# Lateral gene transfer between obligate intracellular bacteria: Evidence from the Rickettsia massiliae genome

Guillaume Blanc, Hiroyuki Ogata, Catherine Robert, Stéphane Audic, Jean-Michel Claverie and Didier Raoult

*Genome Res.* 2007 17: 1657-1664; originally published online Oct 4, 2007;
Access the most recent version at doi:10.1101/gr.6742107



Letter

# Lateral gene transfer between obligate intracellular bacteria: Evidence from the *Rickettsia massiliae* genome

Guillaume Blanc,[1,3] Hiroyuki Ogata,[1] Catherine Robert,[2] Stéphane Audic,[1] Jean-Michel Claverie,[1] and Didier Raoult[2]

[1] *Structural and Genomic Information Laboratory, CNRS-UPR 2589, Institut de Biologie Structurale et Microbiologie, IFR 88, Parc Scientifique de Luminy, 13288 Marseille Cedex 9, France;* [2] *Unité des Rickettsies, CNRS-UMR 6020, IFR 48, Faculté de Médecine, 13385 Marseille Cedex 5, France*

*Rickettsia massiliae* is a tick-borne obligate intracellular α-proteobacteria causing spotted fever in humans. Here, we present the sequence of its genome, comprising a 1.3-Mb circular chromosome and a 15.3-kb plasmid. The chromosome exhibits long-range colinearity with the other Spotted Fever Group *Rickettsia* genomes, except for a large fragment specific to *R. massiliae* that contains 14 *tra* genes presumably involved in pilus formation and conjugal DNA transfer. We demonstrate that the *tra* region was acquired recently by lateral gene transfer (LGT) from a species related to *Rickettsia bellii*. Further analysis of the genomic sequences identifies additional candidates of LGT between *Rickettsia*. Our study indicates that recent LGT between obligate intracellular *Rickettsia* is more common than previously thought.

[Supplemental material is available online at www.genome.org.]

The *Rickettsia* genus, which includes *Rickettsia massiliae*, is a group of arthropod-borne, obligate intracellular alphaproteobacteria (Raoult and Roux 1997). This genus has been classified into three groups, namely, the typhus group (TG), which includes *Rickettsia prowazekii* and *Rickettsia typhi*, the spotted fever group (SFG), which includes *Rickettsia rickettsii* and *Rickettsia conorii*, and the *Rickettsia bellii* group (Stothard et al. 1994; Roux et al. 1997; Blanc et al. 2007). The SFG rickettsiae are further divided in two subgroups: the *Rickettsia conorii* subgroup and the *R. massiliae* subgroup, the latter also including *Rickettsia montanensis, Rickettsia aeschlimannii,* and *Rickettsia rhipicephali* (Roux et al. 1997; Fournier et al. 1998). Natural resistance to rifampin is a phenotypic marker of the *R. massiliae* subgroup (Drancourt and Raoult 1999), whereas rickettsiae belonging to the *R. conorii* subgroup and the TG rickettsiae are naturally susceptible to rifampin (Rolain et al. 1998).

Many SFG rickettsiae infect and multiply in almost all organs of ticks, in particular the salivary glands, which enable the bacteria to be transmitted to vertebrate hosts during feeding (Raoult and Roux 1997). *Rickettsia massiliae* was first isolated from *Rhipicephalus sanguineus* collected in Marseille (France) in 1992 (Beati and Raoult 1993). Since then, this rickettsiae has been commonly detected in *Rhipicephalus* ticks from Central Africa, France, Greece, Mali, Portugal, Spain, Switzerland, and the United States (Matsumoto et al. 2005; Eremeeva et al. 2006). It is also transmitted transovarially, at least in *Rhipicephalus turanicus* (Matsumoto et al. 2005). *R. massiliae* might be commonly associated with these worldwide-distributed ticks. *R. massiliae* has also been recently identified in *Ixodes ricinus* ticks (Fernandez-Soto et al. 2006). At the beginning of the 20th century, SFG rickettsiae were recognized as human pathogens. Although SFG rickettsioses are among the oldest known vector-borne diseases, they are also now recognized as emerging infections of humans (Parola et al. 2005). In Spain where *R. massiliae* is prevalent in ticks, serological studies (Cardenosa et al. 2003) and resistance to rifampin make this bacterium a putative agent of spotted fever. Recently, *R. massiliae* has been reported as a human pathogen (Vitale et al. 2005).

The rickettsiae have comparatively small genomes (1.1–1.5 Mb) that have arisen through a recent and ongoing genome degradation process (Andersson et al. 1998). To date, 10 complete *Rickettsia* genome sequences have been deposited in public databases. They include the genomes of two *R. bellii* strains (GenBank AARC00000000; Ogata et al. 2006), two sequences from the TG rickettsiae (*R. prowazekii* [Andersson et al. 1998] and *R. typhi* [McLeod et al. 2004]), five sequences from the SFG group (*R. conorii* [Ogata et al. 2001], *Rickettsia sibirica* [Malek et al. 2004], *R. rickettsii* [GenBank AADJ00000000], *Rickettsia felis* [Ogata et al. 2005] and *Rickettsia akarii* [GenBank AAFE00000000]), and the sequence of *Rickettsia canadensis* (Eremeeva et al. 2005) for which the phylogenetic position is unclear. Here, we report the sequencing and the primary analysis of the genome of *R. massiliae* strain MTU5 isolated from the *R. turanicus* tick collected on horses in Camargues, France. Recently, our group used seven *Rickettsia* species (including *R. massiliae*) in a large-scale comparative genomic study to investigate the evolutionary processes that shaped the chromosomes of the genus (Blanc et al. 2007). In particular, our results suggested a lack of recent horizontal gene transfer between *Rickettsia* species and other bacteria. In the present analysis, a different approach allowed us to pinpoint horizontal gene transfers between *Rickettsia* species that were not discovered previously.

## Results

### Highlights of the *R. massiliae* genome

The genome of *R. massiliae* is made of two replicons: a 1,360,898-bp circular chromosome, and a 15,286-bp circular plasmid. After

[3] **Corresponding author.**
**E-mail guillaume.blanc@igs.cnrs-mrs.fr; fax 33-4-91-82-54-21.**
Article published online before print. Article and publication date are at http://www.genome.org/cgi/doi/10.1101/gr.6742107.

*R. felis* (Ogata et al. 2005), *R. massiliae* is the second sequenced Rickettsiales harboring a plasmid. We detected 968 protein genes on the chromosome as well as 212 pseudogenes. Only 26 gene products, including 21 transposases, exhibited no reciprocal best-BLAST hit in the other *Rickettsia* species (BLASTP *E*-value $< 1 \times 10^{-5}$). The plasmid is predicted to contain 12 protein genes and a pseudogene, among which seven have homologs on the *R. felis* plasmid and two others are related to transposable elements. Thus, the *R. massiliae* genome appears to encode only a limited number of novel functions with respect to the other *Rickettsia*. A more detailed description of the genomic features of *R. massiliae* is presented in the Supplemental Material.

As expected from their close relatedness (identity >98% in coding sequence), the *R. massiliae* and *R. conorii* chromosomes exhibit a high level of colinearity (Fig. 1). However, a particular 54.6-kb segment retained our attention because it is absent in *R. conorii*. This region contains 44 genes, including 14 *tra*-related genes that probably encode components of a type IV secretion system (T4SS) for conjugal DNA transfer (Ogata et al. 2005, 2006). The relatively high number of specific transposase genes and pseudogenes (36 sequences) suggests that *R. massiliae* underwent a recent expansion of transposons, a phenomenon also encountered in the chromosomes of *R. bellii* (39 sequences) and *R. felis* (66 sequences) but not in the other sequenced *Rickettsia* genomes. Interestingly, the *R. massiliae*, *R. bellii*, and *R. felis* genomes contain several *tra* genes. The recently released genome of *Orientia tsutsugamushi* (Cho et al. 2007), a remote relative of *Rickettsia* species, exhibits also very high numbers of transposase (>400 genes) and *tra* genes (359 genes). Thus, there may be a link between the relative abundance of transposase genes in the *Rickettsiale* genomes and an active process of conjugation enabling the acquisition of foreign DNA.

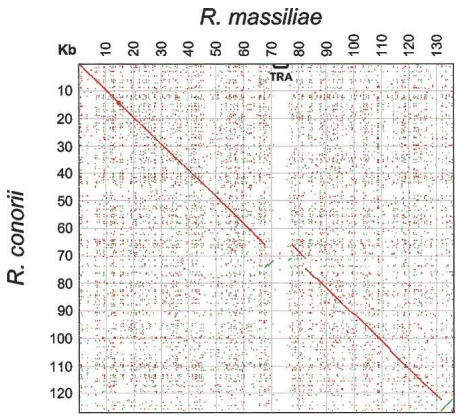


**Figure 1.** Dot-plot comparison of the *R. massiliae* and *R. conorii* genomes. Red and green dots represent forward and reverse BLASTN matches, respectively (*E*-value value $< 1 \times 10^{-5}$). Position of the *tra* cluster is indicated in gray.

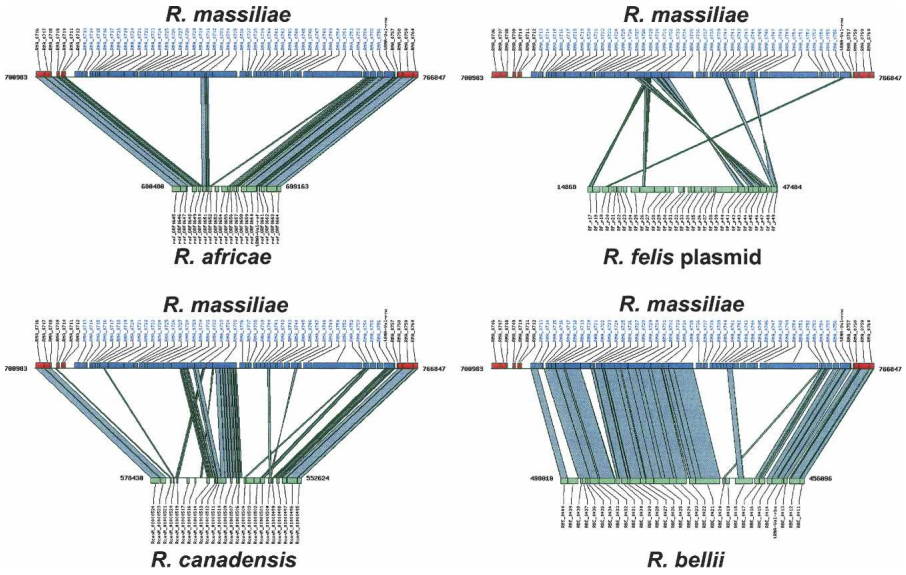


**Figure 2.** Alignments of the *R. massiliae tra* region with the corresponding regions in four *Rickettsia* species. Genomic coordinates are indicated on both sides of the segments. Blue and red rectangles represent the ORFs inside and outside of the *R. massiliae tra* cluster, respectively. Green rectangles show predicted ORFs in the compared species. Genomic segments are connected by gray shading when significant nucleotide similarity is detected using BLASTN (*E*-value $< 1 \times 10^{-5}$).

## The *tra* gene cluster

The *R. massiliae tra* region is most similar to that of *R. bellii* in terms of gene order and content (Fig. 2). The two segments share 25 ORFs, including the entire set of 14 *tra*-related genes and 11 ORFs encoding five unknown proteins; a site-specific DNA adenine methylase, dam-like; an integrase; a signal transduction histidine, kinase-like; and a spoT family protein (Table 1). In addition, both species contain specific genes: the *R. massiliae tra* cluster contains 21 extra genes (genes for 13 transposases, four unknown proteins, a toxin of a toxin–antitoxin system, a phage-associated protein, a protein similar to the DNA polymerase III alpha chain, and an archaeal ATPase family protein), and *R. bellii* contains three specific genes (genes for a leucine-rich protein, a cassette chromosome recombinase B, and *spoT16*). The *tra* genes are absent in the other sequenced *Rickettsia* chromosomes except for *R. canadensis* and *Rickettsia africae*, which contain small subsets of the *R. massiliae tra* cluster genes (Fig. 2; Table 1). The genes surrounding the *tra* region are conserved in sequence and order in most *Rickettsia* species, even in those that do not possess a *tra* cluster region (Supplemental Figs. S1–S7).

In both *R. massiliae* and *R. bellii*, the *tra* cluster is flanked by an integrase gene (tyrosine recombinase family) and a *tRNA-Val* gene. Many integrases specify integration of prokaryotic genetic elements into tRNA or tmRNA genes (Williams 2002). To avoid deactivation of the recipient gene (e.g., tRNA), the foreign genetic element usually carries at its integration site the same segment of gene as the one displaced. After integration, the host tRNA is then split into two pieces lying at each extremity of the integrated genetic element (see Fig. 1 in Williams 2002 for details). Consistent with such an integration process, we identified a 51-bp sequence highly similar to the 3′-extremity of the *tRNA-Val*, located 90 bp upstream from the *R. massiliae* integrase gene (coordinates 707,321–707,371). In *R. bellii*, the 40-bp long displaced *tRNA-Val* segment lies 395 bp upstream from the integrase gene (coordinates: 498,401–498,440). Thus, the *tra* regions in *R.*

**Table 1.** Genes of the *R. massiliae tra* cluster region and their homologs in other *Rickettsia*

| *R. massiliae*[a] | *R. bellii* | *R. africae*[a,b] | *R. canadensis*[a] | *R. felis* plasmid[a] | *R. felis*[a] | *R. rickettsii* | *R. akarii* | Predicted function |
|---|---|---|---|---|---|---|---|---|
| RMA_0713 | RBE_0440 | | | | | | | Integrase |
| RMA_0714 | | | | | | | | Phage-associated protein |
| RMA_0715 | | | | | | | | Toxin of TA system |
| | RBE_0439 | | | | | | | Leucine-rich repeat protein |
| RMA_0716 | RBE_0438 | | | | | | | Unknown |
| RMA_0717 | RBE_0437 | | | | | | | Conjugal transfer protein $TraE_F$ |
| RMA_0718 | RBE_0436 | | | | | | | Unknown |
| RMA_0719 | RBE_0435 | | | | | | | Unknown |
| RMA_0720 | RBE_0434 | | RcanM_01000520 | | | | | F pilus assembly protein $TraB_F$ |
| RMA_0721 | RBE_0433 | | | | | | | Conjugal transfer protein $TraV_F$ |
| RMA_0722 | RBE_0432 | | | | | | | Conjugal transfer protein $TraC_F$ |
| RMA_0723 | | | | | | | | Archaeal ATPase family protein |
| | | | RcanM_01000516 | | | | | Unknown |
| RMA_0724 | RBE_0431 | | | | | | | Conjugal transfer protein $TraW_F$ |
| RMA_0725 | | | | | | | | |
| RMA_0726 | RBE_0430 | | | | | | | Conjugal transfer protein $TraU_F$ |
| RMA_0727 | RBE_0429 | | | | | | | Conjugal transfer protein $TrbC_F$ |
| RMA_0728 | RBE_0428 | | | | | | | Conjugal transfer protein $TraN_F$ |
| RMA_0729 | RBE_0427 | | | | | | | F pilus assembly protein $TraF_F$ |
| RMA_0730 | RBE_0426 | | | | | | | F pilus assembly protein $TraH_F$ |
| RMA_0731 | RBE_0425 | | RcanM_01000514 | pRF46 pRF47 | | | | Conjugal transfer protein $TraG_F$ |
| RMA_0732 | RBE_0424 | | RcanM_01000512 | pRF26 pRF17 | | | | Tetratricopeptide |
| | | | RcanM_01000513 | | | | | repeat-containing protein |
| RMA_0733 | RBE_0423 | raf_ORF0650 | RcanM_01000511 | pRF43 pRF44 | | | | Conjugal transfer protein $TraD_F$ |
| | | | RcanM_01000519 | | | | | |
| RMA_0734 | RBE_0422 | raf_ORF0651 | | pRF41 | | | | Transposase |
| RMA_0735 | RBE_0421 | | RcanM_01000506 | pRF38 pRF39 | | | | Conjugal transfer protein $TraA_{Ti}$ |
| | | | RcanM_01000507 | | | | | |
| | | | RcanM_01000509 | | | | | |
| | | | RcanM_01000510 | | | | | |
| RMA_0736 | | | | | | | | Transposase |
| RMA_0737 | | | | | | | | Transposase |
| RMA_0738 | | | | | | | | Transposase |
| RMA_0739 | | | | | | | | Unknown |
| RMA_0740 | | | | | | | | Transposase |
| RMA_0741 | | | | | | | | Transposase |
| RMA_0742 | | | | | | | | Transposase |
| RMA_0743 | | | RcanM_01000500 | | | | | Transposase |
| RMA_0744 | RBE_0418 | | | | | | | Unknown |
| RMA_0745 | | | | | | | | Transposase |
| RMA_0746 | | | | | | | | Transposase |
| RMA_0747 | | | | | | | | Transposase |
| RMA_0748 | | | RcanM_01000500 | pRF49 | | | | Transposase |
| RMA_0749 | | | | | | | | Transposase |
| RMA_0750 | | | | | | | | Transposase |
| RMA_0751 | | | | | | | | DNA polymerase III alpha chain |
| RMA_0752 | RBE_0420 | | RcanM_01000504 | pRF37 | | | | Conjugal transfer protein $TraD_{Ti}$ |
| | | | RcanM_01000503 | | | | | Proline/betaine transporter |
| | | | RcanM_01000502 | | RF_0385 | | | Unknown |
| | | | | | RF_0386 | | | |
| | | raf_ORF0652 | RcanM_01000501 | | RF_0384 | | | ppGpp hydrolase spoT15 |
| | | | | | RF_0383 | | | Transposase |
| | | | | | RF_0382 | | | Transposase |
| | RBE_0419 | raf_ORF0653 | | | RF_0381 | | | ppGpp synthetase spoT16 |
| | | raf_ORF0654 | | | | | | |
| RMA_0753 | RBE_0417 | raf_ORF0655 | | | RF_0380 | | | ppGpp hydrolase spoT6 |
| | | raf_ORF0656 | | | | | | |
| RMA_0754 | RBE_0416 | raf_ORF0657 | | | | | | Signal transduction |
| | | raf_ORF0658 | | | | | | histidine kinase |
| RMA_0755 | RBE_0415 | raf_ORF0659 | | | | | | |
| RMA_0756 | RBE_0414 | raf_ORF0660 | RcanM_01000498 | pRF19 | | Rick02000935 | Pseudogene | Unknown (*U* gene) |
| | | | RcanM_01000499 | | | | | |
| tRNAVal | tRNAVal | tRNAVal | tRNAVal | | | tRNAVal | tRNAVal | Anticodon: GAC |

[a]Genes comprising several ORFs in a species correspond to split pseudogenes.
[b]The *R. africae* gene annotation is from Blanc et al. (2007) and is available at http://www.igs.cnrs-mrs.fr/mgdb/Rickettsia/rig/.

*massiliae* and *R. bellii* probably result from the integration of a site-specific integrative and conjugative element (Burrus et al. 2002). The organization of the rickettsial *tra* regions is very similar to that of the *tra* clusters of *O. tsutsugamushi* (Cho et al. 2007), including *tra*-associated genes such as transposase, *spoT*, Histidine Kinase, and *Dam*. Overall, 29 of the 44 *R. massiliae tra* cluster genes have homologs in the *O. tsutsugamushi* genome. This points to a probable common origin of those genetic elements.

The nucleotide sequence immediately downstream of the recombined *R. massiliae tRNA-Val* gene (toward the RMA_0756 gene) is part of the original conjugative element that gave birth to the chromosomal *tra* cluster. We searched for similar sequences in the downstream region of the *tRNA-Val* gene in all sequenced *Rickettsia* genomes. Significant matches (BLASTN *E*-value $< 1 \times 10^{-5}$) were identified not only for *R. bellii*, *R. africae*, and *R. canadensis* (as expected) but also for *R. rickettsii* and *R. akarii*, over regions that extend into the first gene (RMA_0756; referred to hereafter as the *U* gene) of the *R. massiliae* conjugative element (Supplemental Fig. S8). Thus, these two latter species appear to contain traces of ancient integrated genetic elements, at least partly related to that of *R. massiliae*. Out of the *Rickettsia* genus, homologs of the *U* gene are only found in the *O. tsutsugamushi* genome (35 copies, BLASTP *E*-value $< 10^{-10}$). In contrast, homologs are present in the *R. massiliae* (RMA_p01) and *R. felis* (pRF19) plasmids and in the *R. bellii* (RBE_1035) and the *R. prowazekii* (pseudogene between *RP708* and *RP709*) genomes. The *R. prowazekii U* pseudogene is located in a 12-kb region specific to *R. prowazekii* (McLeod et al. 2004), which also includes several transposase (*RP710* and *RP715*) and recombinase pseudogenes (*RP711*) that are uniquely found on both the *R. felis* and *R. massiliae* plasmids (Supplemental Fig. S9). In addition, this region contains a fragment of a patatin-like protein gene mostly similar to the *pat2* gene of the *R. felis* plasmid. This region might also result from the integration of a genetic element or plasmid. We also note that homologs of *spoT6*, *spoT16*, *spoT15*, and the unknown protein gene RcanM01000502 (*R. canadensis*) are organized in cluster in the *R. felis* genome, in the same order as in the consensus structure of the *Rickettsia tra* regions (Table 1). However, they lie at a location remote from the *tRNA-Val* gene, suggesting that they do not originate from the same genetic element. Thus, this analysis supports the view that integrated genetic elements are relatively widespread in *Rickettsia*, at least in the form of traces.

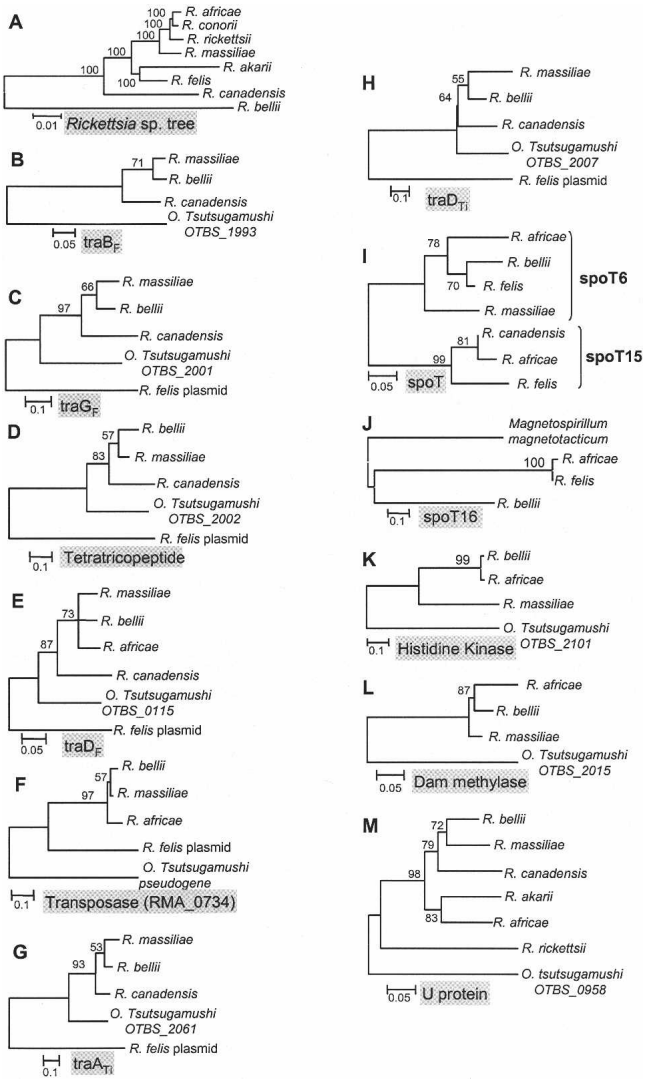


**Figure 3.** Phylogenetic relationships of the *tra* cluster genes. (*A–M*) Phylogenetic trees are constructed from protein sequences using the maximum-likelihood method. Models of substitution that best fit the data are as follows: JTT for *D* and *M*; JTT + Γ for *C*, *E*, and *J*; JTT + I for *K*; WAG + Γ for *G*; WAG + Γ + I for *A*; CpREV for *B*, *F*, and *L*; CpREV + Γ for *I*; RtREV for *H*. The *Rickettsia* species tree (*A*) is constructed from the concatenated alignment of 38 ribosomal proteins (6209 sites). Gene names and branch length scales (number of substitution per site) are shown *below* trees. Bootstrap support values, calculated over 500 alignment replicates, are given *beside* branches. Consecutive ORFs of pseudogenes (Table 1) were concatenated for the analysis. In addition to the *Rickettsia* sequences, the *O. tsutsugamushi* best BLASTP hit was included in the phylogenetic reconstructions when available (*E*-value $< 1 \times 10^{-10}$).

## Phylogenetic study of the *tra* cluster genes

A previous study of the *R. bellii* genome showed that the closest homologs of the *tra* genes outside the *Rickettsia* genus were found in microbes living inside amoebae, including *Legionella pneumophila* and *Protochlamydia amoebophila* (Ogata et al. 2006). Here we present the phylogenetic relationships of the *tra*-cluster genes among the *Rickettsia* species. Phylogenetic reconstructions were conducted for 13 homologous gene families for which sequences were present in at least three distinct *Rickettsia* species and with sufficient overlap (protein alignment >75 sites). The *R. felis* plasmid proteins were generally difficult to align with their rickettsial homologs, suggesting that the plasmid sequences are the most distantly related. This is confirmed in six phylogenetic reconstructions where plasmid proteins are separated by the longest branches from the rest of the rickettsial sequences (Fig. 3C–H). In all cases but in the transposase tree (Fig. 3F), the plasmid proteins are more distant from their rickettsial chromosomal counterparts than are the *O. tsutsugamushi* homologs.

In seven phylogenetic trees, the *R. massiliae* proteins are the closest homologs of the *R. bellii* sequences (Fig. 3C–H,M). This phylogenetic affinity is consistent with the high level of synteny between the two regions but in disagreement with the species phylogeny (Fig. 3A), in which *R. bellii* is the earliest diverging species and *R. massiliae* is anchored within the spotted fever group (together with *R. africae* and *R. rickettsii*). Exceptions to this rule are spoT6 (Fig. 3I), histidine kinase (Fig. 3K), and dam methylase

(Fig. 3L), for which the *R. bellii* sequences are more closely related to *R. africae* or *R. felis* than *R. massiliae*. This phylogenetic proximity is also in disagreement with the species tree. The *U* gene exhibits even more phylogenetic incongruence: the SFG rickettsiae (*R. africae*, *R. massiliae* and *R. rickettsii*) fall in three distinct lineages (Fig. 3M). The position of *R. canadensis* with respect to *R. massiliae* and *R. bellii* is always consistent across the phylogenetic trees of the *tra* proteins and the U protein, in which it stands out as a sister group (Fig. 3B–E,G,H,M).

The striking picture emerging from this analysis is that none of the *tra* cluster proteins seems to present an evolutionary history compatible with a strict vertical inheritance. Only two protein families (histidine kinase and *U* gene) formally rejected ($\alpha = 5\%$) the species tree topology using the Shimodaira–Hasegawa (SH) test (Shimodaira and Hasegawa 1999) (Supplemental Table S1). The poor discrimination of the SH tests between alternative topologies is likely the consequence of the short length of some alignments and/or the lack of strong phylogenetic signal. The departure from the species phylogeny implies that the *tra*-regions were subject to horizontal transfer events between rickettsiae, but other less likely scenarios such as a complex interplay of duplication/differential loss can also be invoked. In this respect, the phylogenetic relationships between the *U* genes (Fig. 3M) are particularly interesting because they suggest that the genetic elements (or traces of thereof) result from different integration events, many of which occurred in the same *tRNA-Val* gene due to the integrase site specificity. Furthermore, the different genes did not produce consistent phylogenies in spite of their collinear organization. However, we note that the phylogenies of the *tra* genes and the transposase (Fig. 3B–H) that lie in the 5′ part of the *tra*-cluster are consistent between each other (albeit with some low bootstrap supports), whereas the genes corresponding to the 3′ extremity of the *tra*-region (*spoT*, histidine kinase, *dam*, and U protein; Fig. 3I–M) exhibit heterogeneous evolutionary histories. This suggests that the original genetic elements result from the assembly of gene modules with different evolutionary trajectories, a phenomenon frequently observed in conjugative genetic elements (Burrus et al. 2002).

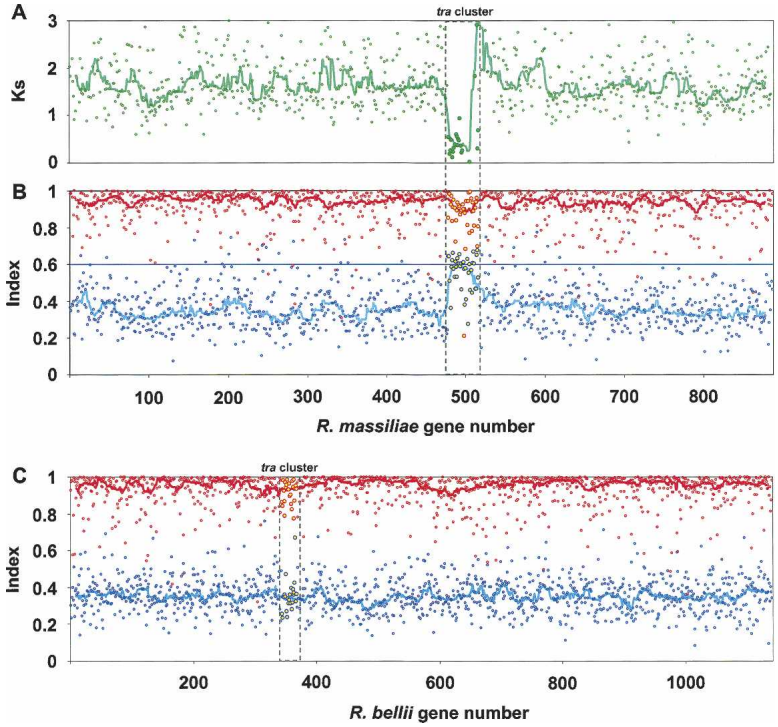


**Figure 4.** Horizontal transfer of the *tra* cluster region. (*A*) $K_s$ values estimated between *R. massiliae* and *R. bellii* reciprocal best-match genes. (*B*) Horizontal transfer index (HTI) and horizontal transfer donor index (HTDI) of the *R. massiliae* genes (red and blue symbols, respectively). (*C*) HTI and HTDI of the *R. bellii* genes. Gene pairs are ordered on the *X*-axis according to the *R. massiliae* (*A* and *B*) or *R. bellii* (*C*) gene orders. Solid circles show the data for genes of the *tra* region. Curves represent moving medians calculated from 20-gene sliding windows. The color of the curve is the same as that of the data points that were used for the calculations (i.e., $K_s$, HTI, HTDI).

## Horizontal transfer between *R. massiliae* and *R. bellii*

We determined the level of synonymous substitutions ($K_s$) between pairs of putative orthologous genes in *R. bellii* and *R. massiliae*. All but three gene pairs falling in the *tra* cluster exhibit a smaller level of divergence ($K_s$ values from 0.01 to 0.69) than the gene pairs in the rest of the genome (overall median $K_s$ value of 1.58; Fig. 4A). This indicates that those *tra* cluster genes were not vertically inherited from the *R. bellii* and *R. massiliae* common ancestor but rather diverged more recently. One gene pair ($traD_{Ti}$) has $K_s = 1.89$, a figure compatible with the overall median $K_s$ of true orthologs. The two remaining genes (histidine kinase and *spoT6* genes) have relatively high $K_s$ values ($K_s = 2.90$ and 5.57, respectively) and are responsible for the highest peak of the median $K_s$ curve. These two genes probably experienced different evolutionary trajectories than the rest of the *tra* cluster, as already evidenced in their respective phylogenetic reconstructions (Fig. 3, cf. I and K). These sequences diverged seemingly earlier than the split between *R. bellii* and *R. massiliae*.

To further investigate the origin of the *tra* cluster genes, we analyzed the nucleotide composition of the *R. massiliae* and *R. bellii* genes using the Bayesian method developed by Nakamura et al. (2004). The *R. massiliae tra* cluster genes have collectively a lower horizontal transfer index (HTI), causing a small drop in the HTI median curve (Fig. 4B, red curve). This indicates that these genes exhibit a poorer fit to the average genome composition. Conversely, the horizontal transfer donor index (HTDI) median curve calculated using the *R. bellii* composition model exhibit a conspicuous peak centered on the *tra* region (Fig. 4B, blue curve). This indicates that those genes have a better fit with the average *R. bellii* gene nucleotide composition than does the rest of the genome. In contrast, the *R. bellii tra* cluster genes present neither sensible deviation from the average *R. bellii* nucleotide composition (HTI curve; Fig. 4C) nor a better fit to the *R. massiliae* model (HTDI curve; Fig. 4C). Altogether, these data strongly suggest that a recent *R. massiliae* ancestor has acquired the *tra* gene cluster through horizontal transfer from another *Rickettsia* species closely related to *R. bellii*.

### The *tra* regions are free from RPE

The *R. massiliae* genome contains 574 identified repeated elements (see Supplemental Material), most (98%) of which belong to the repeated palindromic element (RPE) families (Ogata et al. 2001; Claverie and Ogata 2003). All *Rickettsia* species sequenced so far are littered with RPE sequences throughout their genomes. In *R. massiliae*, there is on average one repeat unit every 2.4 kb, distributed largely homogeneously across the intergenic spacers of the genome and in some functional genes (RPEs constitute most red and green dots out of the main diagonal in Fig. 1). Intriguingly, the *tra* cluster region is located within by far the largest (63.5 kb) RPE-free segment of the genome (the second largest is 15.8 kb). This segment appears as an empty vertical stripe in Figure 1. In *R. bellii*, the *tra* cluster is also part of the largest (45.3 kb) RPE-free region (average density is 1 repeat every 2.9 kb). The *R. africae* and *R. canadensis tra* regions also contain no RPEs. The lack of RPEs is unlikely to result from strong selective pressures that forced the *tra* region to remain RPE-free, as most of the genes are not a priori involved in vital biological processes. Furthermore, the intergenic regions where most of the RPEs normally lie are not smaller in the *tra* region than in the rest of the genome. Thus the absence of RPEs suggests that the primary rickettsial *tra* gene cluster was most likely acquired after the RPE proliferation in *Rickettsia*.

## Discussion

Although *R. massiliae* is one of the earliest diverging species of the SFG rickettsiae, its genome is very similar to the genomes of other SFG rickettsiae, in terms of sequence and gene order. Nevertheless, our analysis unveiled interesting new features, especially the origin of the *tra* cluster first observed in *R. bellii* (Ogata et al. 2006) and the existence of another rickettsial plasmid. The discovery of the *tra* cluster in *R. bellii* was followed by the identification of sex pili-like appendages, supporting the hypothesis that the *tra* genes are indeed involved in conjugal DNA transfer machinery. Those appendages are also present at the surface of *R. massiliae* cells (Supplemental Fig. S10).

As much as horizontal gene transfer (HGT) is a major contributor of genome evolution in free-living bacteria, the absence of (evidence of) HGT has long been considered a fundamental characteristic of genome evolution in obligate intracellular bacteria. The small genomes and isolated lifestyles of obligate intracellular bacteria are thought to reduce the opportunities for gene acquisition. It has also been hypothesized that genetic exchanges are minored in these species because their relatively constant intracellular environment does not select for the genetic diversity promoted by more challenging environments (Moreno 1998). These early views are now progressively changing with the accumulation of genomic data, which provide evidence for a better appraisal of the nature and extent of genetic exchange in intracellular bacteria (Bordenstein and Reznikoff 2005).

The *Rickettsia* genus has been shown to contain different families of mobile elements, such as plasmids (Ogata et al. 2005), transposases, and phage-related genes (Andersson et al. 1998; Ogata et al. 2001, 2005, 2006; McLeod et al. 2004). While a substantial number of HGT was inferred between the early ancestors of Rickettsiales and amoebal parasites (Ogata et al. 2006), our previous study did not identify clear evidence for more recent HGT events in the *Rickettsia* genomes (Blanc et al. 2007). However, this study was principally tailored for identifying HGT between *Rickettsia* sp. and species outside the *Rickettsia* genus. Thus, HGT involving *Rickettsia* species both as donor and recipient were overlooked. In the present study, we demonstrate that *R. massiliae* received a 54.6-kb genetic element containing 44 genes from a species related to *R. bellii*. Furthermore, we also present evidence that other rickettsiae contain traces of genetic elements related to the *R. massiliae tra* cluster or rickettsiae plasmids. Our analysis suggests that these elements were acquired through distinct integration events and horizontal transfers between *Rickettsia* species (Fig. 3). Thus the *R. massiliae tra* cluster may be a member of a larger family of rickettsial conjugative elements that integrate specifically into the *tRNA-Val* gene owing to the integrase site specificity. We also note that the 12-kb *R. prowazekii*-specific DNA region (McLeod et al. 2004) contains genes that are otherwise only found in rickettsial plasmids. This region results probably from the recent (i.e., after the separation with *R. typhi*) integration of a rickettsial plasmid into the genome. Other studies have also presented evidence of evolutionary events related to HGT: the replacement of the ancestral *R. felis* and *R. akari pat1* gene by a sequence closely related to the *R. felis* plasmid *pat2* gene (Blanc et al. 2005) and recombination of *sca* genes between different *Rickettsia* species (Jiggins 2006).

HGT in obligate intracellular bacteria is not limited to the *Rickettsia* genus, although evidence remains scarce in comparison to free-living bacteria. The *Buchnera aphicolada* genomes, estimated to have diverged 50–70 million years ago, exhibit identical genomic architectures and no indication of HGT (Tamas et al. 2002). This genomic stasis is thought to result from the loss of DNA repair and recombinase function (Moran and Wernegreen 2000; Dale et al. 2003). Nevertheless, evidence indicate that the *Buchnera leuABCD* plasmid might experience rare horizontal transmission (Van Ham et al. 2000) and genetic exchange with the *Buchnera* chromosome (Sabater-Munoz et al. 2004). *Wolbachia* is a genus of intracellular bacteria phylogenetically related to *Rickettsia*. The genome sequence from *Wolbachia pipientis* wMel is littered with repetitive and mobile DNA sequences, including numbers of prophage families (Wu et al. 2004). One of these prophage, WO-B, can be horizontally transferred between different *Wolbachia* subgroups when they co-infect the same host (Masui et al. 2000; Bordenstein and Wernegreen 2004; Baldo et al. 2006). Recently, the *O. tsutsugamushi* genome revealed an astonishing number of mobile sequences, including >400 transposase genes and >24 *tra* gene clusters resulting from intragenomic duplications and/or multiple integrations (Cho et al. 2007).

Evidence suggests that the Rickettsiales ancestor lived in amoeba-like ancestral protozoa, which promoted genetic exchanges with other amoebal parasites and the adaptation to intracellular life in eukaryotes (Ogata et al. 2006). The horizontal transfer of the tra-region occurred between two lineages whose extant representatives live in ticks. As far as the life styles of *R. massiliae* and *R. bellii* can be extended to their recent ancestors, the exchange of the *tra* cluster probably occurred between bacteria that co-infected the same tick. Preliminary studies show that *R. bellii* can coexist with other rickettsiae in ticks in the wild (D. Raoult, unpubl.). Thus, the pattern of identifiable HGT in *Rickettsia* validates the "intracellular arena" hypothesis (Bordenstein and Wernegreen 2004), which posits that genetic material can move in and out of communities of obligate intracellular bacteria that co-infect the same intracellular host environment.

Conjugative elements are present in a variety of Gram-positive and -negative bacteria and probably contribute as much as plasmids to the spread of antibiotic resistance genes in some

genera of disease-causing bacteria (Salyers et al. 1995). In the case of *Rickettsia*, the *tra* cluster includes members of the most prominent multigene families, such as the spoT (hydrolase domain), proline/betaine transporter, and toxin–antitoxin genes, which may have important functions with regard to the interaction with the eukaryotic host (Ogata et al. 2005). The *spoT* and proline/betaine transporter genes are often found as adjacent gene pairs in the *Rickettsia* genomes as well as in the consensus structure of the rickettsial *tra* region (Table 1). Toxin and antitoxin genes are also customarily organized in tandem. These gene pairs might have been hitchhiked between co-infecting *Rickettsia* species via ancient rickettsial genetic elements (for which few traces remain in the genomes today), thus promoting the adaptation to the host.

## Methods

### Bacterial purification and DNA extraction

*Rickettsia massiliae* (strain MTU5) was cultivated on L929 cells growing on EMEM (Biowhittaker, Cambrex Biosciences) supplemented with 2 mM glutamine and 4% fetal calf serum (GIBCO, Invitrogen) in 150-cm$^2$ tissue culture flasks at 32°C. The cultures were first harvested and treated by trypsin to a final concentration of 1% followed by a digestion with DNase I. The final purification was performed by two Renografin density gradient centrifugations. The pellet of purified *R. massiliae* was again digested three times by DNase I. Prokaryotic DNA was extracted according to the classical lytic treatment using SDS and proteinase K followed by phenol–chloroform–isoamyl alcohol extraction; DNA was resuspended in TE buffer and visualized on agarose gel stained by ethidium bromide.

### Shotgun sequencing of *R. massiliae* genome

Genomic DNA was fragmented by mechanical shearing using a Hydroshear device (GeneMachine). A preliminary library (3-kb inserts cloned in pcDNA 2.1 with BstXI adaptors) was constructed and validated. The percentage of eukaryotic DNA contamination was calculated from the sequence on 96 clones and was <7%. We generated another library with 10-kb inserts cloned in the pCNS vector (C. Robert, unpubl.). Plasmid clones were sequenced at both ends of the insert with flanking vector sequences as primers. Dye-terminator reactions were analyzed on a capillary ABI3730 (Applied Biosystems). The whole-genome sequence assembly was performed by means of the PHRED and PHRAP software packages. Cloning gaps and sensitive regions were resolved on subclones templates or confirmed by sequencing PCR products; 579 primers were involved in the finishing. The shotgun sequence assembly revealed a plasmid of ~16 kb. Its presence was confirmed by Southern blot. The sequences have been deposited in the GenBank database (accession nos. CP000683 and CP000684).

### Sequence analysis

We predicted protein coding genes (ORFs) using SelfID as previously described (Ogata et al. 2005). tRNA genes were identified using tRNAscan-SE (Malek et al. 2004). Database searches were performed using BLAST programs against SWISS-PROT/TrEMBL (Apweiler et al. 2004), NCBI/CDD database (Marchler-Bauer et al. 2005), and SMART (Ponting et al. 1999). Representation of the BLAST results, as in Figure 2, was performed using an in-house Perl script available upon request from the author. Repeated DNA sequences were identified using RepeatFinder (Volfovsky et al. 2001). To identify *Rickettsia* palindromic elements, we used hidden Markov models (Eddy 1996) based on previously identified RPE sequences.

We used MUSCLE (Edgar 2004) to perform multiple sequence alignments. Maximum-likelihood phylogenetic analyses were conducted using PHYML (Guindon and Gascuel 2003) after manual edition of the alignments to remove poorly aligned regions and sites containing gaps. We used the ProtTest program to select the substitution models that best describe the data using the AIC criterion (Abascal et al. 2005). We applied the Shimodaira–Hasegawa (SH) test (Shimodaira and Hasegawa 1999) as implemented in the TREE-PUZZLE program (Schmidt et al. 2002) to verify whether the maximum-likelihood tree topologies were significantly better at explaining the data than the expected species-tree topology. The true phylogenetic position of the *R. felis* plasmid sequences is unknown and can change from gene to gene because of the composite nature of plasmids. Thus, we omitted the *R. felis* sequence in the SH test. $K_s$ values were estimated from pairs of codon-aligned nucleotide sequences using the CODEML program (Yang 1997). HTI and HTDI values for genes were determined using a computer program developed by Nakamura et al. (2004): Markov chain models for nucleotide composition were constructed for the *R. massiliae* and *R. bellii* genomes and then used to calculate the average posterior probability (i.e., index) for each gene sequence to be compatible with the rest of the genome. Low horizontal index values indicate poor fit with the tested genome. We used a Markov order of 5 with a 96-bp window sliding with a step size of 12 bp.

## Acknowledgments

We thank Prof. Y. Nakamura for providing us with the computer program that computes the horizontal gene transfer index.

## References

Abascal, F., Zardoya, R., and Posada, D. 2005. ProtTest: Selection of best-fit models of protein evolution. *Bioinformatics* **21:** 2104–2105.

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

Apweiler, R., Bairoch, A., Wu, C.H., Barker, W.C., Boeckmann, B., Ferro, S., Gasteiger, E., Huang, H., Lopez, R., Magrane, M., et al. 2004. UniProt: The Universal Protein knowledgebase. *Nucleic Acids Res.* **32:** D115–D119. doi: 10.1093/nar/gkh131.

Baldo, L., Bordenstein, S., Wernegreen, J.J., and Werren, J.H. 2006. Widespread recombination throughout *Wolbachia* genomes. *Mol. Biol. Evol.* **23:** 437–449.

Beati, L. and Raoult, D. 1993. *Rickettsia massiliae* sp. nov., a new spotted fever group *Rickettsia*. *Int. J. Syst. Evol. Microbiol.* **43:** 839–840.

Blanc, G., Renesto, P., and Raoult, D. 2005. Phylogenic analysis of rickettsial Patatin-like protein with conserved phospholipase $A_2$ active sites. *Ann. N.Y. Acad. Sci.* **1063:** 83–86.

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

Bordenstein, S.R. and Reznikoff, W.S. 2005. Mobile DNA in obligate intracellular bacteria. *Nat. Rev. Microbiol.* **3:** 688–699.

Bordenstein, S.R. and Wernegreen, J.J. 2004. Bacteriophage flux in endosymbionts (*Wolbachia*): Infection frequency, lateral transfer, and recombination rates. *Mol. Biol. Evol.* **21:** 1981–1991.

Burrus, V., Pavlovic, G., Decaris, B., and Guedon, G. 2002. Conjugative transposons: The tip of the iceberg. *Mol. Microbiol.* **46:** 601–610.

Cardenosa, N., Segura, F., and Raoult, D. 2003. Serosurvey among Mediterranean spotted fever patients of a new spotted fever group rickettsial strain (Bar29). *Eur. J. Epidemiol.* **18:** 351–356.

Cho, N.-H., Kim, H.-R., Lee, J.-H., Kim, S.-Y., Kim, J., Cha, S., Kim, S.-Y., Darby, A.C., Fuxelius, H.-H., Yin, J., et al. 2007. The *Orientia tsutsugamushi* genome reveals massive proliferation of conjugative

type IV secretion system and host-cell interaction genes. *Proc. Natl. Acad. Sci.* **104:** 7981–7986.
Claverie, J.M. and Ogata, H. 2003. The insertion of palindromic repeats in the evolution of proteins. *Trends Biochem. Sci.* **28:** 75–80.
Dale, C., Wang, B., Moran, N., and Ochman, H. 2003. Loss of DNA recombinational repair enzymes in the initial stages of genome degeneration. *Mol. Biol. Evol.* **20:** 1188–1194.
Drancourt, M. and Raoult, D. 1999. Characterization of mutations in the *rpoB* gene in naturally rifampin-resistant *Rickettsia* species. *Antimicrob. Agents Chemother.* **43:** 2400–2403.
Eddy, S.R. 1996. Hidden Markov models. *Curr. Opin. Struct. Biol.* **6:** 361–365.
Edgar, R.C. 2004. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32:** 1792–1797. doi: 10.1093/nar/gkh340.
Eremeeva, M.E., Madan, A., Shaw, C.D., Tang, K., and Dasch, G.A. 2005. New perspectives on rickettsial evolution from new genome sequences of *Rickettsia*, particularly *R. canadensis*, and *Orientia tsutsugamushi*. *Ann. N.Y. Acad. Sci.* **1063:** 47–63.
Eremeeva, M.E., Bosserman, E.A., Demma, L.J., Zambrano, M.L., Blau, D.M., and Dasch, G.A. 2006. Isolation and identification of *Rickettsia massiliae* from *Rhipicephalus sanguineus* ticks collected in Arizona. *Appl. Environ. Microbiol.* **72:** 5569–5577.
Fernandez-Soto, P., Perez-Sanchez, R., Diaz Martin, V., Encinas-Grandes, A., and Alamo Sanz, R. 2006. *Rickettsia massiliae* in ticks removed from humans in Castilla y Leon, Spain. *Eur. J. Clin. Microbiol. Infect. Dis.* **25:** 811–813.
Fournier, P.E., Roux, V., and Raoult, D. 1998. Phylogenetic analysis of spotted fever group rickettsiae by study of the outer surface protein rOmpA. *Int. J. Syst. Evol. Microbiol.* **48:** 839–849.
Guindon, S. and Gascuel, O. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* **52:** 696–704.
Jiggins, F.M. 2006. Adaptive evolution and recombination of *Rickettsia* antigens. *J. Mol. Evol.* **62:** 99–110.
Malek, J.A., Wierzbowski, J.M., Tao, W., Bosak, S.A., Saranga, D.J., Doucette-Stamm, L., Smith, D.R., McEwan, P.J., and McKernan, K.J. 2004. Protein interaction mapping on a functional shotgun sequence of *Rickettsia sibirica*. *Nucleic Acids Res.* **32:** 1059–1064. doi: 10.1093/nar/gkh254.
Marchler-Bauer, A., Anderson, J.B., Cherukuri, P.F., DeWeese-Scott, C., Geer, L.Y., Gwadz, M., He, S., Hurwitz, D.I., Jackson, J.D., Ke, Z., et al. 2005. CDD: A conserved domain database for protein classification. *Nucleic Acids Res.* **33:** D192–D196. doi: 10.1093/nar/gki069.
Masui, S., Kamoda, S., Sasaki, T., and Ishikawa, H. 2000. Distribution and evolution of bacteriophage WO in *Wolbachia*, the endosymbiont causing sexual alterations in arthropods. *J. Mol. Evol.* **51:** 491–497.
Matsumoto, K., Ogawa, M., Brouqui, P., Raoult, D., and Parola, P. 2005. Transmission of *Rickettsia massiliae* in the tick, *Rhipicephalus turanicus*. *Med. Vet. Entomol.* **19:** 263–270.
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., et al. 2004. Complete genome sequence of *Rickettsia typhi* and comparison with sequences of other rickettsiae. *J. Bacteriol.* **186:** 5842–5855.
Moran, N.A. and Wernegreen, J.J. 2000. Lifestyle evolution in symbiotic bacteria: insights from genomics. *Trends Ecol. Evol.* **15:** 321–326.
Moreno, E. 1998. Genome evolution within the alpha Proteobacteria: Why do some bacteria not possess plasmids and others exhibit more than one different chromosome? *FEMS Microbiol. Rev.* **22:** 255–275.
Nakamura, Y., Itoh, T., Matsuda, H., and Gojobori, T. 2004. Biased biological functions of horizontally transferred genes in prokaryotic genomes. *Nat. Genet.* **36:** 760–766.
Ogata, H., Audic, S., Renesto-Audiffren, P., Fournier, P.-E., Barbe, V., Samson, D., Roux, V., Cossart, P., Weissenbach, J., Claverie, J.-M., et al. 2001. Mechanisms of evolution in *Rickettsia conorii* and *R. prowazekii*. *Science* **293:** 2093–2098.
Ogata, H., Renesto, P., Audic, S., Robert, C., Blanc, G., Fournier, P.E., Parinello, H., Claverie, J.M., and Raoult, D. 2005. The genome sequence of *Rickettsia felis* identifies the first putative conjugative plasmid in an obligate intracellular parasite. *PLoS Biol.* **3:** e248. doi: 10.1371/journal.pbio.0030248.
Ogata, H., La Scola, B., Audic, S., Renesto, P., Blanc, G., Robert, C., Fournier, P.E., Claverie, J.M., and Raoult, D. 2006. Genome sequence of *Rickettsia bellii* illuminates the role of amoebae in gene exchanges between intracellular pathogens. *PLoS Genet.* **2:** e76. doi: 10.1371/journal.pgen.0020076.
Parola, P., Paddock, C.D., and Raoult, D. 2005. Tick-borne rickettsioses around the world: Emerging diseases challenging old concepts. *Clin. Microbiol. Rev.* **18:** 719–756.
Ponting, C.P., Schultz, J., Milpetz, F., and Bork, P. 1999. SMART: Identification and annotation of domains from signalling and extracellular protein sequences. *Nucleic Acids Res.* **27:** 229–232.
Raoult, D. and Roux, V. 1997. Rickettsioses as paradigms of new or emerging infectious diseases. *Clin. Microbiol. Rev.* **10:** 694–719.
Rolain, J.M., Maurin, M., Vestris, G., and Raoult, D. 1998. In vitro susceptibilities of 27 Rickettsiae to 13 antimicrobials. *Antimicrob. Agents Chemother.* **42:** 1537–1541.
Roux, V., Rydkina, E., Eremeeva, M., and Raoult, D. 1997. Citrate synthase gene comparison, a new tool for phylogenetic analysis, and its application for the rickettsiae. *Int. J. Syst. Bacteriol.* **47:** 252–261.
Sabater-Munoz, B., van Ham, R.C.H.J., Moya, A., Silva, F.J., and Latorre, A. 2004. Evolution of the leucine gene cluster in *Buchnera aphidicola*: Insights from chromosomal versions of the cluster. *J. Bacteriol.* **186:** 2646–2654.
Salyers, A.A., Shoemaker, N.B., Stevens, A.M., and Li, L.Y. 1995. Conjugative transposons: An unusual and diverse set of integrated gene transfer elements. *Microbiol. Rev.* **59:** 579–590.
Schmidt, H.A., Strimmer, K., Vingron, M., and von Haeseler, A. 2002. TREE-PUZZLE: Maximum likelihood phylogenetic analysis using quartets and parallel computing. *Bioinformatics* **18:** 502–504.
Shimodaira, H. and Hasegawa, M. 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Mol. Biol. Evol.* **16:** 1114–1116.
Stothard, D.R., Clark, J.B., and 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. Bacteriol.* **44:** 798–804.
Tamas, I., Klasson, L., Canback, B., Naslund, A.K., Eriksson, A.S., Wernegreen, J.J., Sandstrom, J.P., Moran, N.A., and Andersson, S.G. 2002. 50 Million years of genomic stasis in endosymbiotic bacteria. *Science* **296:** 2376–2379.
Van Ham, R.C., Gonzalez-Candelas, F., Silva, F.J., Sabater, B., Moya, A., and Latorre, A. 2000. Postsymbiotic plasmid acquisition and evolution of the repA1-replicon in *Buchnera aphidicola*. *Proc. Natl. Acad. Sci.* **97:** 10855–10860.
Vitale, G., Mansueto, S., Rolain, J.M., and Raoult, D. 2005. *Rickettsia massiliae* human isolation. *Emerg. Infect. Dis.* **12:** 174–175.
Volfovsky, N., Haas, B.J., and Salzberg, S.L. 2001. A clustering method for repeat analysis in DNA sequences. *Genome Biol.* **2:** RESEARCH0027. doi: 10.1186/gb-2001-2-8-research0027.
Williams, K.P. 2002. Integration sites for genetic elements in prokaryotic tRNA and tmRNA genes: Sublocation preference of integrase subfamilies. *Nucleic Acids Res.* **30:** 866–875.
Wu, M., Sun, L.V., Vamathevan, J., Riegler, M., Deboy, R., Brownlie, J.C., McGraw, E.A., Martin, W., Esser, C., Ahmadinejad, N., et al. 2004. Phylogenomics of the reproductive parasite *Wolbachia pipientis* wMel: A streamlined genome overrun by mobile genetic elements. *PLoS Biol.* **2:** e69. doi: 10.1371/journal.pbio.0020069.
Yang, Z. 1997. PAML: A program package for phylogenetic analysis by maximum likelihood. *Comput. Appl. Biosci.* **13:** 555–556.

*Received May 25, 2007; accepted in revised form August 29, 2007.*