Ticks and Tick-borne Diseases 10 (2019) 462–469

Contents lists available at ScienceDirect

Ticks and Tick-borne Diseases

journal homepage: www.elsevier.com/locate/ttbdis

Paradoxical evolution of rickettsial genomes

Awa Diop[a], Didier Raoult[b], Pierre-Edouard Fournier[a,*]

[a] UMR VITROME, Aix-Marseille University, IRD, Service de Santé des Armées, Assistance Publique-Hôpitaux de Marseille, Institut Hospitalo-Uuniversitaire Méditerranée Infection, 19-21 Boulevard Jean Moulin, 13005, Marseille, France
[b] UMR MEPHI, Aix-Marseille University, IRD, Assistance Publique-Hôpitaux de Marseille, Institut Hospitalo-Uuniversitaire Méditerranée Infection, Marseille, France

ARTICLE INFO

Keywords:
Rickettsia
Genomics
Evolution
Virulence
Genome rearrangement
Non-coding DNA
Gene loss
DNA repeats

ABSTRACT

Rickettsia species are strictly intracellular bacteria that evolved approximately 150 million years ago from a presumably free-living common ancestor from the order *Rickettsiales* that followed a transition to an obligate intracellular lifestyle. Rickettsiae are best known as human pathogens vectored by various arthropods causing a range of mild to severe human diseases. As part of their obligate intracellular lifestyle, rickettsial genomes have undergone a convergent evolution that includes a strong genomic reduction resulting from progressive gene degradation, genomic rearrangements as well as a paradoxical expansion of various genetic elements, notably small RNAs and short palindromic elements whose role remains unknown. This reductive evolutionary process is not unique to members of the *Rickettsia* genus but is common to several human pathogenic bacteria. Gene loss, gene duplication, DNA repeat duplication and horizontal gene transfer all have shaped rickettsial genome evolution. Gene loss mostly involved amino-acid, ATP, LPS and cell wall component biosynthesis and transcriptional regulators, but with a high preservation of toxin-antitoxin (TA) modules, recombination and DNA repair proteins. Surprisingly the most virulent *Rickettsia* species were shown to have the most drastically reduced and degraded genomes compared to closely related species of milder pathogenesis. In contrast, the less pathogenic species harbored the greatest number of mobile genetic elements. Thus, this distinct evolutionary process observed in *Rickettsia* species may be correlated with the differences in virulence and pathogenicity observed in these obligate intracellular bacteria. However, future investigations are needed to provide novel insights into the evolution of genome sizes and content, for that a better understanding of the balance between proliferation and elimination of genetic material in these intracellular bacteria is required.

1. Introduction

The genus *Rickettsia* (order *Rickettsiales*, family *Rickettsiaceae*) comprises strictly intracellular α-proteobacteria mostly associated with diverse arthropod vectors around the world (Raoult and Roux, 1997; Stothard et al., 1994). *Rickettsia* species evolved approximately 150 million years ago from a common ancestor of *Rickettsiales* that was presumably free-living, and progressively followed a transition to an obligate intracellular lifestyle that occurred 775–525 million years ago and then to primarily infecting arthropod lineages approximately 525–425 million years ago (El Karkouri et al., 2016; Merhej and Raoult, 2011; Weinert et al., 2009a,b). These bacteria are also well known to infect mammalian hosts, mostly through arthropod bites or arthropod feces infecting scratching lesions. On the basis of their phenotypic properties, vector hosts and phylogenetic organization, *Rickettsia* species were split into three to four groups by different authors (Fig. 1): i) the spotted fever group (SFG, Fig. 1) contains many spotted fever-causing species as well as numerous species of as-yet unknown pathogenicity. SFG rickettsiae are mostly associated with ticks, but also fleas and mites (Diop et al., 2017); ii) the second phylogenetic group, the typhus group (TG, Fig. 1) is only made of *R. prowazekii* and *R. typhi* that cause epidemic and murine typhus, and are associated with human body lice and rat fleas, respectively (Diop et al., 2017); iii) the ancestral group includes *R. bellii* and *R. canadensis*. These species diverged early from SFG and TG rickettsiae, are associated with ticks but do not cause human disease (Fig. 1) (Diop et al., 2017); iv) a fourth group, named transitional group, was proposed by Gillespie et al. to include SFG species phylogenetically close to *R. felis* (Gillespie et al., 2007). However, as these species do not exhibit significant differences with other SFG species except their phylogenetic position, several authors discussed the validity of this latter group (Shpynov et al., 2018).

Rickettsia species cause a range of illnesses, from mild and self-limiting to severe and life-threatening diseases (Diop et al., 2017). Currently, the most common rickettsioses are African tick-bite fever

* Corresponding author.
E-mail address: pierre-edouard.fournier@univ-amu.fr (P.-E. Fournier).

https://doi.org/10.1016/j.ttbdis.2018.11.007
Received 19 February 2018; Received in revised form 8 August 2018; Accepted 9 November 2018
Available online 12 November 2018
1877-959X/

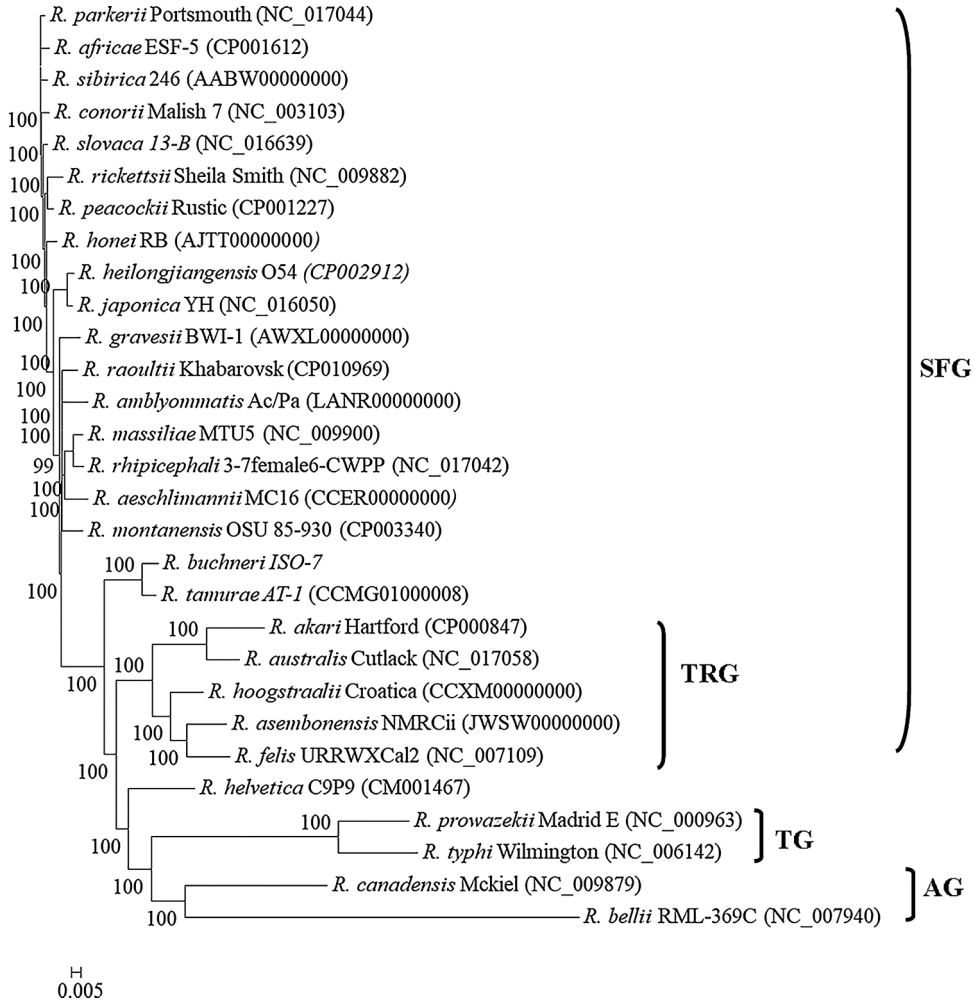


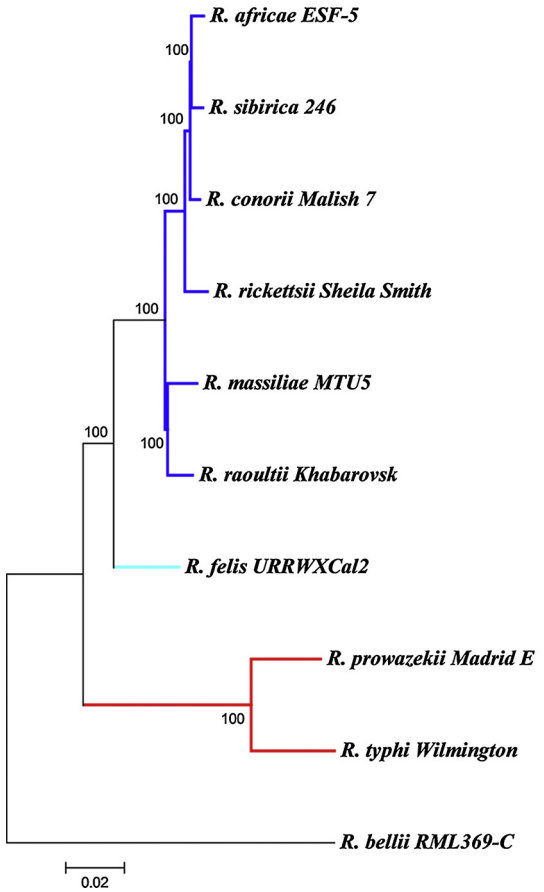
Fig. 1. Phylogenomic tree of 29 *Rickettsia* species based on whole-genome sequence analysis using the Maximum Likehood method within the FastTree software. Genomes were aligned using Mugsy software. Values at the nodes are percentages. Numbers at the nodes represent the percentages of bootstrap values obtained by repeating the analysis 1000 times to generate a majority consensus tree. Only values greater than 90% were reported. AG = Ancestral group; TG = Typhus group; TRG = Transitional group; SFG = Spotted fever group.

caused by *R. africae*, scalp eschar and neck lymphadenopathy (SENLAT) caused by *R. slovaca*, Mediterranean spotted fever (MSF) caused by *R. conorii*, Rocky Mountain spotted fever (RMSF) caused by *R. rickettsii* and murine typhus caused by *R. typhi*. (El Karkouri et al., 2017; Parola et al., 2013; Sahni et al., 2013). *Rickettsia prowazekii*, the historical agent of epidemic typhus, is only rarely encountered currently but has a strong epidemic potential (Parola et al., 2013). Furthermore, recent studies have reported the association of other *Rickettsia* lineages with other reservoirs including protozoa, algae, leeches, plants or insects (Merhej and Raoult, 2011; Murray et al., 2016; Weinert et al., 2009a,b).

In 1998, the first complete *Rickettsia* genome, that of *R. prowazekii* strain Madrid E, was sequenced (Andersson et al., 1998). It was the seventh bacterial genome to be sequenced. Subsequently, the genomes of many *Rickettsia* species have been fully sequenced, allowing a better knowledge of the molecular mechanisms involved in their pathogenicity (Balraj et al., 2009). Genome sequencing also appeared as a potential tool to revolutionize the phylogenetic and evolutionary investigations of prokaryotes, especially endosymbiotic bacteria. Hence, deciphering rickettsial genomes appeared as an efficient tool to understand the evolution of these obligate intracellular bacteria.

2. General features of rickettsia genomes

Rickettsia species have small genome sizes and low G + C contents. SFG and TG rickettsiae exhibit genome sizes of 1.25 to 2.3 Mb, and 1.11 Mb, respectively. They also exhibit G + C contents ranging from 32.2 to 33.0% and 28.9 to 29.0%, respectively. *Rickettsia* species have numbers of predicted protein-coding genes varying between 817 and 2479 and most of them maintain a near perfect chromosomal synteny (Diop et al., 2017), which enabled the identification of an ongoing and progressive genome degradation (Ogata, 2001). Rickettsial genomes contain many functional or unfunctional pseudogenes and possess a high percentage of non-coding DNA (Blanc et al., 2007a,b; McLeod et al., 2004) (Fig. 2). This percentage of non-coding DNA ranges from 16.2% for *R. felis* to 31% for *R. massiliae*. *Rickettsia prowazekii*, the most reduced rickettsial genome contains 24% of non-coding sequence. By comparison, *Chlamydia trachomatis*, another strictly intracellular bacterium, possesses only 10% non-coding DNA (Andersson et al., 1998; Holste et al., 2000; Rogozin et al., 2002).This pseudogenization progressively leads to a genome downsizing and results from a switch from a free-living to an obligate intracellular lifestyle. This progressive reductive evolution has allowed rickettsiae to purge unnecessary and redundant genes mainly involved in metabolisms supplied by eukaryotic host cells (Georgiades and Raoult, 2011; Merhej et al., 2009). Paradoxically to this ongoing genomic reduction, rickettsial genomes exhibit another marker of convergent evolution, *i. e.*, the expansion of genetic elements including small RNAs, tandem repeats, short palindromic elements named rickettsia palindromic elements (RPEs) (Ogata et al., 2002), ankyrin and



Virulence	Size (bp)	protein-coding genes	% coding sequences	RNAs	pseudogenes	RPEs
Milder	1,278,540 pRaf:12,377	1112	78.26	39	246	460
Virulent	1,250,021	1083	77.76	36	200	NA
Virulent	1,268,755	1374	81.5	39	252	559
Virulent	1,257,710	1345	78.5	36	233	NA
Mild	1,360,898 pRma:15,286	968	69	39	286	562
Mild	1,344,605 pRra1:20,840 pRra2:83,219 pRra3:34,583	1180	71.2	39	339	NA
Mild	1,485,148 pRF:6,282 pRFδ:39,268	1444	83.8	39	130	726
Highly virulent	1,111,523	834	76.2	39	181	120
Virulent	1,111,496	838	76.3	39	185	121
Unknown	1,522,076	1429	82.5	40	100	526

Fig. 2. Phylogenomic tree based on 591 core proteins and pathogenic and genomic features, of *Rickettsia* species exhibiting various degrees of pathogenesis. For each genome (downloaded from GenBank), gene prediction was obtained using the Prokka software (Seemann, 2014). The core genome was identified using the ProteinOrtho software (Lechner et al., 2011). Then, the amino acid sequences of 591 proteins (Supplementary Table) conserved in all studied genomes were concatenated for each species and multiple alignment was performed using the Mafft software (Katoh and Standley, 2013, p. 2). Gapped positions were removed. The phylogenetic inferences were obtained using the Maximum Likelihood method and the MEGA software version 6 (Tamura et al., 2013). Branching support was evaluated using the bootstrap method with 1000 replications. Bootstrap values greater than 90% are shown at the nodes. Properties of each species were extracted from the following references (Andersson et al., 1998; Blanc et al., 2007a,b; Guillaume Blanc et al., 2007a,b; El Karkouri et al., 2017, 2016; Fournier et al., 2009; McLeod et al., 2004; Ogata, 2001; Ogata et al., 2006, 2005a). NA = data not available; RPEs = Rickettsia palindromic elements.

tetratricopeptide repeats and gene family duplication mainly ADP-ATP translocases, toxin-antitoxin modules and type IV secretion system (T4SS). Another unexpected property of rickettsial genomes is the presence of plasmids, the first described in obligate intracellular bacteria. The first plasmid was identified in *R. felis* (Ogata et al., 2005a). To date, at least 20 rickettsial plasmids have been described in 11 species. Their number varies from 1 to 4 per species/strain (Baldridge et al., 2007; Blanc et al., 2007a,b; El Karkouri et al., 2016). These findings suggest possible exchanges of genetic material by conjugation, a mechanism that was thought to be absent in obligate intracellular and allopatric bacteria (Georgiades and Raoult, 2011; Merhej et al., 2009; Ogata et al., 2005a).

3. Rickettsia genome in an ongoing convergent evolution

3.1. Ongoing reductive evolution of Rickettsial genomes

Following their adaptation from a free-living to an obligate intracellular lifestyle in eukaryotic cells, rickettsiae underwent genomic changes to fit their specific bottleneck ecosystem, resulting not only in a reducing genome size but also in a specific genomic architecture (Keeling et al., 1994; Sicheritz-Pontén and Andersson, 1997). Comparative genomics revealed that rickettsiae, by taking advantage of host cell metabolites, underwent a genome reductive evolution (Georgiades and Raoult, 2011; Merhej et al., 2009) that occurred through a progressive pseudogenization (Fig. 2) and gene loss of selected biosynthetic pathway components (Andersson et al., 1998; Audia and Winkler, 2006; Fournier et al., 2009; Ogata, 2001; Sakharkar, 2004; Walker, 2005; Wolf and Koonin, 2013). In addition, genomic degradation was detrimental for the G + C content, as it led to an enrichment in A + T, in particular in the high proportion of non-coding DNA (Sakharkar, 2004). However, a great variation in chromosome size, ranging from 1.1 to 2.3 Mb, is observed in rickettsiae (Diop et al., 2017), indicating that some species are at a more advanced stage of reductive genomic evolution (TG rickettsiae) than others (SFG rickettsiae) (Ogata, 2001). In ehrlichiae, a similar genomic reduction is observed, but the G + C content may remain as high as 49.8% in *Anaplasma* species (Dunning Hotopp et al., 2006), suggesting that the reductive process in these bacteria had a lesser impact on the G + C content degradation. Rickettsial genomes are characterized by a high rate of accumulation of

slightly harmful deletions, mutations and insertions (Brynnel et al., 1998). Alternatively, gene loss can also result from the accumulations of small mutations. The formation of internal stop codons within intact genes can occur through the creation of a frameshit by single base mutation, insertion or deletion (Ogata, 2001). This induces the genome degradation resulting from fragmented gene accumulation or gene remnants. An unexpected finding of rickettsial genomics was that the most virulent species had the most reduced genomes (Fournier et al., 2009). Such a finding is not an isolated phenomenon as in *Mycobacterium*, *Streptococcus* spp., *Corynebacterium* spp. and other genera, the highest degree of gene loss is observed in the most virulent species when compared to closely related and milder or nonpathogenic species (Blanc et al., 2007a,b; Merhej et al., 2013; Ogata, 2001). Many of the genes required by free-living bacteria are absent in *Rickettsia* (Bechah et al., 2010) and degraded genes include mostly those involved in the biosynthesis of nutrients (Blanc, 2005; Ogata, 2001; Renesto et al., 2005). For example, *Rickettsia* exhibits few genes for de novo nucleotide synthesis, *i. e.,* only those for conversion of nucleoside monophosphates into all other nucleotides, implying that they take up nucleoside monophosphates from the host (Wixon, 2001). Analysis of *R. conorii* and *R. prowazekii* genomes (Dunning Hotopp et al., 2006; Ogata, 2001) revealed that genes coding glycolytic enzymes and those required for nucleotide or cofactor biosynthesis are totally absent in *R. conorii* and *R. prowazekii* when compared to most genera in the order *Rickettsiales* that have complete glycolytic pathways. Nevertheless, rickettsiae must obtain glycerol-3-phosphate from the host via a glycerol-3-phosphate transporter (Dunning Hotopp et al., 2006). This ATP production profile is similar for *Rickettsia* and mitochondria, as they possess a high number of ATP/ADP translocases, suggesting that they have both evolved from a common ancestor (Andersson et al., 1998; Renesto et al., 2005). In addition, the genome sequencing of *R. prowazekii* revealed a lack of amino acid metabolism such as those for glutamate metabolism (Andersson et al., 1998; Fuxelius et al., 2007). The enzymes involved in the aspartate and alanine metabolism pathways, and those playing a role in the biosynthesis of leucine, valine, isoleucine and aromatic amino acids (tryptophan, tyrosine, phenylalanine) are similarly missing in *Rickettsia* species (Renesto et al., 2005), suggesting the use of host-derived amino acids for their growth, survival and replication. Additionally, all *Rickettsia* species except *R. belli* have a reduced set of folate biosynthesis genes (Fuxelius et al., 2007). In TG rickettsiae all five genes required for the de novo folate biosynthesis are lacking (Hunter et al., 2015). Furthermore, a limited set of genes for LPS and cell wall component biosynthesis, including lipid-A and peptidoglycan, respectively, were identified in *Rickettsia* species (Fuxelius et al., 2007). The rickettsial surface protein-coding genes *rickA* and *sca*2 are another example of genes that were degraded or eliminated by *Rickettsia* species during their specialization. The RickA protein participates in actin polymerization through the activation of Arp2/3 similar to that found in *Listeria monocytogenes* and *Shigella* spp. (Balraj et al., 2008b; Gouin et al., 2004, 1999). While lacking in the TG, *rickA* is present in all AG and SFG rickettsial genomes avalaible (Baldridge et al., 2005; Balraj et al., 2008a, 2008b; Heinzen et al., 1993; Jeng et al., 2004; McLeod et al., 2004; Ogata, 2001; Ogata et al., 2006, 2005a). The absence of *rickA* in *R. prowazekii* is not surprising if we consider its lack of actin motility. In contrast, *R. typhi* exhibits a unique and erratic actin-based motility despite having a nonfunctional RickA protein (McLeod et al., 2004; Reed et al., 2014). In addition, *R. canadensis* expresses RickA but does not exhibit actin-based motility (Heinzen et al., 1993). These data suggest the possible involvement of other actin polymerization mechanisms and that RickA alone may not be sufficient or required for actin-based rickettsial motility. Nevertheless, it was proposed that RickA originated early in rickettsial evolution and may have been lost during the divergence of the TG. Recent research suggests that *Rickettsia* spp. use also Sca2 for actin-based motility with a distinct mechanism compared to RickA. Sca2 was found to be intact in *R. conorii*, absent in *R. prowazekii* and pseudogenized in *R. typhi* (McLeod et al., 2004). In *R. typhi*, Sca2 lacks the FH1 (formin homology 1) domain and contains only a proline-rich tract and a series of five WH2 domains (β-domains) in different locations with a divergence in sequences (Sears et al., 2012). The evolutionary process of genome degradation in rickettsiae led to loss of transcriptional regulator genes with a decreased translational capacity as observed in *R. prowazekii* (Andersson and Kurland, 1998), despite conserved gene sets coding for toxins, toxin-antitoxin (TA) modules and recombination and DNA repair proteins most likely needed for protection against host immune response (Moran, 2002).

The reductive evolution of rickettsial genomes is not only the consequence of gene degradation or loss, but it is also linked to a differential expression level of genes (Diop et al., 2017). Some genes under the influence of evolutionary forces are dormant or repressed while others under this effect are overexpressed. Recent research involving two virulent and two milder SFG rickettsiae demonstrated that the two virulent agents *R. conorii* (MSF) and *R. slovaca* (SENLAT) have the most reduced genome and displayed less up-regulated than down-regulated genes than the milder *R. massiliae* and *R. raoultii* causing MSF and SENLAT, respectively (El Karkouri et al., 2017), that have less reduced genomes. Consequently, to adapt to their specific intracellular environment, *Rickettsia* species were shaped by distinct evolutionary processes. The most pathogenic species are characterized by a strong reductive genomic evolution, with a higher genome degradation rate and accumulation of non-coding DNA than less pathogenic species. These findings suggest that reductive genomic evolution, resulting in protein structural variations, is associated to the emergence of virulence (El Karkouri et al., 2017). It was speculated that the loss of regulator genes, as observed in several intracellular pathogens, is a critical cause of virulence (Darby et al., 2007). This reductive genomic evolution appears to have occurred in several other human pathogens that have no common intracellular ancestor with *Rickettsia* such as *Treponema* spp., *Mycobacterium* spp. or *Yersinia* spp (Merhej et al., 2009; Walker, 2005; Wixon, 2001). Overall, during the course of evolution, rickettsial genomes exhibit a trend toward gene loss rather than acquisition, but strong selective effects co-exist with functional duplication required for survival.

3.2. Gene order, recombination events and “junk DNA” in rickettsial genomes

A comparison of 13 rickettsial genomes (Diop et al., 2017) demonstrated that they exhibit a highly conserved synteny and present few genomic rearrangements, except for *R. bellii* that exhibits little colinearity with other genomes, and *R. felis* that underwent several inversions. In addition, *R. typhi*, underwent a 35-kb inversion close to the replication terminus and a specific 124-kb inversion nearby the origin of replication when compared to *R. prowazekii* and *R. conorii* (McLeod et al., 2004). As in other bacteria, inversions that occured in the origin of replication region are also found in *R. australis*, *R. helvetica* and *R. honei* (Dong et al., 2012a,b; Xin et al., 2012), indicating that this region constitutes a hotspot for genomic rearrangement (Eisen et al., 2000). Homologous intra-chromosomal recombination, the principal mechanism for genomic rearrangement in rickettsiae, occured between repeated sequences or by site-specific recombination. Consequently, duplications, deletions and inversions arose through these structures (Andersson and Kurland, 1998; Krawiec and Riley, 1990). Such events have been observed in Rickettsia spp., in the so-called super-ribosomal protein gene operon. Highly conserved in a broad range of bacteria and archaea, this operon consists of about 40 genes located in seven operons in the same order (Sicheritz-Pontén and Andersson, 1997). Despite their conserved order in many bacteria including *E. coli* and *Bacillus subtilis*, genes in the ribosomal protein gene operon are scattered around the genomes of *Haemophilus influenzae*, *Mycoplasma genitalium* and *R. prowazeki* (Andersson and Kurland, 1998; Fraser et al., 1995). Ribosomal RNA genes in bacterial genomes are normally organized into an operon with a conserved order 16S-23S-5S, and tRNA genes are often found in

the spacer between the 16S and the 23S rRNA genes (Krawiec and Riley, 1990). However, an unusual arrangement of rRNA genes has been observed in all avalaible *Rickettsia* genomes, as the 16S rRNA gene is separated from the 23S and 5S rRNA gene cluster (Andersson et al., 1999; Munson et al., 1993). A similar organization is observed in all members of the order *Rickettsiales* (Dunning Hotopp et al., 2006). The upstream spacer of the rearranged 23S rRNA gene in some *Rickettsia* species contains short repetitive sequences that have been eliminated in other related species, suggesting that the rearrangement of rRNA genes occurred by intra-chromosomal recombination prior to speciation in *Rickettsia* spp. Rickettsial genome analysis highlighted a second major genomic rearrangement in rickettsiae, the elongation factor proteins (tuf and fus) being present in more than one copy in *Rickettsia* genomes (Syvänen et al., 1996). These genes can serve as repeat sequences, and initiate a rapid gene loss through intra-chromosomal recombination (Krawiec and Riley, 1990). In addition, the degree and positions of deletions caused by intra-chromosomal recombination in *Rickettsia* is different among the species, which suggests that the homologous recombination is an ongoing process that may result in an ongoing genes loss under weak or no selection pressure.

When compared to other bacterial genomes, rickettsial genomes have a high percentage of non-coding DNA sequences which also contains many DNA repeat sequences (Holste et al., 2000; Rogozin et al., 2002). Non-coding DNA in rickettsial genomes is traditionally considered as "junk DNA" resulting from gene degradation. *R. prowazekii* and *R. typhi*, the most reduced rickettsial genomes, harbor high rates of non-coding DNA with 24.6 and 23.7%, respectively. However, *R. bellii* exhibits the lowest rickettsial level of non-coding DNA with 14.8% (Diop et al., 2017).

3.3. Paradoxical genomic expansions

From a general point of view, rickettsial genomes are typical of those of symbiotic bacteria, in which the reductive trend is the dominant mode of evolution (Andersson and Andersson, 1999; Georgiades and Raoult, 2011; Merhej et al., 2009; Ogata, 2005). However, despite this reductive evolution, a paradoxical expansion of genetic elements can still occur in rickettsial genomes (Ogata et al., 2002). Genome sequence analysis revealed that rickettsial genome expansion may occur through proliferation of selfish DNA (small non coding RNAs (sRNAs) and rickettsia palindromic elements (RPEs)), gene duplications and horizontal gene transfer (Merhej and Raoult, 2011). Bacterial non-coding RNAs, whose biogenesis is predominantly attributed to either the intergenic regions (trans-acting) or to the antisense strand of an open reading frame (cis-acting) (Schroeder et al., 2015), were well documented in many bacterial taxa including *Enterobacteriaceae*, *Listeria monocytogenes*, *Clostridium perfringens*, *Staphylococcus aureus, Pseudomonas aeruginosa* and *Mycobacterium tuberculosis* (Papenfort and Vanderpool, 2015). sRNAs are classified among the most important post-transcriptional regulators involved in virulence and adaptation depending on the host niche, through transcriptomic regulation (Schroeder et al., 2015). Schroeder et al. (2015) were the first to identify sRNAs in *Rickettsia* species. Twenty to 30% of intergenic regions presumably encode for trans-acting sRNAs (14–191 sRNAs, depending on species). These findings may explain the highly conserved intergenic spacers identified by early comparative studies in *Rickettsia* (Ogata, 2001). More than 1700 trans-acting sRNAs were predicted in 16 genomes of 13 species spanning all rickettsial groups (Schroeder et al., 2015). *Rickettsia prowazekii* was shown to possess stem loop structures after homopolymeric poly(T) stretches in the termination sites where the expression of sRNAs occurs (Woodard and Wood, 2011). Rickettsia palindromic elements (RPEs) were identified in 2002 by Ogata et al. (Ogata et al., 2002). These genetic elements are more abundant in SFG than TG rickettsiae (Fig. 2). In the *R. conorii* genome, a total of 656 RPEs, classified into 8 families, were identified (RPE-1 to RPE-8) and represent 3.2% of the entire genome (Ogata et al., 2002). By comparison, only 10 of the 44 RPE-1 copies described in *R. conorii* were found in the *R. prowazekii* genome. Surprisingly, nine of these 10 RPE-1 copies that are present in *R. prowazekii* are inserted in protein-coding genes, *versus* 19/44 in *R. conorii*. In addition, the RPE-1 s inserted into protein-coding genes have a position compatible with the 3-dimentional fold and function of proteins (Ogata et al., 2000). This process of genomic evolution by inserting RPEs within protein-coding genes was initially thought to be unique to *Rickettsia* species but is also encounterssed in the *Wolbachia* genus (Ogata et al., 2005b; Riegler et al., 2012). Bacteria may use this random strategy to adapt their genetic repertoire in response to selective environmental pressure. The presence of a mobile element inserted in many unrelated genes also suggests the potential role of selfish DNA in rickettsial genome for de novo creation of new protein sequences during the course of evolution, suggesting an implication in the dynamics of genome evolution (Claverie and Ogata, 2003). Moreover, genomic comparison also enabled the identification of several copies of Ankyrin and Tetratricopeptide (TPR)-repeats in rickettsiae. Such repeated elements are frequently found in endosymbionts and assumed to play a role in host-pathogen interaction (Caturegli et al., 2000; Felsheim et al., 2009; Seshadri et al., 2003; Wu et al., 2004). Twenty-two copies of ankyrin- and 11 copies of TPR-repeats were found in *R. felis* (Ogata et al., 2005a). In both species, they were proposed to be linked to pathogenicity. In *Legionella pneumophila*, which exhibits 20 Ankyrin-repeat copies and numerous TPR-repeat copies, these elements are suspected to play a modulatory role in the interactions with the host cytoskeleton and in interferences with the host cell trafficking events, respectively (Cazalet et al., 2004).

In addition to DNA repeat sequences, various gene families are duplicated in rickettsial genomes. Gene duplication was considered as an important source of bacterial adaptation to environmental changes in the host (Hooper, 2003). Following duplication, gene copies can evolve by conserving the same functions or undergoing mutations and becoming non-functional or assuming new functions, thus providing a putative new selective advantage in a new environment (Greub and Raoult, 2003; Walsh, 1995). *Rickettsia prowazekii*, the most reduced and degraded rickettsial genome that lacks the genes encoding the biosynthesis of purines and pyrimidines (Andersson et al., 1998), exhibits five copies of *tlc1* genes. These genes encode ADP/ATP translocases responsible of energy exploitation from host cells (Greub and Raoult, 2003; Renesto et al., 2005). Similar sequences were found in *R. typhi*, *R. rickettsii* and *R. montanensis*. Thus, the duplication of the tlc genes in *Rickettsia* is most likely explained by their important role in maintaining an efficient uptake and transport system of host cytoplasmic. ATP Four to 14 copies of *spoT* genes, involved in stringent response and the adaptation to intracellular environment, were also found in rickettsiae (Ogata et al., 2005a; Renesto et al., 2005; Rovery et al., 2005). The *R. conorii* genome has multiple copies of ampG agent encoding β-lactamase, which may explain the resistance of these bacteria to β-lactam antibiotics (Ogata, 2001). The T4SS, a multiple component, membrane-spanning transporter system containing eight distinct classes such as the MPF-T class (P-T4SSs), is largely found in many rickettsial genomes. Rickettsiae possess an incomplete P-T4SS system (related to systems of the IncP group conjugative plasmid) that is characterized by the lack of *virB*5 but the duplication of the *virB*4, *virB*6, *virB*8 and *virB*9 genes (Gillespie et al., 2016). The *R. prowazekii* genome has six Vir components (*virB*4, *virB*8-*virB*11, *virD*4), and the *virB*4 and *virB*9 were duplicated (Gillespie et al., 2009). Seventeen orthologous surface cell antigen-coding genes (*sca*) were identified in rickettsial genomes (Blanc, 2005). SCA proteins autotransporter proteins that were demonstrated to play roles in mammalian cell infection as well as infection of their arthropod host cells, notably by promoting actin-based motility (Sears et al., 2012). The *R. bellii* genome possesses a set of complete conjugation genes, and pilli like-filaments were observed on the bacterial surface (Ogata et al., 2006). Among 13 tested Rickettsia collection strains, 11 got positive conjugation gene detection. This suggests that

the conjugation elements are widely present among Rickettsia spp (88), and that horizontal gene transfer (HGT) occured at a high rate (Weinert et al., 2009a,b). Within amoebae, HGTs have given the *Rickettsia* ancestor the access to novel gene pools, with possibility to acquire foreign DNA from other intracellular bacteria, thus, in capability of adaptation environment (Ogata et al., 2006). In addition, a RAGE module, considered as a genetic exchange facilitator, was found in multiple copies in the genome from *Rickettsia* endosymbiont of *Ixodes scapularis* (REIS), the largest rickettsial genome described to date (Gillespie et al., 2014, 2012).

Finally, a large number of mobile genetic elements (MGEs) referred to as mobilome are found in rickettsiae despite their reduced genome size. This mobilome, mostly consisting of plasmids, may ensure DNA movement within and between genomes. To date, at least 20 known rickettsial plasmids have been described in 11 species despite their allopatric lifestyle (Diop et al., 2017). Recent phylogenomic analysis revealed that rickettsial plasmids are undergoing reductive evolutionary events similar to those affecting their co-residing chromosomes (El Karkouri et al., 2016). Rickettsial plasmids were thus shaped by a biphasic model of convergent evolution including a strong reductive evolution as well as an increased complexity via horizontal gene transfer and gene duplication and genesis (El Karkouri et al., 2016). The most reduced and virulent rickettsial genomes have probablely lost plasmid(s) during their evolution when compared to the related milder or non pathogenic species (Darby et al., 2007; El Karkouri et al., 2017; Ogata et al., 2005a).

4. Conclusions and perspectives

Rickettsia species are strictly intracellular bacteria that are likely to have evolved from a presumably free-living ancestor and followed a transition to an obligate intracellular lifestyle. To adapt to such a bottleneck lifestyle associated with genetic drift, *Rickettsia* species have been shaped by distinct evolutionary processes resulting not only in differences in genome size, but also in genomic architecture. Generally, rickettsial genomes are small and contain a high ratio of non-coding DNA, which suggests that the reductive trend is their dominant mode of evolution. Comparative sequence analysis has provided important clues on the mechanisms driving the genome-reduction process of *Rickettsia* spp. This phenomenon is marked by a selected loss of genes such as those associated with amino-acid, ATP, LPS and cell wall component biosynthesis with a loss of regulatory genes and a high preservation of toxin-associated proteins and toxin-antitoxin modules. Homologous intra-chromosomal recombination, principal mechanism for genomic rearrangement structures seems play a role in rapid gene loss. Consequently, rickettsiae have evolved under a distinct process including a strong reductive evolution as well as a paradoxical expansion of genetic elements acquired by horizontal gene transfer and gene duplication and genesis. Thus, during the course of evolution, rickettsial genomes had a trend of gene loss rather than gene acquisition or duplication, but these strong selective effects co-exist with functional duplications required for survival. In order to understand the evolution of genome size and content, it is necessary to understand the balance between proliferation and elimination of genetic material in these intracellular bacteria.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.ttbdis.2018.11.007.

References

Andersson, J.O., Andersson, S.G., 1999. Genome degradation is an ongoing process in Rickettsia. Mol. Biol. Evol. 16, 1178–1191. https://doi.org/10.1093/oxfordjournals.molbev.a026208.

Andersson, S.G., Kurland, C.G., 1998. Reductive evolution of resident genomes. Trends Microbiol. 6, 263–268. https://doi.org/10.1016/S0966-842X(98)01312-2.

Andersson, S.G., Stothard, D.R., Fuerst, P., Kurland, C.G., 1999. Molecular phylogeny and rearrangement of rRNA genes in *Rickettsia* species. Mol. Biol. Evol. 16, 987–995. https://doi.org/10.1093/oxfordjournals.molbev.a026188.

Andersson, S.G., Zomorodipour, A., Andersson, J.O., Sicheritz-Pontén, T., Alsmark, U.C.M., Podowski, R.M., Näslund, A.K., Eriksson, A.-S., Winkler, H.H., Kurland, C.G., 1998. The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria. Nature 396, 133–140.

Audia, J.P., Winkler, H.H., 2006. Study of the five *Rickettsia prowazekii* proteins annotated as ATP/ADP translocases (Tlc): only Tlc1 transports ATP/ADP, while Tlc4 and Tlc5 transport other ribonucleotides. J. Bacteriol. 188, 6261–6268. https://doi.org/10.1128/JB.00371-06.

Baldridge, G.D., Burkhardt, N., Herron, M.J., Kurtti, T.J., Munderloh, U.G., 2005. Analysis of fluorescent protein expression in transformants of *Rickettsia monacensis*, an obligate intracellular tick symbiont. Appl. Environ. Microbiol. 71, 2095–2105. https://doi.org/10.1128/AEM.71.4.2095-2105.2005.

Baldridge, G.D., Burkhardt, N.Y., Felsheim, R.F., Kurtti, T.J., Munderloh, U.G., 2007. Transposon insertion reveals pRM, a plasmid of *Rickettsia monacensis*. Appl. Environ. Microbiol. 73, 4984–4995. https://doi.org/10.1128/AEM.00988-07.

Balraj, P., Karkouri, K.E., Vestris, G., Espinosa, L., Raoult, D., Renesto, P., 2008a. RickA Expression is not sufficient to promote actin-based motility of *Rickettsia raoultii*. PLoS One 3, e2582. https://doi.org/10.1371/journal.pone.0002582.

Balraj, P., Nappez, C., Raoult, D., Renesto, P., 2008b. Western-blot detection of RickA within spotted fever group rickettsiae using a specific monoclonal antibody. FEMS Microbiol. Lett. 286, 257–262. https://doi.org/10.1111/j.1574-6968.2008.01283.x.

Balraj, P., Renesto, P., Raoult, D., 2009. Advances in *Rickettsia* pathogenicity. Ann. N. Y. Acad. Sci. 1166, 94–105. https://doi.org/10.1111/j.1749-6632.2009.04517.x.

Bechah, Y., El Karkouri, K., Mediannikov, O., Leroy, Q., Pelletier, N., Robert, C., Medigue, C., Mege, J.L., Raoult, D., 2010. Genomic, proteomic, and transcriptomic analysis of virulent and avirulent *Rickettsia prowazekii* reveals its adaptive mutation capabilities. Genome Res. 20, 655–663. https://doi.org/10.1101/gr.103564.109.

Blanc, G., 2005. Molecular evolution of *Rickettsia* surface antigens: evidence of positive selection. Mol. Biol. Evol. 22, 2073–2083. https://doi.org/10.1093/molbev/msi199.

Blanc, G., Ogata, H., Robert, C., Audic, S., Claverie, J.-M., Raoult, D., 2007a. Lateral gene transfer between obligate intracellular bacteria: evidence from the *Rickettsia massiliae* genome. Genome Res. 17, 1657–1664. https://doi.org/10.1101/gr.6742107.

Blanc, G., Ogata, H., Robert, C., Audic, S., Suhre, K., Vestris, G., Claverie, J.-M., Raoult, D., 2007b. Reductive genome evolution from the mother of Rickettsia. PLoS Genet. 3, e14.

Brynnel, E.U., Kurland, C.G., Moran, N.A., Andersson, S.G., 1998. Evolutionary rates for tuf genes in endosymbionts of aphids. Mol. Biol. Evol. 15, 574–582.

Caturegli, P., Asanovich, K.M., Walls, J.J., Bakken, J.S., Madigan, J.E., Popov, V.L., Dumler, J.S., 2000. *ankA*: an *Ehrlichia phagocytophila* group gene encoding a cytoplasmic protein antigen with ankyrin repeats. Infect. Immun. 68, 5277–5283.

Cazalet, C., Rusniok, C., Brüggemann, H., Zidane, N., Magnier, A., Ma, L., Tichit, M., Jarraud, S., Bouchier, C., Vandenesch, F., Kunst, F., Etienne, J., Glaser, P., Buchrieser, C., 2004. Evidence in the *Legionella pneumophila* genome for exploitation of host cell functions and high genome plasticity. Nat. Genet. 36, 1165–1173. https://doi.org/10.1038/ng1447.

Claverie, J.-M., Ogata, H., 2003. The insertion of palindromic repeats in the evolution of proteins. Trends Biochem. Sci. 28, 75–80. https://doi.org/10.1016/S0968-0004(02)00036-1.

Darby, A.C., Cho, N.-H., Fuxelius, H.-H., Westberg, J., Andersson, S.G.E., 2007. Intracellular pathogens go extreme: genome evolution in the Rickettsiales. Trends Genet. 23, 511–520. https://doi.org/10.1016/j.tig.2007.08.002.

Diop, A., Raoult, D., Fournier, P.-E., 2017. Rickettsial genomics and the paradigm of genome reduction associated with increased virulence. Microbes Infect. https://doi.org/10.1016/j.micinf.2017.11.009.

Dong, X., El Karkouri, K., Robert, C., Gavory, F., Raoult, D., Fournier, P.-E., 2012a. Genomic comparison of *Rickettsia helvetica* and other *Rickettsia* species. J. Bacteriol. 194https://doi.org/10.1128/JB.00299-12. 2751–2751.

Dong, X., El Karkouri, K., Robert, C., Raoult, D., Fournier, P.-E., 2012b. Genome Sequence of *Rickettsia australis*, the agent of queensland tick typhus. J. Bacteriol. 194, 5129. https://doi.org/10.1128/JB.01117-12.

Dunning Hotopp, J.C., Lin, M., Madupu, R., Crabtree, J., Angiuoli, S.V., Eisen, J., Seshadri, R., Ren, Q., Wu, M., Utterback, T.R., Smith, S., Lewis, M., Khouri, H., Zhang, C., Niu, H., Lin, Q., Ohashi, N., Zhi, N., Nelson, W., Brinkac, L.M., Dodson, R.J., Rosovitz, M.J., Sundaram, J., Daugherty, S.C., Davidsen, T., Durkin, A.S., Gwinn, M., Haft, D.H., Selengut, J.D., Sullivan, S.A., Zafar, N., Zhou, L., Benahmed, F., Forberger, H., Halpin, R., Mulligan, S., Robinson, J., White, O., Rikihisa, Y., Tettelin, H., 2006. Comparative genomics of emerging human ehrlichiosis agents. PLoS Genet. 2, e21. https://doi.org/10.1371/journal.pgen.0020021.

Eisen, J.A., Heidelberg, J.F., White, O., Salzberg, S.L., 2000. Evidence for symmetric chromosomal inversions around the replication origin in bacteria. Genome Biol. 1 research0011–1.

El Karkouri, K., Kowalczewska, M., Armstrong, N., Azza, S., Fournier, P.-E., Raoult, D., 2017. Multi-omics analysis sheds light on the evolution and the intracellular lifestyle strategies of spotted fever group *Rickettsia* spp. Front. Microbiol. 8. https://doi.org/10.3389/fmicb.2017.01363.

El Karkouri, K., Mediannikov, O., Robert, C., Raoult, D., Fournier, P.-E., 2016. Genome sequence of the tick-borne pathogen Rickettsia raoultii. Genome Announc. 4https://doi.org/10.1128/genomeA.00157-16. e00157–16.

Felsheim, R.F., Kurtti, T.J., Munderloh, U.G., 2009. Genome sequence of the endosymbiont *Rickettsia peacockii* and comparison with virulent *Rickettsia rickettsii*: identification of virulence factors. PLoS One 4, e8361. https://doi.org/10.1371/

journal.pone.0008361.
Fournier, P.-E., El Karkouri, K., Leroy, Q., Robert, C., Giumelli, B., Renesto, P., Socolovschi, C., Parola, P., Audic, S., Raoult, D., 2009. Analysis of the *Rickettsia africae* genome reveals that virulence acquisition in *Rickettsia* species may be explained by genome reduction. BMC Genomics 10, 166. https://doi.org/10.1186/1471-2164-10-166.
Fraser, C.M., Gocayne, J.D., White, O., Adams, M.D., Clayton, R.A., Fleischmann, R.D., Bult, C.J., Kerlavage, A.R., Sutton, G., Kelley, J.M., Fritchman, R.D., Weidman, J.F., Small, K.V., Sandusky, M., Fuhrmann, J., Nguyen, D., Utterback, T.R., Saudek, D.M., Phillips, C.A., Merrick, J.M., Tomb, J.F., Dougherty, B.A., Bott, K.F., Hu, P.C., Lucier, T.S., Peterson, S.N., Smith, H.O., Hutchison, C.A., Venter, J.C., 1995. The minimal gene complement of Mycoplasma genitalium. Science 270, 397–403.
Fuxelius, H.-H., Darby, A., Min, C.-K., Cho, N.-H., Andersson, S.G.E., 2007. The genomic and metabolic diversity of Rickettsia. Res. Microbiol. 158, 745–753. https://doi.org/10.1016/j.resmic.2007.09.008.
Georgiades, K., Raoult, D., 2011. Genomes of the most dangerous epidemic bacteria have a virulence repertoire characterized by fewer genes but more toxin-antitoxin modules. PLoS One 6, e17962. https://doi.org/10.1371/journal.pone.0017962.
Gillespie, J.J., Ammerman, N.C., Dreher-Lesnick, S.M., Rahman, M.S., Worley, M.J., Setubal, J.C., Sobral, B.S., Azad, A.F., 2009. An Anomalous type IV secretion system in *Rickettsia* Is evolutionarily conserved. PLoS One 4, e4833. https://doi.org/10.1371/journal.pone.0004833.
Gillespie, J.J., Beier, M.S., Rahman, M.S., Ammerman, N.C., Shallom, J.M., Purkayastha, A., Sobral, B.S., Azad, A.F., 2007. Plasmids and rickettsial evolution: insight from Rickettsia felis. PLoS One 2, e266. https://doi.org/10.1371/journal.pone.0000266.
Gillespie, J.J., Joardar, V., Williams, K.P., Driscoll, T., Hostetler, J.B., Nordberg, E., Shukla, M., Walenz, B., Hill, C.A., Nene, V.M., Azad, A.F., Sobral, B.W., Caler, E., 2012. A *Rickettsia* genome overrun by mobile genetic elements provides insight into the acquisition of genes characteristic of an obligate intracellular lifestyle. J. Bacteriol. 194, 376–394. https://doi.org/10.1128/JB.06244-11.
Gillespie, J.J., Kaur, S.J., Rahman, M.S., Rennoll-Bankert, K., Sears, K.T., Beier-Sexton, M., Azad, A.F., 2014. Secretome of obligate intracellular Rickettsia. FEMS Microbiol. Rev. https://doi.org/10.1111/1574-6976.12084. n/a–n/a.
Gillespie, J.J., Phan, I.Q.H., Driscoll, T.P., Guillotte, M.L., Lehman, S.S., Rennoll-Bankert, K.E., Subramanian, S., Beier-Sexton, M., Myler, P.J., Rahman, M.S., Azad, A.F., 2016. The *Rickettsia* type IV secretion system: unrealized complexity mired by gene family expansion. Pathog. Dis. 74https://doi.org/10.1093/femspd/ftw058. ftw058.
Gouin, E., Egile, C., Dehoux, P., Villiers, V., Adams, J., Gertler, F., Li, R., Cossart, P., 2004. The RickA protein of *Rickettsia conorii* activates the Arp2/3 complex. Nature 427, 457.
Gouin, E., Gantelet, H., Egile, C., Lasa, I., Ohayon, H., Villiers, V., Gounon, P., Sansonetti, P.J., Cossart, P., 1999. A comparative study of the actin-based motilities of the pathogenic bacteria *Listeria monocytogenes, Shigella flexneri* and *Rickettsia conorii*. J. Cell. Sci. 112, 1697–1708.
Greub, G., Raoult, D., 2003. History of the ADP/ATP-translocase-encoding gene, a parasitism gene transferred from a *Chlamydiales* ancestor to plants 1 billion years ago. Appl. Environ. Microbiol. 69, 5530–5535. https://doi.org/10.1128/AEM.69.9.5530-5535.2003.
Heinzen, R.A., Hayes, S.F., Peacock, M.G., Hackstadt, T., 1993. Directional actin polymerization associated with spotted fever group *Rickettsia* infection of vero cells. Infect. Immun. 61, 1926–1935.
Holste, D., Weiss, O., Grosse, I., Herzel, H., 2000. Are noncoding sequences of *Rickettsia prowazekii* remnants of ``neutralized'' genes? J. Mol. Evol. 51, 353–362. https://doi.org/10.1007/s002390010097.
Hooper, S.D., 2003. On the nature of gene innovation: duplication patterns in microbial genomes. Mol. Biol. Evol. 20, 945–954. https://doi.org/10.1093/molbev/msg101.
Hunter, D.J., Torkelson, J.L., Bodnar, J., Mortazavi, B., Laurent, T., Deason, J., Thephavongsa, K., Zhong, J., 2015. The *Rickettsia* endosymbiont of *Ixodes pacificus* contains all the genes of de novo folate biosynthesis. PLoS One 10, e0144552.
Jeng, R.L., Goley, E.D., D'Alessio, J.A., Chaga, O.Y., Svitkina, T.M., Borisy, G.G., Heinzen, R.A., Welch, M.D., 2004. A *Rickettsia* WASP-like protein activates the Arp2/3 complex and mediates actin-based motility: *rickettsia* RickA activates the Arp2/3 complex. Cell. Microbiol. 6, 761–769. https://doi.org/10.1111/j.1462-5822.2004.00402.x.
Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software Version 7: improvements in performance and usability. Mol. Biol. Evol. 30, 772–780. https://doi.org/10.1093/molbev/mst010.
Keeling, P.J., Charlebois, R.L., Ford Doolittle, W., 1994. Archaebacterial genomes: eubacterial form and eukaryotic content. Curr. Opin. Genet. Dev. 4, 816–822. https://doi.org/10.1016/0959-437X(94)90065-5.
Krawiec, S., Riley, M., 1990. Organization of the bacterial chromosome. Microbiol. Rev. 54, 502–539.
Lechner, M., Findei\s s, S., Steiner, L., Marz, M., Stadler, P.F., Prohaska, S.J., 2011. Proteinortho: detection of (co-) orthologs in large-scale analysis. BMC Bioinformatics 12, 124.
McLeod, M.P., Qin, X., Karpathy, S.E., Gioia, J., Highlander, S.K., Fox, G.E., McNeill, T.Z., Jiang, H., Muzny, D., Jacob, L.S., Hawes, A.C., Sodergren, E., Gill, R., Hume, J., Morgan, M., Fan, G., Amin, A.G., Gibbs, R.A., Hong, C., Yu, X.-j., Walker, D.H., Weinstock, G.M., 2004. Complete Genome sequence of *Rickettsia typhi* and comparison with sequences of other rickettsiae. J. Bacteriol. 186, 5842–5855. https://doi.org/10.1128/JB.186.17.5842-5855.2004.
Merhej, V., Georgiades, K., Raoult, D., 2013. Postgenomic analysis of bacterial pathogens repertoire reveals genome reduction rather than virulence factors. Brief. Funct. Genomics 12, 291–304. https://doi.org/10.1093/bfgp/elt015.
Merhej, V., Raoult, D., 2011. Rickettsial evolution in the light of comparative genomics. Biol. Rev. 86, 379–405. https://doi.org/10.1111/j.1469-185X.2010.00151.x.
Merhej, V., Royer-Carenzi, M., Pontarotti, P., Raoult, D., 2009. Massive comparative genomic analysis reveals convergent evolution of specialized bacteria. Biol. Direct 4, 13. https://doi.org/10.1186/1745-6150-4-13.
Moran, N.A., 2002. Microbial minimalism: genome reduction in bacterial pathogens. Cell 108, 583–586.
Munson, M.A., Baumann, L., Baumann, P., 1993. Buchnera aphidicola (a prokaryotic endosymbiont of aphids) contains a putative 16S rRNA operon unlinked to the 23S rRNA-encoding gene: sequence determination, and promoter and terminator analysis. Gene 137, 171–178. https://doi.org/10.1016/0378-1119(93)90003-L.
Murray, G.G.R., Weinert, L.A., Rhule, E.L., Welch, J.J., 2016. The phylogeny of *Rickettsia* using different evolutionary signatures: how tree-like is bacterial evolution? Syst. Biol. 65, 265–279. https://doi.org/10.1093/sysbio/syv084.
Ogata, H., 2005. *Rickettsia felis*, from culture to genome sequencing. Ann. N. Y. Acad. Sci. 1063, 26–34. https://doi.org/10.1196/annals.1355.004.
Ogata, H., 2001. Mechanisms of evolution in *Rickettsia conorii* and *R. Prowazekii*. Science 293, 2093–2098. https://doi.org/10.1126/science.1061471.
Ogata, H., Audic, S., Abergel, C., Fournier, P.-E., Claverie, J.-M., 2002. Protein coding palindromes are a unique but recurrent feature in Rickettsia. Genome Res. 12, 808–816.
Ogata, H., Audic, S., Barbe, V., Artiguenave, F., Fournier, P.-E., Raoult, D.M., Claverie, J., 2000. Selfish DNA in Protein-coding Genes of Rickettsia.
Ogata, H., La Scola, B., Audic, S., Renesto, P., Blanc, G., Robert, C., Fournier, P.-E., Claverie, J.-M., Raoult, D., 2006. Genome sequence of *Rickettsia bellii* illuminates the role of amoebae in gene exchanges between intracellular pathogens. PLoS Genet. 2, e76. https://doi.org/10.1371/journal.pgen.0020076.
Ogata, H., Renesto, P., Audic, S., Robert, C., Blanc, G., Fournier, P.-E., Parinello, H., Claverie, J.-M., Raoult, D., 2005a. The genome sequence of *Rickettsia felis* identifies the first putative conjugative plasmid in an obligate intracellular parasite. PLoS Biol. 3, e248. https://doi.org/10.1371/journal.pbio.0030248.
Ogata, H., Suhre, K., Claverie, J.-M., 2005b. Discovery of protein-coding palindromic repeats in Wolbachia. Trends Microbiol. 13, 253–255. https://doi.org/10.1016/j.tim.2005.03.013.
Papenfort, K., Vanderpool, C.K., 2015. Target activation by regulatory RNAs in bacteria. FEMS Microbiol. Rev. 39, 362–378. https://doi.org/10.1093/femsre/fuv016.
Parola, P., Paddock, C.D., Socolovschi, C., Labruna, M.B., Mediannikov, O., Kernif, T., Abdad, M.Y., Stenos, J., Bitam, I., Fournier, P.-E., Raoult, D., 2013. Update on tick-borne rickettsioses around the World: a geographic approach. Clin. Microbiol. Rev. 26, 657–702. https://doi.org/10.1128/CMR.00032-13.
Raoult, D., Roux, V., 1997. Rickettsioses as paradigms of new or emerging infectious diseases. Clin. Microbiol. Rev. 10, 694–719.
Reed, S.C.O., Lamason, R.L., Risca, V.I., Abernathy, E., Welch, M.D., 2014. Rickettsia actin-based motility occurs in distinct phases mediated by different actin nucleators. Curr. Biol. 24, 98–103. https://doi.org/10.1016/j.cub.2013.11.025.
Renesto, P., Ogata, H., Audic, S., Claverie, J.-M., Raoult, D., 2005. Some lessons from *Rickettsia* genomics. FEMS Microbiol. Rev. 29, 99–117. https://doi.org/10.1016/j.femsre.2004.09.002.
Riegler, M., Iturbe-Ormaetxe, I., Woolfit, M., Miller, W.J., O'Neill, S.L., 2012. Tandem repeat markers as novel diagnostic tools for high resolution fingerprinting of Wolbachia. BMC Microbiol. 12, S12.
Rogozin, I.B., Makarova, K.S., Natale, D.A., Spiridonov, A.N., Tatusov, R.L., Wolf, Y.I., Yin, J., Koonin, E.V., 2002. Congruent evolution of different classes of non-coding DNA in prokaryotic genomes. Nucleic Acids Res. 30, 4264–4271.
Rovery, C., Renesto, P., Crapoulet, N., Matsumoto, K., Parola, P., Ogata, H., Raoult, D., 2005. Transcriptional response of *Rickettsia conorii* exposed to temperature variation and stress starvation. Res. Microbiol. 156, 211–218. https://doi.org/10.1016/j.resmic.2004.09.002.
Sahni, S.K., Narra, H.P., Sahni, A., Walker, D.H., 2013. Recent molecular insights into rickettsial pathogenesis and immunity. Future Microbiol. 8, 1265–1288. https://doi.org/10.2217/fmb.13.102.
Sakharkar, K.R., 2004. Genome reduction in prokaryotic obligatory intracellular parasites of humans: a comparative analysis. Int. J. Syst. Evol. Microbiol. 54, 1937–1941. https://doi.org/10.1099/ijs.0.63090-0.
Schroeder, C.L.C., Narra, H.P., Rojas, M., Sahni, A., Patel, J., Khanipov, K., Wood, T.G., Fofanov, Y., Sahni, S.K., 2015. Bacterial small RNAs in the genus *Rickettsia*. BMC Genomics 16. https://doi.org/10.1186/s12864-015-2293-7.
Sears, K.T., Ceraul, S.M., Gillespie, J.J., Allen, E.D., Popov, V.L., Ammerman, N.C., Rahman, M.S., Azad, A.F., 2012. Surface proteome analysis and characterization of surface cell antigen (Sca) or autotransporter family of Rickettsia typhi. PLoS Pathog. 8, e1002856. https://doi.org/10.1371/journal.ppat.1002856.
Seemann, T., 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30, 2068–2069. https://doi.org/10.1093/bioinformatics/btu153.
Seshadri, R., Paulsen, I.T., Eisen, J.A., Read, T.D., Nelson, K.E., Nelson, W.C., Ward, N.L., Tettelin, H., Davidsen, T.M., Beanan, M.J., 2003. Complete genome sequence of the Q-fever pathogen Coxiella burnetii. Proc. Natl. Acad. Sci. 100, 5455–5460.
Sicheritz-Pontén, T., Andersson, S.G., 1997. GRS: a graphic tool for genome retrieval and segment analysis. Microb. Comp. Genomics 2, 123–139.
Shpynov, S.N., Fournier, P.E., Pozdnichenko, N.N., Gumenuk, A.S., Skiba, A.A., 2018. New approaches in the systematics of rickettsiae. New Microbes New Infect. 23, 93–102. https://doi.org/10.1016/j.nmni.2018.02.012.
Stothard, D.R., Clark, J.B., Fuerst, P.A., 1994. Ancestral divergence of *Rickettsia bellii* from the spotted fever and typhus groups of *Rickettsia* and antiquity of the genus *Rickettsia*. Int. J. Syst. Evol. Microbiol. 44, 798–804.
Syvänen, A.-C., Amiri, H., Jamal, A., Andersson, S.G., Kurland, C.G., 1996. A chimeric disposition of the elongation factor genes in Rickettsia prowazekii. J. Bacteriol. 178, 6192–6199.
Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. MEGA6: molecular

evolutionary genetics analysis version 6.0. Mol. Biol. Evol. 30, 2725–2729. https://doi.org/10.1093/molbev/mst197.

Walker, D.H., 2005. Progress in rickettsial genome analysis from pioneering of *Rickettsia prowazekii* to the recent *Rickettsia typhi*. Ann. N. Y. Acad. Sci. 1063, 13–25. https://doi.org/10.1196/annals.1355.003.

Walsh, J.B., 1995. How often do duplicated genes evolve new functions? Genetics 139, 421–428.

Weinert, L.A., Welch, J.J., Jiggins, F.M., 2009a. Conjugation genes are common throughout the genus *Rickettsia* and are transmitted horizontally. Proc. R. Soc. B Biol. Sci. 276, 3619–3627. https://doi.org/10.1098/rspb.2009.0875.

Weinert, L.A., Werren, J.H., Aebi, A., Stone, G.N., Jiggins, F.M., 2009b. Evolution and diversity of *Rickettsia* bacteria. BMC Biol. 7, 6. https://doi.org/10.1186/1741-7007-7-6.

Wixon, J., 2001. Featured organism: reductive evolution in bacteria: *buchnera* sp., *Rickettsia prowazekii* and *Mycobacterium leprae*. Comp. Funct. Genomics 2, 44–48.

Wolf, Y.I., Koonin, E.V., 2013. Genome reduction as the dominant mode of evolution: prospects & overviews. BioEssays 35, 829–837. https://doi.org/10.1002/bies.201300037.

Woodard, A., Wood, D.O., 2011. Analysis of convergent gene transcripts in the obligate intracellular bacterium Rickettsia prowazekii. PLoS One 6, e16537. https://doi.org/10.1371/journal.pone.0016537.

Wu, M., Sun, L.V., Vamathevan, J., Riegler, M., Deboy, R., Brownlie, J.C., McGraw, E.A., Martin, W., Esser, C., Ahmadinejad, N., Wiegand, C., Madupu, R., Beanan, M.J., Brinkac, L.M., Daugherty, S.C., Durkin, A.S., Kolonay, J.F., Nelson, W.C., Mohamoud, Y., Lee, P., Berry, K., Young, M.B., Utterback, T., Weidman, J., Nierman, W.C., Paulsen, I.T., Nelson, K.E., Tettelin, H., O'Neill, S.L., Eisen, J.A., 2004. Phylogenomics of the reproductive parasite *Wolbachia pipientis* wMel: a streamlined genome overrun by mobile genetic elements. PLoS Biol. 2, e69. https://doi.org/10.1371/journal.pbio.0020069.

Xin, D., El Karkouri, K., Robert, C., Raoult, D., Fournier, P.-E., 2012. Genomic comparison of *Rickettsia honei* strain RBT and other *Rickettsia* species. J. Bacteriol. 194, 4145. https://doi.org/10.1128/JB.00802-12.