Emma Griffiths and Radhey S. Gupta*

Department of Biochemistry, McMaster University, Hamilton, Ontario, Canada L8N 3Z5

Received 5 January 2004/Accepted 4 February 2004

The Deinococcus-Thermus group of species is currently recognized as a distinct phylum solely on the basis of their branching in 16S rRNA trees. No unique biochemical or molecular characteristics that can distinguish this group from all other bacteria are known at present. In this work, we describe eight conserved indels (viz., inserts or deletions) in seven widely distributed proteins that are distinctive characteristics of the Deinococcus-Thermus phylum but are not found in any other group of bacteria. The identified signatures include a 7-amino-acid (aa) insert in threonyl-tRNA synthetase, 1- and 3-aa inserts in the RNA polymerase β' subunit, a 5-aa deletion in signal recognition particle (Ffh/SR54), a 2-aa insert in major sigma factor 70 (σ^{70}), a 2-aa insert in seryl-tRNA synthetase (SerRS), a 1-aa insert in ribosomal protein L1, and a 2-aa insert in UvrA homologs. By using PCR primers for conserved regions, fragments of these genes were amplified from a number of Deinococcus-Thermus species, and all such fragments (except SerRS in Deinococcus proteolyticus) were found to contain the indicated signatures. The presence of these signatures in various species from all three known genera within this phylum, viz., Deinococcus, Thermus, and Meiothermus, provide evidence that they are likely distinctive characteristics of the entire phylum which were introduced in a common ancestor of this group. The signature in SerRS, which is absent in D. proteolyticus, was likely introduced after the branching of this species. Phylogenetic studies as well as the nature of the inserts in some of these proteins (viz., σ^{70} and SerRS) also support a sister group relationship between the Thermus and the Meiothermus genera. The identified signatures provide strong evidence for the monophyletic nature of the Deinococcus-Thermus phylum. These molecular markers should prove very useful in the identification of new species related to this group.

The Deinococcaceae are well known for their extreme resistance to UV, desiccation, and ionizing radiation (1, 28) and have been found in canned meat, soil, animal feces, dust, and irradiated medical instruments (27, 31, 35). These nonsporulating, polyploidic cocci are also resistant to hydrogen peroxide and other agents that damage DNA due to a highly efficient DNA repair system (1, 28). Although these species stain gram positive due to a thick layer of peptidoglycan in their cell wall (2, 30, 31), they are structurally similar to the gram-negative bacteria in that they contain an outer membrane (2, 31, 41). This outer layer, however, is unique in that it does not contain lipid A or heptoses typical of other gram-negative bacteria (2, 7, 31). Thermus-Meiothermus species are gram-negative thermophilic rods which have been isolated from thermally polluted streams, industrial and domestic water taps, and hydrothermal vents with neutral to alkaline pH (9, 35). Both the deinococci and the Thermus-Meiothermus groups have an atypical cell wall containing ornithine in place of diaminopimelic acid in their peptidoglycan, although the species of the Thermus-Meiothermus group have few other characteristics in common with the deinococci (2, 30, 31). Because of their unusual radiation resistance characteristics, the deinococci have been of great interest with regard to the bioremediation of sites

* Corresponding author. Mailing address: Department of Biochemistry, HSC-4H2, McMaster University, 1200 Main St. West, Hamilton, Ontario, Canada L8N 3Z5. Phone: (905) 525-9140, ext. 22639. Fax: (905) 522-9033. E-mail: gupta@mcmaster.ca. contaminated with radiation and toxic chemicals (1, 3, 28). There is also much interest in this group due to the production of a number of thermostable enzymes of much biotechnological importance, e.g., *Taq* polymerase. These practical applications have increased the desire to understand the evolutionary relationships of the *Deinococcus-Thermus* group to other bacteria (10, 22, 26, 39, 42, 43).

The Deinococcus, Thermus, and Meiothermus genera have been grouped together as a distinct phylum within Bacteria based on their close clustering in 16S rRNA trees, despite morphological and physiological dissimilarity (2, 25, 31, 42, 44). With the rapid increase in the sequence database entries, it is becoming increasingly imprecise to assign species to different taxonomic groups based on branch patterns alone (25). Unfortunately, there are presently no other criteria or molecular means by which species belonging to this phylum can be unambiguously distinguished from other bacterial phyla (2, 9, 31, 33). We have described a new approach based on conserved indels (i.e., inserts or deletions) found in different proteins that is helpful in distinguishing the major bacterial phyla and to understand the interrelationships among them (13–15, 17). Recently, a large number of conserved indels (or signature sequences) which provide distinctive molecular markers for the identification of proteobacteria, chlamydiae, and cyanobacteria have been described (12, 14, 19).

The present communication describes for the first time a number of conserved indels in widely distributed proteins that are distinctive characteristics of the *Deinococcus-Thermus* phylum. The identified signatures include a 7-amino-acid (aa) insert in Thr-tRNA synthetase (ThrRS), a 5-aa deletion in the signal recognition particle protein Ffh, a 1- and a 3-aa insert in the β' subunit of RNA polymerase RpoC, a 1-aa insert in the ribosomal protein L1, and 2-aa inserts in major sigma factor 70 (σ^{70}) , seryl-tRNA synthetase (SerRS), and UvrA homologs. The sequence information for these proteins was previously available from only a limited number of Deinococcus and Thermus species. As the Meiothermus genus has only recently been established, there is little sequence information currently available for this group (33). We have tested the specificity of the identified signatures by PCR amplifying and sequencing fragments of these genes from additional Deinococcus and Meiothermus species for which no sequence information was available. The presence of these signatures in all of the species examined (with a single exception) provide evidence that they are likely distinctive characteristics of the entire phylum and might be used as molecular markers for this group of species.

MATERIALS AND METHODS

Identification of signature sequences. Deinococcus-Thermus-specific signatures were identified in global multiple sequence alignments by means of visual inspection. Alignments for different proteins were constructed by using the ALIGN PLUS 4 program (Scientific & Educational Software, Durham, N.C.) as described in earlier work (12, 14, 19). To qualify as a useful group-specific signature, any identified indel was required to be uniquely (or mainly) present in the Deinococcus-Thermus-Meiothermus group of species and to be flanked on both sides by conserved regions to ensure that the observed insertion or deletion was not a result of sequencing errors or alignment artifacts.

PCR amplification and sequencing. Cultures of *Meiothermus ruber* (ATCC 35948), *Meiothermus silvanus* (DMSZ 9946), and *Deinococcus grandis* (DSMZ 3963) were generously supplied by Peter Gogarten and Lorraine Olenzenski (36). *Deinococcus proteolyticus* (ATCC 35074) high-molecular-weight DNA was prepared as previously described (6, 16). Oligonucleotide primers, in opposite orientations, were designed for conserved regions in the protein sequences that flanked these signatures based on sequence information from available *Deinococcus-Thermus* and other species. Degeneracy was incorporated into the primers were synthesized at the Molecular Biology Central Facility (MOBIX) of McMaster University, Hamilton, Ontario, Canada.

PCRs. PCR was performed in a Techne Techgene thermocycler. The PCRs had a final volume of 10 μ l, and all primer sets were optimized for Mg²⁺ concentration (in the range of 1.5 to 4 mM) for each DNA strain tested. PCR amplification was carried out over 30 cycles (15 s at 94°C, 15 s at 55 or 45°C, 1 min at 72°C) with an initial 1-min hot start at 94°C and a final extension step (15 s at 94°C, 15 s at 55°C, 7 min at 72°C) (12). The reaction mix also contained 22% dimethyl sulfoxide, which improves PCR performance by lowering the melting temperature of DNA. DNA fragments of the expected size were purified from 0.8% (wt/vol) agarose gels (using a GENECLEAN kit) and subcloned into the plasmid pDRIVE by using a TU cloning kit (Invitrogen). *Escherichia coli* JM109 cells were transformed with the ligated vector and insert, and the inserts from a number of positive clones were sequenced at MOBIX. Sequences of all cloned fragments were run through a BLAST search to ensure that the amplification of different genes are as follows.

(i) σ⁷⁰. The following primers were successful in amplifying 504-bp inserts from *D. grandis*, *D. proteolyticus*, *M. silvanus*, and *M. ruber*: forward, 5'-ACNTA YGCNACNTGGTGGAT-3'; reverse, 5'-GRNGCYTTRTTYTCDATYTG-3', where N represents A, G, C, or T; Y is C or T; R is A or G; and D is A, G, or T.

(ii) Threonyl tRNA synthetase. Fragments 432 bp in length were generated from *D. grandis*, *M. silvanus*, and *M. ruber* genomic DNA with the following primers: forward, 5'-TTCCGSCACWCSCTGGSCCACGTCMTG-3'; reverse, 5'-CCNCKCCARTANGCNCC-3', where S represents C or G, W is A or T, K is G or T, and M is A or C.

(iii) Signal recognition particle Ffh. The following primers were used to amplify a 264-bp fragment from *D. grandis*: forward, 5'-ATHYTNGGNATGG

GNGA-3'; reverse, 5'-CKYTCYTTNACNGTCAT-3', where H represents A, C, or T.

(iv) SerRS. Fragments from *M. silvanus* and *D. proteolyticus* of 234 bp in length were successfully amplified by using forward primer 5'-CACSARTTYCGYAA RGTNGARCAG-3' and reverse primer 5'-CGARCAGGARTGGGTYTCGC GRTC-3'.

(v) RNA polymerase β' subunit RpoC. RpoC gene fragments (645 bp) were amplified by PCR from *M. silvanus*, *M. ruber*, *D. proteolyticus*, and *D. grandis* by using the following primers: forward, 5'-GAYGGNGGNMGNTTYGC-3'; reverse, 5'-CATYTGRTCNCCRTCRAARTC-3'.

(vi) Ribosomal protein L1. A 510-bp fragment was generated from *M. silvanus* and *D. grandis* by using the following primers: forward, 5'-ATGCCTAAGCAC GGCAAGCGTTACC-3'; reverse, 5'-CCGGTCTTGTCGTTGCCGGAACTC-3'.

(vii) Exinuclease ABC subunit A UvrA. A 639-bp fragment was amplified from *M. silvanus* by using forward primer 5'-TGGCYTTYGACACCATCTACGCCG AGG-3' and reverse primer 5'-AGGCGAACTTCTCSGAGWACAGCTCCTC-3'.

Phylogenetic analysis. Phylogenetic analysis on protein sequences was carried out by procedures described in earlier work (6, 20). Multiple alignment of protein homologs from different groups of bacteria was created by using the ALIGN program. The data for the newly sequenced fragments were added to the alignment, and the fragments were all trimmed to the same length as the amplified fragments. Phylogenetic analyses were performed in both the presence and the absence of the signature region to determine its influence on the branching pattern. The aligned sequences were used to generate 100 bootstrapped data sets with the SEQBOOT program, and genetic distances were calculated by PROTDIST by using Kimura's method (23). Neighbor-joining trees from these distances were constructed by the NEIGHBOR program (40). A consensus tree for various bootstrapped sequences was obtained by using the CONSENSE program. All of these phylogenetic programs are part of the PHYLIP software package (version 3.5; J. Felsenstein, University of Washington, Seattle, Wash.).

Nucleotide sequence accession numbers. The sequence data for all of the gene fragments cloned and sequenced in this work have been deposited in the Gen-Bank database under accession numbers AY450950, AY452779, AY453862, AY489057, and AY453858 for *D. grandis*; AY450951 and AY453857 for *D. proteolyticus*; AY450952, AY452780, AY455864, AY489058, AY489059, and AY452782 for *M. sylvanus*; and AY452778, AY452781, and AY453861 for *M. ruber*.

RESULTS

Description of conserved indels that are distinctive of the *Deinococcus-Thermus* group. Conserved indels that are shared by all members of one particular group (group-specific signatures), or are commonly present in species belonging to more than one taxa (main-line signatures), provide powerful means to identify individual taxa in molecular terms and to understand the interrelationships among them (13, 14, 17). Evolutionarily significant indels are generally of defined size, are present at a specific location, and are flanked by conserved regions to ensure their reliability. We describe below a number of conserved indels in widely distributed proteins that are distinctive characteristics of the *Deinococcus-Thermus* group (*Deinococcus, Thermus*, and *Meiothermus*) of species.

In σ^{70} , which plays a central role in the transcription process by conferring promoter specificity to RNA polymerase (5), a 2-aa insert is present in a conserved region in various available *Deinococcus-Thermus* homologs (viz., *Deinococcus radiodurans, Thermus aquaticus*, and *Thermus thermophilus*) but not in any other bacteria. However, variable inserts are present in this region in *Mycoplasma* species (data not shown), which are likely of independent origin. The specificity of this insert for the *Deinococcus-Thermus* phylum was tested by PCR amplifying and sequencing fragments of the σ^{70} gene from four other members belonging to this group for which no sequence infor-

			105	145
	FE. coli	NP_417539	QEMGREPTPEELAERMLMPED	KIRKVLKIAKEPISMETPIG
	H. influenzae	AAC22190	AG	
	X. fastidiosa	NP_298639	-QFAKE-D	M
	Sal. typhimurium	NP_462126		
	Pse. aeruginosa	NP_249267	GD	
	Pas. multocida	NP_246178	ASG	
	V. cholerae	NP_230168	Q	
	Buch. sp.	NP_239892	IS-KI	
	Ral. solanacearum	NP_520336	T-ND-ATK-E	IM
	Nit. europaea	ZP_00002070	T-QE-AVK-EE	IS
Proteobacteria	Nei. meningitidis	NP_274545	T-EDSAKL-Q	IM
	Burk. fungorum	ZP_00031680	~-T-LD-ATK-E	IM
	Rh. sphaeroides	ZP_00006118	H-IKLQLE	-VML
	A. tumefaciens	NP_355127	H-IKLA-LE	-VLV-
	R. prowazekii	NP_221206	N-L-YAT-I-N-LSL-	-VML-N-V-
	Rhodo. rubrum	ZP_00013908	H-IKLQLE	-VL
	C. crescentus	NP_421841	H-IKLALE	-V
	Desulf. desulfuricans	ZP_00130466	LDIDY-I-	-VKL
	Geo. metallireducens	ZP_00081933	ISINL-L-	-VL
	Camp. jejuni	NP_282151	-KD-KDVSVI-KEVGLSV-	-VKQ-1TL-A
	Hel. pylori	NP_206888	N-KDL-VVEVGLSL~	-VKN-1-VTUV-
	Aqu. aeolicus	NP_214029	NTYLDT-VE	-VKM-FSQL
Aquifar	Cyt. hutchinsonii	ZP_00120059	-KYES-DVLEVSTA	EVVDTSGRHVDA-FV
Aquijex,	Bact. thetalotaomicron	NP_810224	NE-R-SDELEI-V-	SDIVSGRHVDA-FV
Chlamydiae,	Chi. muridarum	NP_29/2/8	M T Z C FLOPTD	RV-EIIQHLQAEV-
CFBG Group	Chiam provmeniae	NP_220132	M-T-KG-BLGFTP-	RV-EIIQHLOAEV-
-	Tro pollidum	NP_443633	-KESDT-OOLCWTVE	-VKO-KGV-P
	Len interrogane	NP 712413		-VKM-KNV-PIV-
	Bor burgdorfari	NP 212846	-VI-KD-DSD-LGWELK	-VKT-KSVSRV-L
a i i i	Nostoc sp PCC 7120	NP 489303	KEI-TE-TIE	-L-FTA-S-OLL
Spirochetes,	Syn sp WH 8102	NP 897874	FKEIS-E-TIE	-L-FIA-S-OLL
Cyanobacteria,	Pro. marinus	NP 895100	FKEIS-E-TIE	-L-FIA-S-OLL
GNS Bacteria	Tri. ervthraeum	ZP 00072877	KEI-TS-E-TIE	-L-FIA-S-OLL
	Sy. sp. PCC 6803	NP 442860	R-KEIK-E-TIE	-L-FIA-S-QLL
	Thermosyn, elongatus	NP 681407	KEI-DE-TIE	-L-FIA-S-QLL
	Cfx. aurantiacus	ZP_00020465	-T-QI-DA-GISAG	-V-RT-EASMH-L-L-M-V-
	D. radiodurans	NP_294640	LSA-HIA-GPGW- AA	-VEE-Q-VSQV-L
	D. grandis *	AY450950	LSA-YIA-GPGW- AA	-VEE-Q-VSQV-L
D .	D. proteolyticus *	AY450951	LSA-YIA-GPGW- AA	-VEE-Q-VSQV-L
Demococcus-	The. aquaticus	1L9U_Q	LSYIA-GPGW- AK	RVEETQV-L
Thermus	The. thermophilus	BAA74758	LYIA-GPGW- AK	RVEET Q V - L
	Mei. sylvanus *	AY450952	LYIS-A-GPGW- AK	-VEETFQV-L
	^L Mei. ruber *	AY452778	LSYI-DA-GPGW- AK	-VEETFQV-L
	T. maritima	NP_229250	-KH-ESIKM-GK-PE	KEI-EATL-S
	Myc. tuberculosis	NP_337278	-DLKE-DITPE	-VLEIQQY-RLDQT
	Bif. longum	NP_696589	-DLDRELDVE	-VQE-Q-YGRLHL-
	Troph. whipplei	NP_787627	-DLDGRELDPE	RVVE-Q-YGRLHL-
	Thermobif. fusca	ZP_00057441	-DLKELD-TPE	-VVE-Q-YGRLHL-
	Str. coelicolor	NP_629943	-DLKELD-TPE	-VIE-Q-YGRLHL-
	Cor. efficiens	NP_738414	LQSKE-DIS-E	-VLEIQQY-RLDQT
	Cor. glutamicum	NP_601117	LQSKE-DIS-E	-VLEIQQY-RLDQT
	Bac. subtilis	NP_390399	-DLI-D-DLTPE	-V-EIQV-L
Gram(+)ve	Bac. natodurans	NP_242242	DI TO DI	-V-BIQV-L
Bacteria	Oce. inevensis	NP_692865	DI TONE DI TE	-V-DIQV-L
Dutterin	Ent. faecalis	NP_615241	DI DA IGEDIAR	V-BIQV-L
	Sta. auteus Strap mutané	MF_010330		
	Helio mobilia	DDC44890		RV-RTMOV-I
	Thermo tengcongeneia	ND 623345		-V-RIQV-D
	Clo acetobutylicum	ND 347931	LDT-KT-DV-	-V-EIMOV-L
	Lis innocua	NP 470827	-DLD-SIG-E-DL-TE	-V-EIOV-I
	L. lactis	NP 266709	LD-SIGKELH-AP-	-V-EOV-I
	Fuso, nucleatum	NP 604215	T-KDASILG-EV-	KAIOEMNOLV-
	- appr more duction			····· * ** **

FIG. 1. Partial sequence alignment for σ^{70} proteins showing a 2-aa insert (boxed area) in a conserved region that is uniquely present in *Deinococcus, Thermus*, and *Meiothermus* homologs. Dashes in this and all other alignments indicate identity to the amino acid on the top line (*E. coli* protein). The position of this sequence region in the *E. coli* protein is indicated at the top. The accession numbers of different proteins are provided in the second column. Sequence information for only representative species from different bacterial groups is presented. The sequences marked with an asterisk were cloned and sequenced in the present work. Abbreviations for the species names are as follows: A., *Agrobacterium*; Aqu., *Aquifex*; Bac., *Bacillus*; Bact., *Bacteroides*; Bif., *Bifidobacterium*; Bor., *Borrelia*; Buch., *Buchnera*; C., *Caulobacter*; Camp., *Campylobacter*; Cfx., *Chloroftexus*; Chl., *Chlamydia*; Chlam., *Chlamydophila*; Clo., *Clostridium*; Cor., *Corynebacterium*; Cyt., *Cytophaga*; D., *Deinococcus*; Des., *Desulfovino*; E., *Escherichia*; Ent., *Enterococcus*; Fuso., *Fusobacterium*; Geo., *Geobacter*; H., *Haemophilus*; Hel., *Helicobacter*; Helio., *Heliobacillus*; L., *Lactococcus*; Lep., *Leptospira*; Lis., *Listeria*; M., *Mycoplasma*; Mei., *Meiothermus*; Myc., *Mycobacterium*; Nei., *Neisseria*; Nit., *Nitrosomonas*; Oce., *Oceanobacillus*; Pas., *Pasteurella*; Pse., *Steptonoccus*; Sy., *Synechocystis*; Syn., *Synechococcus*; T., *Thermotoga*; Thermo., *Thermoanaerobacter*; The., *Thermosynechococcus*; Str., *Treponema*; Tri., *Trichodesmium*; Troph., *Tropheryma*; V., *Vibrio*; X., *Xylella*. GNS, green nonsulfur bacteria; Gram(+)ve, gram-positive.

mation was available. Results of these studies, which are included in Fig. 1, show that all four species tested, which included two *Deinococcus* (*D. grandis* and *D. proteolyticus*) and two *Meiothermus* (*M. ruber* and *M. silvanus*) species contained the identified signature. The sequence region which flanked the identified insert (Fig. 1, boxed region) was also found to be distinctive for *Deinococcus-Thermus-Meiothermus* species. Since sequence information for this signature is now available for representatives from all three genera within the *Deinococcus-Thermus* phylum, the shared presence of this insert in all of



FIG. 2. A neighbor-joining distance tree with branch lengths based on σ^{70} sequences. The tree is based on 169 aa positions for which sequence information was available for various species. The bootstrap scores (out of 100) of various nodes which were >50 are indicated. The arrow marks the suggested position where the identified insert in this gene was introduced.

them strongly indicates that it is very likely a distinctive characteristic of the entire phylum.

We also performed phylogenetic analysis based on σ^{70} sequences from different bacteria. For these purposes, 169 aa positions for which σ^{70} sequence information was available from different species were utilized. The sequence alignment data were used to generate 100 bootstrapped data sets, and a consensus neighbor-joining tree was obtained from these data. At the same time, a neighbor-joining distance tree showing branch lengths, shown in Fig. 2, was also constructed. The bootstrap scores for different nodes which were >50 are marked on this tree. As shown in Fig. 2, most bacterial groups are clearly distinguished from each other in the tree (as shown by their high bootstrap score), but their branching orders or interrelationships are not resolved, which is a common problem with phylogenetic trees (13, 25). Importantly, in the present context, all of the Deinococcus-Thermus-Meiothermus species formed a well-defined group, branching together 100% of the time. Within this group, different Deinococcus-Thermus genera (viz., Deinococcus, Thermus, and Meiothermus) formed distinct clusters. Of these genera, Deinococcus was found to be the earliest branching lineage, whereas a closer relationship was seen between the Thermus and Meiothermus genera. A similar relationship among these groups is seen in the 16S rRNA trees (2, 39). It is noteworthy that the insert sequence in *Deinococcus* species consists of two alanine residues, whereas in *Thermus* and *Meiothermus* species, the insert sequence is comprised of one alanine and one lysine residue (i.e., AK), again indicating a closer relationship between these two genera. Thus, the inference from signature sequences is in accordance with results from phylogenetic analysis (2, 6, 39, 42). We have also performed phylogenetic analysis on these sequences after omitting the insert region. The tree obtained in this case was very similar to that in Fig. 2 (results not shown), indicating that the observed relationship is not dependant upon or affected by the presence of the insert.

Aminoacyl-tRNA synthetases play an essential role in protein synthesis by catalyzing the attachment of correct amino acids to the 3'-terminal ends of their cognate tRNA to form the aminoacyl-tRNA, which provide the basic substrate for protein synthesis (21). We have identified a 7-aa insert in a conserved region of ThrRS which is uniquely present in all three available sequences from *Deinococcus-Thermus* species (*D. radiodurans, T. thermophilus*, and *T. aquaticus*) but is not found in any other bacterial homologs (Fig. 3). The specificity of this signature was tested by PCR amplifying fragments of

			72		105
	rE. coli	P00955	RHSCAHLLGHAIKQLWP		HTKMAIGPVIDNGFYYD
	Pse, aeruginosa	Q91099	YYY-		TAVEE
	X. fastidiosa	Q9PFE2	YYY-		EAVAD
	H. influenzae	P43014	•••••F•		DVT-E
	Pas. multocida	P57857	F•		NVTK
	V. cholerae	Q9KMN7	Y-		NATS
	Buch, sp.	P59554	-YQSYNI		LAQI-TSNI-EDC-
	Ral. solanacearum	Q8XZ29	TAY-V-E-Y-		EAQVTE
	Nei, meningitidis	Q9K095	Y.		NAVEE
Proteobacteria	Rh. sphaeroides	ZP 00006108	DLIMAR-VQE		DV-VTVAW
	Rhodo, rubrum	ZP 00013438	DAVMAQ-VQE-Y-		G-QVTA-E
	B. prowazekii	Q92IX4	DATAE-V-E-F-		E-QVTA-EY
	A, tumefaciens	Q8UEL1	DAVMAE-VQE		G-QVTE
	C. crescentus	Q9AAX8	DTV-AE-VQE-F-		G-QVTNVED
	Des. desulfuricans	ZP 00129320	A VMAD - VQR - F-		GV-VTA-E
	Geo. metallireducens	ZP_00079665	TSMAQ-V-E-F-		QA-VTAVE
	Hel, pylori	Q9ZMV3	AQSL-A-Y-		DA-FFVVEE
	LCamp, 1eiuni	NP 281416	S-Y-		EA-FFVED
	-Aqu, aeolicus	067583	LIMAQ-L-E-YGA	٩K	KVHLGV TTEE
Aauifex.	Cb. tepidum	NP 663000	SSMAEE-F-		GA - FGA A - EQ
Chlomydiae	Cvt. hutchinsonli	ZP 00119798	SAE-LEA-Y-		GFGA-ET
Cinalityulae	Chlam, nneumoniae	09Z7A0	TSAQ-VLR		DAIPTH
	Chl. trachomatis	NP 220096	TSI-AQ-VLR		SAQPT Q
	Bor, burgdorferi	NP 212854	IVMAE-VLD-F-		NIP-KD
	Tre, pallidum	NP 219273	LVMAE-VQA-F-		GL-VPY
	len, interrogans	NP 7114210	Q A M - VQN - YK		NANLTVPGFF
Spirochetes.	Thermosyn, elongatus	NP 682992	TTSHI-AM-VQK-F-		KAQVTW-E
Cyanobeeterie	Tri. ervthraeum	ZP 00073179	TTS-VMAM-VQK-F-		KAQVT W- E
Cyaliobacteria,	Svn. sp. WH 8102	NP 897571	MS-VMAM-VQF-		KARVT WTES
GNS Bacteria	Sv. sp. PCC 6803	NP 442489	TTS-VMAM-VQK-F-		KAQVTWTET
	Pro. marinus	ZP 00112884	FIVY-		NI
	Nostoc sp. PCC 7120	NP 488763	TAS-VMAM-VQK-F-		KAQVTW-E
	Cfx, aurantiacus	ZP_00018820	LVMAQ-VLEIF-		DA-IP-E
	-D. grandis *	AY542779	LG-VMSR-VGEYYK	GKGYGAD	AI-RGVS-EW-Q-
	D. radiodurans	NP 295804	LG-VMSQ-VGEYYK	AKGYGPD	AI-RGVY-EW-Q-
D .t	The, thermophilus	P56881	TLV-AQ-V-EFFR	EKGYDPE	SVRLGVEK
Deinococcus	Mei, ruber *	AY452781	TLV-AQ-VRE-YT	ERGYRPE	EV-LE
Thermus Group	Mei. svlvanus *	AY452780	TLVMAQ-VREFFA	AKGFDPD	AV-LE
	-Bif. longum	NP 695905	AT-VMAQ-VQEVY-		NA-LGVKD
	Thermobif. fusca	ZP 00057412	V-AQ-VQE-F-		EA-LGP-E
	Str. coelicolor	NP 625810	TVMAQ-VQE-F-		EA-LGPVKD
	Cor. glutamicum	NP 600883	V-AQ-VQAEF-		GLGA-E
	Mvc. tuberculosis	NP_337191	TV-AQ-VQE-F-		QA-LGP-TD
	Myc. leprae	NP 301410	TV-AQ-VQD-F-		QA-LGP-TD
	Tro, whipplei	NP 787390	V-AQ-VQSIYG		DAKLG FTE
	Fuso, nucleatum	NP 603508	TMAQ-VLR-Y-		DVTE
Crom(+)ve	Lis. innocua	NP 470930	TMAQ-L-R-Y-		DV-FGVA-ES
Gram(1)ve	Bac, subtilis	NP 390773	AAQRIYK		DV-FGVE
Bacteria	Sta. aureus	NP_646443	TMAR-YG		NV-FGVEG
	Ent. faecalis	NP 816480	SMAN-LRRLF-		NI-FGVAS
	Thermo, tengcongensis	NP 623309	TSS-I-AQ-V-R-FK		DVKLA
	Clo. thermocellum	ZP_00060721	TTS-I-AQ-V-R-Y-		DA-LA-E
	Clo, perfringens	NP 563238	TAS-V-AA-V-R-F-		QD-LS
	Oce. ihevensis	NP 693075	SAQR-FN		DV-LGVEE
	Strep. pneumoniae	NP_359065	AFAQ-ARR-F-		DIHLGVA-ED
	Strep. mutans	NP 721923	AFAQ-ARR-F-		DIHLGV - A-QD
	lao laotie	NP 268068			DIHLGVA-QD

FIG. 3. Excerpt from a sequence alignment of threonyl-tRNA synthetase showing a 7-aa insert (boxed areas) that is distinctive of the *Deinococcus-Thermus-Meiothermus* species. Cb., *Chlorobium*; Pro., *Prochlorococcus*. See the legend to Fig. 1 for an explanation of additional abbreviations used.

the ThrRS gene from several additional species, viz., *D. grandis*, *M. silvanus*, and *M. ruber*. The sequences for these species are included in Fig. 3, and all of the sequences were found to contain the identified signature. These results strongly indicate that the identified indel in ThrRS is likely a group-specific signature for the *Deinococcus-Thermus* group. In a phylogenetic tree based on ThrRS sequences (data not shown), all of the *Deinococcus-Thermus* species were found to group together with high affinity (95% bootstrap score), supporting the inference that they form a monophyletic group.

The core subunits of the RNA polymerase (i.e., α , β , and β') are evolutionarily conserved in sequence, structure, and function in all species ranging from bacteria to humans (24, 37). In the β' subunit of RNA polymerase, which is encoded by the *rpoC* gene, we have identified a 1- and a 3-aa insert in conserved regions that are only present in *D. radiodurans*, *T. aquaticus*, and *T. thermophilus* but are not found in any other bacteria or species (Fig. 4). Further studies on this indel were

carried out by cloning and sequencing fragments of the *rpoC* gene from three other *Deinococcus-Thermus* species (*D. grandis, M. ruber*, and *M. silvanus*). Results of these studies, which are included in Fig. 4, show that both of these inserts were present in all of these species, indicating that they are distinctive characteristics of the *Deinococcus-Thermus* phylum. Furthermore, as seen in the case of σ^{70} homologs, the sequence of the 3-aa insert in various *Meiothermus* and *Thermus* species (i.e., KDE) was identical and differed from that seen in the *Deinococcus* species, pointing to a closer relationship between the *Meiothermus* and *Thermus* species.

The L1 protein of the 50S ribosomal subunit has been implicated in the release and removal of deacylated tRNA from the E site (32). Our studies have revealed a conserved 1-aa insert in L1 protein which is present in the available *Deinococcus-Thermus* species (Fig. 5). We have amplified 510-bp fragments of the L1 protein gene from two other members of this group (*M. silvanus* and *D. grandis*), and both were found to contain the indel. The insert in all cases is a lysine residue,

			300	330	393	342
	FE. coli	NP_418415	QEAVDALLDNGRRGRAITGSNK	RPLKSLADM	TIKAAKKMVEREEAV	VWDILDEVIREHPVLL
	H. influenzae	NP_438672	SR		D-I	A
	X. fastidiosa	NP_780180	ST		GLAE	BSV-
	Buch. aphidicola	NP_239874			SI	H
	V. cholerae	NP_229983	S			
	Pse. aeruginosa	NP_252959			LPE	V-A
Ductor	Sal. typhimurium	NP_457916				
Proteo-	Pas. multocida	NP_246675	S		D-1	AD
bacteria	Nei. meningitidis	NP_273191	K-MA		-VDQ-VPE	BIM-
	Ral. solanacearum	NP_521154	K-MA		E NORD	B
	Burk, fungorum	2P_00027918	K-MA		DENQIP-	E B B B B B B B B B B B B B B B B B B B
	Nit. europaea	NP_842033			-V-OK-PPF	M-
	C grasgantus	NP 419322		J	-V-Q- 1 R RID	R
	A tumefaciong	NP 354930	SFVA		-V-0LK-KPE	
	R prowazekii	NP 220532		FS	RA-KPE	V-E
	Geo metallireducens	ZP 00080693	K-MA		SK-RPE	V-EKM-
	Des. desulfuricans	ZP 00129102	SFT-G	S	SL-	BVY-I
	Camp, jejuni	NP 281666	FAN-VK-A	SEI	-V-QI-NKTNE	EC-EVKGM-
	Hel, pylori	AAD38400	V-FSTN-VK-A	SEI -	L-QR-I-QKSNE	EC-Q-ITEGY
	Aqu. aeolicus	NP 214332	IKNPVK Q-G	Y	SHRLQKTPE	EC-EVK
Aquifar	Chl. trachomatis	NP_296964	FH-HPVM-AGN	SE	-RSIQ-GAPE	V-E-I-KG
Ацијел,	Chlam. pneumoniae	BAA98292	FH-HPVM-AGN	SE-	RSIQ-GAPE	V-E-I-KG
Chlamydiae,	Cb. tepidum	NP_661062	FS-KAN-VKTGES	NS-A	SV-SLIDKKDP-	V-EKDG
CFBG	Cyt. hutchinsonii	ZP_00119275	S-FS-KIN-VRAEGN	-AS	-V-SI-D-KDP-	EN-LKGI
Group	Tre. pallidum	NP_218682	FSK-KPK-ASN	IS	NKMLQ-SPK	-FSVVKM-
orvep	Bor. burgdorferi	AAB91502	SS-FSHKRKVVKSS	S-A	NRNLI-Q-VDE	QLKI
	Lep. interrogans	NP_713599	FSKVK-KGN	IS	NSKA-DKE	-F-V-EY-VKM-
Spirochates	Nostoc sp. PCC 7120	B32838	TVV-A-N	S-I	NLIS-NDPS	V-EEGM-
Cua-abaatania	Tri. erythraeum	ZP_00070897	TVV-A-N	S-I	NLIQKGDPN	V-EDGM-
Cyanobacteria,	Pro. marinus	NP_895333	TVV-A-N	S-I	NLIQ-ADDE	V-QEGI
GNS Bacteria	Syn. sp. WH 8102	NP_896707	TTVV-A-N	S-1	NLIQ-ADDE	QV-QDGIM-
	Thermosyn. elongatus	NP_681430	TTVV-A-N		NK-IQ-NDPQ	1V-EEG
	Sy. sp. PCC 6803	NP_441586	T WO KO	-AS-1	N N DA MODE	I-SV-BIGM-
	-CIX. aurantiacus	ND 204635	T CDU NDCC	D . C . D T . T	NO-PIVPDT	PDS A-E EDKV
	D. fadfoddfalls	NP_294635	SPV-NPGS	D S-RT-L	NQ-RLYRDT	RDSEDKV
	The thermorbilug	H1455662	APV-NPGS	D RT-T	NVRR-LORDI	KDE A-EHGKV
Deinococcus-	The amaticus	OARMIE	SPV-NPGS	B BT-T	NVRR-LORDI	KDEA-EHGKV
Thormus	Mei silvanus *	AY455864	TPV-NPGS	D -A-RT-T	NVRR-LORDI	KDEA-EHGKV
Casar	Mei, ruber *	AY453861	TPV-NPGS	D -A-RT-I	NV-S-RL-CSRDI	KDEA-EHGKV
Group	Clo. perfringens	NP 553328	PVPGN	L]	NSRVMPQ	V-EAD
	Clo. acetobutylicum	NP_349740	PVPGN	S	NSRVQNQ	V-ESDM-
	Fuso. nucleatum	NP_602821	KPVVAQ-N	-ES	NML-ESDDK	AVIEDAD
	Thermo. tengcongensis	NP_623838	PGN	S	HSRVRPE	V-EK
	Bif. longum	NP_696374	S-FPVASN	S	NM-SRL-D-GD-E	GV-ES
	Thermobif. fusca	ZP_00058085	F PV PGN	S	NSRARP-	V-ES
	Myc. tuberculosis	NP_335108	SFPVPGN	S-L	NSRQRPQ	V-EA
~ ~ ~ ~	Str. coelicolor	CAD55212	PVPGN	S	NSRGRT-	-Y-V-EA
Gram(+)ve	Tro. whipplei	NP_789030	PVS-T-N	-TS	N-RS-RR-IGDPA	MAKAR
Bacteria	Cor. glutamicum	NP_599734	SFPVPGN	S-L	NSRQRPE	V-E-A-SM-
	Bac. subtilis	CAB11884	PVPGN	SH-	NSRKIVQPE	V-ESK
	Bac. halodurans	NP_240993	PGN	SH-	NI-SRKVQPE	V-EK
	Ent. faecalis	NP_816835	T DV DON	SH-	NU-NKKIG-DE	QQ
	Sta. aureus	NP_645315	T PV PGN	SH-	MCDKIMDDE	V-EDK-
	Degulfit hafmionco	ZD 00007957			N-S-PWPD	V-RP
	Lis innorue	NP 469631			NSRKTMADP	fV-E
	Strep preumonize	NP 359369		SH-	NVRICDER	IK
	Lac. lactis	NP 267956	T-IPAGN	SH-	N-RRKODSD	V-ET-VK
		and the second sec	- 11011			

FIG. 4. Partial sequence alignment of RNA polymerase β' subunit (RpoC) showing 1- and 3-aa conserved inserts (boxed area) that are specific for the *Deinococcus-Thermus-Meiothermus* species. This sequence region is highly divergent in *Thermotoga maritima* (data not shown); hence, it is difficult to infer the presence or absence of the inserts in this species. See the legends to Fig. 1 and 3 for abbreviations used.

indicating that it was introduced only once in a common ancestor of these species.

A 2-aa insert is also found in the exinuclease ABC subunit A homologs (i.e., UvrA protein) of the Deinococcus-Thermus group (Fig. 6). UvrA is one of the two subunits of the damage recognition complex required for nucleotide excision during repair of UV light-induced DNA damage (8). Previously, sequences were available only from two species belonging to this phylum (D. radiodurans and T. aquaticus), and no information existed for the *Meiothermus* group of species. To bridge this gap, we have amplified a 639-bp fragment of the uvrA gene from M. silvanus. The amplified fragment contained the 2-aa insert, providing evidence that this signature is also a distinctive characteristic of the Deinococcus-Thermus group. In addition to the Deinococcus-Thermus species, a 2-aa insert is also present in this position in Borrelia burgdorferi, which may have originated either independently or through lateral gene transfer (LGT).

All sequenced organisms contain an Ffh/SRP54 family mem-

ber, which forms part of the signal recognition particle and coordinates the cotranslational targeting of secretory and membrane proteins to either the membrane of the endoplasmic reticulum or the plasma membrane in bacteria (29). In E. coli, the signal recognition particle is composed of Ffh protein and the 4.5S RNA. A 5-aa deletion is present in a conserved region of the Ffh protein that is only seen in Deinococcus-Thermus homologs but is not found in any other bacteria (Fig. 7). Since sequence information for Deinococcus-Thermus was available only from D. radiodurans and T. aquaticus, we have amplified and sequenced a fragment of the Ffh gene from D. grandis. The fragment from this species was also found to contain the deletion (Fig. 7), indicating that this signature may also be specific for the entire Deinococcus-Thermus phylum. Due to DNA limitation, sequence information for this signature for Meiothermus species was not obtained.

Another signature for the *Deinococcus-Thermus* group is present in the protein SerRS (21). The signature in this case consists of a 2-aa insert in a conserved region that is commonly

			111	152
	rE. coli	15804574	FDVVIASPDAMRVVGO	LGOVLGPRGLMPNPKVGTVTPNVAEA
	H. influenzae	16272460		
	V. vulnificus	27364613	V	TT
	Sal enterica	16762303		**
	Pas multocida	15603607		T
	Pse aeruginosa	15599469	V	D
	Ph epharoidee	22957091		KI N MD VO
	A tumofacione	15000252	RCTM-PLR	K M MD C
	Duch aphidicale	15669253	RCIM-PLR	KMMDG-
Protoobactoria	Buch. aphidicola	21672331	-HTAK1-T-	I
1 Toteobacteria	X. Tastidiosa 9a5c	15839225	YK	TLSQ-PG
	Nel. meningitidis	15676056	I	TI
	Nit. europaea	22955871	LA	IVD-IN-
	Ral. solanacearum	17547756	T	IDT-
	C. crescentus	16124893	RTM-ALR	KQ-
	R. prowazekii	15604013	CTM-SMISS	VARIKLLDIKN-
	Burk. fungorum	22982689	IT	IDT-
	Rhodo. rubrum	22966133	RCTM-GR	KMD
	Geo. metallireducens	23054532	TATM-GK	I-KLFD-GR-
	Des. desulfuricans	23473811	SATV-AL	I-RFDK-
	Hel. pylori	1840150	MTM-AK	V-RIKTMDI-K-
	L _{Camp} . jejuni	15791839	LT-NL-GLK	V-RIKTMDQ-
	Aqu. aeolicus	15606946	AT-EM-PK-AK	RISTTEQ-
	Cb. tepidum TLS	21672993	IV-TV-GOL-K	VARISMDK-
Aquifex,	Cvt. hutchinsonii	23137574	I-TM-TV-AKK	KINAT-DK-
Chlamydiae	Chlam, pneumoniae	15618002	AV-TMEK	KTTTTDK-
Cillainyulae,	Chl. muridarum	15835209	AV-TMEK	KTD-VK-
CFBG	Chl. trachomatis	15605039	AV-TMEK	KNTA-
	-Tre pallidum	15639230	AVI N L R	
	Bor burgdorferi	15594737	V M KD K	- BI - K TO N I KD
	Len interrogans	24216122	ACV-TM-KDK	
	Tri erythraeum	23043769		PIRC
Spirochetes	Bro marinua	22122070	II T M DK AK	KL
Spirocacies,	Pro. marinus	16330010	DLIM-PK-AK	RGDI-S-
Cyanobacteria	Sy. Sp. FCC 0803	10330010	DIM-PKIAR	KQSGADL-A-
	Syn. sp. wh 8102	23134703	LLTM-PK-AK	RIDL-G-
	Noscoe sp. PCC /120	1/232/93	KLTV-PQ-AK	KMSGFDS-
	Thermosyn. elong.	22297840	LLTM-PQ-AK	V-RISAFDLPQ-
	D. grandis *	AY489057	A-V-TV-GAS K	RLAGF-II
Deinococcus-	D. radiodurans	15807039	A-V-TM-AQI K	-ARLLSGAD
Thormus	The. thermophilus	730540	A-V-TV-GAS K	RILAGF-IG
C C	The. aquaticus	348584	A-V-TV-GAS K	RILAGF-IG
Group	Mei. silvanus *	AY489058	A-V-TV-GAS K	RIAGFI
	T. maritima	15643221	ATMII-R	KISSQE
	[Thermobif. fusca	23018372	A-V-TM-GKI-R	RIK-
	Str. coelicolor	21223031	A-V-TL-GKR	RDK-
	Bif. longum	23465853	A-V-TM-GKR	RMD-TK-
	Tro. whipplei	NP_787841	A-ST-EL-AQ	RTAD-GK
	Cor. glutamicum	19551720	ATQ-AKI-R	IARNDK-
	Myc. leprae	15828018	AAV-TQ-AKR	IAR
	Myc. tuberculosis	15607781	AATQ-AKR	IARTADK-
	Sta. aureus Mu50	15923528	V-TM-GEK	RKTMD-KK-
Gram(+)ve	Oce. iheyensis	23097563	K	RKTFE-EK-
Gram(1)ve	Bac. subtilis	2160218	IV-TM-GEK	1-RKTFE-EK-
Bacteria	Bac. halodurans	15612683	IV-TM-AQK	RKTFE-EK-
	Lis. innocua	16799358	IV-TM-GEK	RKTMD-TK-
	Clo. acetobutylicum	15896395	R	RKSFDK-
	Clo. perfringens	18311398	YR	RKSFDN
	Strep. pneumoniae	15900538	TM-ALR	R
	Strep. mutans UA159	24380006	TM-AIP	RN
	Thermo, tengcongensis	20808671	YTM-GR	KLKSFE-EK-
	L. lactis	15673982	P	
	Fuso, nucleatum	19705330	TTM-PKT-P	
	M. genitalium	12044934	LT-TKF-GAL-K	KITTEDI-A-
	U. urealyticum	13358104	I-TNOKM-DILAF	-K

FIG. 5. Sequence alignment of ribosomal L1 protein showing a conserved 1-aa insert (boxed area) that is distinctive of the *Deinococcus-Thermus-Meiothermus* species. Burk., *Burkholderia*. See the legends to Fig. 1 and 3 for additional abbreviations used.

present in the SerRS homologs from various available *Deinococcus-Thermus* species (*D. radiodurans, T. aquaticus*, and *T. thermophilus*) (Fig. 8). By means of PCR amplification, we have obtained sequence information for the SerRS gene from two additional species, viz., *M. silvanus* and *D. proteolyticus*. Interestingly, while the *M. silvanus* homolog contained the signature, this insert was not found in the fragment derived from *D. proteolyticus* (Fig. 8). The most parsimonious explanation for these results is that the insert was introduced in a common ancestor of *D. radiodurans, T. aquaticus, T. thermophilus*, and *M. silvanus* after the divergence of *D. proteolyticus*. However, the possibility that the insert has been lost from *D. proteolyticus* cannot be excluded.

DISCUSSION

In 16S rRNA and various protein trees, the *Deinococcus-Thermus* phylum represents one of the earliest branching groups within the Bacteria (10, 16, 19, 25, 38, 39, 42, 44). In the past, this phylum consisted of only two genera (Deinococcus and Thermus); however, a third genus (Meiothermus) has recently been established (34). According to branch patterns, species belonging to the genus Meiothermus form a sister lineage with Thermus species, forming the order Thermales (family Thermaceae), which clusters together with the distantly related Deinococcales in a single lineage (2, 34, 39). Although Deinococcales shows 77.5 to 81% 16S rRNA sequence similarity with the Thermus-Meiothermus group, and species of the Deinococcus-Thermus group share an A3B murein-type peptidoglycan (L-ornithine as the diamino acid and glycylglycine as the interpeptide bridge) and menaquinone-8 as their major respiratory quinone, these characteristics are not unique to these groups, and they share few other characteristics in common (2, 4, 30, 31). Although a few unique base pairs that appear limited to the genus Deinococcus have been identified in the16S rRNA sequences (2), currently there is no molecular

		150	206
	-E. coli NP 418482	LLAPIIKERKGEHTKTLENLASQGYIRARIDG	EVCDLSDPPKLELQKKHTIEVVVDR
	Pas. multocida NP 246890	VVV-LQI-A	-IAH
	Sal. typhimurium NP_463119		•••••I••
	H. influenzae NP 438418	••••VV-N••••V•I•••I•A-•••	-IAA
	Pse. aeruginosa 15599430	V-RLAVFDEMRAFVV	KLYE-DEVDKS-D
	V. cholerae NP_230048	T-VVAF	-TIII
	X. fastidiosa 28199325	VVRQVF-Q-RAFV-V-V	-LYEIDAV-T-T-RQA-I
	Nei. meningitidis 15794105	I AVR FVDFFAD - QA FA - V - V	YQ-DEVKNIN-DI
	Nit. europaeaATCC 30250376	IVVTGQAELFDE-RAFV-V-L	Y-IDALQKT
Proteobacteria	Rhodo. rubrum 22967393	VARGFK-E-AE-QKK-FS-VKV	TIYEIPEV-A-NKKID
	Rh. sphaeroides ZP_00006403	VRDYK-EFIE-RKFQ-VKVN-	TFHE-EET-DKKFR-D-D
	A. tumefaciens NP_354519	IVRGYK-E-AE-MKK-FQ-VKV	QFYEIA-V-A-DKKYD-D
	C. crescentus NP_421392	VVRDYK-EIADWQKA-FQ-LK	QYYPIE-A-A-DKKFD-D
	R. conorii 15893217	VRGHFKREIMD-KKFQKLIVN-	EID-LDKNNI
	Geo. metallireducens 23055444	S-MVRGYR-E-AQ-RKD-FA-VIV	VQYE-AEEIP-DKND-DI
	Des. desulfuricans ZP_00129420	VMLVEHQT-ADR-KR-KAFV-I-LN-	MT-DEV-AKNS-DL
	Camp. jejuni NP_281533	IYS-L-R-KTYADLRNKVQ	VLVR-DEEIE-AKTKL-I
	LHel. pylori NP_207499	IDKSFNDKS-RLKVFV	VMVR-DEEIH-HKTA
	Aqu. aeolicus 15606094	I-SVRGKFREL-RQIEKWS-VKV	-LRRVIEV-PKNL-I
Aquifex,	Cyt. hutchinsonii 23138505	VV-GHYRELFQQIRKMTKV	-ITEITPKLQVDRY-I-DI-I
CERC Group	Bact. thetaiotaomicron 29345988	LVRSHYKELF-QVRKKLYV	-LREVTHGMDRY-N-DI-K
cr bo oroup	Cb. tepidum TLS 21674507	I-S-LVTGHYRELFDR-IKKL	-YVEM-AGMQRY-S-NL-I
	Tre. pallidum 15639505	VVRGKT-H-VAARKD-FV	ALLH-HERIS-DKS-DI
	Bor.burgdorferi 15595182	-FVRGSS-K-VKILNFN-VNS E	D YLIEDALNLN-HKNII
	1r1. erythraeum 23040104	IVVRGKI-K-L-SSAFV-L-V	VEIAENIDKNHTI-I
Spirocnetes,	Sy. sp. PCC 6803 NP_440772	1VV-GKT-VQL-SS-VFV-VN-	RENIE-KKNQAI-I
Cyanobacteria,	Pro. marinus 33862268	VVRGKI-A-L-SGAE-FA-VN-	RE-A-SIE-DKNHI-N
Green Nonsulfur	Syn. sp. wH 8102 33867052	T WROK T D L OO FW T V	RE-A-NIE-DKNHS-N
Bacteria	Nostoc sp. PCC /120 1/231208	1VVRGKI-R-L-SSFV-I-V	RESIE-DKNIII
Dacteria	Inermosyn. eiongatus 22297728	VV-GRI-R-L-SSE-FV-V	TRO DEEL NIKKY O T
	D padiadupana 15906770		
D .	The aquations 2492564		LITELEEACUS KVE D DI T
Demococcus-	Mai eilvanue * AV480059		
Thermus	Thermotoga maritima NP 228200		TYP-EEV-E-DKN-RVKI
	-Cor efficiens 25028061	VAT-R	0.Y0E.T.K.T.D.D.
	Cor glutamicum 19552593		HOK-ID-D
	Thermohif fusca 23019663	VVVBGY-ELEOS-OAK-ETV	VI VB-DEA-E-KRYET-D-A
	Myc. tuberculosis 15608776	VVVBTFADI FDK-NAS-V-V	V-HP-TKK-ED
	Myc. leprae 15827727	VVVBTFADLEDK-NAS-V-V-D	V-YPKK-ED
	Bif. longum ZP 00120387	IVV-GFVDMI-L-R-DAL	-MTQDITK
	Tro, whipplei AA044393	S-LVRQKFVDLFSR-TTS-FS-VVV-E	- IYL-TEVKKNSRK
	Str. coelicolor 21220443	V-S-LVRFVDLFAD-QTKSV	- TVQNT-KK-E
Gram(+)ve	Ent. faecalis V583 29375356	IVVVKKQ-K-VF-MIQREV-M-V	-TY-V-EA-EKND-AI-I
Bacteria	Thermo. tengcongensis 20808374	IVVRGYS-L-ADIKKSV-VK	IMY-VNEEIDKN
Dacteria	Oce. iheyensis 23099942	IVVSGV-IKQEV-I-V-N	-MREVT-DIQKNSI
	L. lactis 15673800	IVVRTKT-V-MF-RIQKDV-V-V	Y-I-EV-E-DKNNI-I
	Strep. pneumoniae 15900123	I V - RKK Q - KSVI - KVQKD V - V - V	Y-VTEV-E-SKS-Q-N-D
	Strep. mutans 24380212	IVVRRKQ-KAVFDRIQKDV-V	DIM-VAEV-E-SKN-M-N
	Bac. subtilis 16080569	VVSGA-V-VQIRKV-V	-MAEDIEKNSI
	Bac. halodurans 15616156	ILVSGT-V-VDIKKFV-V	-MREVAEEIE-DKNI
	Fuso. nucleatum 34763030	V-DKT-KNIFLFKK-FV-T-V	-ILY-E-EIE-DKNS
	Sta. aureus 21282450	V-AHS-E-LI-DIGKKV-L	- IV - VN - V - T - DKN - N
	Lis. innocua NP_471961	IMVSGKT-KI-EIKKEV-I-V	-IY-IN-EIEI-KNSIII
	^L Clo. perfringens 18309330	IRGT-E-VIKKFV	EIYDLTE-EIKNINA

FIG. 6. Partial sequence alignment of UvrA protein showing a 2-aa insertion in different *Deinococcus-Thermus-Meiothermus* homologs (boxed area). The insert seen in *B. burgdorferi* could either have occurred independently or have been derived by means of LGT. See the legends to Fig. 1 and 3 for abbreviations.

or structural marker known that is distinctive to the entire *Deinococcus-Thermus* phylum which might be used to distinguish or define this group of bacteria from all others.

In the present work, we have identified eight conserved indels in seven widely distributed proteins that are distinctive characteristics of the Deinococcus-Thermus phylum. Based on the work reported here and information available in the databases, information for six of these proteins containing seven signatures (viz., SerRs, ThrRS, σ^{70} , RpoC, UvrA, and ribosomal L1 protein) is available from all three genera within the Deinococcus-Thermus phylum. The sequence information for Ffh/SR54 is currently available from only Deinococcus and Thermus genera, but based on the observation that Meiothermus forms a sister lineage with Thermus species (9, 33, 39), it is expected that this signature will also be found in Meiothermus organisms. Except for the absence of the SerRS insert in D. proteolyticus, the identified signatures are present in all Deinococcus-Thermus species examined but not in other bacteria. These signatures thus provide molecular markers for distinguishing the *Deinococcus-Thermus* phylum from all other bacteria and for identifying new species related to them based simply on the presence or absence of these signatures. The presence of these distinctive signatures also provides strong evidence for the monophyletic nature of the Deinococcus-Thermus phylum as indicated by 16S rRNA trees (38, 42, 44). The most likely explanation for these signatures is that they were introduced in a common ancestor of this lineage and then were passed on to all descendants. This inference is also supported by phylogenetic analysis based on a number of these proteins. The presence of the insert in SerRS in various Deinococcus-Thermus species, but not D. proteolyticus, might be accounted for by two different possibilities. First, it is possible that this insert was introduced in a common ancestor of the other Deinococcus-Thermus species after the branching of D. proteolyticus. Alternatively, this insert may have been introduced in a common ancestor of the entire phylum but then subsequently lost from D. proteolyticus. We favor the first of these possibilities, based on the observation that in phylogenetic trees de-

			215 265
	FE. coli	NP_417101	LGMGDVLSLIEDIESKVDRAQAEKLASKLKKGDGFDLNDFLEQLRQMKNMG
	Sal. typhimurium	NP_461607	KKKKK
	H. influenzae	NP_438280	D-T-D-RIEK
	Buch. aphidicola	NP_240210	N-IMBQS-IQTKHD-NT-IKKI-
	Pas. multocida	NP_246120	N-T-ERI-RSEKM-Q-FN-T-ERI
	V. cholerae	NP_230211	EE
	X. fastidiosa	NP_297366	-DW-QV-QAQEKTAA-VAKK-NMKEQ
	Pse. aeruginosa	NP_252435	ERDQA-QNLDKK-IKERDQ
	Ral. solanacearum	NP_520932	I-A-V-EAQRGMEQA-ITGEKA-IGK
Proteobacteria	Nei. meningitidis	NP_273349	TKIQR
	Nit. europaea	NP_841501	GGEAQRTS-QKER-MK-M-S-KSQQ-FQK
	Rh. sphaeroides	ZP_00006756	IVA-V-KAQETFEAEA-MMKRFQ L-NMLKMELK
	Rhodo. rubrum	ZP_00016149	VA-V-KAAETIEMDRV-KRMME-KE-MLA-II-K
	A. tumefaciens	NP_355632	IVV-KAAENI-AEK-RSM-E-MAKLADGK
	R. proważekii	NP_220563	-DII-FVKKAA-IEET-IS-KY-Q-M-SI-K
	C. crescentus	NP_422447	QVA-V-KA AADLDQAEAERMA-KLAKGKD-LAAQK
	Hel. pylori	NP_207943	M-AIVA-KTA-VLNPNE-KD-SKQ -TFN-IEKV-KL-
	Geo. metallireducens	ZP_00081709	SQERSQ-I-K
	Des desulfuricans	ZP_00130373	TKAQTTIKEEEE-GR-MQRAEFERT-M-R-RKL-
	-Camp. jejuni	NP_281881	M-ELAT-A-KTAAII-EKE-KNQ-IE -NFN-MESI-KL-
	Aqu. aeolicus	NP_214181	LIQV-KAQEVITEDK-QVM-T-VMT-EE-LR-MMIQQ
Aquifex,	Cyt. nutchinsonii	ZP_00119217	VV-RAQQVF-EEETKNQR-NQ -NFEN-IQ-I-K
Chlamvdiae.	Bact. thetaiotaomicron	NP_810514	IVV-RAQEQY-EEE-KR-QK-IA-NQFS-IA-I-K
CERC	Chlam proumonico	NP_662058	IV-FV-KAQEAL-LEKIMAMQ-~-M-NEDFDQ-L-K
Сгво	Chl muridarum	ND 206672	TI UDENDEC CEEDNEE DE CAR THE VEN APRE
	Chl trachomatic	NF_290073	TI VOKMORCIOERENKE DE TAN WYR UN TANDON
	-Por burgdorfari	NF_219527	V V V V V V V V V V V V V V V V V V V
	Tre pallidum	ND 218856	
Spirochetes.	Nostoc sp PCC 7120	NP 485992	
Cyanobacteria	Pro marinus	NP 894182	T V KAQSSF B D MOE IDBAR V FIVVR-ME-
CNS Destaria	Tri. erythraeum	ZP 00071675	T-V-KAOEEE-I-DMOE-II-SAKETK-T-L
GIVS Dacteria	SV. SD. PCC 6803	NP 441626	-NT-V-KAOBAT-VGDVON-TLBATFDTK-M-F
	Thermosyn, elongatus	NP 682763	T-V-KAOEEL-DMSR-TLEAOFDKL
	Cfx. aurantiacus	ZP 00017969	RA-OLY-AEK-MOKR-SEFENS-HRKI-
	rD. grandis *	AY453858	
Deinococcus/	D. radiodurans	NP 295559	GRAQAADLK-MEV-K PE-L-LLRKL-
Thermore	The. aquaticus	AAB58502	AKE LS-EK-MONL-RL-
Inermus	T. maritima	NP_229365	KV-KEL-QEKMK-S-E-FL-AE -T-EKQEKL-
	_[Cor. glutamicum	NP_601261	QA-AVM-QEKVA-QGS-E LT-ED-MLMIRR
	Bif. longum	NP_695507	-DIMTQAQKQF-EEE-R-A-E-ISE-S -G-DDQ-VRKL-
	Thermobif. fusca	ZP_00059090	-DTQAQRTFEEEEVA-MT-ASDED-T-EMMMIRKL-
	Myc. tuberculosis	NP_337496	QA-QVF-AQEA-A-IGA-E LT-EMLAVRK
	Str. avermitilis	NP_823824	-DL-TQA-KTFSQEEMASKK GQDFTLDDF-SEQVR
	Str. coelicolor	NP_629720	-DL-TQA-KTFSQEEMASKK GQDFTLDDF-AEQVR
Gram(+)ve	Thermo. tengcongensis	NP_623080	TKAQAAI-EKK-LEMGQ-ILSKQ -T-ESL
Bacteria	Lis. innocua	NP_471249	KAQTDAEKMKAMEQ-M-DNS MT-DDQ-V-Q
	Sta. aureus	NP_645937	KAQQDQEK-KD-EK-MRESS -T-DD-VL-
	Bac. subtilis	NP_389480	TKAQASEDK-KE-EQ-MRTMS -T-DG-VR
	strep. pyogenes	NP_802303	L-TKASQEY-EKKSLEE-MRENTFID-VQ
	dla anababubulian	NP_692452	MG-V-Q
	LID. acetoputylicum	NP_348380	QAFEKL-
	M proumonico	NP_0/2/08	LMT-V-RA-UVF-KKDLT-TIMFL-K ME-L-IYMQHK
	Euro nucleatum	NP_109/49	LMI-V-KA-EVF-KQSLT-TVMFL-K ME-L-LYMENQ
	L lostia	NP_00428/	M W KAOANY DE CA E MAENE VE V
	H. IGULIS	MP_201112	NACANI-DE-SAE-MAENKIEVD-VT

FIG. 7. Partial sequence alignments of Ffh protein showing a 5-aa deletion (boxed area) that is a unique characteristic of the *Deinococcus-Thermus-Meiothermus* homologs. See the legends to Fig. 1 and 3 for abbreviations.

rived from 16S rRNA sequences, a branch comprised of *D. proteolyticus* and *D. radiophilus* forms the deepest group within the *Deinococcus-Thermus* phylum (2, 39).

LGT is indicated to have played an important role in the evolution of the Deinococcus-Themus group. These organisms are thought to have received genes from a number of other phyla such as the Archaea, Eucarya, and cyanobacteria (11, 26, 36, 43). However, for the various genes studied in the present work, which contain identified signatures, there is no evidence of lateral gene exchange between the Deinococcus-Thermus group and other bacterial phyla, except possibly the UvrA gene in B. burgdorferi. If these genes were subjects of LGTs, one would expect a more random distribution of these signature sequences in which these indels would have been present in other groups of bacteria and at the same time several Deinococcus-Thermus species would be lacking them, which is clearly not the case here. However, in contrast to these genes, a number of genes studied in earlier work contained signature sequences that were commonly shared by cyanobacteria and the Deinococcus-Thermus species, which may be the results of LGTs (13, 18).

We have also previously described many main-line signa-

tures (i.e., indels commonly shared by a number of different bacterial phyla), which provide useful information concerning the phylogenetic placement of the Deinococcus-Thermus group within the bacterial domain (13, 15, 17). The distribution patterns of these signatures in bacterial sequences indicate that the Deinococcus-Thermus phylum has evolved after the divergence of various gram-positive phyla (viz., Firmicutes, Actinobacteria, Clostridia, and relatives) but before the emergence of Aquifex, Chloroflexi, cyanobacteria, spirochetes, the Chlamydia-Cytophaga-Flavobacteria-Bacteroides-green sulfur bacteria group, and proteobacteria (15, 17). The branching of the Deinococcus-Thermus phylum in between the gram-positive bacteria and gram-negative bacteria also accounts for a hitherto puzzling characteristic of Deinococcus. Although all Deinococcus-Thermus species are surrounded by an outer membrane, which is a distinguishing property of the gramnegative bacteria, most species belonging to the genus Deinococcus (all except D. grandis) exhibit positive Gram staining and contain a thick sacculus characteristic of gram-positive bacteria (2, 30, 31, 41). These seemingly contradictory properties are readily explained by the suggested placement of the Deinococcus-Thermus phylum between the gram-positive bac-

Proteobacteria R. coli NP_415413 HQPDKVEMVQIVRP EDSMAALEEMTGHAEKVLQLI Sal. typhimurium NP_459939 D- DKEL- DKEL- N. influenzae NP_43284 D- DKEL- DKEL- Y. fastidiosa NP_29565 EL-SVCK Q8-EGHQRRCTEM- Pas. multocida NP_251302 D- AT-QVLQ Pse. aeruginosa NP_251302 D- AT-YEGL-AN-R V. cholerae NP_230755 LK- -K-YEKL-N Buch. sp. APS NP_240135 QHNI-RA- Nit. europaea NP_840274	LGLPYRKI I LCTGDMGFGACKTYD
Sal. typhimurium NP_459939	$\begin{array}{cccccccccccccccccccccccccccccccccccc$
H. influenzae NP_438284	-NVLS -EVLS -EVLA
X. fastidiosa NP_299565 EL-SVCK- Q8-ECGHQR-RC-TC-EM- Pas. multocida NP_245195 EA -T-QVL-Q Pse. aeruginosa NP_251302 D- AT-YEGL-AN-R V. cholerae NP_230755	$\begin{array}{c}$
Pas. multcoida NP_245195 A- -T-OVL-Q Pse. aeruginosa NP_251302 D- AT-YEGL-NN-R V. cholerae NP_230755 D- AT-YEGL-NN-R Buch. sp. APS NP_240135	$\begin{array}{c} E = & \cdots & VL = & \cdots & S = & \cdots \\ = & E = & \cdots & VL = & \cdots & S = & \cdots \\ = & E = & \cdots & V = & \cdots & S = & \cdots \\ = & E = & F = VVV = & \cdots & S = & \cdots \\ = & E = & F = VVV = & \cdots & S = & \cdots \\ = & E = & V = V = & \cdots & S = & \cdots \\ = & E = & V = V = & \cdots & S = & \cdots \\ = & E = & V = & T = & \cdots & S = & \cdots \\ = & E = & V = & T = & \cdots & S = & \cdots \\ = & E = & V = & T = & \cdots & S = & \cdots \\ = & E = & V = & T = & \cdots & S = & \cdots \\ = & E = & V = & T = & \cdots & S = & \cdots \\ = & E = & V = & T = & \cdots & S = & \cdots \\ = & E = & V = & T = & \cdots & S = & \cdots \\ = & V = & T = & T = & \cdots & S = & \cdots \\ = & V = & T = & T = & T = & \cdots & T = & \cdots \\ = & V = & T = & T = & T = & T = & \cdots \\ = & V = & T = & T = & T = & T = & T = & \cdots \\ = & V = & T =$
Pse. aeruginosa NP_251302 D- AT-YEGL-ANR V. cholerae NP_230755 D- AT-YEGL-ANR Buch. sp. APS NP_240135 Q- L Ral. solanacearum NP_520458 Q- -T-FEDA-HNI-RK- Nei. meningitidis NP_274688 R- -K-YETVNI-KA- Nit. europaea NP_840274 R- -T-YDQ-VTI-K- Burk. fungorum ZP_00029558 A- -T-YDQ-VTI-K- C. crescentus NP_420806 A- -T-YDQ-VTI-K-	E VLA S -E V-T S E F-TVV ST E V-T S-A E F-TML S-A E F-TML K
Proteobacteria V. cholerae NP_230755	
Proteobacteria Buch. sp. APS NP_240135	
Ral. solanacearum NP_520458 Q- -T-FE-DA-HNI-RK. Nei. meningitidis NP_774688 H- -K-YETVNI-KA. Nit. europaea NP_840274 IK. -K-YETVNI-KA. Burk. fungorum ZP_00029558 A- -T-YDQ-VTIK. C. crescentus NP_420806 YLS-TT- DQ-E-EHQR-VECTKK. R. prowazekii NP_221133 GLS-TT- DQ-E-VENDE-TKK.	EF-TVVST ELV-TS-A ELV-TS-A EF-TMLS-A
Net. meningitidis NP_274688	-EV-TS-A -ELVMLS-A -EV-TS-A -EF-TMLK
Nit. europaea NP_840274	-ELVMLS-A -EV-TS-A -EF-TMLK
Витк. rungorum 21_00029558АT-YDQ-VTLК. С. crescentus NP_420806YL-S-TT- DQ-E-EHQR-VEC-TКК- R. prowazekii NP_221133GL-S-TTVMUH-YT-NSC-TL-V-	-EKK
C. crescentus NP_420806YL-S-TT- DQ-E-EHQR-VEC-TKK. R. prowazekii NP 221133GI.S-TT- P-FWDH-VT-NPC-TI-VF	-EF-TMLK
IR. DIOWAZERII NP ZZIIJA $i-S-i^{-}-i-S-i^{-}$	10 10 10 10
	-DA-K
A. tumeraciens NP_354591W-C-L-S-TDA -SAV-EH-R-AC-EKR-	HF-TLTS-R
Competition and the competition of the competition	HVVTSS-A
Get metallieutens $2F_{00050426}$ NL-RFH- SI-Y-ERDLAN-ERC	HVVES-A
Camp. Jejuni NP_201579EL-S-TKQ-DSVRALEC-SDL-SS-	AH-HLMLS-AV-
$\frac{1}{100} = \frac{1}{100} = \frac{1}$	-EH-FVQSLS-SN-1-
M_{L} depidum NP_661506D-K*M* DI*IDDKLVKD-E	VI VE C ION A C
Aquifer (vt butchingonii) ZP_00136591	P IU C C CON P
Characteristic Bact thetalotaomicron NP_813223 = F_S_=L_P_DL H_KSG00_TD_H_CTK	-BILBGS-ISAL-F-
Children Chi	-00181-88-T-KT-
CFBG Chlam pneumoniae NP 44536	
Lep. interrogans NP 714519OL-KFCKFEFEKK-ISSN-KK	-FVSS
Tre. pallidum NP 219084	ET-F-VVEV-AL-AD-VRKW-
Bor. burgdorferi NP 212360SFCFCKA -E-GVIHD-FLSIO-OIFTE	EIVINI-SE-L-SP-Y-K
Spirochetes. Syn. sp. WH 8102 NP 897985NLYWFAH- DH-AE-HAOI-ADAA	VLDLA-IS-OR
Cvanobacteria. Pro. marinus NP 895201NLYWF-H- DHSOHAOITADAA	-EV-E
CNS Roctaria Tri. erythraeum ZP 00075337NKHT-EEEH0KLVADAIA-	KILESLS-A-C
Sy. sp. PCC 6803 NP_440547NL-KL-KE-A-EHQALVADAIA-	EA-C
Nostoc sp. PCC 7120 NP_488016NL-KV-H- ST-FDEKLV-NAIA-	KV-NLT
Thermosyn. elongatus NP_682342NL-KF-HT-A-EH-ALVADFIA-	-KV-ELA-M-C
LCfx. aurantiacus ZP_00019553YMF-T- DQ-YQKLRRDECARR-	F-TKLLST
D. proteolyticus * AY453857 -E-RQYV-CGWF-KLLANGIA-	EVVQNSL-KHLMM-
D. radiodurans NP_295000 -E-RQVVLC-A DQ -EGLKWF-RLLSNGLA-	EV-QNA-KVLM
Demococcus- The. thermophilus 1SRY_AHQYVLTEA SL -A-DR-FQ-LLENEI-R	ELVEVAP-KWRQV-
The. aquaticus P34945HQYVLTEA SL -A-DR-FQ-LLENEI-R	EP-KWRQV-
Mei. silvanus * AY452782RQYVLCIA DV -E-NRWF-R-LANSIA-	EVVEVSL-KYRQVY
T. maritima NP_229180L-WVTTL-WVTTL-WVTT-	-EVVSLTSA
Bit. Longum NP_696787FVYAKQYKEH-HLLAMEQEM-AKV	/EVI-DTAAL-SS-ARKF-
Cor. glutamicum NP_602083PVYCKAEDVHQQLL-MEKEM-AAI	IEVV-DVAGL-AS-ARKF-
Thermobil: Iusca ZP_00057154NFVYAH- DEAHEEHLRLLAWEREM-D-I	IEVVVDIAAL-TS-ARK
Str. coeficior NP_228145FSY-LO-EHQRLEEWEKQW-TS-	EF-V-DVASA-L-SS-ARK
Str. dvermittills NP_2634/0BL-RVCAAP-Q-L-VEC-RR-	E-SVVQ-PAL-S-RM
myc. Luberculosis NP_388495GFVICI- A-AEHEH-RLL-WORM-ARL	EVV-DVAAL-SS-ARKF-
I lotic $MP_{26/654} = -5 \pm -FSI + QQ + RE - HIVSMQ - M \cdot N I lotic MP_{26/654} = U KPA PO VDP K AN NY K$	EIVSDIAAEEL-TS-SRK
$\begin{array}{c} \text{Cram(+)ve} \\ \text{Fise purchastrum} \\ \text{NP} 602020 \\ \text{NP} 602020 \\ \text{Cram(+)ve} \\ \text{NP} 602020 \\ \text{Cram(+)ve} \\ \text{NP} 602020 \\ \text{Cram(+)ve} \\ \text{Cram(+)ve} \\ \text{NP} 602020 \\ \text{Cram(+)ve} \\ Cram($	
U urealyticum NP 077937	P DW I CCA
Bacteria Bacteliis NP 387894	EVMCMIT. A.K
Oce internals NP 50093	
Strep, mutans NP 722195	E
Clo, acetobutylicum NP 346664N-TF-YTKD	HLSK-DDD-CSDSMA
Lis, innocua NP 472218F-K	EVI.SMA-IT-A-K
Thermo, tengcongensis NP 621737	VVATT,T-S-K
Ent. faecalis NP 816887HKFSDA -H-VFE-K-NN-GDT-FK-	V-T-SS-A
Sta. aureus NP 644824RFEORFEO	
M. genitalium NP_072665Q-T-L-KPCKNAINEA-VRDOI-KA-	KF-RLLS-E

FIG. 8. Excerpt from SerRS sequence alignment showing a 2-aa insert (boxed area) that is present in various *Deinococcus-Thermus-Meiothermus* species, except *D. proteolyticus.* This insert was likely introduced in a common ancestor of this group after the branching of *D. proteolyticus.* U., *Ureaplasma.* See the legends to Fig. 1, 3, and 5 for additional abbreviations used.

teria (monoderm bacteria surrounded by a single membrane) and gram-negative bacteria (diderm bacteria bound by both inner and outer membranes), and they indicate that this group of species may represent evolutionary intermediates in the transition between these two structurally distinct groups of bacteria (13).

ACKNOWLEDGMENTS

This work was supported by a research grant from the National Science and Engineering Research Council of Canada.

We are thankful to Peter Gogarten and Lorraine Olenzenski for providing us bacterial strains and Melanine Havers for assistance in the creation of some of the signature sequence files.

REFERENCES

- 1. Battista, J. R. 1997. Against all odds: the survival strategies of *Deinococcus radiodurans*. Annu. Rev. Microbiol. **51**:203–224.
- Battista, J. R., and F. A. Rainey. 2001. Family I. *Deinococcaceae* Brooks and Murray 1981, 356^{VP}, emend. Rainey, Nobre, Schumann, Stackebrandt and da Costa 1997, 513, p. 395–403. *In* D. R. Boone and R. W. Castenholz (ed.), Bergey's manual of systematic bacteriology, 2nd ed. Springer-Verlag, New York, N.Y.

- Brim, H., S. C. McFarlan, J. K. Fredrickson, K. W. Minton, M. Zhai, L. P. Wackett, and M. J. Daly. 2000. Engineering *Deinococcus radiodurans* for metal remediation in radioactive mixed waste environments. Nat. Biotechnol. 18:85–90.
- Brooks, B. W., R. G. E. Murray, J. L. Johnson, E. Stackebrandt, C. R. Woese, and G. E. Fox. 1980. Red-pigmented micrococci: a basis for taxonomy. Int. J. Syst. Bacteriol. 30:627–646.
- Burgess, R. R., and L. Anthony. 2001. How sigma docks to RNA polymerase and what sigma does. Curr. Opin. Microbiol. 4:126–131.
- Bustard, K., and R. S. Gupta. 1997. The sequences of heat shock protein 40 (DnaJ) homologs provide evidence for a close evolutionary relationship between the *Deinococcus-Thermus* group and cyanobacteria. J. Mol. Evol. 45:193–205.
- Counsell, T., and R. G. E. Murray. 1986. Polar lipid profiles of the genus Deinococcus. Int. J. Syst. Bacteriol. 36:202–206.
- Courcelle, J., J. R. Donaldson, K.-H. Chow, and C. T. Courcelle. 2003. DNA damage-induced replication fork regression and processing in *Escherichia coli*. Science 299:1064–1067.
- da Costa, M. S., M. F. Nobre, and F. A. Rainey. 2001. Genus I. *Thermus* brock and freeze, 1969, 295^{AL}, emend. Nobre, Truper and da Costa 1996b, 605, p. 404–414. *In* D. R. Boone and R. W. Castenholz (ed.), Bergey's manual of systematic bacteriology, 2nd ed. Springer-Verlag, New York, N.Y.
- Ferreira, A. C., M. F. Nobre, F. A. Rainey, M. T. Silva, R. Wait, J. Burghardt, A. P. Chung, and M. S. da Costa. 1997. *Deinococcus geothermalis* sp. nov. and *Deinococcus murrayi* sp. nov., two extremely radiation-resistant and slightly thermophilic species from hot springs. Int. J. Syst. Bacteriol. 47:939–947.

- Goldman, B. S., and R. G. Kranz. 1998. Evolution and horizontal transfer of an entire biosynthetic pathway for cytochrome *c* biogenesis: *Helicobacter*, *Deinococcus*, *Archaea* and more. Mol. Microbiol. 27:871–873.
- Griffiths, E., and R. S. Gupta. 2002. Protein signatures distinctive of chlamydial species: horizontal transfers of cell wall biosynthesis genes *glmU* from archaea to chlamydiae and *murA* between chlamydiae and *Streptomyces*. Microbiology 148:2541–2549.
- Gupta, R. S. 1998. Protein phylogenies and signature sequences: a reappraisal of evolutionary relationships among archaebacteria, eubacteria, and eukarvotes. Microbiol. Mol. Biol. Rev. 62:1435–1491.
- Gupta, R. S. 2000. The phylogeny of proteobacteria: relationships to other eubacterial phyla and eukaryotes. FEMS Microbiol. Rev. 24:367–402.
- Gupta, R. S. 2003. Evolutionary relationships among photosynthetic bacteria. Photosynth. Res. 76:173–183.
- Gupta, R. S., K. Bustard, M. Falah, and D. Singh. 1997. Sequencing of heat shock protein 70 (DnaK) homologs from *Deinococcus proteolyticus* and *Thermomicrobium roseum* and their integration in a protein-based phylogeny of prokaryotes. J. Bacteriol. 179:345–357.
- Gupta, R. S., and E. Griffiths. 2002. Critical issues in bacterial phylogeny. Theor. Popul. Biol. 61:423–434.
- Gupta, R. S., and V. Johari. 1998. Signature sequences in diverse proteins provide evidence of a close evolutionary relationship between the *Deinococcus-Thermus* group and cyanobacteria. J. Mol. Evol. 46:716–720.
- Gupta, R. S., M. Pereira, C. Chandrasekera, and V. Johari. 2003. Molecular signatures in protein sequences that are characteristic of cyanobacteria and plastid homologues. Int. J. Syst. Evol. Microbiol. 53:1833–1842.
- Gupta, R. S., and B. Singh. 1994. Phylogenetic analysis of 70 kD heat shock protein sequences suggests a chimeric origin for the eukaryotic cell nucleus. Curr. Biol. 4:1104–1114.
- Ibba, M., A. W. Curnow, and D. Soll. 1997. Aminoacyl-tRNA synthesis: divergent routes to a common goal. Trends Biochem. Sci. 22:39–42.
- Karlin, S., and J. Mrazek. 2001. Predicted highly expressed and putative alien genes of *Deinococcus radiodurans* and implications for resistance to ionizing radiation damage. Proc. Natl. Acad. Sci. USA 98:5240–5245.
- Kimura, M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. 16:111–120.
- Klenk, H. P., and W. Zillig. 1994. DNA-dependent RNA polymerase subunit B as a tool for phylogenetic reconstructions: branching topology of the archaeal domain. J. Mol. Evol. 38:420–432.
- Ludwig, W., and H.-P. Klenk. 2001. Overview: a phylogenetic backbone and taxonomic framework for prokaryotic systematics, p. 49–65. *In* D. R. Boone and R. W. Castenholz (ed.), Bergey's manual of systematic bacteriology. Springer-Verlag, New York, N.Y.
- Makarova, K. S., L. Aravind, Y. I. Wolf, R. L. Tatusov, K. W. Minton, E. V. Koonin, and M. J. Daly. 2001. Genome of the extremely radiation-resistant bacterium *Deinococcus radiodurans* viewed from the perspective of comparative genomics. Microbiol. Mol. Biol. Rev. 65:44–79.
- Masters, C. I., R. G. Murray, B. E. Moseley, and K. W. Minton. 1991. DNA polymorphisms in new isolates of '*Deinococcus radiopugnans*'. J. Gen. Microbiol. 137:1459–1469.
- Minton, K. W. 1994. DNA repair in the extremely radioresistant bacterium Deinococcus radiodurans. Mol. Microbiol. 13:9–15.

- Muller, M., H. G. Koch, K. Beck, and U. Schafer. 2001. Protein traffic in bacteria: multiple routes from the ribosome to and across the membrane. Prog. Nucleic Acid Res. Mol. Biol. 66:107–157.
- Murray, R. G. E. 1986. Family II. *Deinococcaceae* Brooks and Murray 1981, 356^{VP}, p. 1035–1043. *In* P. H. A. Sneath, N. S. Mair, M. E. Sharpe, and J. G. Holt (ed.), Bergey's manual of systematic bacteriology. Lippincott Williams & Wilkins, Baltimore, Md.
- Murray, R. G. E. 1992. The family *Deinococcaceae*, p. 3732–3744. *In A.* Balows, H. G. Truper, M. Dworkin, W. Harder, and K. H. Schleifer (ed.), The prokaryotes. Springer-Verlag, New York, N.Y.
- Nikulin, A., I. Eliseikina, S. Tishchenko, N. Nevskaya, N. Davydova, O. Platonova, W. Piendl, M. Selmer, A. Liljas, D. Drygin, R. Zimmermann, M. Garber, and S. Nikonov. 2003. Structure of the L1 protuberance in the ribosome. Nat. Struct. Biol. 10:104–108.
- 33. Nobre, M. F., and M. S. da Costa. 2001. Genus II. *Meiothermus* Nobre, Truper and da Costa 1996b, 605^{VP}, p. 414–420. *In* D. R. Boone and R. W. Castenholz (ed.), Bergey's manual of systematic bacteriology, 2nd ed. Springer-Verlag, New York, N.Y.
- 34. Nobre, M. F., H. G. Truper, and M. S. da Costa. 1996. Transfer of *Thermus ruber* (Loginova et al. 1994), *Thermus silvanus* (Tenreiro et al. 1995) and *Thermus chliarophilus* (Tenreiro et al. 1995) to *Meiothermus* gen. nov. as *Meiothermus silvanus* comb. nov., and *Meiothermus chliarophilus* comb. nov., respectively, and emendation of the genus *Thermus*. Int. J. Syst. Bacteriol. 46:604–606.
- Nold, S. C., and D. M. Ward. 1995. Diverse Thermus species inhabit a single hot spring microbial mat. Syst. Appl. Microbiol. 18:274–278.
- Olendzenski, L., L. Liu, O. Zhaxybayeva, R. Murphey, D. G. Shin, and J. P. Gogarten. 2000. Horizontal transfer of archaeal genes into the *Deinococcaceae*: detection by molecular and computer-based approaches. J. Mol. Evol. 51:587–599.
- Olsen, G. J., and C. R. Woese. 1997. Archaeal genomics: an overview. Cell 89:991–994.
- Olsen, G. J., C. R. Woese, and R. Overbeek. 1994. The winds of (evolutionary) change: breathing new life into microbiology. J. Bacteriol. 176:1–6.
- Rainey, F. A., M. F. Nobre, P. Schumann, E. Stackebrandt, and M. S. da Costa. 1997. Phylogenetic diversity of the deinococci as determined by 16S ribosomal DNA sequence comparison. Int. J. Syst. Bacteriol. 47:510–514.
- Saitou, N., and M. Nei. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4:406–425.
- Thompson, B. G., and R. G. Murray. 1981. Isolation and characterization of the plasma membrane and the outer membrane of *Deinococcus radiodurans* strain Sark. Can. J. Microbiol. 27:729–734.
- Weisburg, W. G., S. J. Giovannoni, and C. R. Woese. 1989. The *Deinococcus-Thermus* phylum and the effect of rRNA composition on phylogenetic tree construction. Syst. Appl. Microbiol. 11:128–134.
- 43. White, O., J. A. Eisen, J. F. Heidelberg, E. K. Hickey, J. D. Peterson, R. J. Dodson, D. H. Haft, M. L. Gwinn, W. C. Nelson, D. L. Richardson, K. S. Moffat, H. Y. Qin, L. X. Jiang, W. Pamphile, M. Crosby, M. Shen, J. J. Vamathevan, P. Lam, L. McDonald, T. Utterback, C. Zalewski, K. S. Makarova, L. Aravind, and M. J. Daly. 1999. Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1. Science 286:1571–1577.
- 44. Woese, C. R. 1987. Bacterial evolution. Microbiol. Rev. 51:221-271.