List of Supplemental Figures and Tables

Supplemental Table 1
Characteristics of Archaeal Genomes that were employed in Phylogenetic Analyses

	Organism	GenBank Accession No.	Class	Size	% GC content	Temp. range	Reference
Crenarchaeota	*Aeropyrum pernix K1*	BA000002.3	Desulfurococcales	* 1.7	56.3	H	[1]
	Staphylothermus marinus F1	CP000575.1	Desulfurococcales	* 1.6	35.7	H	[2]
	Desulfurococcus kamchatkensis 1221n	CP001140.1	Desulfurococcales	* 1.4	45.3	H	[3]
	Hyperthermus butylicus DSM 5456	CP000493.1	Desulfurococcales	* 1.7	53.7	H	[4]
	Ignicoccus hospitalis KIN4/I	CP000816.1	Desulfurococcales	1.3	56.5	H	[5]
	Metallosphaera sedula DSM 5348	CP000682.1	Sulfolobales	2.2	46.2	T	[6]
	Sulfolobus acidocaldarius DSM 639	CP000077.1	Sulfolobales	2.23	36.7	T	[7]
	Sulfolobus islandicus L.D.8.5	CP001731.1	Sulfolobales	*2.73	35.3	H	[8]
	Sulfolobus islandicus L.S.2.15	CP001399.1	Sulfolobales	*2.7	35.1	H	[8]
	Sulfolobus islandicus M.14.25	CP001400.1	Sulfolobales	2.6	35.1	H	[8]
	Sulfolobus islandicus M.16.27	CP001401.1	Sulfolobales	2.7	35	H	[8]
	Sulfolobus islandicus M.16.4	CP001402.1	Sulfolobales	2.6	35	H	[8]
	Sulfolobus islandicus Y.G.57.14	CP001403.1	Sulfolobales	2.7	35.4	H	[8]
	Sulfolobus islandicus Y.N.15.51	CP001404.1	Sulfolobales	2.84	35.3	H	[8]
	Sulfolobus solfataricus P2	AE006641.1	Sulfolobales	3	35.8	H	[9]
	Sulfolobus tokodaii str. 7	BA000023.2	Sulfolobales	2.7	32.8	H	[10]
	Pyrobaculum aerophilum str. IM2	AE009441.1	Thermoproteales	* 2.2	51.4	H	[11]
	Pyrobaculum arsenaticum DSM 13514	CP000660.1	Thermoproteales	* 2.1	55.1	H	DOE-JGI
	Pyrobaculum calidifontis JCM 11548	CP000561.1	Thermoproteales	2	57.2	H	DOE-JGI
	Pyrobaculum islandicum DSM 4184	CP000504.1	Thermoproteales	* 1.8	49.6	H	DOE-JGI
	Caldivirga maquilingensis IC-167	CP000852.1	Thermoproteales	* 2.1	43.1	H	DOE-JGI
	Thermofilum pendens Hrk 5	CP000505.1	Thermoproteales	1.83	57.6	H	[12]
	Thermoproteus neutrophilus V24Sta	CP001014.1	Thermoproteales	* 1.8	59.9	H	DOE-JGI
Th	*Nitrosopumilus maritimus SCM1*	CP000866.1	Thaumarchaeota	* 1.6	34.2	M	DOE-JGI
	Cenarchaeum symbosium A	N/A	Thaumarchaeota	2	57	M	[13]
Euryarchaeota	*Archaeoglobus fulgidus DSM 4304*	AE000782.1	Archaeoglobi	2.18	48.6	H	[14]
	Haloarcula marismortui ATCC 43049	AY596297.1	Halobacteria	4.27	61.1	M	[15]
	Halobacterium salinarum R1	AM774415.1	Halobacteria	2.66	65.7	T	[16]
	Halobacterium sp. NRC-1	AE004437.1	Halobacteria	2.57	65.9	M	[17]
	Haloquadratum walsbyi DSM 16790	AM180088.1	Halobacteria	* 3.15	47.9	--	[18]
	Halorubrum lacusprofundi ATCC 49239	CP001365.1	Halobacteria	* 3.66	64	M	DOE-JGI
	Natronomonas pharaonis DSM 2160	CR936257.1	Halobacteria	2.75	63.1	--	[19]
	Methanocaldococcus jannaschii DSM 2661	L77117.1	Methanococcales	* 1.76	31.3	H	[20]
	Methanococcus aeolicus Nankai-3	CP000743.1	Methanococcales	* 1.6	30	M	DOE-JGI
	Methanococcus maripaludis C5	CP000609.1	Methanococcales	1.81	33	M	DOE-JGI
	Methanococcus vannielii SB	CP000742.1	Methanococcales	* 1.7	31.3	M	DOE-JGI
	Methanosphaera stadtmanae DSM 3091	CP000102.1	Methanobacteria	1.77	27.6	M	[21]
	Methanothermobacter thermautotrophicus str. Delta H	AE000666.1	Methanobacteria	* 1.8	49.5	T	[22]
	Methanobrevibacter smithii ATCC 35061	CP000678.1	Methanobacteria	* 1.9	31	M	[23]
	Methanopyrus kandleri AV19	AE009439.1	Methanopyri	1.69	61.2	H	[24]
	Methanosphaerula palustris E1-9c	CP001338.1	Methanomicrobia	* 2.9	55.4	M	DOE-JGI
	Methanospirillum hungatei JF-1	CP000254.1	Methanomicrobia	3.54	45.1	M	DOE-JGI
	Methanocorpusculum labreanum Z	CP000559.1	Methanomicrobia	1.8	50	M	[25]
	Methanoculleus marisnigri JR1	CP000562.1	Methanomicrobia	* 2.5	62.1	M	[25]
	Candidatus Methanoregula boonei 6A8	CP000780.1	Methanomicrobia	* 2.5	54.5	M	DOE-JGI
	Methanosaeta thermophila PT	CP000477.1	Methanomicrobia	* 1.9	53.5	T	DOE-JGI
	Methanosarcina acetivorans C2A	AE010299.1	Methanomicrobia	5.75	42.7	M	[26]
	Methanosarcina barkeri str. Fusaro	CP000099.1	Methanomicrobia	* 4.84	39.2	M	[27]
	Methanosarcina mazei Go1	AE008384.1	Methanomicrobia	4.1	41.5	M	[28]

Supplemental Table 1 [continued]
Characteristics of Archaeal Genomes that were employed in Phylogenetic Analyses

Methanococcoides burtonii DSM 6242	CP000300.1	Methanomicrobia	2.58	40.8	M	DOE-JGI
Pyrococcus abyssi GE5	AL096836.1	Thermococcales	* 1.8	44.7	H	[29]
Pyrococcus furiosus DSM 3638	AE009950.1	Thermococcales	1.9	40.8	H	[30]
Pyrococcus horikoshii OT3	BA000001.2	Thermococcales	* 1.7	41.9	H	[31]
Thermococcus gammatolerans EJ3	CP001398.1	Thermococcales	* 2	51.3	H	[32]
Thermococcus kodakarensis KOD1	AP006878.1	Thermococcales	2.09	52	H	[33]
Thermococcus onnurineus NA1	CP000855.1	Thermococcales	* 1.8	51.3	H	[34]
Thermococcus sibiricus MM 739	CP001463.1	Thermococcales	* 1.8	--	--	[35]
*Ferroplasma acidarmanus fer1 ** *	N/A	Thermoplasmata	* 1.87	36.5	M	DOE-JGI
Thermoplasma acidophilum DSM 1728	AL139299.1	Thermoplasmata	* 1.6	46	T	[36]
Thermoplasma volcanium GSS1	BA000011.4	Thermoplasmata	1.58	39.9	T	[37]
Picrophilus torridus DSM 9790	AE017261.1	Thermoplasmata	* 1.5	36	T	[38]
Candidatus Korarchaeum cryptofilum OPF8	CP000968.1	Korarchaeota	1.6	49	T	[39]
Nanoarchaeum equitans Kin4-M	AE017199.1	Nanoarchaeota	0.49	31.6	H	[40]

This table lists all the archaeal genomes that were available at the end of October 2009 and sequence information for whom was used for construction of phylogenetic trees shown in Figure 1 and Supplemental Figures 1A and 1B. Since then genome sequences for two crenarchaea species (viz. *Staphylothermus aggregans* and *Staphylothermus hellenicus*) and a number of euryarchaea have become available. The specificity of various signatures described in this work has been checked for all archaeal genomes available to date and information for various species that contained the indels is included in this work.
Abbreviations for temperature range: H- hyperthermophilic; T- thermophilic; M- mesophilic. Only distinct species are shown using a representative strain.
**** estimated size*** ***** genome sequence for F. acidarmanus is still in progress but included in our phylogenetic trees.***
Th = Thaumarchaeota.
DOE-JGI = genome sequenced by the *Department of Education Joint Genome Institute*

Reference List

1. *Kawarabayasi Y, Hino Y, Horikawa H, Yamazaki S, Haikawa Y, Jin-no K, Takahashi M, Sekine M, Baba S, Ankai A, Kosugi H, Hosoyama A, Fukui S, Nagai Y, Nishijima K, Nakazawa H, Takamiya M, Masuda S, Funahashi T, Tanaka T, Kudoh Y, Yamazaki J, Kushida N, Oguchi A, Kikuchi H, .:* ***Complete genome sequence of an aerobic hyper-thermophilic crenarchaeon, Aeropyrum pernix K1****. DNA Res. 1999,* ***6:****83-52.*

2. *Anderson IJ, Dharmarajan L, Rodriguez J, Hooper S, Porat I, Ulrich LE, Elkins JG, Mavromatis K, Sun H, Land M, Lapidus A, Lucas S, Barry K, Huber H, Zhulin IB, Whitman WB, Mukhopadhyay B, Woese C, Bristow J, Kyrpides N:* ***The complete genome sequence of Staphylothermus marinus reveals differences in sulfur metabolism among heterotrophic Crenarchaeota****. BMC.Genomics 2009,* ***10:****145.*

3. *Ravin NV, Mardanov AV, Beletsky AV, Kublanov IV, Kolganova TV, Lebedinsky AV, Chernyh NA, Bonch-Osmolovskaya EA, Skryabin KG:* ***Complete genome sequence of the anaerobic, protein-degrading hyperthermophilic crenarchaeon Desulfurococcus kamchatkensis****. J.Bacteriol. 2009,* ***191:****2371-2379.*

4. *Brugger K, Chen L, Stark M, Zibat A, Redder P, Ruepp A, Awayez M, She Q, Garrett RA, Klenk HP:* ***The genome of Hyperthermus butylicus: a sulfur-reducing, peptide fermenting, neutrophilic Crenarchaeote growing up to 108 degrees C****. Archaea. 2007,* ***2:****127-135.*

5. *Podar M, Anderson I, Makarova KS, Elkins JG, Ivanova N, Wall MA, Lykidis A, Mavromatis K, Sun H, Hudson ME, Chen W, Deciu C, Hutchison D, Eads JR, Anderson A, Fernandes F, Szeto E, Lapidus A, Kyrpides NC, Saier MH, Jr., Richardson PM, Rachel R, Huber H, Eisen JA, Koonin EV, Keller M, Stetter KO:* ***A genomic analysis of the archaeal system Ignicoccus hospitalis-Nanoarchaeum equitans****. Genome Biol. 2008,* ***9:****R158.*

6. *Auernik KS, Maezato Y, Blum PH, Kelly RM:* ***The genome sequence of the metal-mobilizing, extremely thermoacidophilic archaeon Metallosphaera sedula provides insights into bioleaching-associated metabolism****. Appl.Environ.Microbiol. 2008,* ***74:****682-692.*

7. *Chen L, Brugger K, Skovgaard M, Redder P, She Q, Torarinsson E, Greve B, Awayez M, Zibat A, Klenk HP, Garrett RA:* ***The genome of Sulfolobus acidocaldarius, a model organism of the Crenarchaeota****. J.Bacteriol. 2005,* ***187:****4992-4999.*

8. *Reno ML, Held NL, Fields CJ, Burke PV, Whitaker RJ:* ***Biogeography of the Sulfolobus islandicus pan-genome****. Proc.Natl.Acad.Sci.U.S.A 2009,* ***106:****8605-8610.*

9. *She Q, Singh RK, Confalonieri F, Zivanovic Y, Allard G, Awayez MJ, Chan-Weiher CC, Clausen IG, Curtis BA, De Moors A, Erauso G, Fletcher C, Gordon PM, Heikamp-de Jong I, Jeffries AC, Kozera CJ, Medina N, Peng X, Thi-Ngoc HP, Redder P, Schenk ME, Theriault C, Tolstrup N, Charlebois RL, Doolittle WF, Duguet M, Gaasterland T, Garrett RA, Ragan MA, Sensen CW, Van der OJ:* ***The complete genome of the crenarchaeon Sulfolobus solfataricus P2****. Proc.Natl.Acad.Sci.U.S.A 2001,* ***98:****7835-7840.*

10. *Kawarabayasi Y, Hino Y, Horikawa H, Jin-no K, Takahashi M, Sekine M, Baba S, Ankai A, Kosugi H, Hosoyama A, Fukui S, Nagai Y, Nishijima K, Otsuka R, Nakazawa H, Takamiya M, Kato Y, Yoshizawa T, Tanaka T, Kudoh Y, Yamazaki J, Kushida N, Oguchi A, Aoki K, Masuda S, Yanagii M, Nishimura M, Yamagishi A, Oshima T, Kikuchi H:* ***Complete genome sequence of an aerobic thermoacidophilic crenarchaeon, Sulfolobus tokodaii strain7****. DNA Res. 2001,* ***8:****123-140.*

11. *Fitz-Gibbon ST, Ladner H, Kim UJ, Stetter KO, Simon MI, Miller JH:* ***Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum aerophilum****. Proc.Natl.Acad.Sci.U.S.A 2002,* ***99:****984-989.*

12. *Anderson I, Rodriguez J, Susanti D, Porat I, Reich C, Ulrich LE, Elkins JG, Mavromatis K, Lykidis A, Kim E, Thompson LS, Nolan M, Land M, Copeland A, Lapidus A, Lucas S, Detter C, Zhulin IB, Olsen GJ, Whitman W, Mukhopadhyay B, Bristow J, Kyrpides N:* ***Genome sequence of Thermofilum pendens reveals an exceptional loss of biosynthetic pathways without genome reduction****. J.Bacteriol. 2008,* ***190:****2957-2965.*

13. *Hallam SJ, Konstantinidis KT, Putnam N, Schleper C, Watanabe Y, Sugahara J, Preston C, de la TJ, Richardson PM, DeLong EF:* ***Genomic analysis of the uncultivated marine crenarchaeote Cenarchaeum symbiosum****. Proc.Natl.Acad.Sci.U.S.A 2006,* ***103:****18296-18301.*

14. *Klenk HP, Clayton RA, Tomb JF, White O, Nelson KE, Ketchum KA, Dodson RJ, Gwinn M, Hickey EK, Peterson JD, Richardson DL, Kerlavage AR, Graham DE, Kyrpides NC, Fleischmann RD, Quackenbush J, Lee NH, Sutton GG, Gill S, Kirkness EF, Dougherty BA, McKenney K, Adams MD, Loftus B, Peterson S, Reich CI, McNeil LK, Badger JH, Glodek A, Zhou L, Overbeek R, Gocayne JD, Weidman JF, McDonald L, Utterback T, Cotton MD, Spriggs T, Artiach P, Kaine BP, Sykes SM, Sadow PW, D'Andrea KP, Bowman C, Fujii C, Garland SA, Mason TM, Olsen GJ, Fraser CM, Smith HO, Woese CR, Venter JC:* ***The complete genome sequence of the hyperthermophilic, sulphate-reducing archaeon Archaeoglobus fulgidus****. Nature 1997,* ***390:****364-370.*

15. *Baliga NS, Bonneau R, Facciotti MT, Pan M, Glusman G, Deutsch EW, Shannon P, Chiu Y, Weng RS, Gan RR, Hung P, Date SV, Marcotte E, Hood L, Ng WV:* ***Genome sequence of Haloarcula marismortui: a halophilic archaeon from the Dead Sea****. Genome Res. 2004,* ***14:****2221-2234.*

16. *Pfeiffer F, Schuster SC, Broicher A, Falb M, Palm P, Rodewald K, Ruepp A, Soppa J, Tittor J, Oesterhelt D:* ***Evolution in the laboratory: the genome of Halobacterium salinarum strain R1 compared to that of strain NRC-1****. Genomics 2008,* ***91:****335-346.*

17. *Ng WV, Berquist BR, Coker JA, Capes M, Wu TH, DasSarma P, DasSarma S:* ***Genome sequences of Halobacterium species****. Genomics 2008,* ***91:****548-552.*

18. *Bolhuis H, Palm P, Wende A, Falb M, Rampp M, Rodriguez-Valera F, Pfeiffer F, Oesterhelt D:* ***The genome of the square archaeon Haloquadratum walsbyi : life at the limits of water activity****. BMC.Genomics 2006,* ***7:****169.*

19. *Falb M, Pfeiffer F, Palm P, Rodewald K, Hickmann V, Tittor J, Oesterhelt D:* ***Living with two extremes: conclusions from the genome sequence of Natronomonas pharaonis****. Genome Res. 2005,* ***15:****1336-1343.*

20. *Bult CJ, White O, Olsen GJ, Zhou L, Fleischmann RD, Sutton GG, Blake JA, FitzGerald LM, Clayton RA, Gocayne JD, Kerlavage AR, Dougherty BA, Tomb JF, Adams MD, Reich CI, Overbeek R, Kirkness EF, Weinstock KG, Merrick JM, Glodek A, Scott JL, Geoghagen NS, Venter JC:* ***Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii****. Science 1996,* ***273:****1058-1073.*

21. *Fricke WF, Seedorf H, Henne A, Kruer M, Liesegang H, Hedderich R, Gottschalk G, Thauer RK:* ***The genome sequence of Methanosphaera stadtmanae reveals why this human intestinal archaeon is restricted to methanol and H2 for methane formation and ATP synthesis****. J.Bacteriol. 2006,* ***188:****642-658.*

22. *Smith DR, Doucette-Stamm LA, DeLoughery C, Lee H, Dubois J, Aldredge T, Bashirzadeh R, Blakely D, Cook R, Gilbert K, Harrison D, Hoang L, Keagle P, Lumm W, Pothier B, Qiu D, Spadafora R, Vicaire R, Wang Y, Wierzbowski J, Gibson R,*

Jiwani N, Caruso A, Bush D, Reeve JN, .: ***Complete genome sequence of Methanobacterium thermoautotrophicum deltaH: functional analysis and comparative genomics****. J.Bacteriol. 1997,* ***179:****7135-7155.*

23. *Samuel BS, Hansen EE, Manchester JK, Coutinho PM, Henrissat B, Fulton R, Latreille P, Kim K, Wilson RK, Gordon JI:* ***Genomic and metabolic adaptations of Methanobrevibacter smithii to the human gut****. Proc.Natl.Acad.Sci.U.S.A 2007,* ***104:****10643-10648.*

24. *Slesarev AI, Mezhevaya KV, Makarova KS, Polushin NN, Shcherbinina OV, Shakhova VV, Belova GI, Aravind L, Natale DA, Rogozin IB, Tatusov RL, Wolf YI, Stetter KO, Malykh AG, Koonin EV, Kozyavkin SA:* ***The complete genome of hyperthermophile Methanopyrus kandleri AV19 and monophyly of archaeal methanogens****. Proc.Natl.Acad.Sci.U.S.A 2002,* ***99:****4644-4649.*

25. *Anderson I, Ulrich LE, Lupa B, Susanti D, Porat I, Hooper SD, Lykidis A, Sieprawska-Lupa M, Dharmarajan L, Goltsman E, Lapidus A, Saunders E, Han C, Land M, Lucas S, Mukhopadhyay B, Whitman WB, Woese C, Bristow J, Kyrpides N:* ***Genomic characterization of methanomicrobiales reveals three classes of methanogens****. PLoS.One. 2009,* ***4:****e5797.*

26. *Galagan JE, Nusbaum C, Roy A, Endrizzi MG, Macdonald P, FitzHugh W, Calvo S, Engels R, Smirnov S, Atnoor D, Brown A, Allen N, Naylor J, Stange-Thomann N, DeArellano K, Johnson R, Linton L, McEwan P, McKernan K, Talamas J, Tirrell A, Ye W, Zimmer A, Barber RD, Cann I, Graham DE, Grahame DA, Guss AM, Hedderich R, Ingram-Smith C, Kuettner HC, Krzycki JA, Leigh JA, Li W, Liu J, Mukhopadhyay B, Reeve JN, Smith K, Springer TA, Umayam LA, White O, White RH, Conway dM, Ferry JG, Jarrell KF, Jing H, Macario AJ, Paulsen I, Pritchett M, Sowers KR, Swanson RV, Zinder SH, Lander E, Metcalf WW, Birren B:* ***The genome of M. acetivorans reveals extensive metabolic and physiological diversity****. Genome Res. 2002,* ***12:****532-542.*

27. *Maeder DL, Anderson I, Brettin TS, Bruce DC, Gilna P, Han CS, Lapidus A, Metcalf WW, Saunders E, Tapia R, Sowers KR:* ***The Methanosarcina barkeri genome: comparative analysis with Methanosarcina acetivorans and Methanosarcina mazei reveals extensive rearrangement within methanosarcinal genomes****. J.Bacteriol. 2006,* ***188:****7922-7931.*

28. *Deppenmeier U, Johann A, Hartsch T, Merkl R, Schmitz RA, Martinez-Arias R, Henne A, Wiezer A, Baumer S, Jacobi C, Bruggemann H, Lienard T, Christmann A, Bomeke M, Steckel S, Bhattacharyya A, Lykidis A, Overbeek R, Klenk HP, Gunsalus RP, Fritz HJ, Gottschalk G:* ***The genome of Methanosarcina mazei: evidence for lateral gene transfer between bacteria and archaea****. J.Mol.Microbiol.Biotechnol. 2002,* ***4:****453-461.*

29. *Lecompte O, Ripp R, Puzos-Barbe V, Duprat S, Heilig R, Dietrich J, Thierry JC, Poch O:* ***Genome evolution at the genus level: comparison of three complete genomes of hyperthermophilic archaea****. Genome Res. 2001,* ***11:****981-993.*

30. *Robb FT, Maeder DL, Brown JR, DiRuggiero J, Stump MD, Yeh RK, Weiss RB, Dunn DM:* ***Genomic sequence of hyperthermophile, Pyrococcus furiosus: implications for physiology and enzymology****. Methods Enzymol. 2001,* ***330:****134-157.*

31. *Kawarabayasi Y, Sawada M, Horikawa H, Haikawa Y, Hino Y, Yamamoto S, Sekine M, Baba S, Kosugi H, Hosoyama A, Nagai Y, Sakai M, Ogura K, Otsuka R, Nakazawa H, Takamiya M, Ohfuku Y, Funahashi T, Tanaka T, Kudoh Y, Yamazaki J, Kushida N, Oguchi A, Aoki K, Kikuchi H:* ***Complete sequence and gene organization of the genome of a hyper-thermophilic archaebacterium, Pyrococcus horikoshii OT3****. DNA Res. 1998,* ***5:****55-76.*

32. *Zivanovic Y, Armengaud J, Lagorce A, Leplat C, Guerin P, Dutertre M, Anthouard V, Forterre P, Wincker P, Confalonieri F:* ***Genome analysis and genome-wide proteomics of Thermococcus gammatolerans, the most radioresistant organism known amongst the Archaea****. Genome Biol. 2009,* ***10:****R70.*

33. *Fukui T, Atomi H, Kanai T, Matsumi R, Fujiwara S, Imanaka T:* ***Complete genome sequence of the hyperthermophilic archaeon Thermococcus kodakaraensis KOD1 and comparison with Pyrococcus genomes****. Genome Res. 2005,* ***15:****352-363.*

34. *Lee HS, Kang SG, Bae SS, Lim JK, Cho Y, Kim YJ, Jeon JH, Cha SS, Kwon KK, Kim HT, Park CJ, Lee HW, Kim SI, Chun J, Colwell RR, Kim SJ, Lee JH:* ***The complete genome sequence of Thermococcus onnurineus NA1 reveals a mixed heterotrophic and carboxydotrophic metabolism****. J.Bacteriol. 2008,* ***190:****7491-7499.*

35. *Mardanov AV, Ravin NV, Svetlitchnyi VA, Beletsky AV, Miroshnichenko ML, Bonch-Osmolovskaya EA, Skryabin KG:* ***Metabolic versatility and indigenous origin of the archaeon Thermococcus sibiricus, isolated from a siberian oil reservoir, as revealed by genome analysis****. Appl.Environ.Microbiol. 2009,* ***75:****4580-4588.*

36. *Ruepp A, Graml W, Santos-Martinez ML, Koretke KK, Volker C, Mewes HW, Frishman D, Stocker S, Lupas AN, Baumeister W:* ***The genome sequence of the thermoacidophilic scavenger Thermoplasma acidophilum****. Nature 2000,* ***407:****508-513.*

37. *Kawashima T, Amano N, Koike H, Makino S, Higuchi S, Kawashima-Ohya Y, Watanabe K, Yamazaki M, Kanehori K, Kawamoto T, Nunoshiba T, Yamamoto Y, Aramaki H, Makino K, Suzuki M:* ***Archaeal adaptation to higher temperatures revealed by genomic sequence of Thermoplasma volcanium****. Proc.Natl.Acad.Sci.U.S.A 2000,* ***97:****14257-14262.*

38. *Futterer O, Angelov A, Liesegang H, Gottschalk G, Schleper C, Schepers B, Dock C, Antranikian G, Liebl W:* ***Genome sequence of Picrophilus torridus and its implications for life around pH 0****. Proc.Natl.Acad.Sci.U.S.A 2004,* ***101:****9091-9096.*

39. *Elkins JG, Podar M, Graham DE, Makarova KS, Wolf Y, Randau L, Hedlund BP, Brochier-Armanet C, Kunin V, Anderson I, Lapidus A, Goltsman E, Barry K, Koonin EV, Hugenholtz P, Kyrpides N, Wanner G, Richardson P, Keller M, Stetter KO:* ***A korarchaeal genome reveals insights into the evolution of the Archaea****. Proc.Natl.Acad.Sci.U.S.A 2008,* ***105:****8102-8107.*

40. *Waters E, Hohn MJ, Ahel I, Graham DE, Adams MD, Barnstead M, Beeson KY, Bibbs L, Bolanos R, Keller M, Kretz K, Lin X, Mathur E, Ni J, Podar M, Richardson T, Sutton GG, Simon M, Soll D, Stetter KO, Short JM, Noordewier M:* ***The genome of Nanoarchaeum equitans: insights into early archaeal evolution and derived parasitism****. Proc.Natl.Acad.Sci.U.S.A 2003,* ***100:****12984-12988.*

Supplemental Table 2

(a) Proteins specific for Various Sulfolobales

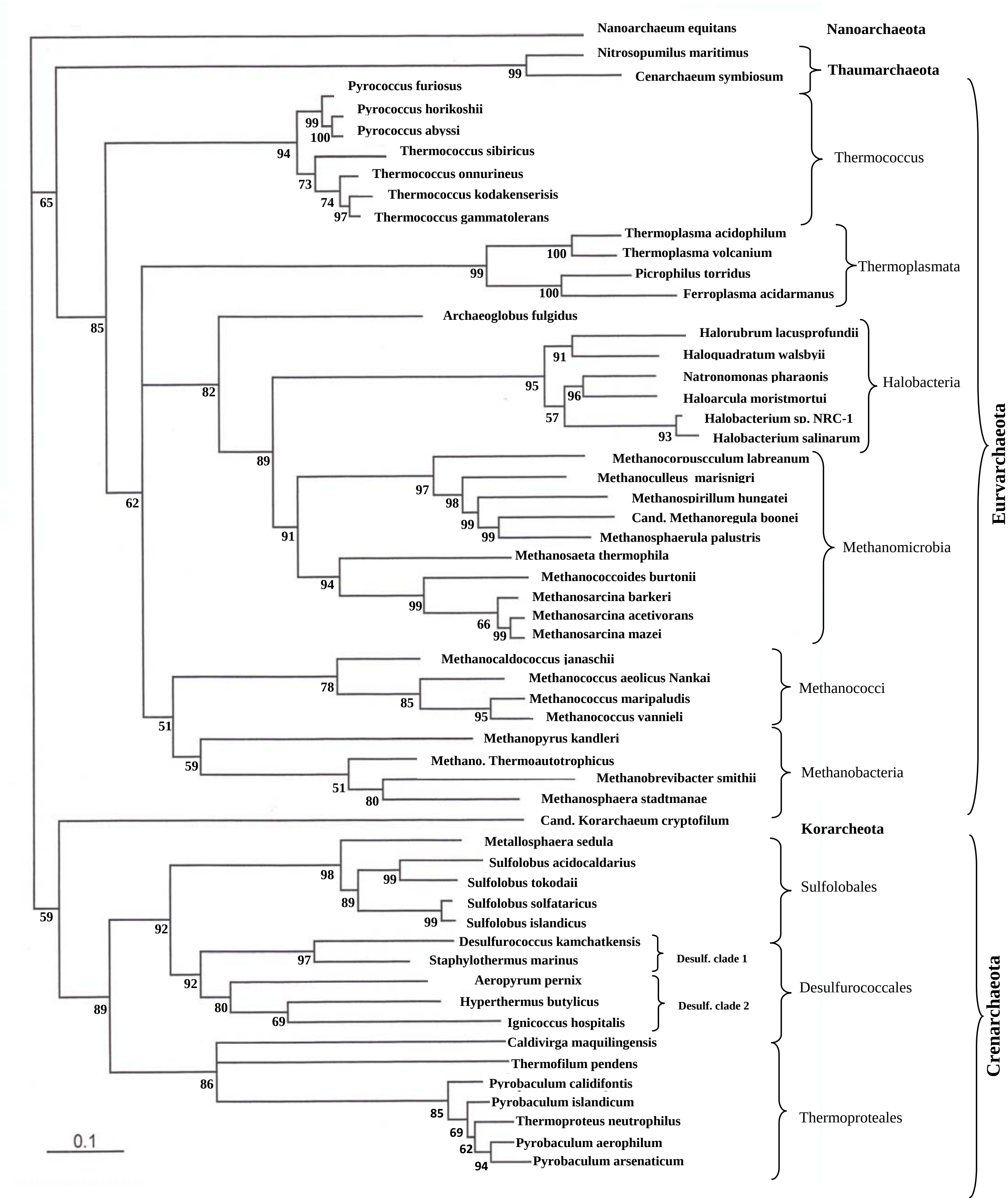
YP_254745	YP_255141	YP_255844	YP_256290
YP_254751	YP_255144	YP_255845	YP_256336
YP_254768	YP_255209	YP_255867	YP_256345
YP_254772	YP_255243	YP_255868	YP_256347
YP_254810	YP_255338	YP_255882	YP_256351
YP_254817	YP_255375	YP_255890	YP_256353
YP_254827	YP_255388	YP_255913	YP_256407
YP_254828	YP_255413	YP_255934	YP_256410
YP_254846	YP_255415	YP_255953	YP_256439
YP_254847	YP_255418	YP_255979	YP_256444
YP_254858	YP_255449	YP_255982	YP_256458
YP_254860	YP_255476	YP_256020	YP_256462
YP_254864	YP_255505	YP_256021	YP_256529
YP_254868	YP_255506	YP_256025	YP_256587
YP_254890	YP_255520	YP_256054	YP_256610
YP_254922	YP_255536	YP_256106	YP_256633
YP_254927	YP_255537	YP_256107	YP_256637
YP_254930	YP_255551	YP_256110	YP_256639
YP_254932	YP_255568	YP_256118	YP_256652
YP_254944	YP_255569	YP_256120	YP_256669
YP_254965	YP_255590	YP_256179	YP_256678
YP_254994	YP_255591	YP_256196	YP_256679
YP_254996	YP_255592	YP_256197	YP_256686
YP_255013	YP_255622	YP_256208	YP_256690
YP_255028	YP_255624	YP_256232	YP_256694
YP_255036	YP_255625	YP_256238	YP_256695
YP_255048	YP_255629	YP_256252	YP_256742
YP_255050	YP_255641	YP_256258	YP_256773
YP_255056	YP_255643	YP_256262	YP_256816
YP_255057	YP_255644	YP_256263	YP_256817
YP_255061	YP_255649	YP_256271	YP_256835
YP_255063	YP_255662	YP_256276	YP_256836
YP_255069	YP_255679	YP_256279	YP_256837
YP_255074	YP_255710	YP_256281	YP_256843
YP_255085	YP_255713	YP_256285	YP_256858
YP_255132	YP_255811	YP_256286	YP_256894
YP_255136	YP_255815	YP_256287	YP_256924
	YP_255817	YP_256288	YP_256940

(b) Proteins specific for Various Sulfolobales Except *Metallosphaera*

YP_255007	YP_255904	YP_256453	YP_256508
YP_255529	YP_256004	YP_256455	YP_256513
YP_255606	YP_256157	YP_256484	YP_256659
YP_255825	YP_256274	YP_256491	YP_256888
YP_255893	YP_256431		

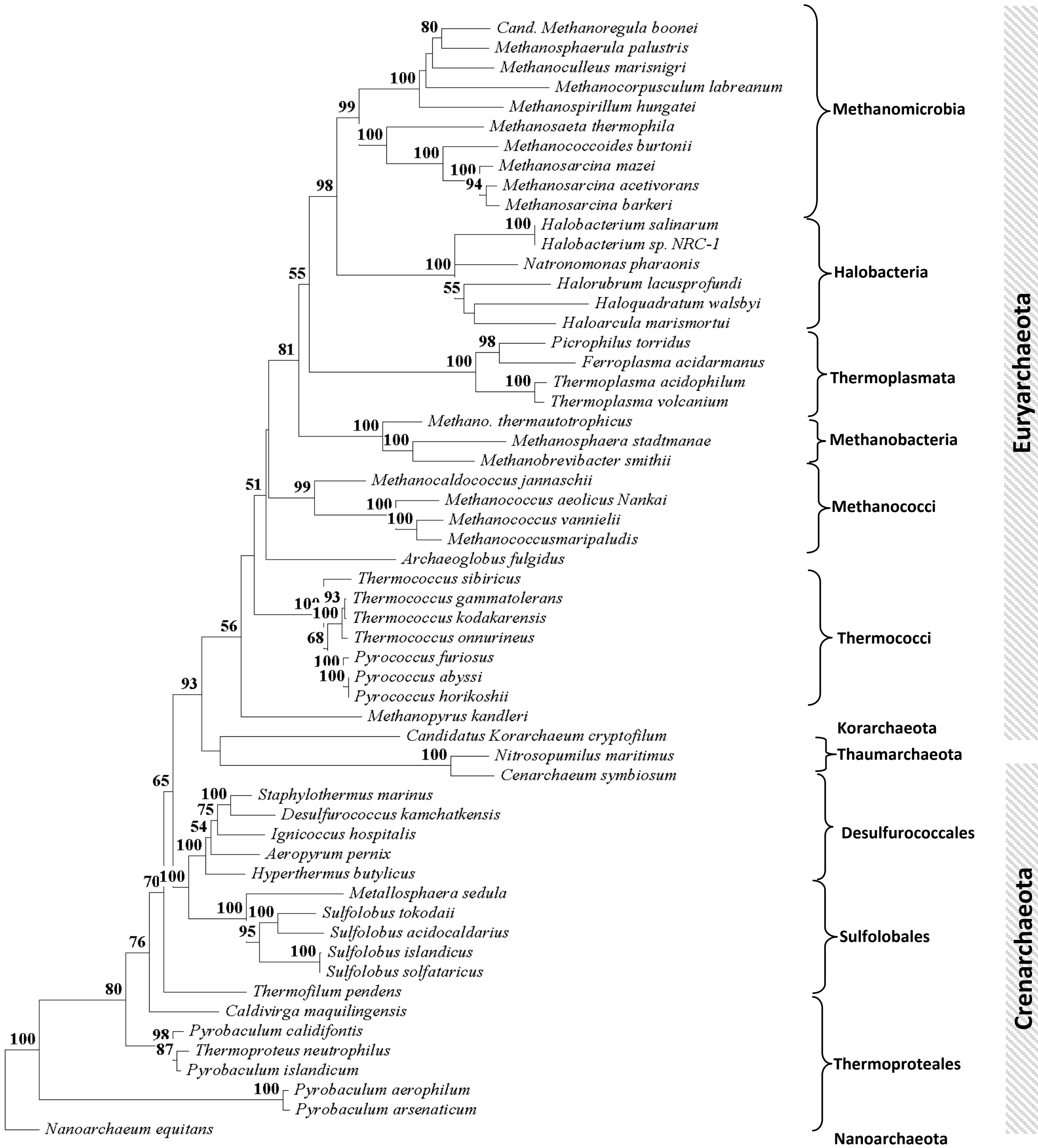
(c) Proteins Specific to a Desulfurococcus clade (Staphylothermus + Desulfurococcus)

	YP_002428054.1	YP_001041471.1	YP_001040913.1
YP_001040050.1	YP_001041538.1	YP_001041470.1	YP_001040937.1
YP_001040055.1	YP_001041513.1	YP_001041466.1	YP_001040968.1
YP_001040080.1	YP_001041496.1	YP_001041429.1	YP_001040978.1
YP_001040096.1	YP_001041481.1	YP_001041411.1	YP_001041007.1
YP_001040127.1	YP_001041479.	YP_001041386.1	YP_001041049.1
YP_001040209.1	YP_001041472.1		YP_001041130
YP_001040242.1			



Supplemental Figure 1A

A maximum-likelihood distance tree for archaeal species based concatenated sequences of 23 conserved proteins. The tree is unrooted and the % of puzzling quartets supporting the branching orders are indicated on various nodes.



Supplemental Figure 1B

A neighbour joining distance tree for the same archaeal species as in Figure 1 (or Sup. Fig. 1A) based on 16S rRNA sequences. The tree is unrooted and the bootstrap scores for nodes which were >50 are shown.

			12				59
Thermoproteaceae	Pyrobaculum aerophilum	NP_558700	GGVNSPARALKH	LP	SPLVAKAASGAYLYT		DRGKLVDYCMAFGAIILGH
	Pyrobaculum calidifontis	YP_001056600	------------	--	T---------P----		---R---F-L---------
	Pyrobaculum islandicum	YP_929640	------------	-A	A-----G---P----		---R---------------
	Thermoproteus neutrophilus	YP_001795261	------------	-A	A----NG---P----		---R---------------
	Caldivirga maquilingensis	YP_001541228	------V--M--	--	T--IV-R-E--F---	I	-GV--I----G--PL----
Desulfurococcales	Aeropyrum pernix	NP_148523	------V--AVK	-	Y-FYV-RGE------	V	-GARI--LVL-Y-PL----
	Hyperthermus butylicus	YP_001013033	------V--AVK	-	Y-FYVER-E-P-I--	V	-GE--I--VL-Y-PL----
	Ignicoccus hospitalis	YP_001435272	------I--AVK	-	Y-FYV-K----T-V-	V	-GVE-I--VLGY-PL----
Sulfolobales	Sulfolobus acidocaldarius	YP_255445	------V--AVK	-	Y-FYVER-E------	A	-NKQ-I--VLGY-PL----
	Sulfolobus islandicus	YP_002829995	------V--AVK	-	F-FYVERGK---I--	V	EGN-FI--VLGY-PL----
	Sulfolobussolfataricus	NP_341738	------V--AVK	-	F-FYVERGK---I--	V	-GK-FI--VLGY-PL----
	Metallosphaera sedula	YP_001190317	------V--AVK	-	F-FYTRS-K----V-	E	-SRR-I-FVLGY-PL----
Thaumarchaeota	Nitrosopumilus maritimus	YP_001581824	S-----V-YFEP		Y-FFT-K-N---IWD	V	-NR--I-F-NGY--LL---
	Cenarchaeum symbiosum	YP_876284	S-----V-YYAP		Y-FF-AS---GSIRD	A	-GR-YT-L-NGY--LL---
	Uncul. marine crenarchaeote HF	ABZ06901	S-----V-FF-P		Y-FFV-KSQ-SSIWD	E	-GKRYI---NGY--LL---
Other Archaea	Methanosphaera stadtmanae	YP_448204	------V--Y-P		Y-FF--S-K-SKI-D	V	-GNEYI---LGY-PLLF--
	Methanosaeta thermophila	YP_843549	---S--V--IRP		Y-FYV-R-E-P--WD	E	-GNSFI---L-Y-PM----
	Methanococcus maripaludis	YP_001548780	------V-SF-P		F-FFV-S-KDCF--D	E	-GNEFI---L-Y-PMV---
	Methanobrevibactersmithii	YP_001273806	---S--V--FEP		Y-FFV-K---SKI-D	V	-GN-YI-H-L-Y-PL----
	Methanococcus vannielii	YP_001323746	------V-SF-P		F-FFV-S-KSCF--D	E	-GNEFI---L-Y-PMV---
	Methanococcus voltae	ZP_02193844	------V--F-P		Y-FFV-S-N-CH-ID	E	-DNTYI---L-Y-PMV---
	Methanopyrus kandleri	NP_613567	---S--V-RFDP		Y-FYVER-E-SR---	V	-GHV-I---L---PL----
	Halomicrobium mukohataei	ZP_03875162	-----SV--TQP		Y-FFVERGD-GHVID	A	-GNRYL-FV-GY-PLL---
	Haloarcula marismortui	YP_137119	-----SV--TRP		Y-FFVEKGD-GHVID	A	-GNRYI-FV-GY-PLL---
	Halogeometricum borinquense	ZP_03998621	-----SV--TQP		Y-FFIERGD--HVVD	A	-GNRYL--VLGY-PLLY--
	Haloquadratum walsbyi	YP_659130	-----SV--TQP		Y-FFIERGD--TVID	A	-GTRYL--V-GY-PLLY--
	Halorubrum lacusprofundi	YP_002567263	-----SV--TMP		H-FFIERGD-GHVID	A	-GNRY--WV-GY-PLLY--
	Halorhabdus utahensis	ZP_03871814	-----SV--IQP		Y-FFVEKGD-GHVID	A	-GNRYL-FV-GY-PLL---
	Methanocorpusculumlabreanum	YP_001029966	---S--V--I-P		Y-FYV-S-R--T-E-	E	-GES---C--GY-PLL---
	Archaeoglobus fulgidus	NP_070069	---S--V--F-P		H-FYTARGK-SKI-D	V	-GNAYI-----Y-PLV---
	Candidatus Methanoregula boonei	YP_001404396	---S--V--I-P		Y-FYTAH-Q--HIT-	V	-GED-I-C---Y-PLL---
	Methanococcoides burtonii	YP_565899	---S--V--I-P		Y-FYTES-N-CRIKD	V	-GNEYI---L-Y-PN----
	Methanocaldococcusjannaschii	AAB98593	------V-YF-P		Y-FFVEK-KDC--FD	V	-GNCYI---L-Y-PMV---
	Methanocaldococcusjannaschii	NP_247585	------V-YF-P		Y-FFVEK-KDC--FD	V	-GNCYI---L-Y-PMV---
	Methanoculleus marisnigri	YP_001046897	---S--V--I-P		Y-FYVER-A-SH-T-	V	-GAD-I-C-LGY-PL----
	Thermoplasma acidophilum	NP_394045	M-----V-YF-D		Y-FYVDN---SRI-D	V	-GNEYI---L-Y-PS----
	Methanosarcina mazei	NP_633767	---S--V--I-P		Y-FYTAS-N-SKIRD	L	-GNEYI---L-Y-PA----
	Methanotherm. thermautotrophicus	NP_275371	---S--V-RFDP		H-FF-AGG--CL-ES	V	-GESYI---L-Y-PL----
	Methanosarcina barkeri	NP_305001	---S--V--I-P		Y-FYTAS-N-SKIKD	L	-GNEYI---L-Y-PAL---
	Thermoplasma volcanium	NP_111154	M-----V-YF-D		Y-FYV-EGK-STILD	V	-GN-YI---LGY-PL----
	Methanospirillum hungatei	YP_503978	---S--V--IQP		Y-FYTVRG--SRI--	V	-GTE-I-C-G-Y-PL----
	Picrophilus torridus	YP_023026	S-----V-FYRP		V-VMFERGR-SRI-D	V	NGKEYI--SLG--PM----
	Methanosarcina acetivorans	NP_615545	---S--V--I-P		Y-FYTAS-D-SKIRD	L	-GNEYI---L-Y-PAV---
	Methanococcus aeolicus Nankai	YP_001325397	-----V-YF-P		F-F-NN-KDYYL VD	V	-GN-YI---L-Y-PAV---

Supplementary Figure 2

Parital sequence alignment of the protein glutamate-1-semialdehyde aminotransferase (GSA-AT) showing two conserved indels that are specific for the Crenarchaeota species

```
                                                                      172                                227
                   Pyrobaculum aerophilum               NP_560669     GCYVCIEGPRFSTKAESRIWRE VF GCDIIGMTLVPEINLARELGMCYGLIALVTDY
                   Pyrobaculum calidifontis             YP_001055334  ---I------------------ -- --------------------------------
Thermoproteales    Pyrobaculum islandicum               YP_930103     ---I------------------ -- --------------------------------
                   Pyrobaculum arsenaticum              YP_001153955  ---------------------- -Y ------------------------M-------
                   Thermoproteus neutrophilus           YP_001794942  ---------------------- -Y --------------------------------
                   Caldivirga maquilingensis            YP_001541390  ---------------------D -Y ------------V----------A-L-VI---
                   Ignicoccus hospitalis                YP_001435687  -T------------------K- -- KA-V--------V---C-AQ---ATV-M----
                   Staphylothermus marinus              YP_001040758  -T-I---------R----V-K- -- KA--------------C-AQI--ATV-MI---
Desulfurococcales  Desulfurococcus kamchatkensis        YP_002428739  -T-I---------R----V-K- -- KA-V--------V---C-AQL--AT--M----
                   Staphylothermus hellenicus           YP_003668091  -T-I---------R------K- -- KA----------V---C-AQT--ATV-MI---
                   Aeropyrum pernix                     NP_148238     -T-----------R----V-KD -- KA--------------C-AQL--ATL-M----
                   Hyperthermus butylicus               YP_001012245  -T-I---------R----V--- -- KA----------V---C-AQL--AT-------
                   Thermosphaera aggregans              YP_003668091  -T-----------R----V-K- -- KA----------V---C-AE---ATVSM----
                   Sulfolobus acidocaldarius            YP_255650     -T-I---------R----V-K- -- KA----------V---C-AQ---ST--MI---
                   Sulfolobus solfataricus              NP_343704     -T-I---------R----T--- -Y KA----------V---C-AQ---AT--M----
Sulfolobales       Sulfolobus islandicus                YP_002828329  -T-I---------R----T--- -Y KA----------V---C-AQ---AT--M----
                   Sulfolobus tokodaii                  NP_376369     -T-I---------R---IV-K- -- KA----------V---C-AE---SV-GM----
                   Metallosphaera sedula                YP_001190788  -T-I---------R---LV-KD -- KA----------V---C-AE---ST--MI---
                   Thermofilum pendens                  YP_920025     -T-I---------R----L--Q  - -A----------V-----AR--FLN--M----
                   C. Korarchaeum cryptofilum           YP_001737931  -V---T--N--E-P--I-AFSI  L -G--V---GC--AS-----AI--AS-SV--N-
                   Nitrosopumilus maritimus             YP_001582735    ---------------KFY-T  T -A----------CQ----AQI--AS-ST----
Thaumarchaeota     Cenarchaeum symbiosum                YP_875728     -T-A---------R---ALF-A  A -A------M---CQ-S--AQ---ASVST----
                   Uncul. marine crenarchaeote HF       ABC71112        -A--------------FYKS  T -S----------CQ----AQI--VS-ST----
                   Halogeometricum borinquense          ZP_03998053   -T-------QY--R---EFY--  Q -W------TI--AK----AE-A-ATVTG----
                   Haloquadratum walsbyi                YP_657042     -T-------Q---R---EFY-S  Q -W------TI--AK----AELA-VT--G----
                   Natrialba magadii                    ZP_03692888   -T-------QY--R---EFY-D  Q -W-VV---TI--AK----AELS-AT--GI---
                   Pyrococcus furiosus                  NP_578582     -T---T-----E-A--I-AY-I  L -G-VV---QC--AI-----E---ATV-I--N-
                   Halorhabdus utahensis                YP_003129928  -V-AGMP--SYE-P--IEMLDT  V -A-AV--ST---VIV-NQ---DVVGMSTI-N-
                   Halorubrum lacusprofundi             YP_002565095  -T-------QY--R---EFYKS  Q -W-LV---AI--AK----AEIA-AT--G----
                   Thermococcus kodakarensis            YP_183895     -T-A-M-----E-R--I-ALKI  L -A-VV---QC--AA-----E---ASV-I--N-
                   Thermococcus sp. AM4                 YP_002582338  -T-A-M-----E-R--I-ALKI  L -A-VV---QC--AV-----E---ASV-I--N-
                   Thermococcus barophilus              YP_002579188  -T-A-M-----E-R--I-AL-I  L -A-VV---QC--AI-----E---SSV-I--NF
                   Thermococcus onnurineus              YP_002306510  -T-A-M-----E-R--I-ALKI  L -A-VV---QC--AA-----E---ASV-I--NF
                   Pyrococcus abyssi                    NP_126754     -T---T-----E-A--IKAY-I  L -G-VV---QC--AI-----EI--ATV-I--N-
                   Pyrococcus horikoshii                NP_143046     -T---T-----E-A--I-AY-I  L -G-VV---QC--AA-----E---ATV-I--N-
                   Methanobrevibacter smithii           YP_001273238  -T---TQ----E-P--IKMFKL  L -G-LV---G---VT----RAI--NS-CI-SN-
                   Methanotherm. thermautotrophicus     NP_276708     -V---T-----E-P--I-MFQT  L -GTVV---GL--AV-----E---AS-C--SN-
                   Methanopyrus kandleri                NP_613308     AV--AT-----E-P--I-AF-K  L -G--V---GF--VV-----EI--ASVC-C-N-
                   Methanosphaera stadtmanae            YP_447683     -I-IAT-----E-G--IQFYKL  I -GKVV---G---VV--K-KQ---SS-CA--N-
                   Picrophilus torridus                 YP_023720     -T-----------R----MF-Q    FG--------------D--A----M--T----
                   Pyrococcus horikoshii                NP_142133     -T-----------R----MF-Q    FA-V--------V----------VN-ST----
Other Archaea      Pyrococcus furiosus                  NP_577745     -T-I---------R----MF-Q    FA-V--------V----------VN-ST----
                   Pyrococcus abyssi                    NP_125816     -T-I---------R----MF-Q    FA-V--------V----------VN-ST----
                   Thermococcus kodakarensis            YP_184308     -T-----------R---FMF-Q    FAH---------V----------VN--T----
                   Thermococcus sp. AM4                 YP_002581816  -T-----------R---FMF-Q    YAH--------------------VN--T----
                   Thermococcus onnurineus              YP_002308038  -T-----------R---FMF-Q    YAH--------------------VN--T----
                   Aciduliprofundum boonei              YP_002578821  -T-----------R----MF-Q    FA----------VQ-----E---VN-STI---
                   Thermoplasma acidophilum             NP_393811     -T-IT--------R---KMF-Q    FA----------VS--G--AL--SV--SI---
                   Thermoplasma volcanium               NP_110957     -T-IT--------R---KMF-Q    FA----------VG--G--AL--SV--SI---
                   Thermococcus barophilus              YP_002580476  -T-----------R---AMF-Q    FAH---------VV---------VN--TI---
                   Methanococcus maripaludis            NP_987305     -V---T-----E--T-IQ-YKN    WGHVV---AY--VV----ME---TS-CN-SN-
                   Methanosaeta thermophila             YP_843644     -V---T---H-E-H-QIKML-S    FG-VV---GY--VV----R-L--AS-C-A-N
                   Methanocaldococcusjannaschii         NP_247024     -V---T-----E--K-IA-YKN    WG-VV---GY--VV-----E---VSLCNI-N-
                   Methanococcoides burtonii            YP_564942     -T---T-----E-R--IAMMST    FG-VV---GL--VV-----EL--AS-CTI-N
                   Methanococcus vannielii              YP_001323697  -V---T-----E--T-IQMYKP    LGHVV---AY--VV----MEI--TS-CN-SN-
                   Methanococcus aeolicus Nankai        YP_001325119  -T---T-----E----INFYKT    IG--V---AY--VV-----QL---S-CT-SN-
                   Methanosarcina mazei                 NP_634403     -I-A-T-----E-R--I-MMSQ    FA-VV---GI--VV--K---L--ASL-I--N
                   Methanosarcina barkeri               YP_307042     -I-A-T-----E-R--I-MMSQ    FA-VV---G---VG--K--SL--ASL-I--N
                   Methanosarcina acetivorans           NP_616344     -V-A-T-----E-R--I-MMSQ    FA-VV---G---VV--K--SL--ASLSI--N
```

Supplemental Figure 3

Partial sequence alignment for 5'-methylthioadenosine phosphorylase II (MTAP II) showing a 2 aa insert that is present in various Crenarchaeotes except *T. pendens.* A 1 aa insert is also present in the same position in *T. pendens*, Korarchaeota, various Thaumarchaeota and a number of other archaea.

			250		297
Thermoproteales	Pyrobaculum aerophilum	NP_560393	KTDLTHPDPSVRDWVAFRIIDTS	KTPH	PLTGDKYIVWPTYNFAVSIDD
	Pyrobaculum islandicum	YP_929892	----S------------------	----	--V------------------
	Pyrobaculum calidifontis	YP_001056028	----S------------------	----	--V------------------
	Pyrobaculum arsenaticum	YP_001153553	----A-----------------T	----	---------------------
	Thermoproteus neutrophilus	YP_001795120	----S-------------V----	----	-----R-V-------------
	Caldivirga maquilingensis	YP_001541346	----SS----I-------V--AH	-Y--	-RV--------------A---
	Thermoproteus tenax	CAF18456	---------------------VE	RN--	--V-A--H-------------
	Thermofilum pendens	YP_920220	---PA--N------I---VL--E	----	-------R--------CAV--
Desulfurococcales	Ignicoccus hospitalis	YP_001435694	----R------------------	-H--	-------VL-------AGV--
	Desulfuro. kamchatkensis	YP_002429047	----Q-------------V----	----	-V--E-----------AGV--
	Staphylothermus marinus	YP_001040833	----S-----------A------	----	-IV-----L-----L-AA---
	Hyperthermus butylicus	YP_001012747	---PR------------------	RY--	--V--R-VA-------AAV--
	Aeropyrum pernix	NP_148534	---PR--N--L-----M-----E	-H--	--V-SR-L----------V--
	Staphylothermus helenicus	YP_003669627	----S-----------A----TS	----	P-V-----LW----LA—IDD
	Thermosphaera aggregans	YP_003650597	----Q------------------	RTPH	P-V-E-F---------GVDD
Sulfolobales	Metallosphaera sedula	YP_001192194	--NPED----KI---ML-----K	RN--	-IA---FR--------TAV--
	Sulfolobus tokodaii	NP_377381	----SL----Q----LL----VK	-N--	-I-----WI-------SAL--
	Sulfolobus solfataricus	NP_341660	----SD----QI---ML-----A	-N--	-RV-S--W--------SI---
	Sulfolobus acidocaldarius	YP_256131	----KL----Q----LL-V-NVK	-S--	-IE----W--------SA---
	Sulfolobus islandicus	YP_002838328	--N-MD----QI---ML-----T	-N--	-RV-S--Y--------SI---
Methanomicrobia	Methanosarcina mazei	NP_633773	---IE-K--AL---G----RKM-	H-R	-EI-N------LLD--GA-E-
	Methanosarcina barkeri	YP_305007	---IK-K--AL---G----RKM-	H-R	-EI-N--V---LLD--GA-E-
	Methanosarcina acetivorans	NP_615551	---IE-K--AL---G----RKM-	H-R	AEI-N--V---LLD--GA-E-
	Methanococcoides burtonii	YP_566939	---IE-K--AM--FG----VK-A	H-R	-EVD---V---LLD-EGA-E-
Thermococci	Thermococcus onnurineus	YP_002306571	----K--N-A----P-L----NP	S	-R-----R---L----SA---
	Thermococcus barophilus	YP_002579495	----N--N-A----P-L----NP	D	-R-----R---L----SA---
	Thermococcus sp. AM4	YP_02582833	----K--N-A----P-L----NP	N	-R-----R---L----SA---
	Thermococcus kodakarensis	YP_183821	----N--N-A----P-L----NP	N	-R--N--R---L----SA---
	Pyrococcus horikoshii	NP_143530	----N--N-A----P-L----NP	N	-R--N--R---L----SA---
	Pyrococcus furiosus	NP_579482	----N--N-A----P-L--V-EP	N	-R--N--R---L----SA---
	Pyrococcus abyssi	NP_126175	----N--N-A----P-L--V-NP	N	-RA-N--R---L----SA---
	Aciduliprofundum boonei	YP_002577198	----Q--N-AI---A-----KDV	E	--RL-N----Y-LM--S-A--
Other Archaea	Cenarchaeum symbiosum	118576798	RG-MESQNTVM--P-L---MEER		--L--RHR---S-DM--A-E-
	Nitrosopumilus maritimus	161527819	RG-MKADNAVM--P-LL----EK		YTV----RI--S-D---A-E-
	Nanoarchaeum equitans	41615091	------K--G---FPIA--VENP		-RI-Y-PV F-L-----V---
	Thermoplasma volcanium	NP_111605	----N--N------I------AK		-R-----R-F-MMS-S-AV--
	Picrophilus torridus	YP_023656	----N--N--I---I----SYKR		-----E-C-Y-MM--S-A---
	Archaeoglobus fulgidus	NP_069098	---MR-K--AI---------KE-		--V----V-Y--LD-ESA-E-
	Methano. thermautotrophicus	NP_275194	----N--N-AI----SM--VEAE		-R--TR-R-Y-MM--S-AV--
	Methanocaldococcus jannaschii	NP_248380	---IK-KN--I--FPI--VEK-P		-R-----C-Y-LM--S-PV--
	Thermoplasma acidophilum	NP_394401	----K--N------I-----EAR		-RVD---R-Y-MMS-S-AV--
	Methanobrevibacter smithii	ZP_03704661	---IN-KN-AI-----M--VEET		-RL-N--R-Y-MM--S-AV--
	Methanococcus voltae	ZP_02193119	---I--KN--I--FPI---EH-P		-KN-T--V-Y-LM-LS--V--
	Methanosaeta thermophila	YP_843880	---IK-K--AM---AG---VTK-		--V-ER-R---LLD-ESAVE-
	Methanosphaera stadtmanae	YP_447954	---IN-KN-AI-----M--VNQE		-RL-N--KIY-MM--S-TV--
	Methanopyrus kandleri	NP_614041	--EVD----A--E-I----VEEE		-M--SR-L----M----AV--
	Methanococcus vannielii	YP_001322859	---IG-KN--I--FPI---EK-P		-KN-T--H-Y-LM-LS-T---
	Methanococcus maripaludis	YP_001549692	---IA-KN--I--FPI---ER-P		-KN-T--H-Y-LM-LS-TV--
	Methanococcus aeolicus Nankai	YP_001325299	---IE-KN--I--FPI---EN-P		-K--T--R-Y-LM-LS--V--
	C. Korarchaeum cryptofilum	YP_001737504	---MS--NVTM--PGI--V-EAE		-IH----R-Y-V-D-S--VM-

Supplemental Figure 4

Partial sequence alignment of glutamyl-tRNA synthetase (GluRS) showing a 4 aa insert found in all Crenarchaeota. In the same position, a 3 aa insert is also present in Methanomicrobia and a 1 aa insert in Thermococci.

			54		90
Sulfolobales	Sulfolobus solfataricus	15899532	GIGCSGKIPHFIRL	PA	SGVHTLHGRALTFAIGIKLAN
	Metallosphaera sedula	146303307	-------T---F--	-I	-------------------S-
	Sulfolobus tokodaii	15922631	-----------F-T	-I	----------IA--T----S-
	Sulfolobus sp.	1565184	-----------F-T	-I	----------IA--T----S-
	Sulfolobus acidocaldarius	70608016	-------M---VN-	-V	G----------A--T------
Thermoproteales	C. Micrarchaeum acidiphilum	255513384	-----------VS-	-I	----------MA--L---V--
	Caldivirga maquilingensis	159041358	-----------MNI		E--------SIPY--------
	Pyrobaculum aerophilum	18313553	-------T--Y-NA		G----------P--Q---I--
	Pyrobaculum calidifontis	126458733	-----SRL---VKT		TS--AI----IPY-M------
	Pyrobaculum arsenaticum	145590403	-----SRL---V-T		TS--AI----IPY-L------
	Pyrobaculum islandicum	119872678	-----SRL---V-T		TS--AI----IPY-L------
	Thermofilum pendens	119719765	----H-R-SDYMKT		N-F--I---V-PL-T------
	Thermoproteus neutrophilus	171184614	-----SRL--YV-A		TS---I----IPY-L------
	Thermoproteus tenax	41033655	-----SRL--YV-T		ISA--I----IP----L----
Desulfurococcales	Hyperthermus butylicus	124027855	-------T--Y-NA		N-F-------IP--T------
	Aeropyrum pernix	14601863	-----S-T----NV		N----I---GIA--T------
	Ignicoccus hospitalis	156937865	------R-SG--KV		DSM--T---PVA--T------
	Desulfurococcus kamchatkensis	218883845	-----ARMAFYVSY		DSA-V-----IP--T-V----
	Staphylothermus marinus	126466332	----H--LANY-NV		NSI--I---V-PMMT---AV-
	Thermosphaera aggregans	296242995	-------T----NV		N---------IPF-T----S-
Thaumarchaeota	uncultured marine crenarchaeote	A167044494	-------TS---NV		Y-I------V----Q-A-I--
	Nitrosopumilus maritimus	161527922	-------TS-Y-NT		Y--------V----Q-A-I--
	Cenarchaeum symbiosum	118575508	-------TS-YVNV		Y-I------V----Q-A----
Other Archaea	Ferroplasma acidarmanus	257077174	-------T--YLNI		A-T-------IPY-V-V----
	Methanopyrus kandleri	20094166	-----SRL-GYVK-		DSL--T---P-A--T------
	Methanosphaera stadtmanae	84490171	-----SR--GYVKC		DSL--T----IA--T-L-I--
	C. Parvarchaeum acidiphilum	269986533	-----S-L--LVNV		N--------PIP--E------
	Picrophilus torridus	48478071	-------T--YLNI		A-A-------IPY-V-V--S-
	Thermoplasma acidophilum	10640051	-------T--YVNA		A-A-------IPV-V-V--T-
	Thermoplasma volcanium	13541678	-------T--YVNA		A-A-------IPV-V-V--T-
	Halomicrobium mukohataei	257389127	--------GTYMHS		YAL-GV-----PVG--V----
	Thermococcus kodakarensis	57640751	---QAA-M--Y-NV		N-Y-------IPI-T-V-AS-
	Thermococcus barophilus	254170697	---QAA-M--Y-NV		H-Y-------IPI-TAV-AS-
	Thermococcus gammatolerans	240103433	-----SRV-G-VNF		D-L--T-----A--T------
	Thermococcus sp. AM4	254173809	-----SRV-G-VNF		D-L--T-----A--T------
	Thermococcus onnurineus	212223731	-----SRV-G-VNF		D-L--T-----A--T------
	Thermococcus sibiricus	242399442	-----SRV-GYVNF		D-L--T-----A--T---M--
	Pyrococcus abyssi	14520724	-----SRV-G-VNF		D-L--T-----A--T------
	Pyrococcus furiosus	18978144	-----SRV-G-VNF		D-L--T-----A--T------
	Pyrococcus horikoshii	14591430	-----SRV-G-VNF		D-L--T-----A--T------
	Methanobrevibacter smithii	261350217	-----SRV-GYVKC		DSL--T-----S--T-L-VS-
	Archaeoglobus fulgidus	11498079	------R-TQYLDF		ET---T-----A--T-----
	Methanothermo. thermautotrophicus	15679052	-----SR--GYVNC		DSL--T---PIS--T-L--G-
	Methanospirillum hungatei	88601412	-----SRA-GY-LT		DSL--T----IA--T-V-M--
	Can. Korarchaeum cryptofilum	8170290876	------WT-LYF-S		DVA--------P----L---
	Methanococcus maripaludis	45357868	---QA--M--Y--V		N-F-------IPA-TAV-AT-
	Methanocaldococcus infernus	241904815	-----SRV-GYFN-		DSL--T-----A--T---VS-
	Methanococcus aeolicus Nankai	150401889	-----SR--GYFEC		DSI--T-----A--T-L-TF-
	Natronomonas pharaonis	76802656	-------LSSYFES		Y-F--------PI-RAA----
	Halobacterium sp. NRC-1	15789708	--------GTYM-S		YAL-GV-----PVGT-V----
	Haloarcula marismortui	55377568	--------GTYMHS		YAL-GV-----PVG--V-M--
	Natrialba magadii	224819551	-------LNSYLDT		Y-F--I-----PV-RAA----
	Halogeometricum borinquense	227879998	--------GTYMHS		YAL-GV-----PVGA-V----
	Halorubrum lacusprofundi	222479321	-------LNSYLDT		Y-F--I---S-PV-RAA----
	Haloquadratum walsbyi	110667494	--------GTYMHS		YAL-GV-----PVGT-V----
	Halorhabdus utahensis	257051981	--------GTYMHS		YAL-GV-----PVGT-V----
	Halomicrobium mukohataei	257388208	-------LNSYFNS		Y-F--I---S-PV-RAA----
	Ferroglobus placidus	260569185	----H---ADY-E-		PSF-VI---VIP-LT-L----
	Can. Methanoregula boonei	154150272	-----SRA-GY-VT		DSL--T-----A--T-V-M--
	Can. Methanosphaerula palustris	1219852543	-----SRA-GY-QT		DSL--T-----A--T-V---
	Methanoculleus marisnigri	126178793	-----SRA-GY-LT		DSL--T-----A--T-V-M-

Supplemental Figure 5

Partial sequence alignment of 2-oxoglutarate ferredoxin oxidoreductase (OGOR) subunit beta showing a 2 aa insert that is only found in Sulfolobales and the partially-sequenced *C. Micrarchaeum acidiphilum.*

			211		250
Sulfolobales	Sulfolobus tokodaii	NP_378065	VIAFLDGDHGGDLILKELL	A	NNVKIDYVARAPTGKEVEEL
	Sulfolobus acidocaldarius	YP_254795	-------------------	S	S-I-L---------------
	Sulfolobus islandicus	YP_002830080	-------------------	S	------F-----V-R-----
	Sulfolobus solfataricus	NP_341650	-------------------	S	------F-----I-R-----
	Metallosphaera sedula	YP_001192216	------------M---D-I	N	A-----------V-R-----
	Nanoarchaeum equitans	NP_963628	-TL-I---R--E-V-RN--	-	AGAD-----V--P-R-----
Thaumarchaeota	Nitrosopumilus maritimus	NP_13453639	-V------RA-GF-----K		SV-P---EL--DNDV-----
	Uncul. marine crenarchaeote HF	167044857	-V------RA-GF-----K		SV-N--IEL--DP-L-----
	Cenarchaeum symbiosum	11887585	-V--M---RS-GF-----K		TL-SV-ESI--DPDT-----
Desulfurococcales	Staphylothermus marinus	YP_001040897	--V----------------		RTI-V-FI----E-R-----
	Staphylothermus hellenicus	YP_003669566	--V----------------		RTI-V-FI----E-R-----
	Desulfurococcus kamchatkensis	YP_002428985	-V---------M--V-		R-A----I----P-R-----
	Aeropyrum pernix	NP_148382	--L-V---R--E----NV-		PQM-V-------E-R---S-
	Ignicoccus hospitalis	YP_001434652	---V---R--LMV--N--		SQAD--------P-----D-
	Hyperthermus butylicus	YP_001013084	---V----------W---		KVAD-----K--P-------
	Thermosphaera aggregans	YP_003650464	ML-V------------VL		R-L---F-----A-R-----
Thermoproteales	Pyrobaculum calidifontis	YP_001055941	-I---K--EMV----M		KVAH--------P-----Q-
	Pyrobaculum aerophilum	NP_559773	-TV-I---K--E-V-R---		KVAHV--I----P-----Q-
	Pyrobaculum arsenaticum	YP_001154089	-TV-I---K--E-V-----		KVAHV--I----P-----Q-
	Pyrobaculum islandicum	YP_930479	-TV-I---K--EIV-R--I		KVAH---I----P-----Q-
	Thermofilum pendens	YP_919911	--L-V---R--E--ARNVI		-AL-V-F-----P-R---D-
	Thermoproteus neutrophilus	YP_001794694	-TV-I---K--E-V-R---		KVAH---I----P-----Q-
	Caldivirga maquilingensis	YP_001540008	T---T---R--EM--R---		KVAD---I----P-----Q-
Other Archaea	Archaeoglobus fulgidus	NP_070724	-T------R----------		QVAEV-------E-----D-
	Aciduliprofundum boonei	YP_002577885	-T-----------------		QVAD--FI----P-Y-----
	C. Korarchaeum cryptofilum	YP_001737620	---V---R--EMV-RT--		EQTD--------E--V----
	Thermococcus barophilus	YP_002579483	-T--T---R--E-------		QVAD--------E-------
	Thermococcus kodakarensis	YP_183823	-T--T---R--E-------		QVADV-------E-------
	Pyrococcus furiosus	NP_579454	-T--T---R--E-------		QVAD--------E-------
	Pyrococcus horikoshii	NP_143542	-T--T---R--E-------		QVAD--------E-------
	Pyrococcus abyssi	NP_126166	-T--T---R--E-------		QVAD--------E-------
	Thermococcus onnurineus	YP_002306570	-T--T---R--E-------		QVADV-------E-------
	Thermococcus gammatolerans	ACS34278	-T--T---R--E-------		QVADV-------E-------
	Thermococcus sp. AM4	YP_002582832	-T--T---R--E-------		QVADV-------E-------
	Methanopyrus kandleri	NP_613902	-T--V---R--E-------		QVADV----K--K-------
	Methanotherm. thermautotrophicus	NP_276028	-T--V---R--E-------		QVGE----T---K-----D-
	Thermoplasma volcanium	Q97CE5	-TV-----R--------M-		QVAEV-F-----P-T-----
	Methanococcus maripaludis	YP_001329750	--T---R--E-------		QTCD--------Y-----
	Picrophilus torridus	YP_023395	-TL-V---------I--M-		QVADV-FI----P-T-----
	Methanococcus voltae	ZP_02192327	-T---R--E-------		QTCDV-------Y------
	Methanocaldococcus jannaschii	NP_248201	-TV-T---R--E-------		QVCDV-F-----P-------
	Thermoplasma acidophilum	Q9HLX7	-TV-----R--------M-		QVAEA-F-----P-T-----
	Methanococcus vannielii	YP_001323114	--T---R--E-------		QTCD----S---Y-----
	uncultured methanogenic	YP_685076	-TV-T---R--E-------		QVADV-------D--G----
	Methanosphaera stadtmanae	YP_448376	-T------R----------		QIGD----T---R-L---Y-
	Methanobrevibacter smithii	ZP_03703689	--V---R--E-------		QIGDV--IT---K-----D-
	Methanococcoides burtonii	YP_566409	-T--T---R--K--I----		QVAD---I----D--S--D-
	Methanosaeta thermophila	YP_843152	--T---R--E--V----		QVAD--------E--C----
	Methanosarcina acetivorans	NP_619489	-T--T---R--E--IR---		QVAD--------D--C--D-
	Methanosarcina mazei	NP_633319	-T--T---R--E--IR---		QVAD--------D--C--D-
	Natrialba magadii	ZP_03693707	-T------R------E--S		QVGD-----F--S-TS----
	Halorhabdus utahensis	ZP_03870019	-TT-----R---------A		QVGSV----F--A-RA--D-
	Methanosarcina barkeri	YP_304511	-T--T---R--E--IR---		QVAD--------D--C--D-

Supplemental Figure 6

Partial sequence alignment of DNA primase showing a 1 aa conserved insert that is found only in Sulfolobales and Nanoarchaeota.

			106		150
Thermoproteales	Pyrobaculum aerophilum	NP_559473	TAASHDGIASPYVS	YT	LSQKLQTYGKIVASPVAIIADTSVILSAP
	Pyrobaculum arsenaticum	YP_001154030	-S------------	-A	----MASH-------I------T---N--
	Pyrobaculum calidifontis	YP_001055459	-S------------	FS	-T-RFSK----SS----------I-----
	Pyrobaculum islandicum	YP_931183	-S------------	--	-A--MIK----M---I--V----I--N--
	Thermoproteus neutrophilus	YP_001794096	-S----------I-	-I	-AE-VKR---V----I--V----I-----
	Thermoproteus tenax	CAP05013	-S------------	-V	-NRR-SER---AV--T--L---TI-MG--
	Thermofilum pendens	YP_920632	--I----F---I-A	L	KDPEGNPLSIFTRP-A-VLV-LA-VSR--
	Caldivirga maquilingensis	YP_001541777	-V----------I-	FLM	QDDLNKLGV--VRKT-L---V--GIVAE-
Desulfurococcales	Staphylothermus marinus	YP_001040946	-----------FA-		-KGTSKPTSTPTTT-Y-----IDM-SK--
	Aeropyrum pernix	NP_147296	-V------T--FS-		-KGFDKPISRPAKA-E---I-VD--AE--
	Ignicoccus hospitalis	YP_001435758	--T--------FS-		-KGLKGPTSVKTKT-I-VV--IE--AR--
	Desulfurococcus kamchatkensis	YP_002428982	--P--------FA-		-KGTSRPTSIYTRT-F------DL-ST--
	Staphylothermus hellenicus	YP_003669513	-----------FA-		-KGTDKPTSTPTTT-P------D--SK--
Sulfolobales	Metallosphaera sedula	YP_001192317	--P-----T--FA-		IKGLGKPVSVKAKM-Y-----IN-LS---
	Sulfolobus acidocaldarius	YP_255321	-SP-----T--FAA		IKGLGKPISIKAKE-L---S-VEILA---
	Sulfolobus islandicus	YP_002832130	-SP-----T--FA-		IKGLQKPVSVKAKE-L---V-IEILSLS-
	Sulfolobus solfataricus	NP_342265	-SP-----T--FA-		IKGLQKPVSVKAKE-L-----IEILSLS-
	Sulfolobus tokodaii	NP_376213	--P-----T--FA-		IKGLGKPTSVKAKG-I-----IN-LA---
Other Archaea	Nitrosopumilus maritimus	YP_001583063	-------VS--F--		VKSDKPHSIVATA-LGVFV-IDI-KK--
	Cenarchaeum symbiosum	YP_876473	-----------F--		IKGDKPHSITATA-LGVFV-IG--RK--
	Archaeoglobus fulgidus	NP_070502	-T---------VA-		FKENGKPISISTNP-S-V---LNI-KNC-
	Halorhabdus utahensis	ZP_03870432	--------V-GRA-		VPEDGTRHSVAANP---VV---GILAD--
	Methanopyrus kandleri	NP_614513	-S---------FA-		IRREGRP-SEPAQA-L--L--IE--RE--
	Methanosarcina barkeri	YP_303855	----------SRA-		IVDNGRSTSAQAQA-I-V----EI-SA--
	Methanosphaera stadtmanae	YP_448028	-T-A---M---RA-		IKNDKG-VSLKANA-F-L----TI-SQ--
	Methanospirillum hungatei	YP_502604	--------I-GRAT		-PTETGSVSVGAHP-I-VV---GI-SQ--
	Pyrococcus furiosus	NP_579111	-T--------ANA-		IKGLGTKTSIKARP---V---IRI-K---
	Pyrococcus horikoshii	NP_143340	-T--------ANA-		IKGLNVKTSIKAKP-I-V---ID--KT--
	Thermococcus kodakarensis	YP_183202	-T--------ANA-		IKDLGSKTSVKAVP---V---VK--KT--
	Thermoplasma volcanium	BAB60393	-SP--------RA-		-RRNGIS-SEEG-M-IGV------MIK--
	Picrophilus torridus	YP_023628	-TP--------RA-		IYDGKSV-SEEATM-S--V----IMVL--
	Thermoplasma acidophilum	NP_394614	--P--------RA-		-RRNGIS-SEEG-M-IGV----AIMIK--

Supplementary Figure 7

Partial sequence alignment of glycerol dehydrogenase showing an insert that is mainly found in the Thermoproteales.

(A)

			207		246
Thermoproteales	Thermoproteus neutrophilus	171186023	PEINLARELGMCYGLIALVTDYDVWV	P	HQPVTAEAVEKMM
	Pyrobaculum aerophilum	18314002	-----------------------I--	-	-------------
	Pyrobaculum calidifontis	126459056	-----------------------I--	-	-------------
	Pyrobaculum arsenaticum	145591953	--------------M--------I--	-	----------R--
	Pyrobaculum islandicum	119872096	------------------------L-	-	-------------
	Caldivirga maquilingensis	159042138	--V----------A-L-VI----I--	-	-------L---T-
	Thermofilum pendens	119719530	--V-----AR--FLN--M-------A		EK----HE-ARV-
Desulfurococcales	Aeropyrum pernix	14601697	------C-AQL--ATL-M-------A		DR-----E--RV-
	Hyperthermus butylicus	124026925	--V---C-AQL--AT----------A		ER-----E-ARV-
	Desulfuro. kamchatkensis	218884357	--V---C-AQL--AT--M-------A		EK-----E-V-T-
	Staphylothermus marinus	126465649	------C-AQI--ATV-MI------A		EK-----E-IRV-
	Ignicoccus hospitalis	156937891	--V---C-AQ---ATV-M-------A		ER----QE-IET-
	Staphylothermus hellenicus	297526067	-V---C-AQT--ATV-MI------A		EK-----E--RV-
Sulfolobales	Sulfolobus solfataricus	15899099	--V---C-AQ---AT--M------FA		EI-----E-TRV-
	Sulfolobus islandicus	227826550	--V---C-AQ---AT--M------FA		EI-----E-TRV-
	Metallosphaera sedula	146303472	--V---C-AE---ST--MI-----FA		EV-----E-T-V-
	Sulfolobus acidocaldarius	70606780	--V---C-AQ---ST--MI-----FA		EL-----E-T-V-
	Sulfolobus tokodaii	15920700	--V---C-AE---SV-GM------FA		DI-----E-T-V-
	Sulfolobus solfataricus	241769285	--V---C-AQ---AT--M------FA		EIT----E-TRV-
Other Archaea	Nitrosopumilus maritimus	161528909	--CQ----AQI--AS-ST-------A		EK----KE-
	Cenarchaeum symbiosum	118575985	--CQ-S--AQ---ASVST-----A-A		EKA---KE-
	Picrophilus torridus	48478014	------D--A----M--T-------S		EK--DTNE-I-I-
	Ferroplasma acidarmanus	257075852	------A--S----ML-T-----A-K		DEA-E-SD-MQI-
	Thermoplasma volcanium	13541269	--VG--G--AL--SV--SI------S		EK--D-RE-MTI-
	Thermoplasma acidophilum	16081464	--VS--G--AL--SV--SI------S		TK--D-RE-MNI-
	Thermococcus gammatolerans	240102193	--V----------VN--T-------A		EK--D-QE-L-V-
	Thermococcus kodakarensis	57641830	--V----------VN--T-------A		DK--D-QE-L-V-
	Pyrococcus horikoshii	14590069	--V----------VN-ST-------A		EK--D-QE-LRV-
	Pyrococcus furiosus	18976388	--V----------VN-ST-------A		EK--D-QE-LRV-
	Thermococcus sibiricus	242399954	--VV---------TN--A-------A		EK--D-QE-I-V-
	Aciduliprofundum boonei	254167286	--VQ-----E---VN-STI------A		EK---T-E-MRV-

(B)

			70		111
Thermoproteales	Pyrobaculum aerophilum	18313250	EVVRIFAATANPLKVIVAEVEP	E	RRGVVGVVDGHSPLGVETE
	Pyrobaculum islandicum	119872532	----------------------	-	-------------------
	Pyrobaculum calidifontis	126460257	--------------I-------	-	----I--------------
	Pyrobaculum arsenaticum	145591729	-----Y----------------	-	-------------------
	Thermoproteus neutrophilus	171185911	----------------------	-	----I-------------D
	Caldivirga maquilingensis	159041605	--T--L---G---SI----IA-	-	----------Y------
	Thermofilum pendens	119719593	--AS--C-----VQ-----TQQ		G---M--I--FTVV---S-
	C. Parvarchaeum acidiphilum	269987071	---N-Y---------L-YDDGM	D	G-S-I--I--KK---T-RD
Sulfolobales	Sulfolobus tokodaii	15921805	-------------Q-L---TDQ		G---I-----YT---I---
	Sulfolobus islandicus	227827754	-------------Q-----TDQ		G--II-----YT-----S-
	Sulfolobus solfataricus	15897561	------V------Q-----TDQ		G---I-----YA-----S-
	Sulfolobus acidocaldarius	70607363	-------------Q-L--QTDQ		G--II-----FT-----GD
	Metallosphaera sedula	146304745	-----H-------Q-V---T-Q		G---L--I--MT-----S-
Desulfurococcales	Staphylothermus marinus	126465427	---N-YV-----VQ-----TDQ		G-A-I-----YT-----GD
	Staphylothermus hellenicus	297526290	---N-YV-----VQ-----TDQ		G-A-I-----YT-----GD
	Aeropyrum pernix	118431815	---S-D------V--V--DLGE		Q-ALL--A--FK-----G-
	Desulfuro. kamchatkensis	218884112	---SLYV--G--VQ---G-TNQ		G-VIL-----YP-----S-
	Hyperthermus butylicus	124028089	---T-----G---Q---V-TKQ		G---------FT------D
	Thermosphaera aggregans	296242823	-I-QLYV-----VQ----DTDQ		GSAII-----YP-----S-
Other Archaea	C. Korarchaeum cryptofilum	170290759	-IAT-H-------------FGD		G--IL--I--GK-----
	Aciduliprofundum boonei	254167706	-IT--MV-S----Q--IGKT-Q		GNAIL-I---YA-----D-
	Picrophilus torridus	48478083	--GT---------S---Y-GSN		GN-IL--I--Y----I-D-
	Thermoplasma acidophilum	10640696	--GS---------QI--YKG-R		GN--L--I--Y--V---SD
	Thermoplasma volcanium	13541328	--GG-------S--I-IYKG-N		GN--L--I--M--V---SD
	Ferroplasma acidarmanus	257075597	-LGA--------I---L-KT-N		TNA-I--A--Y--V----D
	Ferroglobus placidus	260570537	A-CNVY--S---ME-----T-L		G-A-L-----QAVNAI-N-
	Pyrococcus furiosus	18976620	A-AMVYG-SE--FQ-----T-L		G-A-I-----KAATKI---
	Pyrococcus horikoshii	14591660	A-AM-YG-SE--FQ-----T-L		G-A-I-----KAANKI--D
	Thermococcus kodakarensis	57642047	A-AMVYG-SE--FQ-----TDL		G-S-L-----KAANKI---
	Thermococcus sibiricus	242399831	A--MVYG-SE---Q--I--T-L		G-A-L-I---KAANKI-N
	Archaeoglobus fulgidus	11499998	A-CNV-V-S---IEI----T-L		G-A-L-A---AAVTRI-N-

Supplementary Figure 8

Partial sequence alignments of the proteins (A) 5'-methylthioadenosine phosphorylase II (MTAP-II) and (B) hypothetical protein PAE2307 showing two conserved indels that are specific for the Thermoproteales.

			3		37
Thermoproteales	Thermoproteus neutrophilus	YP_001794920	HLRRTTSPYWWPIPRKA	G	GVWAVRPSPGPHSLAYS
	Pyrobaculum calidifontis	YP_001056630	---KSLA----------	-	--------T-----E--
	Pyrobaculum islandicum	YP_930081	------A---------I	-	---V--S----------
	Pyrobaculum aerophilum	NP_560647	-----LA-A-------K	-	---V-------------
	Pyrobaculum arsenaticum	YP_001153933	----SLA----------	-	---V-K-------F---
	Methanopyrus kandleri	NP_614503	-VK-LAA--A------E	-	-KF-P--Y----TMDT-
	Nanoarchaeum equitans	NP_963760	---SLAV-RT--V---S	R	-Y-I-K-------KE
Thermoproteales	Thermofilum pendens	YP_919656	----SVA-PF---K--E		Y--T-K--------LS-
	Caldivirga maquilingensis	YP_001541639	----S---E---VE--G		AP-SIK------PANS-
Desulfurococcales	Ignicoccus hospitalis	YP_001435993	--K-LAA-AF---H--E		A----K-R----PIEE-
	Hyperthermus butylicus	YP_001013487	--KTLAA-KF--VRQR-		-IFT-K------PIER-
	Staphylothermus marinus	YP_001041037	--KALVA-KF---L--E		YK---K------
	Thermosphaera aggregans	YP_003650084	--KSIAA-R----L--E		YK---K------AIARS
	Staphylothermus hellenicus	YP_003669419	--KALAA-RF---L--E		YK---K------
	Aeropyrum pernix	NP_147175	--KALAA-RF---L--E		YK---K-R----PAEK-
	Desulfurococcus kamchatkensis	YP_002428856	--KAINA-RY---L--E		YR-V-KS-----PISR-
Other Archaea	C. Korarchaeum cryptofilum	YP_001737990	--KSLAA-GYY--AKRE		R--V-K------AID
	Methanococcus vannielii	YP_001323240	--K-LAA-AN-QL---E		RTFT---------MDK-
	Methanococcus maripaludis	YP_001549307	--K-LAA-AN-QL---V		KAFT---------MDK-
	Methanococcusvoltae	ZP_02193082	--K-LAA-AN-Q----V		RTFT---A--S-AMDK-
	Thermococcus gammatolerans	YP_002960356	--K-LAA-TS-Y-H--E		YK-----R-----MQT-
	Thermococcus onnurineus	YP_002306460	--K-LAA-NQ-Y-E---		YK-----R-----MKT-
	Thermococcus sp. AM4	ZP_04879475	--K-LAA-TS-Y-H---		YK---------NMQT-
	Thermococcus kodakarensis	YP_183942	--K-LAA-TS-Y-H---		YK----------MKT-
	Pyrococcus furiosus	NP_579541	--K-LAA-TS-Y-E---		YK-----R----NMRT-
	Pyrococcus abyssi	NP_126020	--K-LAA-TS-Y-E---		YK-----R----NMRT-
	Pyrococcus horikoshii	NP_143603	--K-LAA-SS-Y-E---		YK-----R----NMRT-
	Thermococcus sibiricus	YP_002993716	--K-LAA-TQ-Y-E--V		FK---------DMRT-
	Methanocaldococcusinfernus	ZP_04790068	--K-LAA-VR-EL---T		HKFT---L--A-PMSE-
	Methanocaldococcusfervens	YP_003128003	--K-LAA-VR-EL---V		HKFT---L--A-PMSE-
	Methanocaldococcusvulcanius	YP_003246568	--K-LAA-VR-EL---V		HKFT---L--A-PMSE-
	Methanocaldococcus jannaschii	NP_247443	--K-LAA-VR-EL---I		HKFT---L--A-PMSE-
	Methanococcus aeolicus Nankai	YP_001325584	--K-LPA-KN-Q-S--T		NKYTT---A---AMGS-
	Halogeometricum borinquense	ZP_03999508	-QK-LSA-NS---E--E		-TFT-KA-A---
	Methanosaeta thermophila	YP_844122	-RK-V-I-RS-L----T		SK-IAT-R-----AEE-
	Archaeoglobus fulgidus	NP_070738	-QK-LSA-KTYK----V		SK-V-K------NKE
	Natronomonas pharaonis	YP_331164	-QK-LAV-KS--VE--E		EAYTTKAAA---
	Methanospirillum hungatei	YP_503665	--K-VV--KS-G----T		DKFVTKT-----
	Methanosarcina mazei	NP_634160	-QK-LSV-KS-KVG--G		NK-ISTTR-----Q-R-
	Methanosphaera stadtmanae	YP_447927	--K-FKA-KM---HK-E		E--TTKT-A---A-EE-
	C. Methanosphaerula palustrs	YP_002465551	--K-LVA-RA-H-AK-T		NTFISKTA----
	Halorubrum lacusprofundi	YP_002567082	-QK-LAV-NS--VE--T		NTFT-KAGA---
	Haloquadratum walsbyi	YP_658540	-QK-LSA-NS---E--T		ETFT-KAGA---
	Methanosarcina acetivorans	NP_616029	-QK-LSV-KS-KVGK-G		NK-ISTTR-----Q-R-
	C. Methanoregula boone	YP_001403705	--K-LNA-DS-H-TK-T		TKFVAKTA----NAN
	Haloarcula marismortui	YP_136223	-QK-LSV-NS--VE--T		ATFT-KAGA---
	Halobacterium sp. NRC-1	NP_280468	-QK-LSV-KS--VE--T		ETFTAKAGA---
	Methanoculleus marisnigri	YP_001046491	Y-K-LVA-GS-H--K-V		QKFVMKTA----N
	Ferroglobus placidus	ZP_05839485	-QK-LSA-KTIKV---V		SK-I-K------NKE
	Halorhabdus utahensis	YP_003131204	-QK-LSV-NS---E--T		ETFT-KADA---
	Halomicrobiummukohataei	YP_003177661	-QK-LSV-KS--VE--T		ETFT-KADA---
	Natrialba magadii	ZP_03694687	-QK-LSV-KS--VE--T		D-FT-KAGA---
	Methanococcoides burtonii	YP_564787	-QK-ISV-KS-Q-SK-S		NK-ITSTR----
	Methanobrevibacter smithii	YP_001273322	--KTYKA-KT---HP-E		DT-T-K--A-S--IEN-

Supplemental Figure 9

Partial sequence alignment of 30S ribosomal protein S4e showing a 1 aa insert in most Thermoproteales except the early branching *T. pendens* and *C. maquilingensis*. The insert is also found in *M. kandleri* and Nanoarchaeota.

			209		247
Desulfurococcales	Aeropyrum pernix	118431712	PDALRTRIVVTMNAREL	I	QVFFPLRMCTRAQWEIRHIAW
	Staphylothermus helenicus	003669387	-Q-VK---L--MN----	L	E--L-----S-------M---
	Staphylothermus marinus	126465975	-Q-VK---L-S------	L	ES-L-----SH------Y---
	Desulfurococcus kamchatkensis	218884431	-Q-VK--LL-S------	-	E--I---T-A-------S---
	Hyperthermus butylicus	124027729	-Q-V-----F-------	L	ES-L-----SH------ML--
	Thermosphaera aggregans	003650146	-Q-VK---L-------	L	-V-------S-------M---
Thermoproteales	Pyrobaculum aerophilum	18312917	-N-TAVLFA-Q------	L	LN-A---CAYA--A----VC-
	Pyrobaculum arsenaticum	145590945	-N-TAVLFA-Q------	L	LN-A---CAYT--A----VC-
	Pyrobaculum islandicum	119873240	-N-TAVLFA-Q------	L	LN-V---CAYA--A----VC-
	Pyrobaculum calidifontis	126459117	-N-TAVLFA-Q------	L	LN-V---CAAT--A----VC-
	Thermoproteus neutrophilus	171185235	-N-TAVLFA-Q------	L	LN-V---CAYA--A----VC-
	Caldivirga maquilingensis	159041713	-N-MAS--W-Q-----F	F	LN-I---TGLG-FH---LV--
	Ignicoccus hospitalis	156938209	-Q-I--SML--A-L--W		LHIIE--DHPK------GV--
Sulfolobales	Sulfolobus islandicus	227828043	-NGVN-N----------		YN--A--L-S-------A---
	Sulfolobus solfataricus	15897261	-NGVN-N----------		YN--S--L-S-------A---
	Sulfolobus acidocaldarius	70608052	-NGVN-N----------		YN--G--L-S-------A---
	Sulfolobus tokodaii	15922543	-NGVN-N----------		YN--A--L-S-------A---
	Metallosphaera sedula	146304642	-NGVN-NL---------		YN--G--L---------AV--
Other Archaea	Pyrococcus abyssi	14521502	-Q-V-SK------L---		KH--G--L-E-------EV--
	Pyrococcus horikoshii	14590631	-QSVK-K------L---		KH--G--L-E-------EV--
	Pyrococcus furiosus	18976912	-Q-VK-K-I----L---		KH--G--L-E-------EV--
	Thermococcus onnurineus	212223583	-Q-V--K------L---		KH-LG--A-E-------DV--
	Thermococcus kodakarensis	57640805	-Q-V--K------L---		KH-LG--A-E-------EV--
	Thermococcus gammatolerans	240103208	-Q-V--K------L---		KH--G--L-E-------EV--
	Thermococcus barophilus	254171500	-Q-V--K------L---		KH--G--T-E------NYV--
	Thermococcus sibiricus	242399559	-Q-V--K------L---		KH-LGF-A-A------N----
	Thermococcus sp. AM4	254173575	-Q-VK-K------L---		KH--G--L-E-T---L-EV--
	Aciduliprofundum boonei	254166653	-QGIE-K-II-------		RH-IS--CDPS------AM-
	Halorhabdus utahensis	257053951	-IGTKVN--M-----M-		MHIGDM-AAAD------EM
	Haloarcula marismortui	55377933	-IGTKVN--MS----M-		MHVADM-AAAD------EM
	Halobacterium sp. NRC-1	15790361	-IGTEVNV-I-L-P-S-		MHVADM-AAAD------EL
	Halobacterium salinarum	169236095	-IGTEVNV-I-L-P-S-		MHVADM-AAAD------EL
	Nanoarchaeum equitans	41614811	-IGTK-FVTTAAD-E-W		LHIMKV-C-MD------IT--

Supplemental Figure 10

Partial sequence alignment for thymidylate synthase (ThyX) showing a 1 aa insert that is specifically found in various Thermoproteales and Desuflurococcales.

(A)

			154		190
Desulfurococcales	Aeropyrum pernix	14600751	YGTRDWVLGLEIVLPDEEGTIL	RV	GCRTLKCRQGYD
	Hyperthermus butylicus	124028309	-------NQ-----A--N--RM	TI	----V-------
Other Crenarchaeotes	Caldivirga maquilingensis	159040855	---K-----A-V---TG-V-FM		--------A---
	Pyrobaculum aerophilum	18313049	---K------RV-TGRGDLLK-		----Y---N---
	Pyrobaculum islandicum	119872843	--PSH------A-IGTG-IIRV		----V---E---
	Pyrobaculum islandicum	119873321	---K------RV-TGRGDTLK-		----Y---N---
	Pyrobaculum arsenaticum	145592285	---K------RV-TGRGDVLKV		--K-F---N---
	Pyrobaculum calidifontis	126459831	--PSH---A--A-IGTGDVIRV		--S-V---E--N
	Thermoproteus neutrophilus	171185735	---K------RV-TGRGDVLR-		----Y---N---
	Sulfolobus acidocaldarius	70606138	---K----A-KV---NGKVVKI		-EPFT-N-G---
	Sulfolobus tokodaii	15920870	--------A-KV---NGKVIK-		-EPLR-N-G---
	Metallosphaera sedula	146303221	--------S-RV---TG-VVQV		-EPFR-N-A---
	Metallosphaera sedula	146303889	--------A--V---SGKVIRV		-EPFV-N-A---
Euryarchaeotes	Archaeoglobus fulgidus	11498414	---V-------V-TA-G-VLNI		-S---------N
	Picrophilus torridus	48478117	---KE-----------GSIIRT		-EY---RSA---
	Methanocella paludicola	282162951	---K-Y-----V---SGDIINT		--KAP-TSS---
	Thermoplasma acidophilum	16081649	--VK-------V--A-GTLVKF		-NKA--RSM---
	Uncul. methanogenic archaean	147919591	---K-Y--D--V----G-IIHT		--KAP-TSA---
	Thermoplasma volcanium	13541427	--VK--I---DV--A-GTMVKF		-NK---RTL---
	Methanococcoides burtonii	91772953	--A--Y---M-V---TG-IARF		-S----NSA--
	Ferroplasma acidarmanus	257075624	--AKE----M-V---TG-IIRT		-EY---RSA---
	Methanohalophilus mahii	294496597	---KRY--D--V--A-GSIIRT		-S----MAS---
	Ferroglobus placidus	288930739	--I-NY---M-V---SG-ILN-		-GKVI-NVT---
	Methanosaeta thermophila	116754940	---KSY--D--V--A-GTVIRT		-SKV--SSA---

(B)

			139		165
Desulfurococcales	Aeropyrum pernix	118431258	ASLPYMIPAPP	LQLG	FLGTFGAVINMR
	Desulfurococcus kamchatkensis	218884346	IEG-IL-----	I---	---------F-K
	Staphylothermus hellenicus	288788819	IEG--F-----	I---	-I--L----S-K
	Staphylothermus marinus	126466160	IEG--F-----	I---	-I--L----S-K
Thaumarchaeota	Cenarchaeum symbiosum	118576865	TTW--F--GI-	IY	GIP----F-QS-
	Uncul. marine cren. HF4000-APKG10F15	ABZ10057	TTW--F--GV-	VF	GIP------QSK
	Uncul. marine cren. HF4000-ANIW141M18	ABZ08007	TTW--F--GV-	VF	SIP----L-QSK
	Nitrosopumilus maritimus	161527799	TTW--F--GL-	VI	GIP----F-QS-
Euryarchaeotes	Natronomonas pharaonis	76801782	-----F--V-T		-I-------S-K
	Natrialba magadii	289581186	-----F--V-T		LI--M----K-K
	Halorhabdus utahensis	257053045	-----F--V--		PI--M--I-R--
	Halogeometricum borinquense	227883247	-----FL-L-T		L---M----R-N
	Haloterrigena turkmenica	284166739	-----F--V-T		II--M----KLK
	Haloarcula marismortui	55378407	-----FL-F-N		V---L----S-N
	Halobacterium sp. NRC-1	15790876	-T---F--F-S		LF--M----R--
	Halogeometricum borinquense	227881082	V----L--FIV		PF--L--I-Q--
	Haloferax volcanii	292655997	V----L--FIF		PF--L--I-R--
	Picrophilus torridus	48477315	----FF--F-Y		II-----FVSL-
	Thermoplasma acidophilum	10640687	----FF--F-V		GI-----F-SL-
	Thermoplasma volcanium	13541085	----FF--F-V		GI-----F-SL-
	Ferroplasma acidarmanus	257077390	----FF--F-Y		LI-----FVS
	Aciduliprofundum boonei	254167031	V---FF----T		I---L--F-SI-
	Methanococcoides burtonii	91773435	T----F--F-T		II--M----KH-
	Methanosarcina barkeri	73668262	-----F--F-T		-I--M--L-RY-
	Methanosarcina acetivorans	20089141	T----F--F-T		-I--M----RY-
	Methanosarcina mazei	21227625	T----F--F-T		-I--M----RYK
	Pyrococcus horikoshii	14590262	STF--F--F-S		-I--L----RVK
	Thermococcus barophilus	254170941	STF--F--F-S		-I--M----RVK
	Thermococcus gammatolerans	240103175	-TF--F--F-N		L---L----RVK
	Pyrococcus furiosus	18976764	STF--F--F-S		-I--L--I-RVK
	Pyrococcus abyssi	14521814	STF--F--F-S		-I--L----RVK
	Archaeoglobus fulgidus	11497673	T----F--F-T		II--L----RH-
	Ferroglobus placidus	288930510	T----F--F-T		II--L--I-KH-
	Archaeoglobus profundus	284161803	T----F--F-S		II--L----KH-

Supplemental Figure 11

Partial sequence alignment of (A) D-lactate dehydrogenase sequences showing a 2 aa conserved insert that is found only in Desulfurococcales. (B) Partial sequence alignment of a putative membrane associated metallopeptidase (MMP) showing a 4 aa insert that is specific for Desulfurococcales. A 2 aa insert is also present in the same position in various Thaumarchaeota. Homologs of this protein were not detected in other crenarchaeota.

Group	Organism	GI	41		103
Thaumarchaeota	Nitrosopumilus maritimus	161527863	GRSALKRFNKADVNIVERLINKTMH	FGKKYAKN	TGRMTGKKTKVLNTVRTAFTIIELKTGKNP
	Un. marine cr. HF4000_APKG4H17	167044062	-------------------V--M--	--------	----A---FHTF-V-K---E--HQ------
	Un. marine cr. HF4000_APKG3E18	167043880	-------------------V--M--	--------	----A---IHAY-V-K---E--YQ------
	Un. marine cr. HF4000_APKG5C13	167044141	-------------------V--M--	--------	----A---IHAY-V-K---E--YQ------
	Un. marine cr. HF4000_ANIW93124	167042226	-------------------V--M--	--------	----A---IHTY-V-K---E--YQ------
	Cenarchaeum symbiosum	118575639	------------A------A--M--	--------	----A---MGSI---K---E--N----R--
Crenarchaeota	Pyrobaculum arsenaticum	145590268	--YQK---G--RIP-------LM-R		P--NA---H-AY-I-KR--DLVYY------
	Pyrobaculum calidifontis	126458634	--YQKR--G--HIP-----V-LL-R		P--N----H-AY-I-KR--E--YY---R--
	Pyrobaculum aerophilum	18312139	--YQNT--G--RIP-------LM-R		P--N----H-AY-I-KR--DL-YY------
	Pyrobaculum islandicum	119872169	--YQKH--G--RIP-------LM-R		P--N----H-AY-I-KR--D--YY---Q--
	Thermofilum pendens	119719186	--H-K---A--R-S-----A-NL-R		P--NA---MLA--I-KR--E----------
	Thermoproteus neutrophilus	171186448	--YQNY--G--RIP-------LM-R		P--NA---H-AY-I-KR--D-VYY---Q--
	Caldivirga maquilingensis	159042486	RK-G--MIP-I--MM-QL-K		P--NA---Q-AYKILK---D--YVA-----
	Aeropyrum pernix	118431614	--HEKR--G--E-P-----M--L-R		P--NG---HLAY-I-K---D--YFE--E--
	Staphylothermus marinus	126465709	--HEHR--G--E-P--------I-R		P--N----HLAY-I-KK--DL-Y----E--
	Desulfurococcus kamchatkensis	218884680	--HEH---G--L-P-------NV-R		P--NM---HLAY-I-KKS-EL-Y----E--
	Desulfurococcus mobilis	410439	--HEHR--G-TH-P--------V-R		P--NM---HLAYHI-KK--D--HI---E--
	Ignicoccus hospitalis	156937044	--HEKR--G--E-P--------L-R		P--NM---A-LYSV-ER--DL-Y-M--Q--
	Hyperthermus butylicus	124028426	--HEHR--G--E-P--------L-R		P--NM---HLAY-I-KQ--E--Y-R--E--
	Sulfolobus islandicus	227828179	--HEHR--G--K-P-------QI-R		P--NK---HLAY-I-KL--D--Y----Q--
	Sulfolobus acidocaldarius	70606489	--HEHR--G-SRIP-------NL-R		P--NK---MLAY-I-K-T-D--AV---Q--
	Sulfolobus tokodaii	15920459	--HEHR--G--KLP-------NV-R		P--NK---MLAY-I-K---D--A----Q--
	Sulfolobus solfataricus	15897165	--HEHR--G--K-P-------QI-R		P--NK---HLAY-I-KL--D--Y----Q--
	Metallosphaera sedula	146302825	--HEHR--G--R-S-------EL-R		P-KNK---ILAY-I-KAT-EL-NAR--Q--
Other Archaea	Ferroglobus placidus	260570712	--H-KEP-G-QK-F--------V-R		K--N----ILAY-I-KE--E---KR-K---
	Halobacterium sp. NRC-1	15791384	--H-Q-Q-K-SEIS-----A-RL-K		--ANA---QQA-KI--D--D-VHER-DE--
	Aciduliprofundum boonei	254168611	--HVK-HLG--Q---------NL-R		-EKY----MSAY-V-KK--E---K--KQ--
	Methanocaldococcus infernus	241905629	--FTKRQ-E--KM------V--V-R		REEN----L-A-KI-ER--E---QR-KQ--
	Methanopyrus kandleri	20094118	--H-K---A--E-P-------RV-R		-EKN----HLAY-I-KR--D--HER--E--
	Archaeoglobus fulgidus	11499477	--H-NVP-A-QK-F--------V-R		K-HN----ILAY-I-KE--E---K--K---
	Methanocella paludicola	282163893	--LGC-Q-G--EMP--------L-R		NEFN--E-I-CY-I--D--D--NRR-KQ--
	Methanosaeta thermophila	116753359	--H-KQQ-K-SELHV-------M-R		KEKN--Q-QIAYRI-EE--D--HSR-KE--
	Methanococcoides burtonii	91773153	-KH-RQQ---S-IT-----V-NV-R		NAQN----QIA-RI-DE--D-VNS--K---
	Methanospirillum hungatei	88604062	--LTRQE-H--NML-------QL-R		-EVN----QLAIRI--D--E-VHQ--K---
	C. Methanoregula boonei	154149811	-KFSRQE----NMM-------RL-Q		-ENN----QLAIGI--D--EL-NK--KR--
	C. Methanosphaerula palustris	19850954	-KL-GTD---SNML-------RL-Q		-EQN----QLAIRI--D--G--AK--K---
	Haloquadratum walsbyi	110669264	--H-A-Q-E-SKISV------RL-Q		-EDN--H-Q-TM-I-ED--DT-HNR-EE--
	Methanosarcina mazei	21228368	-KH-RQQ---SEIS-----A-NL-R		-ETN----QVT-RA-EE--D--NR--QQ--
	Methanosarcina barkeri	73671115	-KH-RQQ---SEIS-----A-NL-R		-EIN----Q-T-RA-EE--D-VSK--KQ--
	Methanococcus aeolicus Nankai	150401213	--NSGRM-E--KM------V-NL-R		AEQN--Q-QVT-SI-EE--E---KR-NE--
	Methanocorpusculum labreanum	124486427	--LTQQQ-T-SEMA-------RL-Q		QEYN----MMCTKM-MD--DV-NK--KQ--
	Methanosarcina acetivorans	20090122	-KH-RQQ---SEIS-----A-NL-R		-ETN----QVT-RA-EE--D-VNK--KQ--
	Methanotherm. thermautotrophicum	15679069	--HVKRQ-A-SK-S--------V-R		-E-NS---N-AYKI-QE--E--NRR-KE--
	Methanosphaera stadtmanae	84490155	--HVKRQ-A-SK-S-----M--I-R		-E-NS---N-AYSI-ED-LE--NN--KQ--
	Methanocaldococcus vulcanius	261403640	--YTKRQ-E--KM------V--V-R		REEN----L-A-KI-EE--E---KR-KQ--
	Methanocaldococcus fervens	256811209	--YTKRQ-E--KM------V--V-R		REEN----L-A-KI-EN--E---KR-KQ--
	Methanocaldococcus jannaschii	15669236	--YTKRQ-E--KM------V--V-R		REEN----L-A-KI-EN--E---KR-KQ--
	Methanococcus voltae	163799207	--FSK-M-D-SKM-----VL--L-A		EQTN----YEA-AVMEE-LE---KR-KE--
	Methanococcus maripalud	150402537	--NSK-S-D-SKM------A--L-A		NQEN----HET-AI-EE-L----NR-KE--
	Thermoplasma volcanium	17367594	-INM-V-------L-R		-EKW----YSAYKITKD--D--AD--KQ--
	Ferroplasma acidarmanus	257076706	-RN--TI---L-NL-R		-EKW----YSAYKVMSQ--D--AT--KQ--
	Methanococcus vannielii	150399429	--NSK-M-D-NKMHV----A--L-A		-QVN----NE--SIIEE-L--V-NR-KE--
	Halorubrum lacusprofundi	222478539	--H-S-Q-K-SEISL---M--RL-Q		-DEN----QQATRI--D--D--HER-EE--
	Halorhabdus utahensis	257053450	--H-K-Q-Q-SEIS-------RL-Q		-EEN----QQSTRL--E--E-VHNN-DE--
	Natrialba magadii	224819327	--H-S-Q-K-SEIS----F--RL-Q		-EDN----QQT--F--E--EL-HER-EE--
	Haloarcula marismortui	55379090	--H-D-Q-K-SEIS-------RL-Q		-DEN----QLATSI-TE--ELVHER-DE--
	Natronomonas pharaonis	76800717	--HSE-Q-K-SEIS-------RL-Q		-DEN----QQATKI--E--D-VAER-DE--
	Halogeometricum borinquense	227880472	--H-S-Q-Q-SEIS-------RL-Q		SEDN--H-Q-TMKI-KD--E-VHER-DE--
	Halococcus morrhuae	133993	--H-G-Q-Q-SEIS-------RL-Q		-DEN----QQATRI--N--DL-AER-EES-
	Halomicrobium mukohataei	257386577	--HSE-Q-K-SEIS-------RL-Q		-DEN----QLATTI-KD--ET-HER-EE--
	Methanobrevibacter smithii	148642960	--HVKRQ-A-SK-S-I---M--I-R		-HINS---N-AY-I-KD-LD--NKR-K---
	Methanoculleus marisnigri	126180146	-KLVGQQ---SNML--------M-Q		-ENN----ELTIRI--D--E-VNK--K---

Supplemental Figure 12

Partial sequence alignment of Ribosomal protein S7 showing an 8 aa insert that is specific for the Thaumarchaeota. In the same position larger inserts of 16 aa insert are present in various Thermococcales, a 12 aa insert is present in Korarchaeum and a 11 aa insert in Nanoarchaeum (not shown).

```
Nitrosopumilus maritimus         MQNLNLIAGILIIASPILFSMIAYPDSVAWSWNEGRGGYLFALVFVVAEL
Uncul. Crenarchaeote-4B7         MENKSVIMAILLVLSPILFTMAIYPDSFSLSWNQGRGGFLFAAAFIAAEL
Un. marine Cren-HF4000_ANIW93E5  MMHKNTILAMLLIASPILFVFVAYSDTFSMSWNQGRGGFLFGLAFIVAEI
Un. marine Cren-HF4000_APKG9P22  MMHKNTILAMLLIASPILFVFVAYSDTFSMSWNQGRGGFLFGLAFIVAEI
Un. marine Cren-HF4000_APKG3K8   MMHKNTILAILLIASPFLFVFVAYSDTFSMSWNQGRGGFLFGLAFIVAEI
Cenarchaeum symbiosum            --------------------------------------------------


Nitrosopumilus maritimus         VGLKIVISKKRLLAVIPIALLTISYLISLEYGLRDYIVESAEQFDVQ---
Un. marine Crenaeote-4B7         IGLKFVIPKKRFFYCLPLIALTMTYFVSLQFGLRDYIMSLVDVFGV----
Un. marine Cren-HF4000_ANIW93E5  VGIKFVVSKNRLIFVIPLVIATILYFVALDFGLHDYILNAAPAFNVVGCE
Un. marine Cren-HF4000_APKG9P22  VGIKFVVSKNRLIFGIPLVIATILYFVALDFGLHDYILNAAPAFNVVGCE
Un. marine Cren-HF4000_APKG3K8   VGIKFVVSKNRLIFGIPLAIATILYFVALDFGLHDYILNAAPAFNVVGCE
Cenarchaeum symbiosum            ----MGVSTRRLLATIPLAAAVIAYIIAVENGLREQIEAAAPDFGVE---
                                     : :...*::  :*:   .: *::::: **:: *   .   *.*

Nitrosopumilus maritimus         ------LIYSWTWMWDFVVMAIFIVVALSIFFGKRWIRIAPAGPIFLTGT
Un. marine Crenaeote-4B7         ------LEYSWEWLFDFTVMAIFVTASLAILFGRKWIRIGPAGPIFLAGN
Un. marine Cren-HF4000_ANIW93E5  VANPQGCIYSWGWLWDFVIITIFVISAAIIMFGKKWIRIVIAGPVFLGGS
Un. marine Cren-HF4000_APKG9P22  VANPQGCIYSWGWLWDFVIITIFVISAAIIMFGKKWIRIVIAGPVFLGGS
Un. marine Cren-HF4000_APKG3K8   VANPQGCIYSWGWLWDFVIITIFVISAAVILFGKKWIRIVIAGPVFLGGS
Cenarchaeum symbiosum            ------LVYSWIWLWDFAVMTAFVSVALAIYFGKKWVRIAPAGPIFLGGS
                                         *** *::**.::: *:  :  * **::*:**  ***:** *.

Nitrosopumilus maritimus         AIILSLDAFFPYDTLGPLQYIVPYFVQANVWVITALELGTAVARDNVMFL
Un. marine Crenaeote-4B7         AIILALDSFFPYDTLGPLQYIVPYFVQANVWVITALDLGVATARDNLMFL
Un. marine Cren-HF4000_ANIW93E5  AIILSLDTFFPFDTLGPLQYFVPYLVETNVWVINALELGIATGRDNLMFL
Un. marine Cren-HF4000_APKG9P22  AIILSLDTFFPFDTLGPLQYFVPYLVEANVWIINALELGIATGRDNLMFL
Un. marine Cren-HF4000_APKG3K8   AIILSLDTFFPFDTLGPLQYFVPYLVEANVWIINALELGIATGRDNLMFL
Cenarchaeum symbiosum            AAILSLDAFFPYDSLGPLQFFVPYLVELNVWLIGLFDLGMATANDNLMFL
                                 * **:**:***:*:*****::***:*: ***:*  ::** *...**:***

Nitrosopumilus maritimus         RGDHGSMALQVFWPSAGVHSIIIFSLVIGAFMLKMNIPRARKSMYFVLGI
Un. marine Crenaeote-4B7         NGDFGPFALQVFWPSAGVHSIIIYSLVMMAFLIKMRIPRNRKIIYFILGI
Un. marine Cren-HF4000_ANIW93E5  KGDHGPFVLQVFWPSAGVHSIIIYSLVMMAFLLKMNIQRNRKALYFGLGI
Un. marine Cren-HF4000_APKG9P22  KGDHGPFVLQVFWPSAGVHSIIIYSLVMMAFLLKMNIQRNRKALYFGLGI
Un. marine Cren-HF4000_APKG3K8   SGDHGPFVLQVFWPSAGVHSIIIYSLVMMAFLLKMNIQRNRKALYFGLGI
Cenarchaeum symbiosum            RGDHGAFALQVFWPSAGVHSIIIYSLVMMAFLLKMNVPRGRKAAWFAIGI
                                  **.*.:.**************:***: **::**.: * **  :* :**

Nitrosopumilus maritimus         IGTITVNLIRIFSLSWYALKVTTDPVAWEEYHKIAGEIMFLPWLFAFILV
Un. marine Crenaeote-4B7         LGTIFVNLIRIFSLSVYALKVTTDPEAWEEFHSVAGEIMFLPWLFIFLLI
Un. marine Cren-HF4000_ANIW93E5  IGTIIINLIRIFSLSVFALKVSTNPVEFEEYHSIAGEIMFLPWLFVFLLV
Un. marine Cren-HF4000_APKG9P22  IGTIIINLIRIFSLSVFALKVSTNPVEFEEYHSIAGEIMFLPWLFVFLLV
Un. marine Cren-HF4000_APKG3K8   IGTIIINLIRIFSLSVFALKVSTNPVEFEEYHSIAGEIMFLPWLFVFLLV
Cenarchaeum symbiosum            IGTIAVNVARIFALSWYALKVTTDAAKWEEFHSVAGEIMFLPWLFAFLVI
                                 :*** :*: ***:** :****:*:.  :**:*.:*********** *:::

Nitrosopumilus maritimus         VILIESRRLKKQEERDGTKNNS
Un. marine Crenaeote-4B7         VTVVETKRLKRIESEKTSKNNL
Un. marine Cren-HF4000_ANIW93E5  VTMIETKRMKKKETSVQK----
Un. marine Cren-HF4000_APKG9P22  VTMIETKRMKKKETSVQK----
Un. marine Cren-HF4000_APKG3K8   VTVIETKRMKKKETSVQK----
Cenarchaeum symbiosum            VMVVESRRLTRLKAQSQK----
                                 * ::*::*:.: :    .
```

Supplemental Figure 13

Amino acid sequence alignment of a protein of unknown function (Nmar_0641; Gene ID. 161528149) that is uniquely found in various Thaumarchaeota species. The asterisks (*) indicate positions that are completely conserved, whereas other conserved residues are identified by period (.) and semicolon (:).A 9-10 aa conserved insert that distinguishes uncultured crenarchaea sequences HF4000-ANIW93E5, HF4000-APKG9P22 and HF4000-APKG3K8 from other Thaumarchaeota is also present in this sequence alignment.