Genomic, proteomic, and transcriptomic analysis of virulent and avirulent *Rickettsia prowazekii* reveals its adaptive mutation capabilities

Yassina Bechah,[1] Khalid El Karkouri,[1] Oleg Mediannikov,[1] Quentin Leroy,[1] Nicolas Pelletier,[1] Catherine Robert,[1] Claudine Médigue,[2] Jean-Louis Mege,[1] and Didier Raoult[1,3]

[1]*Unit for Research on Emergent and Tropical Infectious Diseases (URMITE), Centre National de la Recherche Scientifique–Institut de Recherche pour le Développement, Faculty of Medicine, University of the Mediterranean, 13005 Marseille, France;* [2]*Genoscope, Centre National de Séquençage, Laboratoire de Génomique Comparative, 91057 Evry cedex, France*

Rickettsia prowazekii, the agent of epidemic typhus, is an obligate intracellular bacterium that is transmitted to human beings by the body louse. Several strains that differ considerably in virulence are recognized, but the genetic basis for these variations has remained unknown since the initial description of the avirulent vaccine strain nearly 70 yr ago. We use a recently developed murine model of epidemic typhus and transcriptomic, proteomic, and genetic techniques to identify the factors associated with virulence. We identified four phenotypes of *R. prowazekii* that differed in virulence, associated with the up-regulation of antiapoptotic genes or the interferon I pathway in the host cells. Transcriptional and proteomic analyses of *R. prowazekii* surface protein expression and protein methylation varied with virulence. By sequencing a virulent strain and using comparative genomics, we found hotspots of mutations in homopolymeric tracts of poly(A) and poly(T) in eight genes in an avirulent strain that split and inactivated these genes. These included *recO*, putative methyltransferase, and exported protein. Passage of the avirulent Madrid E strain in cells or in experimental animals was associated with a cascade of gene reactivations, beginning with *recO*, that restored the virulent phenotype. An area of genomic plasticity appears to determine virulence in *R. prowazekii* and represents an example of adaptive mutation for this pathogen.

[Supplemental material is available online at http://www.genome.org. The sequence data from this study have been submitted to GenBank (http://www.ncbi.nlm.nih.gov/Genbank) under accession no. CP001584. The microarray data from this study have been submitted to the NCBI Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo) under accession nos. GSE16123 and GSE15630.]

Epidemic typhus, caused by *Rickettsia prowazekii*, is a devastating disease with a mortality rate reaching 30% (Zinsser 1935; Bechah et al. 2008a). It has been reported to have killed millions of people during wars, including the Napoleonic (Raoult et al. 2006) and the First and the Second World Wars (Raoult and Roux 1999). The typical transmission of epidemic typhus involves human beings and lice (Raoult and Roux 1999). Epidemic typhus is characterized histologically by generalized vasculitis with increased vascular permeability, edema, mononuclear cell infiltrations, and activation of inflammatory mechanisms. *R. prowazekii* is known to target endothelial cells in vivo.

The prototypic virulent strain of *R. prowazekii* is Breinl (Ormsbee et al. 1978). Reference avirulent strain Madrid E was obtained after serial passages of virulent strain Madrid in eggs (Fox et al. 1954). It has been widely used as a vaccine (Fox et al. 1957). Its genome has been sequenced; it is small (1,111,523 bp) and contains large amounts (up to 24%) of noncoding DNA (Andersson et al. 1998). It was demonstrated that when Madrid E is cultivated in cell cultures (Ignatovich 1975) or when inoculated to animals (Balayeva and Nikolskaya 1972) instead of chick embryos, the virulence of bacterium is restored after several passages. This new strain that regained virulence (revertant strain) was named Evir (for virulent) (Balayeva and Nikolskaya 1972); however, no differences in protein pattern were detected when comparing parental Madrid E strain and Evir (Balayeva et al. 1992). In 1999, a new virulent strain of *R. prowazekii*, Rp22, was isolated from a patient (Birg et al. 1999).

Comparative genomic microarray study revealed highly conserved genome content between Breinl and Madrid E strains (only ~3% variation) (Ge et al. 2004).

Methylation has been proposed to play a major role in bacterial virulence. Indeed, the lysine methylation profile of surface proteins is different in Breinl and Evir compared with Madrid E (Rodionov et al. 1991; Ching et al. 1993; Turco and Winkler 1994). A methytransferase gene is inactivated in Madrid E (Zhang et al. 2006), and the *metK* gene, which codes for S-adenosylmethionine synthetase, is split in Madrid E but complete in Breinl (Andersson et al. 1998). Furthermore, restricted growth of Madrid E compared with Breinl and Evir is found in macrophages (Gambrill and Wisseman 1973; Turco and Winkler 1982).

Recently, we established an infection of BALB/c mice with Breinl that mimics human disease (Bechah et al. 2007). This murine model offers the opportunity to test the virulence of different *R. prowazekii* strains. This can also be assessed by increased migration of peripheral blood mononuclear cells (PBMCs) across endothelial cell monolayers (Bechah et al. 2008b). Here, we compare

[3]**Corresponding author.**
E-mail didier.raoult@medecine.univ-mrs.fr; fax 33-4-91-38-77-72.
Article published online before print. Article and publication date are at http://www.genome.org/cgi/doi/10.1101/gr.103564.109.

the virulence phenotypes of various *R. prowazekii* strains with their genotypes and gene expression profiles by an integral approach to identify factors associated with virulence (Fig. 1).

Results

R. prowazekii phenotypes

We have succeeded in generation of a new strain of *R. prowazekii* from avirulent Madrid E-M (from Gamaleya Institute in Moscow, Russia). After 3 mo of culture on L929 cells, the growth rate of parental Madrid E strain has changed dramatically. The bacteria began to grow as fast as virulent strains, including Evir. Considering this change, we decided to call this hypothetically new strain Erus ("E" is for Madrid E, and "rus" is for Russia from which Madrid E came to our laboratory).

Having compared the properties of all strains, we found four different phenotypes for *R. prowazekii* strains (Supplemental Table 1). Two different phenotypes identified in cell culture Madrid E grew very slowly in cells compared with all other strains (Supplemental Table 1). A revertant (Erus) was obtained after 3 mo of culture of Madrid E in L929 cell line (Fig. 1), as previously reported (Ignatovich 1975). We found three levels of pathogenicity when inoculating mice using 10^5 *R. prowazekii* cells (Supplemental Table 1). Rp22, as was previously described for Breinl (Bechah et al. 2007), caused bacteremia for 10 d; its DNA was detected in the liver, lungs, and brain, and it caused histological lesions in these organs (Supplemental Fig. 1). Evir caused bacteremia, but no lesions were observed in the sampled organs. Erus did not cause bacteremia. After intradermic inoculation in guinea pigs with the same inocula, wild strains (Breinl, Rp22) and Evir caused erythematous skin lesions, but Erus did not. Histologically, lesions showed intense inflammatory infiltrate at the inoculum site consisting mainly of macrophages and lymphocytes (Supplemental Fig. 2). Finally, the infection of endothelial cells with Breinl, Rp22, and Evir, but not Erus, significantly ($P < 0.05$) increased leukocyte transmigration compared with uninfected endothelial cells (Supplemental Fig. 3). Our results show that Erus is avirulent and Evir is less pathogenic than wild strains, as proposed previously by Balayeva and Nikolskaya (1970).

Transcriptional profile of infected human endothelial cells

The total number of genes modulated in response to Rp22 ($n = 67$) or Erus ($n = 65$) infection is comparable (Fig. 2A). However, distinct clusters of genes were identified for each strain (Fig. 2B). Indeed, 16 genes were up-regulated by Erus (Supplemental Table 2A); seven genes were interferon I (IFN-I)-inducible. Eighteen genes were specifically up-regulated in response to Rp22, including eight associated with antiapoptotic responses (Supplemental Table 2B). An antiapoptotic profile was further confirmed as we found, by reverse transcription PCR, in which proapoptotic genes *BCL2*, caspase 8, and *NAIP* were markedly down-regulated in Rp22-infected cells (Supplemental Fig. 4). In contrast, Evir modulated the expression of only 30 genes. Among the modulated genes, there is induction of expression of *HAS1* (a gene involved in antiapoptotic response) as Rp22 and of *OASL* and *IFI44* (two genes of IFN-I-inducible gene pathway) as Erus.

R. prowazekii comparative genomics

We sequenced the genome of a virulent strain (Rp22) of *R. prowazekii*, and its main characteristics (GenBank accession no. CP001584) are shown in Table 1 and Figure 3. The genome of Rp22

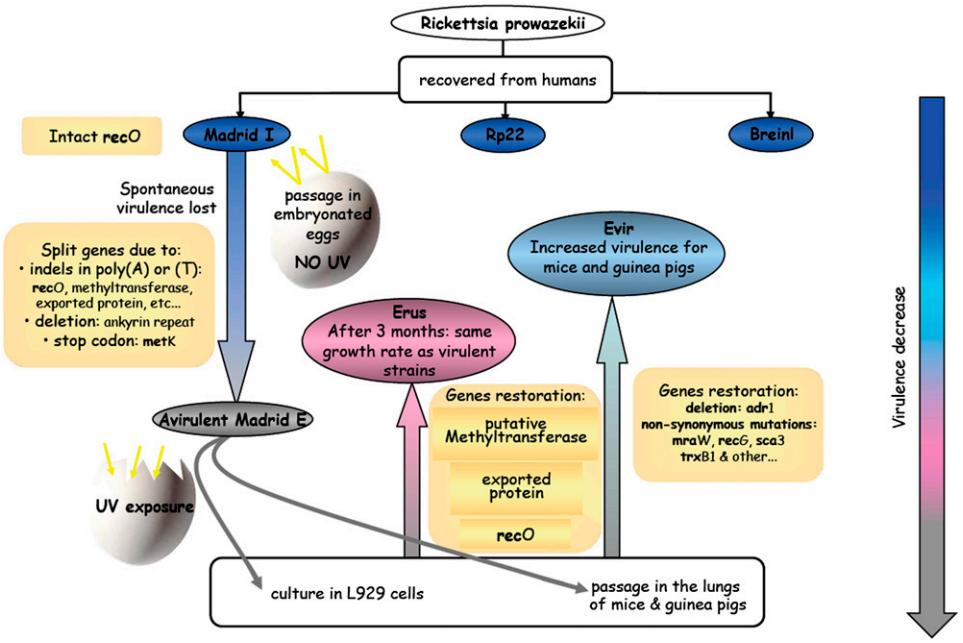

Figure 1. Scheme of *R. prowazekii* strains origin and evolution. The Breinl strain and the most recent isolate, Rp22, are considered highly virulent. The Madrid I strain was isolated in 1941 during the Madrid outbreak of epidemic typhus. After passages in embryonated eggs, Madrid I has lost its virulence and has been used under the name of Madrid E as a vaccine in humans since 1944. When it was inoculated to small rodents, Madrid E recovered some virulence (Evir). From Madrid E (300–600 passages in eggs), we have recently generated a new isolate by cultivating them in L 929 cells (Erus). (Breinl and Rp22) Virulent for humans and animals and replicates efficiently in L929 cells; (Evir) virulent for animals and replicates efficiently in L929 cells; (Erus) avirulent for animals, but replicates with L929; (Madrid E) avirulent for humans and animals and grows slowly in L929 cells.

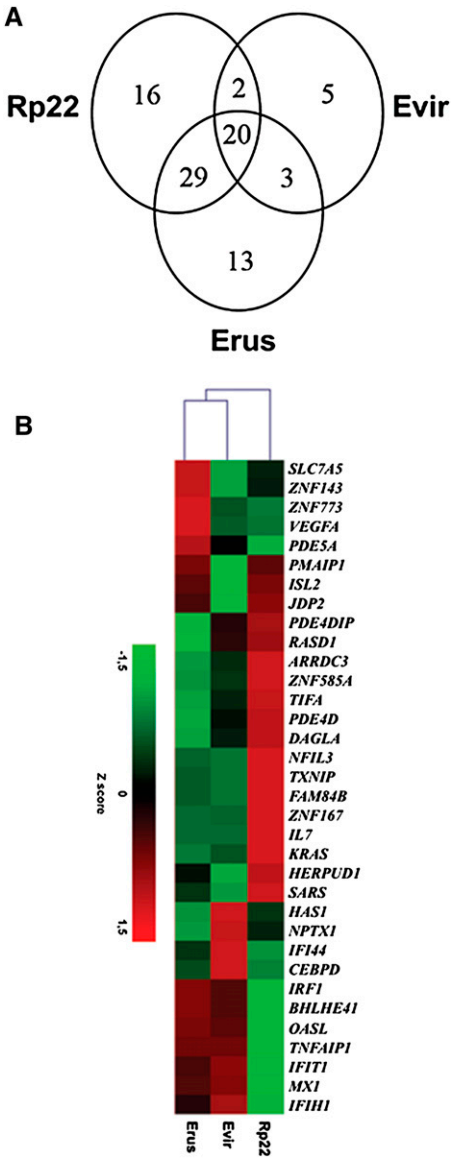

Figure 2. Transcriptional response of human endothelial cells infected by different strains of *R. prowazekii*. (*A*) Venn diagram illustrating the number of genes of endothelial cells differentially expressed in response to different *R. prowazekii* strains as compared with uninfected cells. (*B*) Transcriptional response analyzed by RNA microarrays. Modulated genes (fold change ≥ 2) were compared by unsupervised hierarchic clustering analysis.

consists of a single circular chromosome of 1,111,612 bp, with a G + C content of 29%. The RNAs from Rp22 and Madrid E exhibit identical sequences except for the 16S rDNA and a tRNA Val (Supplemental Fig. 5). We found only 81 genes differing between the two strains and classified them into four categories: 64 genes (including one split gene) with nonsynonymous mutations or gaps; 11 genes complete in Rp22 but split in Madrid E; four genes split in Rp22 but complete in Madrid E, and two genes split in both genomes at different places (Table 1). Indeed, significantly more genes are split in Madrid E (including *metK*, *recO*, and a putative methyltransferase) than in Rp22 (Table 2). Homopolymeric tracts of poly(A) and poly(T) were found significantly more often with split genes in both strains. Among the 15 genes split in either Rp22 or Madrid E, 10 showed insertion/deletion in homopolymeric poly(A) or poly(T) tracts of variable lengths (3–10 bp) (Supplemental Table 3; Fig. 4). The evolution of the Madrid E genome was found to be more rapid than that of Rp22. Using *R. typhi* as an outgroup, Madrid E showed a slightly higher K_a (nonsynonymous substitutions)/K_s (synonymous substitutions) ratio (0.378) and a longer phylogenetic branch length (0.00264) than Rp22 (0.336, 0.00122, respectively) (Supplemental Table 5; Supplemental Fig. 6).

The genes differing between these two strains were then amplified, sequenced, and compared with five additional *R. prowazekii* strains: Breinl, Madrid E USA (propagated in eggs in a laboratory in the Naval Medical Research Hospital), Madrid E-M (from the Gamaleya Institute), Evir, and Erus. This allowed the identification of four genotypes (Table 2) corresponding to the four different phenotypes. The Madrid E genotype possessed three genes differing from Erus. These genes may be responsible for phenotype differences between Erus and Madrid E in cell culture. A putative methyltransferase (RP027-28) was split in Madrid E and restored in Erus. This gene has been proposed as a candidate gene for the difference of virulence between Madrid E and Evir (Zhang et al. 2006) . However, Erus is not pathogenic; therefore, we conclude that the presence of this gene did not fully restore the virulence. A gene that repairs DNA, *recO* (RP548m-48m), was also restored in Erus. The third restored gene (RP061) had no identified function. Interestingly, early sampling during Erus transition from Madrid E showed that *recO* and Rp061 were the first restored genes (data not shown).

In a second step, we compared Erus with less pathogenic Evir and with virulent strains. No split gene differences were found; however, we found a deletion in the adhesin gene *adr1* (Table 2). We also found nonsynonymous mutations in 14 genes (Supplemental Table 6). They include surface protein (Sca3), thioredoxin (TrxB1), and genes coding for a manosyltransferase (RP340) and a putative acyltransferase (RP804). Finally, in a third step of comparing less pathogenic Evir and wild strains, we found six split genes in Evir (Table 2), including *metK* (S-adenosylmethionine synthetase) and an ankyrin repeat-containing protein. We found an insertion in *asmA* (coding for outer membrane protein assembly) in avirulent and less pathogenic strains. We also identified nonsynonymous mutations in several genes, including those coding for surface proteins of the Sca family (Sca1, Sca4, Sca5, and Sca6).

R. prowazekii gene expression

Transcriptome analysis revealed that 77 genes were differentially expressed in Rp22, as compared with Erus, when cultured in L929 cells (Fig. 5). Functional category classification identified genes

Table 1. Comparison analysis of wild strain (Rp22) and vaccine-attenuated strain (Madrid E) genomes

Species	Rp22	Madrid E
Genome size	1,111,612	1,111,523
G + C content (%)	29	29
Genes	966	966
tRNAs	33	33
rRNAs	3	3
Other RNAs	3	3
Complete and conserved genes	878	878
Nonsynonymous mutation or gapped genes	64	64
Split in one strain only	4	11
Split (same) in both strains	7	7
Split (different) in both strains	2	2

PCRs were applied for genes that were found to be different between the two sequence genomes.

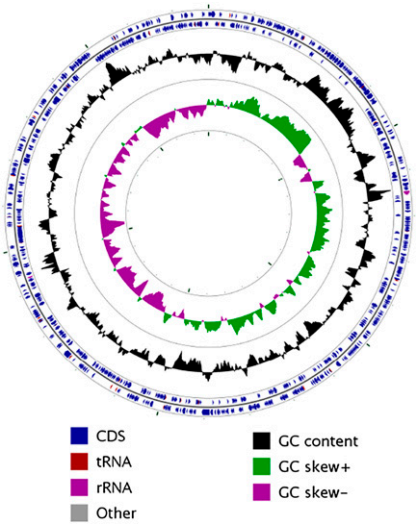

Figure 3. Circular representation of *R. prowazekii* Rp22 genome. The circles show the G-C content and skews, CDSs, rRNA, tRNA, and miscellaneous RNA (other).

mainly implicated in translation, replication, cell wall/membrane biogenesis, and post-transcriptional modifications. Thirty-three genes were classified as function unknown, general function prediction only, or not found in the COG (Clusters of Orthologous Groups) database (Supplemental Table 7). Three genes (*rOmpB*, mannosytransferase, and *asmA*) with expression differences between Rp22 and Erus had different genetic sequences (non-synonymous mutations for the two first genes and 18 coding base pair insertions for *asmA*) (Supplemental Table 6). These three genes are implicated in cell wall membrane biogenesis and are down-regulated in Rp22. In addition, Rp22 up-regulates the expression of an ankyrin repeat-containing protein, a heat shock protein (HtpG), and a peptidoglycan-associated lipoprotein precursor. The increased production of the latter was further confirmed.

Proteomic analysis showed that six proteins were down-regulated in Erus compared with Rp22 (Supplemental Table 8). The six proteins included two putative methyltransferases, translation-associated GTPase, peptidoglycan-associated lipoprotein precursor, heat shock protein (Hsp22), and preprotein translocase. Non-synonymous mutations were identified in genes coding the first three proteins. In contrast to Erus, Rp22 down-regulated the protein thioredoxin reductase (TrxB1) (Supplemental Table 8). However, this experimental result should be considered carefully, because the *trx*B1 gene has a nonsynonymous mutation (E/K) that can result in the shift of pI in a two-dimensional gel. In addition, MetK protein was produced only in Rp22 as predicted by genomic study (Table 2). Lysine methylation in both strains was also studied (Fig. 6). Surprisingly, the methylation of lysine in Erus was closer to that of Rp22 (Fig. 6A–D) than to Madrid E (Rodionov et al. 1991; Ching et al. 1993; Turco and Winkler 1994). Among all proteins only putative methyltransferase (RP789) was found in four lysine methylated isoforms in Rp22 (Fig. 6C,E) and in two isoforms in Erus (Fig. 6D,F).

Altogether, transcriptomic and proteomic analysis revealed that major differences concerned surface-exposed proteins, post-translational modification (methylation), and modifications associated with modulation of stress response. *R. prowazekii* modulated its surface proteins to escape immune recognition.

Table 2. Summary of the differences among four *R. prowazekii* genotypes

RP22 ID	ME MaGe	COG functional categories	Annotation	Brein, Rp22[a]	Evir	Erus	ME	Rp22/ME
rpr22_CDS759	RP777-777	Coenzyme metabolism	*metK*; S-adenosylmethionine synthetase	TAT	TAG	TAG	TAG	Complete/split
rpr22_CDS740	RP757-RICPRO828	Function unknown	Unknown	—	A	A	A	Complete/split
rpr22_CDS769	RP787-88	Not in COGs	Similarity to putative invasin, adhesin and agglutinin 1	I	D	D	D	Complete/split
RPR22_0486	n.i.[b]	Not in COGs	Hypothetical protein	A	—	—	—	Complete/split
rpr22_CDS010	RP012	Not in COGs	Similarity to polysaccharide deacetylase precursor	I	D	D	D	Complete/split
rpr22_CDS699	RP716	Not in COGs	Ankyrin repeat	I	D	D	D	Complete/split
rpr22_CDSx325	RICPRO325	Not in COGs	ABC transporter ATP-binding protein	I	D	D	D	Split/complete
rpr22_CDSx885	RICPRO885	Not in COGs	Hypothetical protein	—	A	A	A	Split/complete
rpr22_CDS340	RP347	Cell envelope biogenesis, outer membrane	*asmA*; Outer membrane assembly protein	D	I	I	I	~/ANQKPK
rpr22_CDS551	RP573	Function unknown	Exported protein of unknown function	D	I	I	I	~/NQ
rpr22_CDS808	RP827	Cell envelope biogenesis, outer membrane	*adr1*; *Rickettsia* adhesin	I	I	D	D	TKTI/~
rpr22_CDS024	RP027-28	General function prediction only	Putative methyltransferase	—	—	—	A	Complete/split
rpr22_CDS529	RP548m-48m	DNA replication, recombination, and repair	*recO*; DNA repair protein	—	—	—	T	Complete/split
rpr22_CDS059	RP061	Not in COGs	Exported protein of unknown function	—	—	—	T	Complete/split

(ME) Madrid E; (I) insertion of bases; (D) deletion of bases.
[a]Rp22 and Breinl yield the same results.
[b]CDSs not identified in MaGe and/or RickBase, but found to be present in *R. prowazekii* genomes.

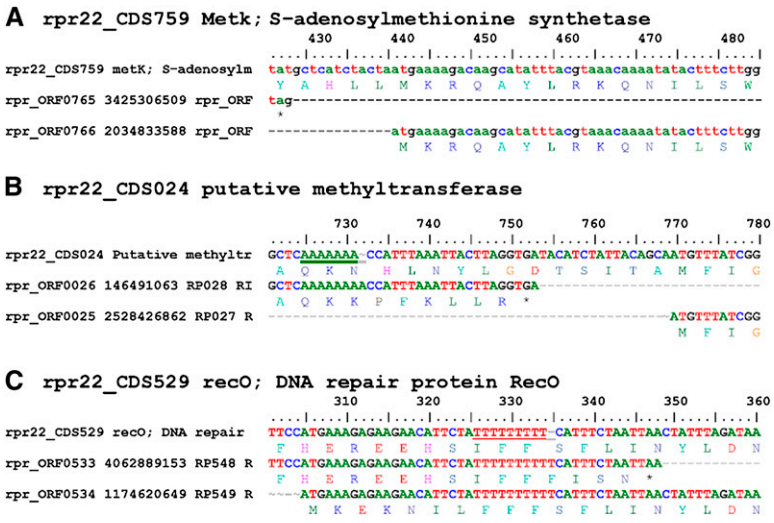

Figure 4. Sequence alignments of three *R. prowazekii* genes compared between Rp22 (rpr22) and Madrid E (rpr) strains. Split genes in Madrid E were due to stop codon (*A*), poly(A) (*B*), or poly(T) (*C*).

Discussion

There are no genetic manipulation techniques currently available for *Rickettsia*, which makes the study of its virulence difficult. *R. prowazekii* strains exhibit various degree of virulence (Turco and Winkler 1982), and a integral approach may help to better understand mutations involved in the different steps of virulence (Renesto et al. 2005). On the basis of our experimental models, we identified four different phenotypes (Fig. 1; Supplemental Table 1). This was critical, as Evir has been widely considered to be a fully restored virulent mutant (Balayeva and Nikolskaya 1973). Moreover, we confirmed that Madrid E grown in L929 cells easily generates a revertant.

To survive in a human host, *R. prowazekii* needs to avoid destruction and premature death. We found that the virulent strain inhibits IFN type I-inducible genes and promotes antiapoptotic genes to prevent death. Manipulation of IFN type I has been reported for viruses (Kato et al. 2006) and for bacteria surviving in the cytosol of infected cells such as *Listeria monocytogenes* (Leber et al. 2008). The control of apoptosis by bacteria of the order *Rickettsiales* has already been documented (Ge and Rikihisa 2006). Host manipulation by *R. prowazekii* may be mediated by ankyrin repeat-containing proteins that were considered to be key factors of host manipulation in other intracellular bacteria (Pan et al. 2008). Post-translational modifications may also play a role in the host–*R. prowazekii* relationship as was shown previously in other host–parasite systems (Polevoda and Sherman 2007). The relation between post-translational modifications, especially methylation, and virulence of *R. prowazekii* strains has been previously suggested (Rodionov et al. 1991; Ching et al. 1993; Turco and Winkler 1994; Chao et al. 2004). Here, we report only subtle differences in methylation of proteins, mainly concerning methyltransferase (RP789) itself, when comparing Rp22 with Erus. However, proteomic analysis showed that both methyltransferases (RP789 and RP527) are overproduced by Rp22, distinguishing it from Erus. Altogether, our data confirm that methylation plays a role in the virulence of *R. prowazekii* but that incompletely explains the loss or gain of virulence.

Surface-exposed proteins also seem to play a role in the variation of *R. prowazekii* virulence. The Sca family of proteins, including rOmpA (Sca0), rOmpB (Sca5), and the recently identified adhesins (Adr1 and Adr2), are major surface-exposed proteins critical in immunity (Blanc et al. 2005) and *Rickettsia*-eukaryotic cell interaction (Pizarro-Cerda and Cossart 2006; Renesto et al. 2006). We found multiple nonsynonymous mutations in *sca* family genes that possibly affect their functions among *R. prowazekii* strains. In addition, we found that the *adr1* gene was conserved among virulent and less pathogenic strains but altered (with a deletion) in Madrid E.

We identified three stages in genome degradation and restoration. Surprisingly, several genes split in the vaccine strain were able to revert to the wild type, confirming that this process is reversible. Gene repair in this process may be dependent on *recO* reparation, as it is the first identified step associated with virulence reversion. This gene codes for a protein involved in a protein complex (RecFOR) that plays a major role in DNA reparation (Chow and Courcelle 2004). We believe that its inactivation may act as a trigger in the loss of virulence of *R. prowazekii* when grown in eggs and that its inactivation favors the rapid alteration of other genes. This type of gene inactivation has been previously described in mutator clones (Caporale 2003).

The RecFOR epistatic group of proteins is implicated in the reparation of ultraviolet (UV)-damaged DNA (Chow and Courcelle 2004). Knock-out of a single coding gene may significantly complicate or even block DNA replication. We hypothesize that when rickettsiae systematically grew in chick embryos, the nearly complete absence of UV radiation and following relatively low amount of UV-damaged DNA make *recO* underused. Accidental knockout of a *recO* gene that usually may be lethal is tolerated in this situation. The beginning of cultivation of Madrid E in cell lines or animals, where UV exposure is significantly higher, will select for the rare occasional revertants (with intact *recO*). Our hypothesis is that the UV-based negative selection may play a role in case of restoration of multiple damaged genes in *R. prowazekii* via reactivation of the RecFOR system.

We believe that this may be relevant to the selection of *R. prowazekii* vaccine strain (Madrid E) in embryonated eggs, as this process has been considered an irreversible step to genome degradation (Dale et al. 2003; Lescot et al. 2008). In our work we found that, surprisingly, the restoration of *recO* gene was rapidly followed by the restoration of other genes. Restored genes, including those coding for post-translational modifications, may play a role in pathogenicity. The reparation of degraded genes after *recO* restoration may explain how the virulence may be restored in several steps.

We believe that we identified hotspot zones of plasticity in the genome of *R. prowazekii*. Poly(A) are common among AT-rich endosymbionts, and polymerase infidelity has been proved to impair and rescue gene functions (Tamas et al. 2008). We observed that revertants can be easily obtained when genes are split because of indels in poly(A) tracts. In *R. prowazekii,* three genes that are critical for pathogenicity in humans and mice were split because of this phenomenon. The reversion phenomenon is easily induced, and this explains why mutants and revertants have been produced in vitro. On the basis of our data, we think that these mutations

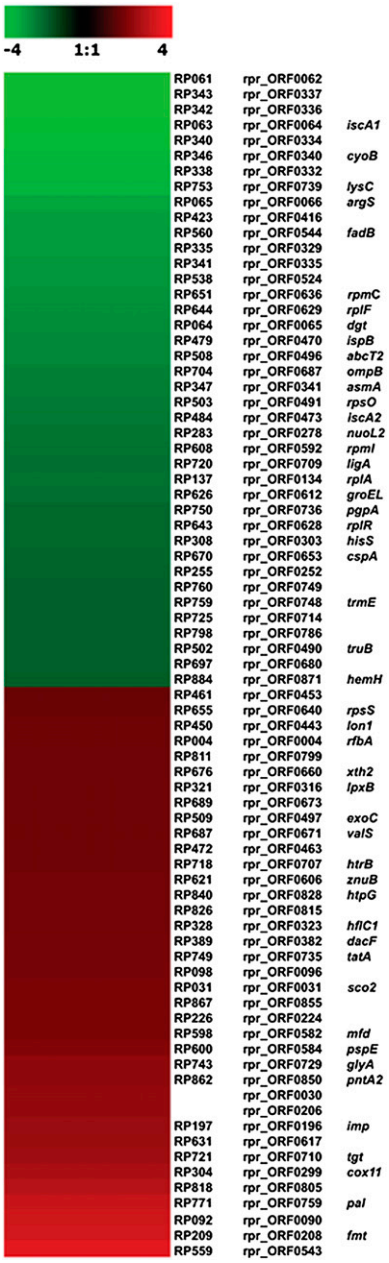

Figure 5. Transcriptional profile of Rp22 vs. Erus cultured on L929 cells and analyzed by RNA microarray. Regulated genes (fold change ≥ 1.5) in Rp22 as compared with Erus were represented using cluster analysis software.

resulting in a premature stop codon are unstable, and we prefer to use the name "split gene" instead of the previously used terms "pseudogene" or "pseudo-pseudogenes" (Baranov et al. 2005). *R. prowazekii* therefore is exemplifying the adaptive mutation concept (Rosenberg 2001), i.e., mutations formed in response to an environment. The reversal mutation of the *recO* gene occurring in a hotspot is rapidly followed by a cascade of gene restoration.

In conclusion, we found that several genes are involved in virulence differences and are mainly associated with surface proteins and post-translational modification. A key factor in the generation of the avirulent mutant is the knock out of the *recO* gene, a homopolymeric tract that is critical in accelerating genome degradation generating a mutator clone because surprisingly, its restoration permitted the restoration of putative pseudogenes.

Methods

R. prowazekii preparation

The strains used in this study are described in Figure 1, including Breinl strain; Rp22 strain, a clinical isolate that had 10 passages on L929 cells; and Evir strain, a putative virulent revertant of Madrid E selected in mice. Initially grown in embryonated eggs (300–600 passages), the vaccine Madrid E strain was cultured for 90 d in L929 cells. Bacteria grew very slowly in the cell culture, so only two passages have been performed. When, after three months, bacterial replication reached the growth rate of other strains, we called this strain "Erus."

R. prowazekii strains were propagated in L929 cells as previously described (Bechah et al. 2007). For in vivo and genomic studies, bacteria were harvested, pelleted, and purified by renographin density gradient centrifugation. For transcriptomic study, infected cells were lysed with 1% trypsin, and cell debris were eliminated by successive centrifugations at 110*g*. Bacteria were stored at −80°C before RNA extraction. For proteomic study, infected cells were lysed by sonication, and the bacterial pellet was purified in a sucrose gradient and a discontinuous renographin gradient. Bacteria were sonicated, and protein extracts were precipitated using the PlusOne 2-D Clean-Up Kit as described previously (Kowalczewska et al. 2006). They were stored at −80°C until cyanine labeling and/or isoelectric focusing (IEF).

Animal infection

BALB/c mice (7-wk-old females) were inoculated intravenously with different strains of *R. prowazekii*. For each strain, eight mice

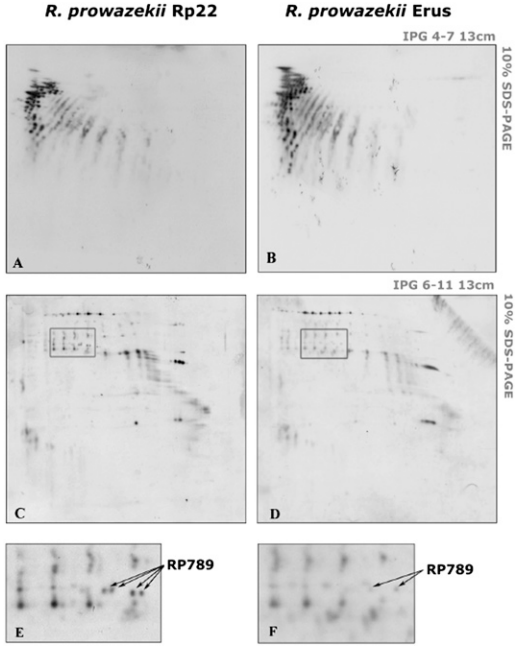

Figure 6. Lysine methylation in Rp22 (*A*,*C*,*E*) and Erus (*B*,*D*,*F*) strains. Two-dimensional Western blots were performed, and lysine methylation was visualized in pH range 4.0–7.0 (*A*,*B*) and pH range 6.0–11.0 (*C*,*D*). The zone that differed among both bacterial strains is boxed in *C* and *D*. Arrows in *E* and *F* indicate methylated spots of Rp22 and Erus strains that were identified as putative methyltransferase (RP789).

were inoculated intravenously with 5×10^5 bacteria and eight mice with 10^5 bacteria, as recently described (Bechah et al. 2007). Eight mice were also injected with 1.5×10^6 Erus bacteria/mouse. The clinical status of mice was recorded twice daily. Blood was collected and stored at −20°C for PCR. Mice were sacrificed at days 4, 6, and 10 post-infection (p.i.). Livers, lungs, and brains were aseptically excised. Half of the samples were stored at −80°C (for PCR) while the rest were fixed in 4% formalin and then embedded in paraffin for histological studies.

Six-week-old female Hartley guinea pigs were inoculated intradermally with 10^6 bacteria of each strain after shaving in separate back areas. The apparition of inflammatory lesions, which were defined as erythematous lesions without ulceration or necrosis of a size >1 mm and <2 mm, was recorded daily. A cutaneous biopsy was taken 7 d p.i. for histological examination. The 3-μm sections of paraffin-embedded specimens from guinea pigs or mice were stained with hematoxylin and eosin stain.

Bacteremia and tissue infection were determined as previously described (Bechah et al. 2007).

PBMC migration across endothelial cells

Confluent cells of the murine lung microvascular EC line L2 were seeded in gelatin-coated inserts (8-μm pore size, Costar) in 24-well plates (5×10^4 cells per well) and cultured until tight confluence. They were then infected with *R. prowazekii* organisms (bacterium-to-cell ratio of 50:1) for 6 h, as recently reported (Bechah et al. 2008b). Murine PBMCs were added to infected endothelial cells (4×10^5/well) for 24 h, and the percentage of PBMCs that had migrated across endothelial cell monolayers was determined by optical enumeration.

Microarrays of endothelial cells

Human endothelial cells (HMEC-1 cell line) were infected with *R. prowazekii* (50:1 bacterium-to-cell ratio) for 6 h at 37°C. Total RNA was extracted using an RNeasy Mini kit (Qiagen). RNA (300 ng) was reverse transcribed into cDNA with the M-MLV reverse transcriptase (RT). Labeled cRNA was synthesized from cDNA using T7 RNA polymerase and cyanine 3-labeled CTP (Cy-3) fluorescent dyes with the One-Color Low RNA Input Linear Amplification KitPLUS (Agilent). Fluorescent cRNAs were fragmented and hybridized onto 4 × 44k human wholegenome microarrays (Agilent) for 17 h at 65°C (Agilent). Slides were scanned at a 5-μm resolution with a G2505B DNA microarray scanner (Agilent). Agilent Feature Extractor Software A.9.1.3 was used for image analysis. The microarray grid was automatically placed, outlier spots were flagged, and feature intensities and backgrounds were accurately determined. Further data processing was performed using Resolver software 7.1 (Rosetta Inpharmatics), and its error model-based transformation pipeline was used to map replicate reporters to genes and to normalize inter-array data sets. The fold change (FC) was calculated using pairs Erus vs. Control, Evir vs. Control, and Rp22 vs. Control. Only genes that had an absolute FC over 2.0 ($P < 0.01$) for at least one of them were considered to be regulated and kept for further statistical analysis. Gene families were determined using numerous databases, including SOURCE (http://smd.stanford.edu/cgi-bin/source/sourceSearch) and Babelomics Fatigo+ (http://babelomics2.bioinfo.cipf.es/fatigoplus/cgi-bin/fatigoplus.cgi). The data have been submitted to the NCBI Gene Expression Omnibus (GEO) (http://www.ncbi.nlm.nih.gov/geo/) under accession numbers GPL4133 and GSE16123 for the platform of the microarray and the experimental data set, respectively.

Genome sequencing, annotation, and comparison of *R. prowazekii* strains

The genome of the Rp22 strain was sequenced using the wholegenome shotgun approach with a final coverage of 12× (Sanger method) as previously reported (Raoult et al. 2003). Genomic sequences were then assembled into contigs using *phred*/*phrap*/*consed* (Gordon et al. 1998) (http://www.phrap.org/phredphrapconsed.html), and all gaps were closed using PCR amplification and sequencing with specifically designed primers. Potential coding sequences (CDSs) were predicted using AMIGene (http://www.genoscope.cns.fr/agc/tools/amigene/Form/form.php). Split genes (either frameshifts with an in-frame stop codon or a change of their putative initiation/termination codon) or nonpredicted genes were detected and corrected manually where appropriate using Artemis (http://www.sanger.ac.uk/Software/Artemis/), BLASTN, and/or NUCmer from the MUMmer package. Assignment of protein functions was performed by searching the RickBase and MaGe databases (Vallenet et al. 2006; Blanc et al. 2007) using BLASTP. The COG functional categories of proteins were assigned using COGnitor (Tatusov et al. 1997). Ribosomal RNAs, tRNAs, and other RNAs were identified using BLASTN or tRNAscan-SE.

The Rp22 genome was then compared with that of *R. prowazekii* strain Madrid E (GenBank accession number NC_000963). Reciprocal-best BLAST matching was used to compare CDSs, proteins, and RNAs. Genes that were conserved, with or without synonymous mutations, were distinguished from those exhibiting nonsynonymous mutations, insertion/deletion(s), and/or a split state. Subsequently, discriminatory genes were compared among the *R. prowazekii* strains Breinl, Evir, Erus, Madrid E USA (propagated in eggs in a laboratory in the Naval Medical Research Hospital), and Madrid E-M (from Gamaleya Institute) and aligned using CLUSTALW. For each gene, differences were checked using amplification and sequencing with primers targeting flanking sequences.

The K_a/K_s ratio was used to estimate genomic divergence between *R. prowazekii* strains Rp22 and ME. The pairwise K_a/K_s ratios between *R. prowazekii* strains and *R. typhi* were obtained from the concatenated nucleotide alignment of 44 orthologous protein-coding genes using the K-Estimator software (Comeron 1999). The amino acid sequences of these 44 orthologs were concatenated for each genome, and multiple alignment was performed using Mafft software (Katoh et al. 2002). This was used to infer a Neighbor Joining tree using the MEGA 3.1 software (Kumar et al. 2004).

Study of bacterial transcriptome

Bacterial RNA was extracted using lysozyme and the RNeasy Mini kit (Qiagen) as recommended by the manufacturer, and DNA was digested by DNase. cDNA was synthesized from 500 ng of RNA with a random primer and the M-MLV RT. cDNA from Erus was used as reference DNA, whereas the cDNA from Rp22 strain was referred to as test DNA. Reference cDNA was amplified and labeled with cyanine 3-labeled CTP (Cy3) using the Bioprime DNA Labeling System (Invitrogen), whereas test cDNAs were labeled with cyanine 5-CTP (Cy5). Fluorescent cDNAs were fragmented and hybridized, as described above, onto *Rickettsia* Genome microarrays (Agilent Technologies) comprising probes specific for all genes and spacers from *R. prowazekii* strain Madrid E. Slides were scanned with XDR range at a 5-μm resolution. Feature Extractor 7.1 (Agilent Technologies) was used for image analysis. Data filtering and normalization were then performed using the Midas module of TM4 (La et al. 2007). The background-subtracted signals were normalized by the local subtraction method, and intensity signals were normalized by the global lowess method (La et al.

2007). Normalized signals were used for analysis with the TMev module of TM4. A *t*-test based on *P*-value permutation with a Bonferroni correction was used. Only genes with a fold change >1.5 ($P < 0.05$) were considered to have significant differential expression. The data have been submitted to the NCBI GEO (http://www.ncbi.nlm.nih.gov/geo/) under accession numbers GPL8427 and GSE15630 for the platform of the microarray and the experimental data set, respectively.

Real time RT-PCR

qRT-PCR was performed as previously described (Bechah et al. 2008b). The primer sequences are listed in Supplemental Table 4. Reverse transcriptase was omitted in the negative control. The FC in target gene cDNA relative to the housekeeping gene was determined by the $2^{-\Delta\Delta Ct}$ method (Bechah et al. 2008b). Only gene expression with a FC of >1.5 was considered to be modulated.

Differential gel electrophoresis

Differential gel electrophoresis (DIGE) was performed as follows: Each protein sample (50 μg) was labeled with 400 pmol of Cy3 or Cy5, and Cy2 was used as internal calibrator as previously described (Alban et al. 2003). IEF was performed according to the manufacturer's protocol (Ettan IPGphor II, GE Healthcare) using two pH ranges (4–7 and 6–11) of Immobiline DryStrips. For the second dimension, proteins were resolved by 10% SDS polyacrylamide gel electrophoresis (SDS-PAGE) (Ettan IPGphor II, GE Healthcare), and the gels were digitized using a Typhoon Trio scanner (Ettan IPGphor II, GE Healthcare) at wavelengths of 532 nm for Cy3, 633 nm for Cy5, and 488 nm for Cy2, and scanned at a 100-μm resolution. Protein quantification and statistical analysis were carried out using the Decyder software program (GE Healthcare). Spots with twofold or greater changes were considered significantly different and excised for identification by mass analyses (MALDI-TOF/TOF Bruker Ultraflex II spectrometer, Bruker Daltonics).

Methylation studies

For methylation studies, resolved proteins were transferred onto nitrocellulose membranes (Bio-Rad) using a semidry transfer unit (Hoefer Scientific). Membranes were then blocked in TBS (20 mM Tris-HCl at pH 7.5, 150 mM NaCl) supplemented with 0.1% Tween-20 and 5% bovine serum albumin for 1.5 h before incubation with rabbit antibodies directed against methyl-lysine (Biomol GmbH) diluted at 1:400. After 1 h, membranes were washed and probed with 1:1,000 horseradish peroxidase-conjugated goat anti-rabbit IgG (GE Healthcare). Spots were visualized using an ECL kit (GE Healthcare). Reactive spots were identified as above.

Acknowledgments

We thank R. Allen for providing us with DNA of the Madrid E strain, I. Tarasevitch for giving us the Evir strain, and S. Audic for his valuable comments. We also thank S. Azza, Y. Lepriol, T. N'Guyen, and H. Parrinello for technical assistance in proteome, transcriptome, and sequencing experiments. We thank Isabelle Combe for helping in formatting the manuscript.

References

Alban A, David SO, Bjorkesten L, Andersson C, Sloge E, Lewis S, Currie I. 2003. A novel experimental design for comparative two-dimensional gel analysis: Two-dimensional difference gel electrophoresis incorporating a pooled internal standard. *Proteomics* **3:** 36–44.

Andersson SGE, Zomorodipour A, Andersson JO, Sicheritz-Pontén T, Alsmark UCM, Podowski RM, Näslund AK, Eriksson AS, Winkler HH, Kurland CG. 1998. The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria. *Nature* **396:** 133–140.

Balayeva NM, Nikolskaya VN. 1970. Stability of increased virulence of vaccine strain E of *Rickettsia prowazekii*. *J Hyg Epidemiol Microbiol Immunol* **8:** 36–38.

Balayeva NM, Nikolskaya VN. 1972. Enhanced virulence of the vaccine strain E of *Rickettsia prowazeki* on passaging in white mice and guinea pigs. *Acta Virol* **16:** 80–82.

Balayeva NM, Nikolskaya VN. 1973. Analysis of lung culture of *Rickettsia prowazedi* E strain with regard to its capacity of increasing virulence in passages on the lungs of white mice. *J Hyg Epidemiol Microbiol Immunol* **17:** 294–303.

Balayeva NM, Eremeeva ME, Ignatovich VF, Dmitriev BA, Lapina EB, Belousova LS. 1992. Protein antigens of genetically related *Rickettsia prowazekii* strains with different virulence. *Acta Virol* **36:** 52–56.

Baranov PV, Hammer AW, Zhou J, Gesteland RF, Atkins JF. 2005. Transcriptional slippage in bacteria: Distribution in sequenced genomes and utilization in IS element gene expression. *Genome Biol* **6:** R25. doi: 10.1186/gb-2005-6-3-r25.

Bechah Y, Capo C, Grau GE, Raoult D, Mege JL. 2007. A murine model of infection with *Rickettsia prowazekii*: Implications for pathogenesis of epidemic typhus. *Microbes Infect* **9:** 898–906.

Bechah Y, Capo C, Mege JL, Raoult D. 2008a. Epidemic typhus. *Lancet Infect Dis* **8:** 417–426.

Bechah Y, Capo C, Raoult D, Mege JL. 2008b. Infection of endothelial cells with virulent *Rickettsia prowazekii* increases the transmigration of leukocytes. *J Infect Dis* **197:** 142–147.

Birg ML, La Scola B, Roux V, Brouqui P, Raoult D. 1999. Isolation of *Rickettsia prowazekii* from blood by shell vial cell culture. *J Clin Microbiol* **37:** 3722–3724.

Blanc G, Ngwamidiba M, Ogata H, Fournier PE, Claverie JM, Raoult D. 2005. Molecular evolution of *Rickettsia* surface antigens: Evidence of positive selection. *Mol Biol Evol* **22:** 2073–2083.

Blanc G, Ogata H, Robert C, Audic S, Suhre K, Vestris G, Claverie JM, Raoult D. 2007. Reductive genome evolution from the mother of Rickettsia. *PLoS Genet* **3:** e14. doi: 10.1371/journal.pgen.0030014.

Caporale LH. 2003. Natural selection and the emergence of a mutation phenotype: An update of the evolutionary synthesis considering mechanisms that affect genome variation. *Annu Rev Microbiol* **57:** 467–485.

Chao CC, Wu SL, Ching WM. 2004. Using LC-MS with de novo software to fully characterize the multiple methylations of lysine residues in a recombinant fragment of an outer membrane protein from a virulent strain of *Rickettsia prowazekii*. *Biochim Biophys Acta* **1702:** 145–152.

Ching WM, Wang H, Davis J, Dasch GA. 1993. Amino acid analysis and multiple methylation of lysine residues in the surface protein antigen of *Rickettsia prowazekii*. In *Techniques in Protein Chemistry* (ed. R. Angeletti), pp. 307–314. Academic Press, San Diego, CA.

Chow KH, Courcelle J. 2004. RecO acts with RecF and RecR to protect and maintain replication forks blocked by UV-induced DNA damage in *Escherichia coli*. *J Biol Chem* **279:** 3492–3496.

Comeron JM. 1999. K-Estimator: Calculation of the number of nucleotide substitutions per site and the confidence intervals. *Bioinformatics* **15:** 763–764.

Dale C, Wang B, Moran N, Ochman H. 2003. Loss of DNA recombinational repair enzymes in the initial stages of genome degeneration. *Mol Biol Evol* **20:** 1188–1194.

Fox JP, Everritt MG, Robinson TA, Conwell DP. 1954. Immunization of man against epidemic typhus by infection with avirulent *Rickettsia prowazeki* (strain E); observations as to post-vaccination reactions, the relation of serologic response to size and route of infecting dose, and the resistance to challenge with virulent typhus strains. *Am J Hyg* **59:** 74–88.

Fox JP, Jordan ME, Gelfand HM. 1957. Immunization of man against epidemic typhus by infection with avirulent *Rickettsia prowazeki* strain E. IV. Persistence of immunity and a note as to differing complement-fixation antigen requirements in post-infection and post-vaccination sera. *J Immunol* **79:** 348–354.

Gambrill MR, Wisseman CL Jr. 1973. Mechanisms of immunity in typhus infections. I. Multiplication of typhus rickettsiae in human macrophage cell cultures in the nonimmune system: Influence of virulence of rickettsial strains and of chloramphenicol. *Infect Immun* **8:** 519–527.

Ge Y, Rikihisa Y. 2006. *Anaplasma phagocytophilum* delays spontaneous human neutrophil apoptosis by modulation of multiple apoptotic pathways. *Cell Microbiol* **8:** 1406–1416.

Ge H, Chuang YY, Zhao S, Tong M, Tsai MH, Temenak JJ, Richards AL, Ching WM. 2004. Comparative genomics of *Rickettsia prowazekii* Madrid E and Breinl strains. *J Bacteriol* **186:** 556–565.

Gordon D, Abajian C, Green P. 1998. *Consed*: A graphical tool for sequence finishing. *Genome Res* **8:** 195–202.

Ignatovich VF. 1975. Enhancement of the antigenic activity and virulence of the vaccine strain E of *Rickettsia prowazeki* by passages in cell culture. *Acta Virol* **19:** 481–485.
Kato H, Takeuchi O, Sato S, Yoneyama M, Yamamoto M, Matsui K, Uematsu S, Jung A, Kawai T, Ishii KJ, et al. 2006. Differential roles of MDA5 and RIG-I helicases in the recognition of RNA viruses. *Nature* **441:** 101–105.
Katoh K, Misawa K, Kuma K, Miyata T. 2002. MAFFT: A novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res* **30:** 3059–3066.
Kowalczewska M, Fenollar F, Lafitte D, Raoult D. 2006. Identification of candidate antigen in Whipple's disease using a serological proteomic approach. *Proteomics* **6:** 3294–3305.
Kumar S, Tamura K, Nei M. 2004. MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. *Brief Bioinform* **5:** 150–163.
La MV, Crapoulet N, Barbry P, Raoult D, Renesto P. 2007. Comparative genomic analysis of *Tropheryma whipplei* strains reveals that diversity among clinical isolates is mainly related to the WiSP proteins. *BMC Genomics* **8:** 349. doi: 10.1186/1471-2164-8-349.
Leber JH, Crimmins GT, Raghavan S, Meyer-Morse NP, Cox JS, Portnoy DA. 2008. Distinct TLR- and NLR-mediated transcriptional responses to an intracellular pathogen. *PLoS Pathog* **4:** e6. doi: 10.1371/journal.ppat.0040006.
Lescot M, Audic S, Robert C, Nguyen TT, Blanc G, Cutler SJ, Wincker P, Couloux A, Claverie JM, Raoult D, et al. 2008. The genome of *Borrelia recurrentis*, the agent of deadly louse-borne relapsing fever, is a degraded subset of tick-borne *Borrelia duttonii*. *PLoS Genet* **4:** e1000185. doi: 10.1371/journal.pgen.1000185.
Ormsbee RA, Peacock MG, Gerloff R, Tallent G, Wike D. 1978. Limits of rickettsial infectivity. *Infect Immun* **19:** 239–245.
Pan X, Luhrmann A, Satoh A, Laskowski-Arce MA, Roy CR. 2008. Ankyrin repeat proteins comprise a diverse family of bacterial type IV effectors. *Science* **320:** 1651–1654.
Pizarro-Cerda J, Cossart P. 2006. Bacterial adhesion and entry into host cells. *Cell* **124:** 715–727.
Polevoda B, Sherman F. 2007. Methylation of proteins involved in translation. *Mol Microbiol* **65:** 590–606.
Raoult D, Roux V. 1999. The body louse as a vector of reemerging human diseases. *Clin Infect Dis* **29:** 888–911.
Raoult D, Ogata H, Audic S, Robert C, Suhre K, Drancourt M, Claverie JM. 2003. *Tropheryma whipplei* twist: A human pathogenic actinobacteria with a reduced genome. *Genome Res* **13:** 1800–1809.
Raoult D, Dutour O, Houhamdi L, Jankauskas R, Fournier PE, Ardagna Y, Drancourt M, Signoli M, La VD, Macia Y, et al. 2006. Evidence for louse-transmitted diseases in soldiers of Napoleon's Grand Army in Vilnius. *J Infect Dis* **193:** 112–120.
Renesto P, Azza S, Dolla A, Fourquet P, Vestris G, Gorvel JP, Raoult D. 2005. *Rickettsia conorii* and *R. prowazekii* proteome analysis by 2DE-MS: A step toward functional analysis of *Rickettsial* genomes. *Ann N Y Acad Sci* **1063:** 90–93.
Renesto P, Samson L, Ogata H, Azza S, Fourquet P, Gorvel JP, Heinzen RA, Raoult D. 2006. Identification of two putative rickettsial adhesins by proteomic analysis. *Res Microbiol* **157:** 605–612.
Rodionov AV, Eremeeva ME, Balayeva NM. 1991. Isolation and partial characterization of the Mr 100 kD protein from *Rickettsia prowazekii* strains of different virulence. *Acta Virol* **35:** 557–565.
Rosenberg SM. 2001. Evolving responsively: Adaptive mutation. *Nat Rev Genet* **2:** 504–515.
Tamas I, Wernegreen JJ, Nystedt B, Kauppinen SN, Darby AC, Gomez-Valero L, Lundin D, Poole AM, Andersson SG. 2008. Endosymbiont gene functions impaired and rescued by polymerase infidelity at poly(A) tracts. *Proc Natl Acad Sci* **105:** 14934–14939.
Tatusov RL, Koonin EV, Lipman DJ. 1997. A genomic perspective on protein families. *Science* **278:** 631–637.
Turco J, Winkler HH. 1982. Differentiation between virulent and avirulent strains of Rickettsia prowazekii by macrophage-like cell lines. *Infect Immun* **35:** 783–791.
Turco J, Winkler HH. 1994. Cytokine sensitivity and methylation of lysine in *Rickettsia prowazekii* EVir and interferon-resistant *R. prowazekii* strains. *Infect Immun* **62:** 3172–3177.
Vallenet D, Labarre L, Rouy Z, Barbe V, Bocs S, Cruveiller S, Lajus A, Pascal G, Scarpelli C, Medigue C. 2006. MaGe: A microbial genome annotation system supported by synteny results. *Nucleic Acids Res* **34:** 53–65.
Zhang JZ, Hao JF, Walker DH, Yu XJ. 2006. A mutation inactivating the methyltransferase gene in avirulent Madrid E strain of *Rickettsia prowazekii* reverted to wild type in the virulent revertant strain Evir. *Vaccine* **24:** 2317–2323.
Zinsser H. 1935. *Rats, lice, and history.* Broadway House, London, UK.

Received November 26, 2009; accepted in revised form February 11, 2010.