Oscillatoriales|Ar.sp..2649-ARTHRO_630020-ZP_09783613.1-F1
2218|Oscillatoriales|Cr.ep.2218-Cri9233_3522-YP_007143857.1-F1
1313|Anaerolineales|An.the.1313-ANT_31120-YP_004175738.1-F1
745|Solirubrobacterales|Co.wre.74-Cwroe_3136-YP_002394930.1-F1
2131|Nitrosopriaceae|Le.fer.2131-LFMI04_0250-YP_006765455.1-F1
207|Nitrosopriaceae|Ca.Nit.207-NIDE2350-YP_003797988.1-F1
2050|Ignavibacteriales|Me.ros.2050-MROS_2123-YP_006528386.1-F1
1917|Ignavibacteriales|Ig.alb.1917-IALB_2550-YP_005847521.1-F1
1558|Planctomycetales|Bl.aci.1558-Sirac_0308-YP_007200449.1-F1
2168|Clostridiales|De.sp..2168-DCF50_p1553-YP_006913542.1-F1
1343|Clostridiales|Sy.gly.1343-Sgly_0590-YP_004264954.1-F1
1572|Clostridiales|De.deh.1572-Desde_3530-YP_006431695.1-F1
1573|Clostridiales|De.dic.1573-Dasdi_2877-YP_007221685.1-F1
1661|Clostridiales|De.aci.1661-Desaci_3832-YP_009498945.1-F1
1529|Clostridiales|De.mer.1529-Desmer_3922-YP_006623929.1-F1
1576|Clostridiales|De.ori.1576-Desor_4895-YP_00497281.1-F1
1291|Clostridiales|Th.mar.1291-Tmar_2098-YP_004102906.1-F1
3585|Clostridiales|Th.sub.3585-ThesuDRAFT_00772-ZP_11319823.1-F1
1291|Clostridiales|Th.mar.1291-Tmar_1874-YP_004102899.1-F1
589|Clostridiales|Sy.the.589-STH1537-YP_075366.1-F1
153|Clostridiales|Th.pot.153-TherJR_0111-YP_003638905.1-F1
532|Thermoanaerobacterales|Mo.the.532-Moth_0742-YP_429603.1-F1
418|Thermoanaerobacterales|Ca.hyd.418-CHY_0987-YP_359814.1-F1
1028|Clostridiales|Ca.Des.1028-Daud_1792-YP_001717920.1-F1
938|Clostridiales|Pe.the.938-PTH_2116-YP_001212686.1-F1
1416|Clostridiales|De.hua.1416-Desku_1821-YP_004517186.1-F1
1049|Clostridiales|De.ace..1049-Dtox_0655-YP_003190194.1-F1
1427|Clostridiales|De.rum.1427-Desru_1203-YP_004544768.1-F1
644|Clostridiales|De.red.644-Dred_2440-YP_001113775.1-F1
1398|Thermoanaerobacterales|Th.nar.1396-Thena_0028-YP_004436811.1-F1
1119|Thermoanaerobacterales|Co.pro.1119-COPRO5265_0293-YP_002246660.1-F1
74|Solirubrobacterales|Co.wre.74-Cwroe_0059-YP_003391870.1-F1
1718|Bacillales|Al.aci.1718-TC41_1351-YP_005517807.1-F1
1639|Clostridiales|Su.aci.1639-Sulac_3239-YP_005258357.1-F1
153|Clostridiales|Th.pot.153-TherJR_1517-YP_003640271.1-F1
2188|Thermoanaerobacterales|Th.pha.2188-Tph_c11070-YP_008918828.1-F1
645|Clostridiales|Ne.mod.645-HM1_2247-YP_001680814.1-F1
174|Clostridiales|Sy.lip.174-Slip_0978-YP_003702322.1-F1
596|Clostridiales|Sy.wol.596-Swol_0874-YP_753863.1-F1
1507|Clostridiales|Ro.hom.1507-RHOM_07070-YP_004838461.1-F1
2469|Clostridiales|Eu.rec.2469-EUR_13670-YP_007772579.1-F1
2403|Clostridiales|Bu.Bd.2403-CIY_20970-YP_007819857.1-F1
216|Clostridiales|Bu.pro.216-bpr_I1385-YP_003836705.1-F1
1079|Clostridiales|Eu.eli.1079-EUBELI_00835-YP_002930290.1-F1
763|Clostridiales|Cl.phy.763-Cphy_2690-YP_001559790.1-F1
160|Clostridiales|Cl.len.160-Clole_2438-YP_004309330.1-F1
1563|Clostridiales|Cl.cla.1563-Clocl_1941-YP_005048467.1-F1
1863|Clostridiales|Cl.the.1863-Cthe1313_1730-YP_005688238.1-F1
1209|Clostridiales|Pe.CI.1209-CLOST_1714-YP_003936739.1-F1
2134|Clostridiales|Cl.aci.2134-Ciuri_c15790-YP_006789434.1-F1
752|Clostridiales|Al.ore.752-Clos_1506-YP_001513045.1-F1
592|Clostridiales|Al.met.592-Amet_2698-YP_001320509.1-F1
414|Thermoanaerobacterales|Th.lan.414-TTE1417-NP_623035.1-F1
1405|Thermoanaerobacterales|Ma.aus.1405-Mahau_0879-YP_004462899.1-F1
2250|Thermoanaerobacterales|Te.ace.2250-TEPIRE1_14120-YP_007272065.1-F1
997|Natranaerobiales|Na.the.997-Nther_1418-YP_001917889.1-F1
142|Bacillales|Ky.lus.142-Blus_1517-YP_003589374.1-F1
1007|Bacillales|Pa.sp..1007-Pjdr2_3432-YP_003012151.1-F1
1520|Bacillales|Th.com.1520-Theco_2081-YP_007213201.1-F1
1569|Bacillales|Pa.ter.1569-HPL003_18255-YP_005078616.1-F1
18|Bacillales|Pa.sp..18-GYMC19_4187-YP_003244201.1-F1
1448|Bacillales|Pa.muc.1448-KNP414_05796-YP_004644190.1-F1
624|Bacillales|Ba.cla.624-ABC2582-YP_176078.1-F1
441|Bacillales|Oc.ihe.441-OB2543-NP_693464.1-F1
2151|Bacillales|Am.xyl.2151-AXY_14790-YP_008845373.1-F1
173|Bacillales|Ba.sal..173-Bsal_1765-YP_003699838.1-F1
70|Bacillales|Ba.cel.70-Bcell_2479-YP_004095467.1-F1
83|Bacillales|Ba.pse.83-BpOF4_02525-YP_003425464.1-F1
403|Bacillales|Ba.hal.403-BH2970-NP_243836.1-F1
2032|Bacillales|So.al.2032-SSIL_2974-YP_006463543.1-F1
2285|Bacillales|Ba.sub.2285-O663_1689-YP_007428825.1-F1
264|Bacillales|Ba.coa.264-Bcoa_0034-YP_004558040.1-F1
1062|Bacillales|An.fla.1062-Aflv_1717-YP_002316083.1-F1
1017|Bacillales|Ge.sp..1017-GWCH70_1134-YP_002940260.1-F1
296|Bacillales|Ge.sp..296-O4MC1_2665-YP_003989072.1-F1
1849|Halanaerobiales|Ha.pra.1849-Hpral_0535-YP_005835849.1-F1
795|Halanaerobiales|Ha.ons.795-Hons_07600-YP_002508512.1-F1
2256|Halanaerobiales|Ha.hal.2256-Halha_0999-YP_007319053.1-F1
215|Halanaerobiales|Ac.ara.215-Acear_0961-YP_003827656.1-F1
3266|Clostridiales|De.alk.3266-DaloDRAFT_2606-ZP_03738017.1-F1
3952|Lactobacillales|La.aci.3952-LaciK1_01010000704-ZP_09456453.1-F1
1508|Lactobacillales|La.rum.1508-LRU_15850-YP_004832763.1-F1
5352|Lactobacillales|La.mal.5352-LMA_00033-ZP_14787397.1-F1
4298|Lactobacillales|La.vin.4298-LvinD2_01010000584-ZP_11097602.1-F1
3396|Lactobacillales|Er.gal.3396-EGBG_02603-ZP_05650968.1-F1
2992|Lactobacillales|Ca.sp..2992-CATT_09520-ZP_02184110.1-F1
4189|Clostridiales|Ca.acs.4189-CAAU_1693-ZP_10506442.1-F1
4694|Clostridiales|Ca.Art.4694-SFB6_0113-ZP_12608994.1-F1
1510|Clostridiales|Ca.Art.1510-RATSFB_0470-YP_004856958.1-F1
575|Clostridiales|Cl.bei.575-Clbei_4307-YP_001311373.1-F1
4373|Clostridiales|Cl.sp..4373-A370_05180-ZP_11367018.1-F1
2288|Clostridiales|Cl.sac.2288-Capa_c45869-YP_007457596.1-F1
3498|Clostridiales|Cl.but.3498-CLP_3492-ZP_04526462.1-F1
1073|Clostridiales|Cl.bot.1073-CLL_A0800-YP_001885001.1-F1
3358|Clostridiales|Cl.sp..3358-CSBG_02025-ZP_10975999.1-F1
218|Clostridiales|Cl.cal.218-Clocel_1693-YP_003843201.1-F1
4236|Clostridiales|Cl.ars.4236-CarbS_01010001464-ZP_10775614.1-F1
6799|Clostridiales|Cl.pas.6799-FS02_07703-ZP_20960035.1-F1
346|Clostridiales|Cl.ace.346-CA_C2220-NP_348837.1-F1
3435|Clostridiales|Cl.bot.3435-CLD_B1199-ZP_04862179.1-F1
823|Clostridiales|Cl.nov.823-NT01CX_1864-YP_877937.1-F1
6329|Clostridiales|Cl.spo.6329-IYC_15668-ZP_16260094.1-F1
2625|Clostridiales|Cl.car.2625-CLCAR_2223-ZP_06855161.1-F1
202|Clostridiales|Cl.lju.202-CLJU_c09430-YP_003779113.1-F1
941|Clostridiales|Cl.klu.941-CKL_2129-YP_001385512.1-F1
1455|Thermotogaceae|Th.the.1455-Theth_0692-YP_004658873.1-F1
715|Thermotogaceae|Th.lat.715-Tlat_0631-YP_001470281.1-F1
1027|Thermotogaceae|Th.nea.1027-CTN_1883-YP_002535425.1-F1
1538|Synergistales|Th.lie.1538-Tlie_0084-YP_004031927.1-F1
40|Synergistales|Th.aci.40-Taci_1310-YP_003317823.1-F1
2030|Synergistales|An.mob.2030-Anamo_0288-YP_006443273.1-F1
2169|Methanosarcinales|Me.pay.2169-Mpay_2388-YP_005824169.1-F1
2258|Methanosarcinales|Me.hol.2258-Metho_1431-YP_007331068.1-F1
183|Methanosarcinales|Ma.eva.183-Mevan_0987-YP_003726581.1-F1
1441|Methanosarcinales|Me.zhi.1441-Mzhil_0577-YP_004615663.1-F1
365|Methanosarcinales|Me.bar.365-Mbar_A0984-YP_304535.1-F1
454|Methanosarcinales|Me.maz.454-MM_1325-NP_633349.1-F1
447|Methanosarcinales|Me.ace.447-MA0014-NP_614988.1-F1
1323|Thermococcales|Py.aby.1323-PAB1332-NP_127228.1-F1
384|Thermococcales|Py.hor.384-PH0484-NP_142462.1-F1
1392|Thermococcales|Py.sp..1392-PNA2_1134-YP_004424054.1-F1
2003|Thermococcales|Py.sp..2003-Py04_0637-YP_006354214.1-F1
1443|Thermococcales|Py.yay.1443-PYCH_15450-YP_004624483.1-F1
1016|Thermococcales|Th.cmi.1016-TCIN_1198-YP_002307583.1-F1
619|Thermococcales|Th.kod.619-TK0635-YP_183048.1-F1
4335|Thermococcales|Th.pl.4335-TpiA_01010000183-ZP_11216393.1-F1
4335|Thermococcales|Th.pl.4335-TpiA_01010002000-ZP_11215427.1-F1
972|Methanococcales|Me.mar.972-MmarC6_1726-YP_001549772.1-F1
1479|Methanococcales|Me.mar.1479-GYY_05355-YP_004742679.1-F1
872|Methanococcales|Me.mar.872-MmarC5_0734-YP_001097259.1-F1
885|Methanococcales|Me.van.885-Mevan_0220-YP_001327742.1-F1
175|Methanococcales|Me.vol.175-Mvol_0949-YP_003707579.1-F1
220|Methanococcales|Me.oki.220-Metok_1419-YP_004577162.1-F1
148|Methanococcales|Me.inf.148-Metin_0774-YP_003616404.1-F1
19|Methanococcales|Ma.vul.19-Metvu_0258-YP_003246607.1-F1
1159|Methanococcales|Me.fer.1159-Mefer_0940-YP_003127959.1-F1
4136|Nitrosopumilaceae|Ca.Nit.4136-CNitB_01010000627-ZP_10370360.1-F1
2157|Nitrososphaeraceae|Ca.Nit.2157-Ngar_c10180-YP_006861617.1-F1
76|Archaeoglobales|Ar.pro.76-Arcpr_1374-YP_003401086.1-F1
366|Archaeoglobales|Ar.ful.366-AF1040-NP_069873.1-F1
10|Archaeoglobales|Fe.pla.10-Ferp_1075-YP_003435509.1-F1
518|Methanosarcinales|Me.bur.518-Mbur_0361-YP_565106.1-F1
447|Methanosarcinales|Ma.ace.447-MA3066-NP_617959.1-F1
454|Methanosarcinales|Me.maz.454-MM_0328-NP_632352.1-F1
1277|Methanocellales|Me.arv.1277-LRC571-YP_684827.1-F1
597|Methanomicrobiales|Ma.hun.597-Mhun_0115-YP_501607.1-F1
3874|Methanomicrobiales|Me.tar.3874-MettaDRAFT_1613-ZP_09043565.1-F1
784|Methanomicrobiales|Me.mar.784-Memar_0943-YP_001048858.1-F1
2055|Methanomicrobiales|Ma.bou.2055-BN140_1881-YP_006545520.1-F1
245|Methanomicrobiales|Me.pet.245-Mpet_2066-YP_003695252.1-F1
1090|Methanomicrobiales|Ma.pal.1090-Mpal_1856-YP_003468884.1-F1
907|Methanomicrobiales|Me.boo.907-Mboo_1336-YP_001404497.1-F1
2240|Methanomicrobiales|Me.for.2240-Memar_1519-YP_007249059.1-F1
1367|Archaeoglobales|Ar.ver.1367-Arcve_1463-YP_004342181.1-F1
6940|Halobacteriales|Ha.cor.6940-C464_07060-ZP_21611826.1-F1
6939|Halobacteriales|Ha.cal.6939-C463_04404-ZP_21607846.1-F1
6932|Halobacteriales|Ha.tab.6932-C472_13402-ZP_21587259.1-F1
6937|Halobacteriales|Ha.ast.6937-C461_10863-ZP_21602856.1-F1
903|Halobacteriales|Ha.lac.903-Hlac_2263-YP_002568909.1-F1
6933|Halobacteriales|Ha.sac.6933-C471_05401-ZP_21588917.1-F1
4240|Halobacteriales|Ha.sal.4240-HSB1_00050-ZP_10768866.1-F1
2301|Halobacteriales|Na.mos.2301-Nmlp_2689-YP_007487554.1-F1
723|Halobacteriales|Na.pha.723-NP21T2A-YP_326738.1-F1
6962|Halobacteriales|Ha.jap.6962-C444_02752-ZP_21629284.1-F1
1048|Halobacteriales|Ha.muk.1048-Hmuk_0249-YP_003176003.1-F1
1088|Halobacteriales|Ha.ura.1088-Huta_0942-YP_003120858.1-F1
6929|Halobacteriales|Ha.car.6929-C475_09810-ZP_21574619.1-F1
6930|Halobacteriales|Ha.par.6930-C474_21071-ZP_21581287.1-F1
6905|Halobacteriales|Ha.bor.6905-C490_18329-ZP_21479973.1-F1
6949|Halobacteriales|Ha.lar.6949-C455_02432-ZP_21641781.1-F1
6945|Halobacteriales|Ha.sp..6945-C459_08010-ZP_21628233.1-F1
6950|Halobacteriales|Ha.gib.6950-C454_16646-ZP_21648210.1-F1
6955|Halobacteriales|Ha.med.6955-C439_04345-ZP_21664368.1-F1
6954|Halobacteriales|Ha.muc.6954-C440_13589-ZP_21662827.1-F1
3642|Halobacteriales|Ha.pau.3642-D2U20509_04507-ZP_08043284.1-F1
392|Halobacteriales|Ha.sp..392-VNG0971G-NP_279913.1-F1
3870|Halobacteriales|Ha.sp..3870-HatDL1DRAFT_2005-ZP_09032296.1-F1
6914|Halobacteriales|Ha.gre.6914-C480_05732-ZP_21514263.1-F1
4068|Halobacteriales|Ha.lac.4056-HlacA_01010000015-ZP_09946097.1-F1
1437|Halobacteriales|Ha.xan.1437-Halxa_3515-YP_004598005.1-F1
6927|Halobacteriales|Ha.sal.6927-C477_14778-ZP_21567476.1-F1
6911|Halobacteriales|Ha.inn.6911-C493_20932-ZP_21504131.1-F1
6925|Halobacteriales|Ha.asi.6925-C479_02141-ZP_21580320.1-F1
2243|Halobacteriales|Na.rub.2243-Halru_1286-YP_002784033.1-F1
6921|Halobacteriales|Ha.hul.6921-C483_16973-ZP_21544474.1-F1
100|Halobacteriales|Na.mag.100-Nmag_3149-YP_003481263.1-F1
6923|Halobacteriales|Ha.aeg.6923-C481_12644-ZP_21551503.1-F1
6924|Halobacteriales|Ha.aeg.6924-C482_02813-ZP_21553745.1-F1
1437|Halobacteriales|Ha.xan.1437-Halxa_2555-YP_004597052.1-F1
1525|Halobacteriales|Ha.gre.1525-Natgr_0515-YP_007176217.1-F1
4056|Halobacteriales|Ha.lac.4056-HlacA_01010001369-ZP_09948734.1-F1
6919|Halobacteriales|Ha.ath.6919-C485_04576-ZP_21533053.1-F1
2054|Halobacteriales|Ha.sp..2054-NJ7G_3526-YP_006542821.1-F1
6915|Halobacteriales|Ha.ver.6915-C489_03681-ZP_21517517.1-F1
1524|Halobacteriales|Na.pal.1524-Natpe_2681-YP_007281371.1-F1
6928|Halobacteriales|Ha.lim.6928-C476_04305-ZP_21569986.1-F1
6910|Halobacteriales|Ha.ban.6910-C494_12641-ZP_21498467.1-F1
6909|Halobacteriales|Na.sul.6909-C495_11219-ZP_21494803.1-F1
6908|Halobacteriales|Na.tib.6908-C496_07188-ZP_21489336.1-F1
6911|Halobacteriales|Ha.inn.6911-C493_03557-ZP_21500700.1-F1
6927|Halobacteriales|Ha.sal.6927-C477_20464-ZP_21568589.1-F1
77|Halobacteriales|Ha.tur.77-Htur_0958-YP_003402525.1-F1

Supplemental Figure 2. Compressed form of Supplemental Figure 1 emphasizing later LGT events. Color scheme as in Supplemental Figure 1: Bacteria are in black, Thaumarchaeota in blue, and Euryarchaeota in red. F1 systems resulting from secondary LGT in Methanosarcinales are in bold.



Supplemental Figure 3. Comparison of the chemoreceptor array (left, white arrows) and polar organelle (right, white arrow, PM denotes Polar Membrane) in *Campylobacter jejuni*. Scale bars 100 nm. Right panel from (Brock and Murray, 1988).



Supplemental Table 1. Sequence similarity of ACF chemotaxis systems in Methanomicrobiales.

Best BLAST hit of the CheA from ACF systems present in Archaea excluding hits to themselves. The sequences were subject to BLAST (Camacho et al., 2009) similarity search against two custom databases built from archaeal and bacterial genomes in MiST.

		Best hit in Bacteria		
Organism	**Locus**	**Organism**	**locus**	**E-value**
Methanospirillum hungatei	Mhun_0494	*Desulfomonile tiedjei*	Desti_0040	1E-155
Methanoregula formicicum	Metfor_1885	*Desulfomonile tiedjei*	Desti_0040	3E-172
Methanosphaerula palustris	Mpal_1329	*Desulfomonile tiedjei*	Desti_0040	0.0
Methanospirillum hungatei	Mhun_0989	*Syntrophus aciditrophicus*	SYN_00431	4E-167
Methanoculleus marisnigri	Memar_0238	*Chthoniobacter flavus*	CfE428DRAFT_6501	6E-150

		Best hit in Archaea		
Organism	**Locus**	**Organism**	**locus**	**E-value**
Methanospirillum hungatei	Mhun_0494	*Pyrococcus yayanosii*	PYCH_15450	2E-070
Methanoregula formicicum	Metfor_1885	*Pyrococcus yayanosii*	PYCH_15450	2E-066
Methanosphaerula palustris	Mpal_1329	*Methanocaldococcus infernus*	Metin_0774	5E-066
Methanospirillum hungatei	Mhun_0989	*Archaeoglobus profundus*	Arcpr_1374	3E-064
Methanoculleus marisnigri	Memar_0238	*Pyrococcus sp.*	Py04_0537	3E-071

Supplemental Table 2. Summary of archaeal species imaged by ECT in the current study.

	Location Isolated	Observed Cell Diameter*	MCP Genes	Membrane-Bound Arrays?	Distance from Membrane to Baseplate	Cytoplasmic Arrays?	Distance between Baseplates
Halobacteriales							
Halobacterium salinarum	Salted fish	550-800 nm	18	Yes	35 nm	No	-
Methanomicrobiales							
Methanoregula formicica	Brewery effluent treatment sludge	300-400 nm	6	No	-	Yes	27 nm
Methanospirillum hungatei	Sewage sludge	350-400 nm	27	Yes	32 nm	No	-
Thermococcales							
Thermococcus kodakarensis	Solfatara	1-1.5 µm	5	Yes	30 nm	No	-

*Cell diameter measured between inner membranes

Supplemental References

Brock, F.M., and Murray, R.G. (1988) The ultrastructure and ATPase nature of polar membrane in Campylobacter jejuni. *Can J Microbiol* **34**: 594-604.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., and Madden, T.L. (2009) BLAST+: architecture and applications. *BMC Bioinformatics* **10**: 421.

Eddy, S.R. (2011) Accelerated Profile HMM Searches. *PLoS Comput Biol* **7**: e1002195.

377 Marucci, E.A., Zafalon, G.F., Momente, J.C., Neves, L.A., Valencio, C.R., Pinto, A.R. et
378 al. (2014) An efficient parallel algorithm for multiple sequence similarities calculation
379 using a low complexity method. *Biomed Res Int* **2014**: 563016.
380 Stamatakis, A. (2014) RAxML version 8: a tool for phylogenetic analysis and post-
381 analysis of large phylogenies. *Bioinformatics* **30**: 1312-1313.
382 Ulrich, L.E., and Zhulin, I.B. (2010) The MiST2 database: a comprehensive genomics
383 resource on microbial signal transduction. *Nucleic Acids Res* **38**: D401-407.
384 Wuichet, K., and Zhulin, I.B. (2010) Origins and diversification of a complex signal
385 transduction system in prokaryotes. *Sci Signal* **3**: ra. 50.
386
387