

Conserved indels in essential proteins that are distinctive characteristics of Chlamydiales and provide novel means for their identification

Emma Griffiths,¹ Astrid K. Petrich² and Radhey S. Gupta¹

Correspondence
Radhey S. Gupta
gupta@mcmaster.ca

Department of Biochemistry and Biomedical Sciences¹ and Department of Pathology and Molecular Medicine², McMaster University, Hamilton L8N 3Z5, Canada

All known chlamydiae are either proven human or animal pathogens or possess such potential. Due to increasing reports of chlamydiae diversity in the environment, it is important to develop reliable means for identifying and characterizing Chlamydiales species. The identification of environmental chlamydiae at present relies on their branching pattern in 16S rRNA trees, as well as 16S/23S consensus motifs which display variability. At present, no reliable molecular signatures are known which are unique to all Chlamydiales species. Besides the rRNAs, sequence information for different Chlamydiales is not available for any other gene sequence. In this report, a number of molecular signatures are described that consist of conserved inserts and deletions (indels), in widely distributed proteins [RNA polymerase α subunit (RpoA), elongation factor (EF)-Tu, EF-P, DNA gyrase B subunit (GyrB) and lysyl-tRNA synthetase (LysRS)], that are distinctive characteristics of all available chlamydiae homologues (from *Chlamydiaceae* species and *Parachlamydiae* sp. UWE25) and not found in any other bacteria. Using PCR primers for highly conserved regions in these proteins, the corresponding fragments of these genes from *Simkania negevensis*, *Waddlia chondrophila*, and in a number of cases for *Neochlamydia hartmanellae*, covering all families within the phylum Chlamydiae, have been cloned and sequenced. The shared presence of the identified signatures in these species provides strong evidence that these molecular signatures are distinctive characteristics of the entire order Chlamydiales and can be used to reliably determine the presence of chlamydiae or chlamydia-related organisms in environmental samples. The sequence information for these protein fragments was also used to determine the interrelationships among chlamydiae species. In a phylogenetic tree based on a combined dataset of sequences from RpoA, EF-Tu, EF-P and GyrB, the environmental chlamydiae (i.e. *Simkania*, *Waddlia* and *Parachlamydia*) and the traditional *Chlamydiaceae* (i.e. *Chlamydomphila* and *Chlamydia*) formed two distinct clades. Similar relationships were also observed in individual protein phylogenies, as well as in a 16S rRNA tree for the same species. These results provide evidence that the divergence between the traditional *Chlamydiaceae* species and the other chlamydiae families occurred very early in the evolution of this group of bacteria.

Received 21 March 2005

Revised 17 May 2005

Accepted 27 May 2005

INTRODUCTION

The chlamydiae group of species, most of which are important human and animal pathogens, form a distinct phylum within the Bacteria (Fields & Barnes, 1992; Kalayoglu & Byrne, 2001). The phylum Chlamydiae contains only a single class (Chlamydiae) and order (Chlamydiales). The order Chlamydiales is currently composed of four families, the *Chlamydiaceae*, which contains all of the traditional chlamydiae species, and three more recently identified

families of chlamydiae-related organisms: the *Parachlamydiaceae* (composed of *Parachlamydia* and *Neochlamydia*), *Waddliaceae* and *Simkaniaceae* (Amann *et al.*, 1997; Birtles *et al.*, 1997; Kahane *et al.*, 1999; Rurangirwa *et al.*, 1999; Horn *et al.*, 2000, 2004; Corsaro *et al.*, 2003). Chlamydiae species, until a few years ago, were generally believed to be intracellular parasites of only a limited host range. However, recent studies provide evidence that they are also present in a wide variety of hosts, including fish, insects, turtles and amoebae (Ossewaarde & Meijer, 1999; Everett, 2000; Horn & Wagner, 2001, 2004; Corsaro *et al.*, 2003) (see also www.chlamydiae.com). In recent ecological sampling studies, novel chlamydiae-related sequences have also been identified from a wide variety of environments, including

The GenBank/EMBL/DDBJ accession numbers for the clones obtained in this work are AY845401, AY845403, AY845404, AY845405, AY845406, AY845408, AY845410, AY882436, AY845412 and AY845413.

freshwater, soil, sewage and water conduit systems (Horn & Wagner, 2001; Corsaro *et al.*, 2002, 2003; Kahane *et al.*, 2004; Collingro *et al.*, 2005). As a result, it is now recognized that the diversity of chlamydiae, in terms of both their numbers and their clinical involvement, is at present significantly underestimated (Ossewaarde & Meijer, 1999; Fritsche *et al.*, 2000; Corsaro *et al.*, 2003; Thao *et al.*, 2003). The detection of novel chlamydiae-related species (which are generally non-culturable) in environmental samples is at present based entirely on the PCR amplification of 16S rRNA sequences and the observed clustering of a sequence in the rRNA trees with those of the known chlamydiae species (Rurangirwa *et al.*, 1999; Everett *et al.*, 1999b; Horn *et al.*, 2000, 2004; Corsaro *et al.*, 2002; Thao *et al.*, 2003). Although consensus sequences for chlamydiae species are found in the 16S and 23S rRNA, these regions display sequence variability. Because of this, and the fact that these motifs do not clearly distinguish chlamydiae species from all other bacteria, assessing the membership of more-divergent species based on these criteria may prove difficult. Since the branching patterns of species in phylogenetic trees are also known to be affected by a large number of variables (Felsenstein, 1988; Gupta, 1998; Moreira & Philippe, 2000; Baldauf, 2003), it is important to develop other reliable markers for identification of chlamydiae species.

The task of identifying novel molecular characteristics that can provide reliable markers for chlamydiae is facilitated by the availability of genomic sequences. The complete genome sequences of five chlamydiae species, *Chlamydomphila* (*Chlam.*) *pneumoniae* (four different strains: AR39, CWL029, J139, TW-183) (Read *et al.*, 2000), *Chlamydomphila caviae* GPIC (Read *et al.*, 2003), *Chlamydia* (*Chl.*) *trachomatis* (serovar D) (Stephens *et al.*, 1998), *Chlamydia muridarum* (Read *et al.*, 2000) and an environmental chlamydia, *Parachlamydia* sp. UWE25 (Horn *et al.*, 2004), are now available. The comparative analyses of these genomes provide a powerful means for identifying novel molecular characteristics that are distinctive to different groups of bacteria (Hatch, 1998; Kalman *et al.*, 1999). In an earlier study, we reported a number of conserved inserts and deletions (indels) in widely distributed proteins that were uniquely shared by *Chlamydiaceae* species (Griffiths & Gupta, 2002). However, most of these indels were small (1 aa), and for most of them sequence information from other chlamydiae families was not available. In the present work, which extends this earlier study, we have identified a number of prominent chlamydiae-specific conserved indels in essential proteins which are found in all bacteria, and in some cases in species from all three domains (Olsen & Woese, 1997). The proteins which contain chlamydiae-specific signatures include RNA polymerase α subunit (RpoA), elongation factor Tu (EF-Tu), DNA gyrase β subunit (GyrB), elongation factor P (EF-P) and lysyl-tRNA synthetase (LyrRS). Sequence information for most of these proteins (RpoA, EF-Tu, EF-P and GyrB) was obtained from species (*Simkania*, *Waddlia* and *Neochlamydia*) covering all of the known chlamydiae families. The presence of these

signatures in all of these species provides evidence that they are distinctive characteristics of the phylum Chlamydiales. For the *Simkania*, *Waddlia* and *Neochlamydia* species, apart from the 16S and 23S rRNA, very little sequence information is available in the databases. Hence, we have carried out a phylogenetic analysis based on a concatenated dataset of sequences from these genes (RpoA, GyrB, EF-P and EF-Tu), to determine the interrelationships among different chlamydiae species.

METHODS

Culturing of chlamydiae and DNA extraction. *Simkania negevensis* (ATCC VR1471), *Waddlia chondrophila* (ATCC 1470) and *Neochlamydia hartmannellae* (ATCC 50802) were obtained from ATCC. *Sim. negevensis* and *Wad. chondrophila* were propagated in monolayers of HEp2 cells in a similar manner to that described for *Chlam. pneumoniae* in earlier work (Roblin *et al.*, 1992; Mahony *et al.*, 2000). After incubation of infected cell cultures for 72 h at 37 °C in a 95% air/5% CO₂ atmosphere, the bacteria were harvested by disrupting the HEp2 cells with sterile glass beads, followed by sonication and centrifugation at 250 g for 5 min (Mahony *et al.*, 2000). DNA from these samples was extracted using the blood protocol of the Qiagen DNA Minikit according to the manufacturer's specifications. DNA from *Neo. hartmannellae* was directly extracted from the cells obtained from ATCC. The identity of different Chlamydiales was confirmed using 23S rDNA PCR amplification followed by DNA sequence analysis (Everett *et al.*, 1999b). The complete genomes of *Chl. trachomatis*, *Chl. muridarum*, *Chlam. pneumoniae*, *Chlam. caviae* and *Parachlamydia* sp. UWE25 are available from the NCBI database (www.ncbi.nlm.nih.gov).

Identification of chlamydiae-specific signatures. Multiple sequence alignments for a large number of proteins have been created in our earlier work (Gupta, 1998, 2000, 2004; Griffiths & Gupta, 2001, 2002, 2004; Gupta *et al.*, 2003) (see also www.bacterialphylogeny.com). To search for chlamydiae-specific signatures, these alignments were visually inspected to identify any indel that was uniquely present in all available chlamydiae homologues and flanked by conserved sequences. The indels which were not flanked by conserved regions, and/or which were not present in all chlamydiae, were omitted from further consideration in this study. The chlamydiae specificity of potentially useful indels was further evaluated by carrying out additional BLAST searches (Altschul *et al.*, 1997) on short sequence segments (usually 60–100 aa) containing the indel and the flanking conserved regions. The purpose of these BLAST searches was to obtain sequence information from all available species to ensure that the identified signatures were only present in the chlamydiae homologues. The sequence information for various useful signatures, which were chosen for further investigation, was compiled into signature files, such as those shown in Figs 1–5.

PCR amplification and sequencing. Oligonucleotide primers, in opposite orientations, were designed for conserved regions in the sequence alignment that flanked the signatures, based on the sequences of the genes/proteins from available chlamydiae species. Degeneracy was incorporated into the primers to account for nucleotide variability at different sites in the alignment. The primers were synthesized at MOBIX, McMaster University. The sequences of various PCR primers that were used to amplify different gene fragments are given in Table 1. PCR was performed in a Techne Techgene thermocycler under the following conditions. Each reaction had a final volume of 10 μ l, and a Mg²⁺ concentration of 4 mM was used during amplification with all primer sets for each DNA strain tested. Also, 2% DMSO (v/v, final concentration) was

Table 1. Sequences of PCR primers

The unusual bases in these primers are as follows: N=A, T, C or G; Y=T or C; R=A or G; V=G, C or A; W=A or T; S=G or C; H=A, C or T; D=G, A or T; K=G or T.

Gene	Primer sequence	Fragment size
EF-Tu, <i>Waddlia</i>	Forward: 5' GCNCARATGGAYGGNGCNAT 3' Reverse: 5' GGCATNARRAANGGYTTRTC 3'	375 bp
EF-Tu, <i>Neochlamydia</i> and <i>Simkania</i>	Forward: 5' GCNGAYTAYGTNAARAAYATG 3' Reverse: 5' TCVCCNGGCATNACCATWTC 3'	810 bp
Gyrase B (all chlamydiae)	Forward: 5' CCNGGNATGTAYATHGGNGA 3' Reverse: 5' TTNGTYTGCCYTCRAAYCG 3'	900 bp
EF-P, <i>Waddlia</i>	Forward: 5' GTNAARCCNGGNAARGG 3' Reverse: 5' ACDCGRGASTCRTARCTYC 3'	456 bp
RpoA, <i>Waddlia</i>	Forward: 5' GGNATGGGNCAYACNYTNGG 3' Reverse: 5' TGCCNACNCKNGTRTCYTC 3'	522 bp
RpoA, <i>Simkania</i>	Forward: 5' TTAGAAGCTCCWGCWATYATYTC 3' Reverse: 5' CGATCRAAAHYGTATCYTGACCHAC 3'	477 bp
LysRS, <i>Waddlia</i>	Forward: 5' CAYAAAYCCNGARTTYAC 3' Reverse: 5' AGNCGRTCDATNCCDATNCC 3'	669 bp

added to each reaction mixture in order to facilitate primer binding and enhance PCR specificity. The PCR amplification was carried out over 30 cycles (15 s at 94 °C, 15 s at 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 45 °C, 7 min at 72 °C). DNA fragments of the expected size were purified from 0.8% (w/v) agarose gels (using a GeneClean kit, Qiagen), and subcloned into the plasmid pDRIVE using a TU cloning kit (Invitrogen). *Escherichia coli* JM109 cells were transformed with the ligated plasmid, and the inserts from a number of positive clones were sequenced through the MOBIX facility. Sequences of all cloned fragments were BLAST searched to ensure that they were from a novel source. The sequence data for various clones obtained in this work have been deposited in GenBank under the following accession numbers: AY845401, AY845403, AY845404, AY845405, AY845406, AY845408, AY845410, AY882436, AY845412 and AY845413.

Phylogenetic analysis. Multiple sequence alignments of different proteins (RpoA, EF-Tu, GyrB and EF-P) for which sequence information was available from various chlamydiae species were created using the ALIGN PLUS4 program (Scientific and Educational Software). All of these alignments were trimmed to the same length as the amplified fragments. Because the lengths of individual gene fragments were small, to increase the reliability of the phylogenetic analyses, these sequences were concatenated to create a larger alignment consisting of characters from the same groups of species. Any gapped regions, or other regions where low sequence conservation was observed, were removed from this combined alignment. Phylogenetic analysis on the combined dataset was carried out in both the presence and absence of the signature regions to determine their influence on the branching pattern. Genetic distances between bootstrapped datasets (100 times) of aligned sequences were calculated after correcting for multiple amino acid substitutions. Neighbour-joining trees from these distances were constructed and a consensus tree was obtained (Felsenstein, 1988). All of the phylogenetic programs used to construct the tree are part of the TREECON for Windows software package (Van de & De Wachter, 1994). The tree was rooted using the sequence data for *Thermus aquaticus*. A

phylogenetic tree for the same species was also constructed for the full-length 16S rRNA sequences (downloaded from the RDP-II database) (Maidak *et al.*, 2001) for comparison purposes.

RESULTS

Identification and characterization of Chlamydiales-specific indels

Our examination of sequence alignments for different proteins identified a number of prominent conserved inserts and deletions, in widely distributed proteins, that were limited to the available chlamydial homologues. These new chlamydiae-specific signatures included a 15–17 aa insert in the RNA polymerase α subunit (RpoA), a 2 aa insertion and a 2 aa deletion in elongation factor Tu (EF-Tu), a 6 aa insert in DNA gyrase B (GyrB), a 1 aa insert in elongation factor P (EF-P) and a 6 aa insert in lysyl-tRNA synthetase (LysRS). The sequence information for these proteins was only available from the four *Chlamydiaceae* species and *Parachlamydia* sp. UWE25, whose genomes have been sequenced (Stephens *et al.*, 1998; Read *et al.*, 2000, 2003; Horn *et al.*, 2004). To determine whether these signatures are distinctive characteristics of chlamydiae, sequence information for these genes/proteins from other chlamydiae families (*Simkania*, *Waddlia* and *Neochlamydia*) was required. This was obtained by means of PCR amplification and sequencing of the corresponding fragments. A brief description of these signatures, the proteins in which they are found, and the results of our studies are presented below.

The core subunits of RNA polymerase, α , β and β' , are highly conserved in species from all three domains (Olsen & Woese, 1997). In the α subunit (RpoA), which is encoded

		110		179
Proteobacteria	<i>E. coli</i>	BAB38326	DGPMPQTRFHILLGRQVGVPIYIIVFLNKDMV	DDEELLELVEMEVRELLSQYDFPG DD TPIVRRSALKAL
	<i>Pos. multocida</i>	AAK03441	-----	-----Q--
	<i>H. influenzae</i>	AAC22292	-----	-----Q--
	<i>Nei. meningitidis</i>	AAF40598	-----A-----M-----	--A-----D--S-----C--Q--
	<i>Ral. solanacearum</i>	CAJ16750	-----A-----	--A-----K-----
	<i>A. lumefaciens</i>	CAA67992	-----A-----A-V-----V-Q	--A-----L-----S-----I--K--A--
	<i>C. crescentus</i>	AAK25161	-----A-----ALV-M--V--	-----S-Q-----I--TK--A-V
	<i>Bru. melitensis</i>	AAL51923	-----A-----A-V-----Q	--A-----L-----K-E--E I--K--A--
	<i>Hel. pylori</i>	AAJ06711	-----S-----H-V-----Q--	--Q-----A-E-----A--R--
	<i>Camp. jejuni</i>	CAB75108	-----S-----V-M-A-----	--A-----S-----I--S-----
Chlamydiales	<i>Chl. trachomatis</i>	AAA62238	--A--X--A--V--I--I--	SE E-A--VD--LA--EEKGYK--C--I--
	<i>Chl. muridarum</i>	AAF39428	--A--X--A--V--I--I--	SE E-A--VD--LS--EEKGYK--C--I--
	<i>Chlam. pneumoniae</i>	3AA98284	--A--X--A--V--I--I--	SQ E-A--ID--LS--EEKGYK--C--I--
	<i>Chlam. caviae</i>	NP_829562	--A--X--A--V--I--I--	SQ E-A--VD--LS--EEKGYK--C--I--
	<i>Wad. chondrophila</i>	AY845401	--A-----AH-MQ--K-V-----L	GE G-Q--D--L-LQ--EAKGYEN A--
	<i>Para. sp. UWE25</i>	YP_007594	--A-----A-MQ--A-V-----V-L	GE S-S--D--LH--ESKGYIK S--R--
	<i>Sim. negevensis</i>	AY845408	--A--X--A--V--I--I--	SQ E-A--VD--LS--EEKGYK--C--I--
	<i>Neo. hartmanellae</i>	AY845412	--A-----A-MQ--A-V-----V-L	AE S-Q--D--LH--ESKGYKD V--R--
	<i>Aqu. aeolicus</i>	AAC06403	-----V-A--N--V-M-----	-----L-----K-EY--E V-V--G--
	<i>Fib. succinogenes</i>	CAA54194	-----AH--K-V-M-----	--A-I-D-----K--D--N-----
Other Gram negative Bacteria	<i>Bact. forsythus</i>	3AA88139	-----A--N--RLV-M-----	-----M--DM--F--D--N-----G--
	<i>Por. gingivalis</i>	3AA88134	-----A--N--RLV-M-----	-----M--DM--F--D--N-----G--
	<i>Co. vibrioforme</i>	CAA54322	-----A--N--ALV-----V	A-P-----L-----TE-G-----I--K--N--
	<i>Pir. marina</i>	AAG35793	-----X-V-A-----VC-----L	-----D--L-----K-E--C-V--AS-P-
	<i>Tre. pallidum</i>	AAC65172	--V--X--L--A--S--V--L--	--P-----E--LA-AG-G-3R ET --K--F--
	<i>Sy. sp.</i>	3AA18321	-----AK--KLV-----X	-----L-----D-----I--A--I
	<i>Nostoc sp. PCC 7120</i>	BAB76036	-----AK--KLV-----E-M	E-A-----L-L-----TE-E-D--I--G-Q--
	<i>Cf. aurantiacus</i>	CAA54193	-----A--Q--A-V-----V-M	--P-----L-L-----K-G--E I--T-RN--
	<i>D. radlodus</i>	AAF11600	-----A--V-M-----V	-----K-E--L-V-K--R--
	<i>The. aquaticus</i>	CAA46998	-----A--V-M-----V	--P--D-----D--N--E--E V-V--L--
Gram positive Bacteria	<i>T. maritima</i>	AAA27415	-----V-A--F--M--I--T--	--P--TD-----D--GY--E V-V--V--
	<i>Str. coelicolor</i>	1091582	-----X-V-A-----V-A--A--	-----I--L-----E-E--V-V-KV--
	<i>Brev. linens</i>	CAA54192	-----V-A-----V-A--S--	-----F--D--SQ--D--N A-V-PV--
	<i>Cor. glutamicum</i>	CAA54323	-----V-A-----L-A-----	E--II--AEQ-YDE EA --HI--
	<i>M. pneumoniae</i>	AB95825	--SV-----A--RMV-----IA	T--VQ--AE--D--TS-G-D- KN --Y--
	<i>Cic. perfringens</i>	BAB82127	-----SSR--DH-V--A--	-----E-N-----I-V-K--V--
	<i>Lis. innocua</i>	CAC98028	-----S-----V-M-----	-----I--D--TE-E--I-V-K--
	<i>Bac. subtilis</i>	CAB11889	-----SKN-----V-----	-----D--E-----V-V-K--
	<i>Strep. pyogenes</i>	AAK33586	-----S--KHL--M--V-L--	-----I--D--E-----L-V-Q--
	<i>Sta. aureus</i>	BAB56710	-----S-N--ATV-----V--	-----E--E-----V-V-TA--

Fig. 2. Partial sequence alignment of EF-Tu sequences, showing two different signatures (an insert and a deletion, boxed) that are distinctive of the chlamydiae species. Sequence information for *Wad. chondrophila*, *Sim. negevensis* and *Neo. hartmanellae* was obtained in this work. Additional abbreviations: *Brev.*, *Brevibacterium*; *Camp.*, *Campylobacter*; *Fib.*, *Fibrobacter*; *Hel.*, *Helicobacter*; *M.*, *Mycoplasma*; *Neo.*, *Neochlamydia*; *Por.*, *Porphyromonas*; *The.*, *Thermus*; *Tre.*, *Treponema*.

separated from the first by a conserved region of 24 aa (Fig. 2). These inserts are again specific for chlamydiae and not found in any other bacteria. Upon sequencing of a 0.4 kb fragment from *Wad. chondrophila*, and 0.8 kb fragments from *Sim. negevensis* and *Neo. hartmanellae*, both these signatures were found to be present in these species (see Fig. 2). These signatures are thus again indicated to be distinctive characteristics of various chlamydiae, which were likely introduced in a common ancestor of this lineage.

DNA gyrase is a type II topoisomerase, which makes transient double strand breaks in DNA to allow another DNA strand to pass through (Levine *et al.*, 1998). DNA gyrase function is essential in bacteria for maintaining the appropriate levels of supercoiling in the chromosome and for transcription and replication processes. The enzyme consists of two subunits, A and B, that combine to form a functional A₂B₂ complex. The B subunit of DNA gyrase (GyrB) contains a 6 aa insert that is uniquely present in various available chlamydia homologues (Fig. 3). The specificity and distinctiveness of this insert were again tested by PCR amplification of 0.9 kb fragments of *gyrB* from *Wad.*

chondrophila, *Sim. negevensis* and *Neo. hartmanellae*. The cloning and sequencing of these fragments revealed that the indicated signature insert was present in all three species (see Fig. 3), confirming that it is a reliable molecular marker for all chlamydiae.

We have previously described a 1 aa insert in EF-P (Griffiths & Gupta, 2002) that is found in all *Chlamydiaceae* species, including *Chlamydomonas abortus* and *Chlamydia suis*, which were sequenced in our previous work, as well as *Sim. negevensis*. This insert is also present in *Parachlamydiae* sp. UWE25, and its sequence is completely conserved in all species. Further work that we have carried out indicates that this insert is also present in *Wad. chondrophila* (Fig. 4), indicating that in a similar manner to the inserts in RpoA, EF-Tu and GyrB, this insert is also a distinctive characteristic of various chlamydiae species.

Another signature that is specific for various chlamydiae was identified in the enzyme lysyl-tRNA synthetase, which catalyses the correct attachment of lysine to its cognate tRNA in the protein synthesis process (Woese *et al.*, 2000). The chlamydial LysRS proteins (from the known

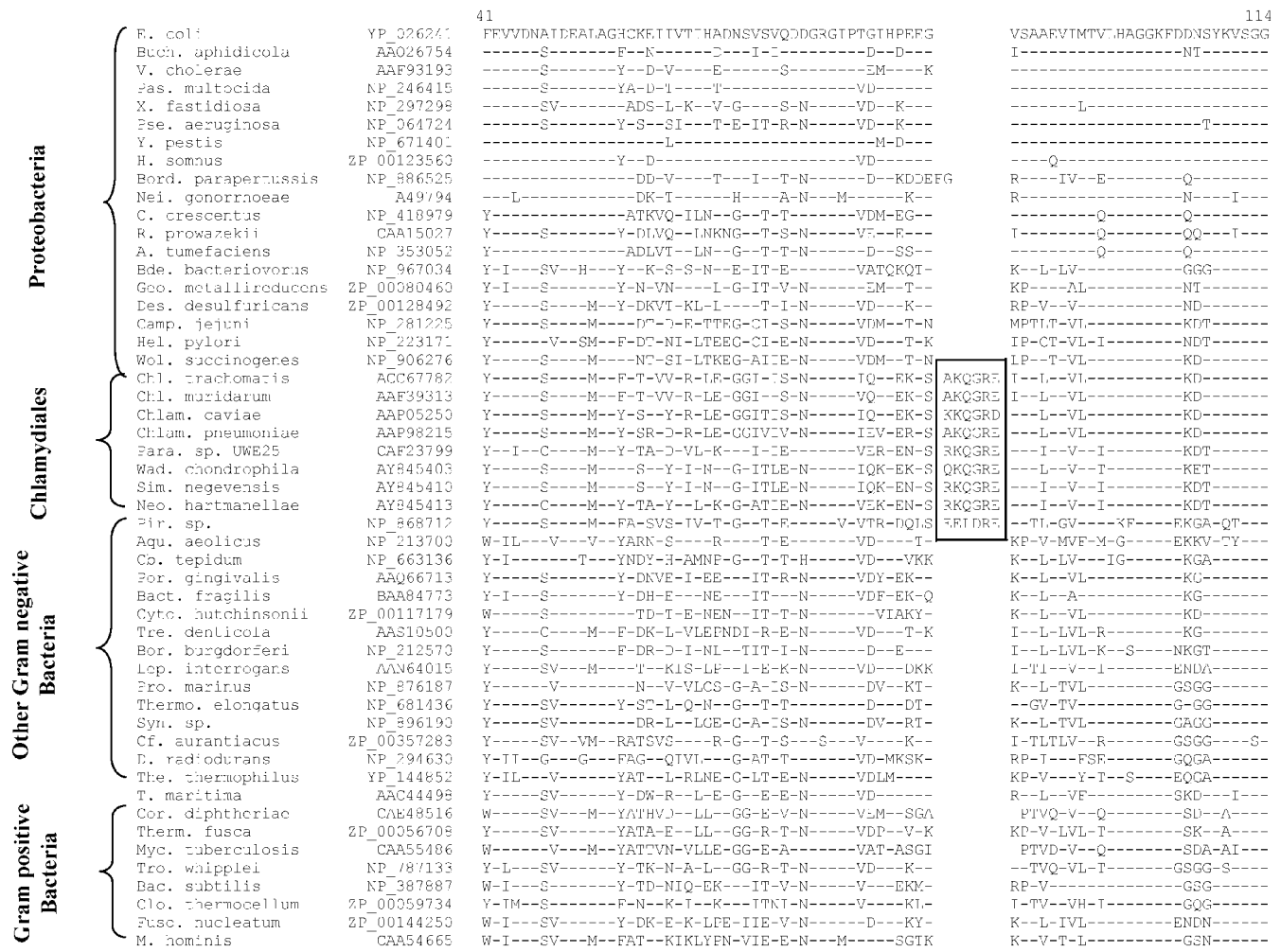


Fig. 3. Excerpt from a partial alignment of gyrase B (GyrB) homologues showing a 6 aa insertion that is distinctive of the Chlamydiales. The sequences for *Wad. chondrophila*, *Sim. negevensis* and *Neo. hartmanellae* were obtained in this work. Additional abbreviations: *Bde.*, *Bdellovibrio*; *Bor.*, *Borellia*; *Buch.*, *Buchnera*; *Fuso.*, *Fusobacterium*; *Geo.*, *Geobacter*; *Pro.*, *Prochlorococcus*; *R.*, *Rickettsia*; *Thermo.*, *Thermosynechococcus*; *Wol.*, *Wolinella*; *V.*, *Vibrio*; *Y.*, *Yersinia*.

Chlamydiaceae and *Parachlamydia*) were found to contain a 6 aa insert which is unique to this group of species (Fig. 5). Using the PCR primers described in Table 1, we were also successful in amplifying a 0.6 kb *LysRS* fragment from *Wad. chondrophila* that was found to contain the indel. However, our attempts to PCR-amplify the corresponding fragment from *Sim. negevensis* and *Neo. hartmanellae* using this set of primers were unsuccessful, indicating that the primer region may have undergone mutational changes in these species. Since species representing other chlamydiae families contain this insert, it is likely that this indel will also be found in *Simkania* and *Neochlamydia*.

Phylogenetic analysis of chlamydiae species based on concatenated protein sequences

Our understanding of the evolutionary relationships among chlamydiae is based solely on 16S and 23S rRNA sequences

(Everett *et al.*, 1999a; Bush & Everett, 2001; Corsaro *et al.*, 2003; Horn *et al.*, 2004). The sequence information for different Chlamydiales species is currently not available for any protein-coding genes. Since in our work we have sequenced fragments for a number of protein-coding genes from different chlamydiae families, it was of interest to carry out a phylogenetic analysis based on these protein sequences and compare it with the 16S rRNA tree. As the amplified fragments generated in this work were small, to obtain a reliable phylogenetic tree we have combined (i.e. concatenated) sequence information for GyrB, RpoA, EF-Tu and EF-P proteins from *Wad. chondrophila*, *Sim. negevensis*, *Parachlamydia* sp. UWE25, *Chl. trachomatis*, *Chl. muridarum*, *Chlam. pneumoniae*, *Chlam. caviae* and *Thermus aquaticus*. This concatenated dataset containing 529 aa positions was used to construct a neighbour-joined tree which was bootstrapped (100 times) for statistical analysis (Felsenstein, 1988). *Thermus aquaticus* was used as an

		144	186	
Proteobacteria	<i>E. coli</i>	729406	GDTAGTGG	
	<i>Bru. melitensis</i>	AAL51508	-Q--ASSY	
	<i>Pas. multocida</i>	AAK02184	---	
	<i>Y. pestis</i>	CAC8921	-----V--I-----	
	<i>H. influenzae</i> Rd	1169480	-----R-----	
	<i>V. cholerae</i>	AAF95801	---Q---R-----A---G---	
	<i>X. fastidiosa</i>	AAF85271	---S-G---E---R---NQD-----	
	<i>Pse. aeruginosa</i>	AAG06239	---S-KVM--T-R-NN--ELQ-SA-CE--DS-EI---T---K--	
	<i>Z. mobilis</i>	5932375	-Q--SSSY--I-DN-VR-M--PHITA-TR-I--VNTQ--K-	
	<i>Nei. meningitidis</i>	CAB84395	---S-KVM--T-R-VG-TETQ-MSYIEN-DKVEI---T--FRK-	
	<i>Ca. crescentus</i>	AAK22706	-S--VE--YTLN-----NE-D-LIIN-GD-S-I--	
	<i>A. tumefaciens</i>	AAG09299	-Q--SSSY--I--N-IRTM--PHIDA-IRVVIA-EDNS--E-	
	<i>Sin. meliloti</i>	CAC41714	-Q--SSSY--V--N-VRTL--PHI-A-TRVVIA-ED-S--E-	
	<i>R. prowazekii</i>	Q92DT7	-A-VTSSY--I-EN-IK---QYLA--K-V-K-DDMT--E-	
	<i>Camp. jejuni</i>	CAB75187	---Q-SNK-----E-----QI-FH-LE---R---VR---IE-	
	<i>Hel. pylori</i> 26695	2494272	---SSASK-----E---Q--FH-LE--I---N-ETE-LEKE	
	<i>Geo. sulfurreducens</i>	NP_952802	---S-SDS--VE--YTLR--P-IEE--M-VI-----ST-	
	Chlamydiales	<i>Desulf. vulgaris</i>	YP_010884	---VSNVT--K-E--I--Q--I--NQ-DR-----R--LG-
<i>Chlam. pneumoniae</i>		4377215	---SGRV L---VTN--KIM--I-IDE--LV-----T-S-E--	
<i>Chlam. caviae</i>		NP_829736	---SGRV L---VTN--KIM--I-IEE--V-I---T-S-E--	
<i>Chlam. abortus</i>		AY038588	---SGRV L---VTN--KIM--I-IEE--V-----T-S-E--	
<i>Chl. muridarum</i>		AAF39011	---SGRV L---TN--KIM--I-IEE--V-----T-S-E--	
<i>Chl. trachomatis</i>		3329213	---SGRV L---TN--KIM--I-IEE--V-----T-S-E--	
<i>Chl. suis</i>		AY038589	---SGRV L---TN--KIM--I-IEE--V-----T-S-E--	
<i>Para. sp. UWE25</i>		YP_008528	---SGRV L---I-ES--K-Q--I--EQ--K-----T-----	
<i>Wad. chondrophila</i>		AY845404	---SGKV L---ETES--K-QI-I-IEE--V-I---T	
<i>Sim. negevensis</i>		AF468693	---SGRV L---ITE--KIM--I	
<i>Bact. fragilis</i>		AAC26328	---TNTL---VES--T-R---ISE--T-EI---D-S--G-	
Other Gram negative Bacteria		<i>Por. gingivalis</i>	NP_905456	---TNTL---VE--E-R---I-E--L-E---D-S-IG-
	<i>Aqu. aeolicus</i>	3913565	---QAG-T---K-E---IQ--F--KE-DIV-----T-S--E-	
	<i>Tre. pallidum</i>	3322817	---VSGAT--V-TE--L--R---IKQ--K-LIN-ETN--QE-	
	<i>Bor. burgdorferi</i>	3913564	---VTNAM--NI--N--L--A---INV-DKVLINSETK--AE-	
	<i>Sy. sp. PCC6803</i>	2494276	---TG-T---I-E---Q-M---IAQ--R-----D-S-LG-	
	<i>Syn. PCC7942</i>	2494275	---TG-T---KVE--Q-M---ISV--R--I---NDS-LG-	
	<i>Anab. sp.</i>	2494269	---TG-T---IVE---T-M---ISQ--R--I---DDK-LG-	
	<i>Nostoc sp. PCC 7120</i>	BAB76757	---TG-T---IVE---T-M---ISQ--R--I---DDK-LG-	
	<i>D. radiodurans</i>	AAF09709	---VSG-T---E-----Q-----EQ-TDV-----T-Q-L--	
	<i>T. maritima</i>	Q9X284	---VSG--V-E--LKIT--Y-IEV-DK-----T---G-	
	Gram positive Bacteria	<i>Myc. leprae</i>	CAC30030	---RSSA-T---E--QIN---INT-DKL---S-D-S-LG-
		<i>Cor. glutamicum</i>	2494271	---RSTG-T---E--EIQ---IET-N-L-----D-S-L-
<i>Str. coelicolor</i>		CAB93371	---RSTG-T---E--HQIN---ITT--K-----TSD-LG-	
<i>Bac. subtilis</i>		1000357	---SG-T---KTE--L--N--F--NE-DTLV-N-SD-S--	
<i>M. pneumoniae</i>		AAB95773	-N--NMAR--R-R-VS-YELD--Q-IRT-DK-VIS-ID-S-RE-	
<i>L. lactis</i>		AAK04790	-A-VTGS----ME--L--N--D--EA-TKLEIN-QT---LK-	
<i>Lis. innocua</i>		CAC96623	---SSG-S---E--L--Q--F--NE-DKLVIN-TEAA---	
<i>Strep. pneumoniae</i>		AAK99196	-A-VTGS----ME--L--N--D-IEA-QKLVIN-AE-T----	
<i>Clo. acetobutylicum</i>		NC_348713	-N-TSNVM---E-N--Q-----NE--T-R-----ME-	
<i>Sta. aureus</i>	BAB45450	---TGAT--S--VE--YTLN-----NE-D-LIIN-GD-S-I--		

Fig. 4. Partial sequence alignment for EF-P homologues showing a completely conserved 1 aa (leucine) insertion that is present in all Chlamydiales species. The sequence for *Wad. chondrophila* was obtained in the present work. Additional abbreviations: *Anab.*, *Anabaena*; *L.*, *Lactococcus*; *Sin.*, *Sinorhizobium*; *Z.*, *Zymobobilis*.

outgroup to root the tree (Fig. 6A). The resulting topology showed that the Chlamydiales appear to have evolved in two distinct groups, the *Chlamydiaceae* and the chlamydiae-like organisms. Within the *Chlamydiaceae*, the genus *Chlamydia* appears to have diverged later than *Chlamydomphila*. Among the chlamydiae-like species, *Simkania* branches the earliest, followed by a bifurcating clade including *Waddlia* and *Parachlamydia*. All branches were resolved with significant bootstrap scores, and the observed branching pattern was not affected by the presence or absence of the signature inserts (not shown). These findings were also corroborated by individual protein phylogenies based on these protein sequences (data not shown). It should be mentioned that the amino acid sequences of the insert in the EF-Tu protein (Fig. 2) also distinguish the clades consisting of *Chlamydia*

(insert sequence SE), *Chlamydomphila* (insert sequence SQ) and *Waddlia/Parachlamydia* (insert sequence GE). We also constructed a neighbour-joined phylogenetic tree based on full-length 16S rRNA sequences (Maidak *et al.*, 2001) (Fig. 6B). This tree showed a similar topology to that seen in the concatenated protein tree, and branches separating the chlamydiae-like species from the *Chlamydiaceae* were resolved with high statistical support (bootstraps > 90 %). 16S rRNA trees constructed by other investigators have resulted in topologies very similar to that obtained with our concatenated protein dataset (Everett *et al.*, 1999a; Bush & Everett, 2001; Corsaro *et al.*, 2003; Horn *et al.*, 2004). Thus, it is likely that the chlamydiae-like species (*Waddlia*, *Parachlamydia* and *Simkania*) form a distinct clade within the phylum Chlamydiae.

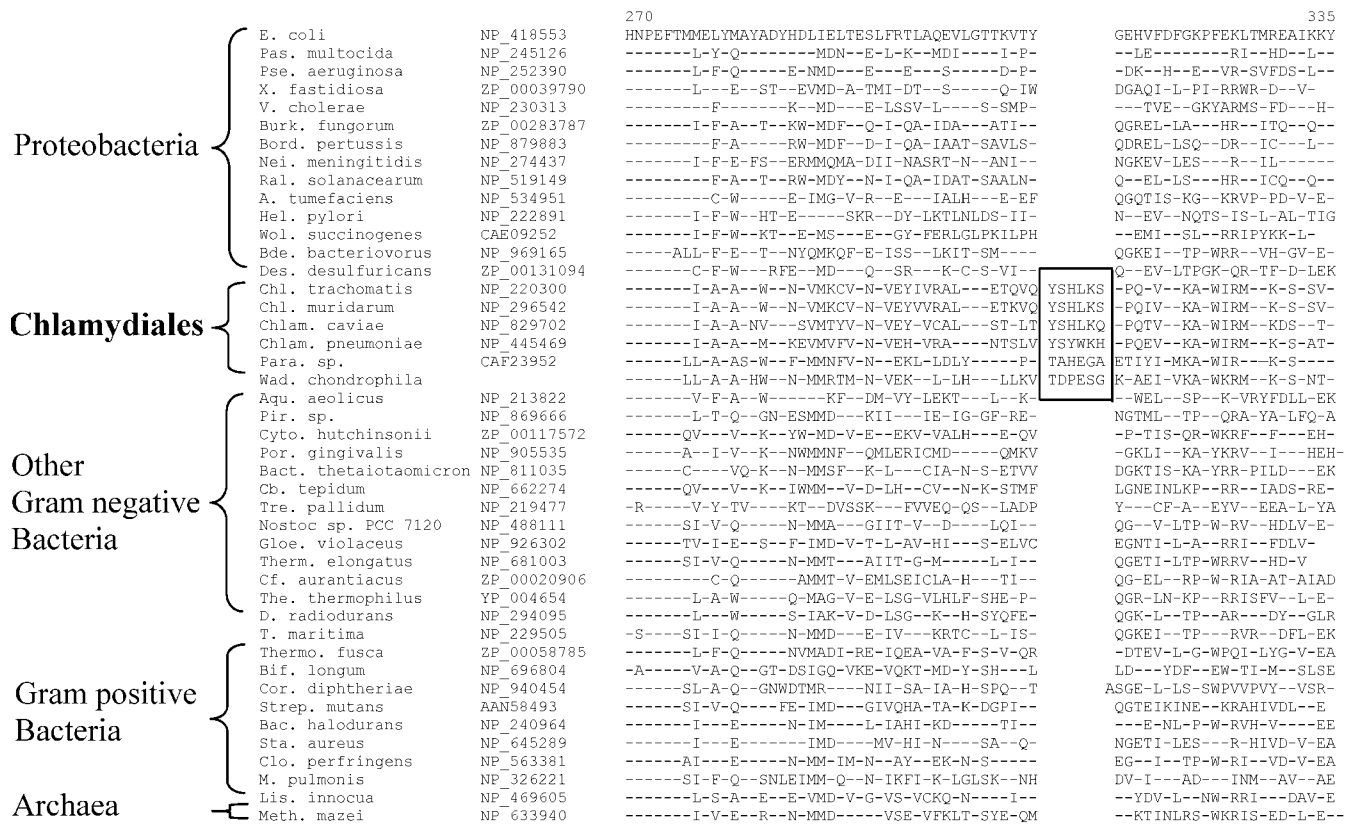


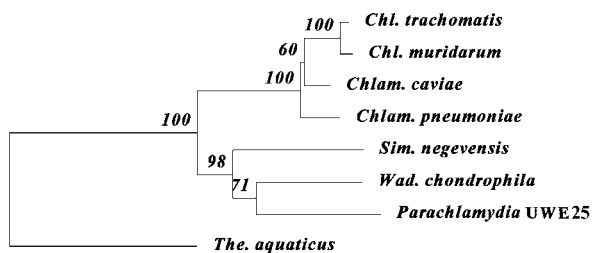
Fig. 5. Excerpt from a lysyl-tRNA synthetase sequence alignment showing a 6 aa insert that is distinctive of the Chlamydiales. The sequence for *Wad. chondrophila* was obtained in the present work. Additional abbreviations: *Gloe.*, *Gloeobacter*; *Meth.*, *Methosarcina*.

DISCUSSION

Our understanding of the chlamydiae has expanded greatly in recent years. Until a few years ago, the phylum Chlamydiae

consisted of a single family, the *Chlamydiaceae*, which contained a single genus with four species: *Chl. trachomatis*, *Chlamydia pecorum*, *Chlam. pneumoniae* and *Chlamydo-phila psittaci* (Fields & Barnes, 1992; Schachter & Stamm, 1999; Kalayoglu & Byrne, 2001). The host organisms for these intracellular parasites were believed to be mainly vertebrates, with only a few exceptions. However, in 1999 a major revision to the chlamydiae taxonomy was proposed by Everett *et al.* (1999a) to account for a great deal of new information that had accumulated in the preceding years. The revised proposal divided the existing *Chlamydiaceae* family into two genera, *Chlamydia* and *Chlamydo-phila*, which now contained a total of nine species (Everett *et al.*, 1999a). However, a more significant aspect of this revision was the recognition of three additional families

(A) Combined protein sequences



(B) 16S rRNA

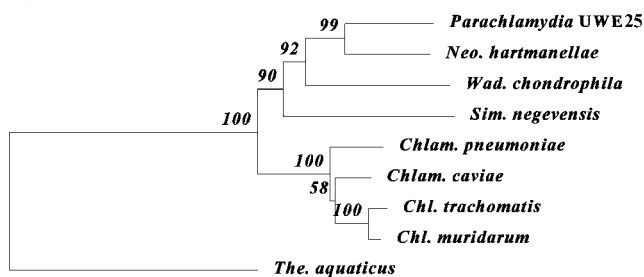


Fig. 6. Branching order of chlamydiae species based on (A) a concatenated dataset of four protein sequences (RpoA, EF-Tu, EF-P and GyrB) and (B) 16S rRNA sequences. The trees shown are neighbour-joined consensus trees obtained from 100 bootstrap samples and they were rooted using the *Thermus aquaticus* sequence as the outgroup. The bootstrap scores for different branches are shown on the nodes. For abbreviations, see legends to Figs 1–5.

of chlamydia-related organisms, the *Waddliaceae*, *Parachlamydiaceae* and *Simkaniaceae*, whose host range varies from humans, rodents, birds, marsupials, insects and turtles to amoeba (as an endosymbiont) (Everett, 2000; Horn & Wagner, 2004). The chlamydiae-related species have now also been detected in numerous other species, including hydra, bivalves, isopod crustacea, spiders, reptiles and fish (Corsaro *et al.*, 2003; Draghi *et al.*, 2004; Kostanjsek *et al.*, 2004; Collingro *et al.*, 2005). Another major development that has taken place in recent years is the finding of chlamydiae-related sequences in a wide variety of environments, including freshwater ponds, soil, sewage and water conduit systems (Corsaro *et al.*, 2002, 2003). As a result of these findings, it is now widely recognized that the diversity of chlamydiae, in terms of both their numbers and their clinical involvement, is currently significantly underestimated (Ossewaarde & Meijer, 1999; Fritsche *et al.*, 2000; Horn & Wagner, 2001; Corsaro *et al.*, 2003; Thao *et al.*, 2003). Because of the zoonotic potential of environmental chlamydiae to humans and animals (Greub & Raoult, 2002; Corsaro & Venditti, 2004; Horn & Wagner, 2004), it is important to develop reliable means for identifying chlamydiae species in different hosts and environments.

All well-studied chlamydiae species are characterized by a unique two-stage life cycle (Fields & Barnes, 1992; Kalayoglu & Byrne, 2001). However, for most new reports of chlamydiae-like species (sequences), which are from non-culturable sources, it is difficult to obtain information in this regard. The primary means used for identifying novel chlamydiae species in different hosts and environmental samples involves PCR amplification using primers based on 16S rRNA sequences, and the clustering of any amplified sequence with the known chlamydiae species in the rRNA trees (Rurangirwa *et al.*, 1999; Everett *et al.*, 1999b; Horn *et al.*, 2000, 2004; Corsaro *et al.*, 2002; Thao *et al.*, 2003). Although this method has proved very useful in advancing our understanding of chlamydial diversity, the branching patterns in phylogenetic trees are known to be affected by a large number of factors (Felsenstein, 1988; Gupta, 1998; Moreira & Philippe, 2000; Baldauf, 2003). Other than their branching pattern in the tree, no distinctive molecular signature for chlamydiae is present in the 16S rRNA sequences. Hence, the availability of other specific and reliable molecular markers for confirming the presence of chlamydiae species in environmental samples should be very useful.

In the present work, we have described several distinctive molecular markers for chlamydiae species in highly conserved proteins. These markers consist of prominent inserts and deletions in highly conserved regions of a number of important proteins (RpoA, EF-Tu, GyrB, EF-P and LysRS) that are unique to various chlamydiae homologues and not found in any other bacteria. The sequence information for most of these proteins (except LysRS) has been obtained from *Waddlia* and *Simkania*, and for EF-Tu and GyrB from *Neochlamydia* as well. The shared presence of

these signatures in all of the chlamydiae families strongly indicates that these signatures were introduced in a common ancestor of the Chlamydiales and that they are distinctive characteristics of the entire phylum. The proteins in which these chlamydiae-specific signatures are present are all involved in essential functions related to information-transfer processes. Most of these proteins are ubiquitous in bacteria, and some (RpoA, EF-Tu and LysRS) are present in species from all three domains. Because of their essential functions, the primary structures of these proteins are highly conserved. The PCR primers for EF-Tu and gyrase B used in our work were successful in amplifying the corresponding gene fragments from *Sim. negevensis*, *Wad. chondrophila* and *Neo. hartmanellae*. Because these primers are based on sequence information for *Chlamydiaceae* and *Parachlamydia* sp. UWE25, they are also expected to work for these species. The PCR primers for RpoA and EF-P also successfully amplified the corresponding fragments from *Sim. negevensis* and *Wad. chondrophila*, but they were not successful with *Neo. hartmanellae* under the experimental conditions used in the present study. However, based on the fact that these sequences are highly conserved, it should be possible to design other primers that would work in cases where PCR amplification was not successful in the present study. These signatures and the PCR primers provide novel means, in addition to the 16S rRNA-based primers, for identifying and confirming the presence of chlamydiae-related sequences in environmental samples. Because of the chlamydiae specificities of these signatures, if any of the amplified sequences are found to contain these signatures, this will provide reliable evidence for the presence of a chlamydiae-like organism in a given sample.

Phylogenetic analysis based on a combined dataset of protein fragments from RpoA, EF-Tu, Ef-P and GyrB proteins indicates that the environmental chlamydiae (*Simkania*, *Waddlia* and *Parachlamydia*) and the traditional *Chlamydiaceae* (*Chlamydomphila* and *Chlamydia*) species form distinct clades in the resulting tree. Similar relationships were also noted in individual protein phylogenies as well as in a tree for these species based on the 16S rRNA sequences. Thus, it is highly likely that the chlamydiae-like species (*Waddlia*, *Parachlamydia* and *Simkania*) have diverged from the traditional *Chlamydiaceae* species at a very early stage in the evolution of this group of bacteria. If this is the case, then the chlamydiae-like species may contain a very different range of metabolic, infective and virulence capabilities, as evidenced by the large difference in genome size between *Parachlamydia* sp. UWE25 and all other sequenced *Chlamydiaceae* (Hatch, 1998; Kalman *et al.*, 1999; Read *et al.*, 2000, 2003; Horn *et al.*, 2004).

ACKNOWLEDGEMENTS

The work was supported by a research grant from the Canadian Institute of Health Research to R. S. G.

REFERENCES

- Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein databases search programs. *Nucleic Acids Res* **25**, 3389–3402.
- Amann, R., Springer, N., Schonhuber, W., Ludwig, W., Schmid, E. N., Muller, K. D. & Michel, R. (1997). Obligate intracellular bacterial parasites of acanthamoebae related to *Chlamydia* spp. *Appl Environ Microbiol* **63**, 115–121.
- Baldauf, S. L. (2003). Phylogeny for the faint of heart: a tutorial. *Trends Genet* **19**, 345–351.
- Birtles, R. J., Rowbotham, T. J., Storey, C., Marrie, T. J. & Raoult, D. (1997). Chlamydia-like obligate parasite of free-living amoebae. *Lancet* **349**, 925–926.
- Bush, R. M. & Everett, K. D. (2001). Molecular evolution of the *Chlamydiaceae*. *Int J Syst Evol Microbiol* **51**, 203–220.
- Collingro, A., Poppert, S., Heinz, E., Schmitz-Esser, S., Essig, A., Schweikert, M., Wagner, M. & Horn, M. (2005). Recovery of an environmental chlamydia strain from activated sludge by co-cultivation with *Acanthamoeba* sp. *Microbiology* **151**, 301–309.
- Corsaro, D. & Venditti, D. (2004). Emerging chlamydial infections. *Crit Rev Microbiol* **30**, 75–106.
- Corsaro, D., Venditti, D. & Valassina, M. (2002). New chlamydial lineages from freshwater samples. *Microbiology* **148**, 343–344.
- Corsaro, D., Valassina, M. & Venditti, D. (2003). Increasing diversity within Chlamydiae. *Crit Rev Microbiol* **29**, 37–78.
- Draghi, A., Popov, V. L., Kahl, M. M., Stanton, J. B., Brown, C. C., Tsongalis, G. J., West, A. B. & Frasca, S., Jr. (2004). Characterization of “*Candidatus piscichlamydia salmonis*” (order Chlamydiales), a chlamydia-like bacterium associated with epitheliocystis in farmed Atlantic salmon (*Salmo salar*). *J Clin Microbiol* **42**, 5286–5297.
- Everett, K. D. (2000). Chlamydia and Chlamydiales: more than meets the eye. *Vet Microbiol* **75**, 109–126.
- Everett, K. D., Bush, R. M. & Andersen, A. A. (1999a). Emended description of the order Chlamydiales, proposal of Parachlamydiaceae fam. nov. and Simkaniaceae fam. nov., each containing one monotypic genus, revised taxonomy of the family Chlamydiaceae, including a new genus and five new species, and standards for the identification of organisms. *Int J Syst Bacteriol* **49**, 415–440.
- Everett, K. D., Hornung, L. J. & Andersen, A. A. (1999b). Rapid detection of the Chlamydiaceae and other families in the order Chlamydiales: three PCR tests. *J Clin Microbiol* **37**, 575–580.
- Felsenstein, J. (1988). Phylogenies from molecular sequences: inference and reliability. *Annu Rev Genet* **22**, 521–565.
- Fields, P. I. & Barnes, R. C. (1992). The genus *Chlamydia*. In *The Prokaryotes*, pp. 3691–3709. Edited by A. Balows, H. G. Trüper, M. Dworkin, W. Harder & K. H. Schleifer. New York: Springer.
- Fritsche, T. R., Horn, M., Wagner, M., Herwig, R. P., Schleifer, K. H. & Gautam, R. K. (2000). Phylogenetic diversity among geographically dispersed Chlamydiales endosymbionts recovered from clinical and environmental isolates of *Acanthamoeba* spp. *Appl Environ Microbiol* **66**, 2613–2619.
- Greub, G. & Raoult, D. (2002). Parachlamydiaceae: potential emerging pathogens. *Emerg Infect Dis* **8**, 625–630.
- Griffiths, E. & Gupta, R. S. (2001). The use of signature sequences in different proteins to determine the relative branching order of bacterial divisions: evidence that *Fibrobacter* diverged at a similar time to *Chlamydia* and the *Cytophaga-Flavobacterium-Bacteroides* division. *Microbiology* **147**, 2611–2622.
- Griffiths, E. & Gupta, R. S. (2002). Protein signatures distinctive of chlamydial species: horizontal transfer of cell wall biosynthesis genes *glmU* from Archaeobacteria to Chlamydiae, and *murA* between Chlamydiae and *Streptomyces*. *Microbiology* **148**, 2541–2549.
- Griffiths, E. & Gupta, R. S. (2004). Distinctive protein signatures provide molecular markers and evidence for the monophyletic nature of the *Deinococcus-Thermus* phylum. *J Bacteriol* **186**, 3097–3107.
- Gupta, R. S. (1998). Protein phylogenies and signature sequences: a reappraisal of evolutionary relationships among Archaeobacteria, Eubacteria, and Eukaryotes. *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. (2004). The phylogeny and signature sequences characteristics of *Fibrobacters*, *Chlorobi* and *Bacteroidetes*. *Crit Rev Microbiol* **30**, 123–143.
- Gupta, R. S., Pereira, M., Chandrasekera, C. & Johari, V. (2003). Molecular signatures in protein sequences that are characteristic of Cyanobacteria and plastid homologues. *Int J Syst Evol Microbiol* **53**, 1833–1842.
- Hashimoto, T. & Hasegawa, M. (1996). Origin and early evolution of eukaryotes inferred from the amino acid sequences of translation elongation factors 1alpha/Tu and 2/G. *Adv Biophysics* **32**, 73–120.
- Hatch, T. (1998). *Chlamydia*: old ideas crushed, new mysteries bared. *Science* **282**, 638–639.
- Horn, M. & Wagner, M. (2001). Evidence for additional genus-level diversity of Chlamydiales in the environment. *FEMS Microbiol Lett* **204**, 71–74.
- Horn, M. & Wagner, M. (2004). Bacterial endosymbionts of free-living amoebae. *J Eukaryot Microbiol* **51**, 509–514.
- Horn, M., Wagner, M., Muller, K. D., Schmid, E. N., Fritsche, T. R., Schleifer, K. H. & Michel, R. (2000). *Neochlamydia hartmannellae* gen. nov., sp. nov. (Parachlamydiaceae), an endoparasite of the amoeba *Hartmannella vermiformis*. *Microbiology* **146**, 1231–1239.
- Horn, M., Collingro, A., Schmitz-Esser, S. & 10 other authors (2004). Illuminating the evolutionary history of chlamydiae. *Science* **304**, 728–730.
- Kahane, S., Everett, K. D., Kimmel, N. & Friedman, M. G. (1999). *Simkania negevensis* strain ZT: growth, antigenic and genome characteristics. *Int J Syst Bacteriol* **49**, 815–820.
- Kahane, S., Platzner, N., Dvoskin, B., Itzhaki, A. & Friedman, M. G. (2004). Evidence for the presence of *Simkania negevensis* in drinking water and in reclaimed wastewater in Israel. *Appl Environ Microbiol* **70**, 3346–3351.
- Kalayoglu, M. V. & Byrne, G. I. (2001). *Chlamydia*. In *The Prokaryotes: an Evolving Electronic Resource for the Microbiological Community*, 3rd edn, release 3.7, <http://141.150.157.117:8080/prokPUB/chaprender/jsp/showchap.jsp?chapnum=335>. Edited by M. Dworkin. New York: Springer.
- Kalman, S., Mitchell, W., Marathe, R. & 7 other authors (1999). Comparative genomes of *Chlamydia pneumoniae* and *C. trachomatis*. *Nat Genet* **21**, 385–389.
- Kostanjsek, R., Lapanje, A., Rupnik, M., Strus, J., Drobne, D. & Avgustin, G. (2004). Anaerobic bacteria in the gut of terrestrial isopod Crustacean *Porcellio scaber*. *Folia Microbiol (Praha)* **49**, 179–182.
- Levine, C., Hiasa, H. & Marians, K. J. (1998). DNA gyrase and topoisomerase IV: biochemical activities, physiological roles during chromosome replication, and drug sensitivities. *Biochim Biophys Acta* **1400**, 29–43.

- Mahony, J. B., Chong, S., Coombes, B. K., Smieja, M. & Petrich, A. (2000).** Analytical sensitivity, reproducibility of results, and clinical performance of five PCR assays for detecting *Chlamydia pneumoniae* DNA in peripheral blood mononuclear cells. *J Clin Microbiol* **38**, 2622–2627.
- Maidak, B. L., Cole, J. R., Lilburn, T. G. & 7 other authors (2001).** The RDP-II (Ribosomal Database Project). *Nucleic Acids Res* **29**, 173–174.
- Moreira, D. & Philippe, H. (2000).** Molecular phylogeny: pitfalls and progress. *Int Microbiol* **3**, 9–16.
- Olsen, G. J. & Woese, C. R. (1997).** Archaeal genomics: an overview. *Cell* **89**, 991–994.
- Ossewaarde, J. M. & Meijer, A. (1999).** Molecular evidence for the existence of additional members of the order *Chlamydiales*. *Microbiology* **145**, 411–417.
- Read, T. D., Brunham, R. C., Shen, C. & 21 other authors (2000).** Genome sequences of *Chlamydia trachomatis* MoPn and *Chlamydia pneumoniae* AR39. *Nucleic Acids Res* **28**, 1397–1406.
- Read, T. D., Myers, G. S. A., Brunham, R. C. & 18 other authors (2003).** Genome sequence of *Chlamydophila caviae* (*Chlamydia psittaci* GPIC): examining the role of niche-specific genes in the evolution of the Chlamydiaceae. *Nucleic Acids Res* **31**, 2134–2147.
- Roblin, P. M., Dumornay, W. & Hammerschlag, M. R. (1992).** Use of HEp-2 cells for improved isolation and passage of *Chlamydia pneumoniae*. *J Clin Microbiol* **30**, 1968–1971.
- Rurangirwa, F. R., Dilbeck, P. M., Crawford, T. B., McGuire, T. C. & McElwain, T. F. (1999).** Analysis of the 16S rRNA gene of micro-organism WSU 86-1044 from an aborted bovine foetus reveals that it is a member of the order *Chlamydiales*: proposal of *Waddliaceae* fam. nov., *Waddlia chondrophila* gen. nov., sp. nov. *Int J Syst Bacteriol* **49**, 577–581.
- Schachter, J. & Stamm, W. E. (1999).** *Chlamydia*. In *Manual of Clinical Microbiology*, pp. 795–806. Edited by P. R. Murray, E. J. Baron, M. A. Tenover & R. H. Tenover. Washington, DC: American Society for Microbiology.
- Stephens, R. S., Kalman, S., Lammel, C. & 9 other authors (1998).** Genome sequence of an obligate intracellular pathogen of humans: *Chlamydia trachomatis*. *Science* **282**, 754–759.
- Thao, M. L., Baumann, L., Hess, J. M., Falk, B. W., Ng, J. C., Gullan, P. J. & Baumann, P. (2003).** Phylogenetic evidence for two new insect-associated *Chlamydia* of the family *Simkaniaceae*. *Curr Microbiol* **47**, 46–50.
- Van de, P. Y. & De Wachter, R. (1994).** TREECON for Windows: a software package for the construction and drawing of evolutionary trees for the Microsoft Windows environment. *Comput Appl Biosci* **10**, 569–570.
- Woese, C. R., Olsen, G. J., Ibba, M. & Soll, D. (2000).** Aminoacyl-tRNA synthetases, the genetic code, and the evolutionary process. *Microbiol Mol Biol Rev* **64**, 202–236.