



**HAL**  
open science

# The Rhizome of Life: The Sympatric *Rickettsia felis* Paradigm Demonstrates the Random Transfer of DNA Sequences

Vicky Merhej, Cedric Notredame, Manuela Royer-Carenzi, Pierre Pontarotti, Didier Raoult

► **To cite this version:**

Vicky Merhej, Cedric Notredame, Manuela Royer-Carenzi, Pierre Pontarotti, Didier Raoult. The Rhizome of Life: The Sympatric *Rickettsia felis* Paradigm Demonstrates the Random Transfer of DNA Sequences. *Molecular Biology and Evolution*, 2011, 28 (11), pp.3213-3223. 10.1093/molbev/msr239 . hal-01105200

**HAL Id: hal-01105200**

**<https://hal.science/hal-01105200>**

Submitted on 16 Feb 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# The Rhizome of Life: The Sympatric *Rickettsia felis* Paradigm Demonstrates the Random Transfer of DNA Sequences

Vicky Merhej,<sup>1</sup> Cedric Notredame,<sup>2</sup> Manuela Royer-Carenzi,<sup>3</sup> Pierre Pontarotti,<sup>3</sup> and Didier Raoult<sup>\*1</sup>

<sup>1</sup>Unité de Recherche en Maladies Infectieuses et Tropicales Emergentes (URMITE) CNRS-IRD UMR6236-198, Université de la Méditerranée, Faculté de Médecine, Marseille, France

<sup>2</sup>Centre for Genomic Regulation, Pompeu Fabra University, Barcelona, Spain

<sup>3</sup>Evolution Biologique et Modélisation UMR6632 CNRS IFR48, Université Aix-Marseille, Marseille, France

\*Corresponding author: E-mail: didier.raoult@gmail.com.

Associate editor: Helen Piontkivska

## Abstract

The intracellular flea symbiont, *Rickettsia felis*, may meet other organisms intracellularly such as *R. typhi*. We used a single-gene phylogenetic approach of the 1375 *R. felis* genes to look for horizontal transfers that occurred as a result of the bacterial promiscuity with other organisms. Our results showed that besides genes that are linked to the Spotted Fever Group, 165 genes have a different history and are linked to other *Rickettsia* such as *R. bellii* (107 genes), *R. typhi* (15 genes), or to other bacteria such as *Legionella* sp. and *Francisella* sp. or to eukaryotes. Among these genes, we identified 73 individual genes and 34 spatial clusters containing 2–4 adjacent genes, a total of 79 genes, with evidence of en bloc transfer. We described 13 chimeric genes resulting from gene recombination with sympatric *R. typhi*. The transferred DNA sequences present different sizes and functions, suggesting that the horizontal transfer in *R. felis* is random and neutral within its specific host. Our study shows that the strict intracellular bacteria *R. felis* exhibits a mosaic genome. We therefore developed a new representation for the evolutionary history of *R. felis* showing its different putative ancestors in the form of a rhizome.

**Key words:** bacterial rhizome, sympatry, *Rickettsia felis*, gene recombination, horizontal gene transfer.

## Introduction

According to the Darwinian principle of “descent with modification through natural selection,” all species arise and develop through the natural selection of small, inherited variations that increase an individual’s ability to compete, survive, and reproduce. Organismal descents have been represented as a tree-like process (Darwin 1859). Later on, biologists have developed genetic methods to map representatives from all groups of living organisms to a tree referred to as the “tree of life” (Woese 1987). In this framework, all living organisms are considered to be descended from a common ancestor; nonvertical transfers are incorporated in this model, but they are considered to be rare enough not to challenge the notion of a core organism that effectively represents the whole organism in the tree of life. This simple vision has been challenged considerably by the emergence of large-scale sequencing, and many approaches have emerged (Doolittle and Baptiste 2007) that propose replacing the tree-like pattern of evolutionary relationships with more complex models such as the “synthesis of life” (Baptiste et al. 2004), the “forest” of evolutionary trees (Puigbo et al. 2009; Schliep et al. 2010), the genetic network (Puigbo et al. 2010; Popa et al. 2011), and the “rhizome of life” (Raoult 2010). Instead of the dichotomous topology of a phylogenetic tree, the different roots of a rhizome model illustrate the multiple sources of the genetic information of bacteria, with inheritance occurring not only vertically but also laterally. Indeed, comparative genomics approaches have revealed that horizontal

gene transfer (HGT) is a driving force in the innovation and evolution of bacterial genomes (Ochman et al. 2000; Jain et al. 2003). Bacteria evolve not only through the modification of inherited genetic information but also through the acquisition of sequences from other organisms. These horizontal transfer events account for a significant fraction of single-gene phylogenies incongruences observed when doing phylogenomic surveys (Doolittle 1999a; Lerat et al. 2003; Santoyo and Romero 2005; Puigbo et al. 2009).

*Rickettsia* spp. are obligate intracellular  $\alpha$ -proteobacteria that are small and rod shaped. They are best known as agents for vector-borne diseases in humans (Raoult and Roux 1997). The genus *Rickettsia* has traditionally been subdivided into three major groups on the basis of molecular genetic methods: 1) the Typhus Group (TG), which includes the etiological agents of epidemic and murine typhus, *Rickettsia prowazekii* and *R. typhi*, respectively, 2) the Spotted Fever Group (SFG), which includes the etiological agents of Mediterranean spotted fever (*R. conorii*), Rocky Mountain spotted fever (*R. rickettsii*), African tick bite fever (*R. africae*), flea-borne spotted fever (*R. felis*), and rickettsial pox (*R. akari*) (Merhej and Raoult 2010) and 3) the “Ancestral Group,” containing *R. bellii* and *R. canadensis* that are associated with ticks (Stothard et al. 1994; Stothard and Fuerst 1995). Later, *R. akari* and *R. felis* were placed in a separate group called the “Transitional Group” because they present phenotypic and genotypic characteristics that are intermediate of the SFG and the TG (Gillespie et al. 2007).

Nevertheless, single- and multigene analyses have often yielded conflicting tree topologies with many ambiguities, especially with regard to the phylogenetic position of *R. canadensis* and the phylogenetic relationships within the SFG. Advances in molecular tools and biological approaches have uncovered other groups that include species with a broad host range and diverse effects on the host (Perlman et al. 2006; Weinert et al. 2009; Merhej and Raoult 2010). *Rickettsiae* with unknown pathogenic effects such as *R. bellii* seems to form a distinct group with other species that act as symbionts for different hosts including insects (Perlman et al. 2006; Merhej and Raoult 2010). The phylogeny from core gene concatenation of the 11 available sequenced genomes has produced a robust phylogeny that showed 4 groups in the *Rickettsia* genus: SFG, TG, *R. bellii* group, and *R. canadensis* group (Blanc, Ogata, Robert, Audic, Suhre, et al. 2007; Fournier et al. 2009; Merhej et al. 2009). A polyphasic taxonomic approach helped to define four different subgroups within SFG: the *R. rickettsii* subgroup (same as *R. conorii* subgroup), the *R. massiliae* subgroup, the *R. helvetica* subgroup, and the *R. akari* subgroup containing *R. akari* and *R. felis* (Roux and Raoult 1999) (supplementary file 2, fig. S1, Supplementary Material online).

Evidence for HGT events has been presented for *R. bellii* (Ogata et al. 2006) and *R. massiliae* (Blanc, Ogata, Robert, Audic, Claverie, and Raoult 2007), and amoebae have been suggested as a site of this genetic exchange. Recently, these findings have led to reconsider the intracellular environment of bacteria with specific attention to the type of speciation according to an allopatric or sympatric model of evolution (Moliner et al. 2010; Boyer et al. 2011). As defined for eukaryotes, the allopatric speciation assumes that reproductive isolation between former populations of a species evolves as a consequence of physical isolation through an extrinsic barrier (Mayr 1942). While in sympatric speciation, the reproductive isolation between former populations of a species evolves without any extrinsic barrier (Johannesson 2001). Indeed, the obligate intracellular habitat of *Rickettsia* contributes to the physical isolation of these bacteria and subsequently to an allopatric evolutionary lifestyle with limited rate of gene exchange. Nevertheless, the host of an obligate intracellular may harbor different organisms that share the same habitat, leading to a sympatric lifestyle. Thereby the intracellular habitat of amoeba may offer a considerable opportunity for gene exchange, leading to a sympatric evolution of rickettsial species, like *R. bellii*, that have the capacity to live within it (Ogata et al. 2006; Boyer et al. 2009; Moliner et al. 2010). Likewise, insects and other rickettsial hosts that can be coinfecting by different intracellular bacteria may represent a melting pot for genetic exchange (Thomas and Greub 2010).

*Rickettsia felis* has a reduced genome of 1.48 Mb that contains a considerable proportion of mobile genetic elements termed the “mobilome” (Smets and Barkay 2005), including 2 plasmids, 61 genes encoding transposases, and 8 phage-related proteins (Ogata et al. 2005). Moreover, in arthropods, *R. felis* exhibits a sympatric lifestyle and may coexist with *R. typhi* in fleas (Eremeeva et al. 2008). Altogether, these features suggest the possibility for HGT and

gene recombination (Frost et al. 2005), probably facilitated by coinfection of *R. felis* hosts. The genome organization of *R. felis* is very close to other rickettsia (with the exception of *R. bellii*), and we suspected that most gene exchange was carried out through recombination rather than transposition. We examined the origin of every gene in the genome of *R. felis* using a single-gene phylogenetic approach. When evidence was found for HGT, we studied the neighboring sequences to evaluate en bloc transfer. Furthermore, we carried out an in-depth study of gene recombination between *R. felis* and the sympatric *R. typhi*. We also described the mosaic gene repertoire of *R. felis*, including genes shared by all or some *Rickettsia* spp., genes with no homologs in *Rickettsia* and genes with no significant similarities in the NCBI database “ORFans.” Finally, we determined the possible origin of all *R. felis* genes, including those that were vertically inherited, duplicated, converted and horizontally acquired.

## Materials and Methods

### Data

At the beginning of the study, the sequences of 11 rickettsial genomes were available. The rickettsial genomes were retrieved from the Genbank database. The species used were *R. prowazekii* strain Madrid E (NC\_000963) (Andersson et al. 1998), *R. typhi* strain Wilmington (NC\_006142) (McLeod et al. 2004), *R. conorii* strain Malish 7 (NC\_003103) (Ogata et al. 2000, 2001), *R. felis* URRWXC2 (NC\_007109) (Ogata et al. 2005), *R. akari* strain Hartford (NC\_009881), *R. rickettsii* strain ‘Sheila Smith’ (NC\_009882) (Ellison et al. 2008), *R. africae* (NC\_012633) (Fournier et al. 2009), *R. massiliae* MTU5 (NC\_009900) (Blanc et al. 2007), *R. bellii* RML369-C (NC\_007940) (Ogata et al. 2006), *R. canadensis* strain MacKiel (NC\_009879), and *R. slovacica* (CP002428).

### Search for Orthologs and Homologs of *R. felis* Open-Reading Frames

Orthologous relationships between the 11 rickettsial proteomes were determined using all-against-all reciprocal best-hit (RBH) analysis (Tatusov et al. 1997; Bork et al. 1998). All pairwise sequence comparisons among the proteins encoded in the 11 complete genomes were performed, and for each protein, the best hit (BeT) in each of the other genomes was detected. We set an *E* value cutoff of equal to or less than  $10^{-5}$  and required that the size of high scoring pairs be at least 60% of the query size (coverage). In the BeT graph thus defined, we looked for highly connected components in which each sequence is the RBH of at least two other sequences (Tatusov et al. 2003). Sequences belonging to such a cluster are said to belong to the same cluster of orthologous groups of proteins (COGs). COGs that were detected in all of the 11 genomes formed the core genes of the studied *Rickettsia* spp. In the case of two copies of orthologous genes in the same genome, we included the open-reading frames (ORFs) with the best bit score in the corresponding cluster of orthology.

Sequence similarity searches were performed with BlastP at the National Center for Biotechnology Information

(Bethesda, MD; <http://blast.ncbi.nlm.nih.gov/Blast.cgi>). The 100 best matches with an *E* value less than  $10^{-4}$  and sequence similarity higher than 30% were selected for phylogenetic analysis. ORFs for which we identified no similar sequences in the nonredundant (NR) protein sequence database were compared with a metagenomic database using BlastP software. Conserved domain search was performed with Reverse PSI-Blast (RPS-Blast; Marchler-Bauer et al. 2003) (<http://www.ncbi.nlm.nih.gov/Structure/cdd/>). The coverage of RPS-Blast alignments between the query sequences and conserved protein domains was used to infer putative domains to the protein query.

### Phylogenetic Analysis

Individual gene sequences were aligned using the MUSCLE algorithm (Edgar 2004), using the nucleotide data set of the core gene and the amino acid data set for the other genes. Phylogenetic analyses were performed using the neighbor-joining (NJ) (Saitou and Nei 1987) and maximum parsimony (MP) methods within PHYLIP package version 3.5c (Felsenstein 1993) and maximum likelihood (ML) within PHYML (Guindon and Gascuel 2003) with 100 bootstrap replicates. The NJ and ML were calculated with a variety of implemented substitution matrices (Kimura, F84 and HKY for nucleotides and JTT and WAG for amino acid sequences). We compared the topologies of the phylogenetic trees inferred from the core genes. Therefore, the approximately unbiased (AU) and Shimodaira–Hasegawa (SH) topology tests as implemented in CONSEL (Shimodaira and Hasegawa 2001; Shimodaira 2002) were used to compare ORF and core-gene trees. Finally, Cochran's test (Cochran 1950) was used to evaluate the significance of differences among gene tree topologies; the significance level was set to 5%. Moreover, the incongruence between tree topologies was estimated using the Robinson–Foulds metric (Robinson and Foulds 1981), which consists of counting the number of bipartitions (splits) that differ between two trees. We used the T-coffee implementation for this measure that makes it possible to efficiently compare and enumerate topologies (Notredame et al. 2000). This approach made it possible to identify the most common topologies and the most frequent splits observed across the dataset.

### Recombination Analysis

We hypothesized that aberrant gene tree topologies showing *R. felis* outside the SFG, especially those showing it within the TG, could be the result of horizontal transfer events through transposition or intergenomic recombination, and the corresponding genes were subjected to further analysis. We conducted a pairwise genome comparison of *R. felis* versus *R. conorii* and *R. akari*; indeed, *R. felis* genes with no corresponding partners in the SFG spp. might be highly suggestive of horizontal transfer through transposition. Recombination was analyzed using the RDP, GENECONV, BOOTSCAN, MAXCHI, CHIMAERA, SISCAN, and 3SEQ methods implemented in the program RDP3BETA30 (Martin et al. 2010). For all the gene trees showing *R. felis* in a phylogenetic clade

with TG, we checked manually for recombination hot spots. Therefore, we examined the gene alignments of *R. felis*, *R. conorii*, *R. akari*, and *R. typhi* looking for significant clusters of polymorphic sites. The polymorphic sites consist of regions (of more than 10 nucleotides) that were more similar between *R. felis* and *R. typhi* than between *R. felis* and *R. conorii* (or between *R. felis* and *R. akari*). The detected regions with possible gene recombination were analyzed using an alternative to McNemar's test to compare paired proportions: the Liddell's exact test (Liddell 1983). *P* values < 0.05 were considered significant.

### Statistical Analysis of Transferred Sequences

All data were analyzed with the R statistical software package [<http://cran.r-project.org/>] using chi-squared test for independence and the unpaired two-tailed Welch's *t*-test for samples with size of 30 or more per group. *P* values < 0.05 were considered significant.

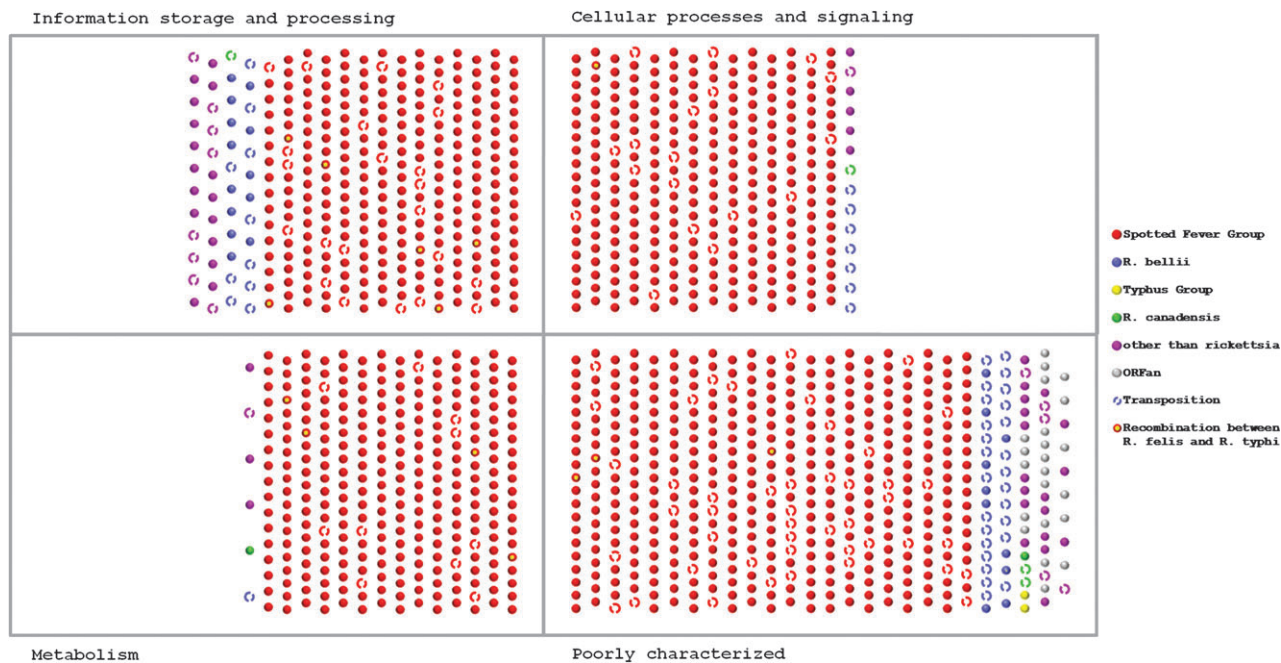
## Results

### Classification of *R. felis* Genes

*Rickettsia felis* has a genome of 1.48 Mb with a considerable proportion of mobile genetic elements (Smets and Barkay 2005) including plasmids, genes encoding transposases, and phage-related proteins (Ogata et al. 2005). We compared the 1,375 chromosomal protein-coding genes in *R. felis* to the NR database using BlastP software.

There were 1,299 ORFs (94.5%) that exhibited significant sequence similarity with other *Rickettsia* spp. (i.e., their BlastP *E* value was lower than  $10^{-4}$ , and they had more than 30% identity). Of these, 1,225 ORFs exhibited homologs in *Rickettsia* from the SFG, whereas 74 ORFs had no homolog in the SFG. Of these 74 ORFs, 65 had homologs in *R. bellii*. There were 76 ORFs in *R. felis* with no homolog in the *Rickettsia* spp. studied here (supplementary file 1, Supplementary Material online). Among these ORFs, 16 matched preferentially with sequences from  $\beta$ -proteobacteria, including the plasmid of *Burkholderia phymatum*, 9 with  $\alpha$ -proteobacteria, 9 with  $\gamma$ -proteobacteria, 5 with the Bacteroidetes/Chlorobi group, and 3 with  $\delta$ -proteobacteria and with sequences from Actinobacteria, Cyanobacteria, Spirochetes, Firmicutes, uncultured bacteria, archae, and Eukaryota. The 76 ORFs with no homolog in the *Rickettsia* spp. mainly encode genes with poorly characterized functions and transposases. These ORFs may have been lost by the other *Rickettsia* spp. or may have been acquired through HGT. This HGT hypothesis is supported by the fact that these ORFs had homologs in phages and prophages (22 ORFs) and plasmids (26 ORFs). Finally, 22 ORFs had no homologous sequences in the NR database or paralogs in *R. felis* and could be defined as ORFan singletons (Boyer et al. 2010). We found homologs for three of these ORFans (rfeORF0778, rfeORF0822, rfeORF0823) in an environmental metagenomic database, which classified them into a "MetaORFan" category (Boyer et al. 2010).

This complex taxonomic distribution of homologs suggests for *R. felis* a genome with a mosaic structure, made of a complex assembly of genes from putative diverse origins (fig. 1).



**Fig. 1.** Schematic representation of the mosaic structure of the *Rickettsia felis* genome. Each dot represents an ORF. ORFs were classified into the different functional categories of COGs: information storage and processing, cellular processes and signaling, metabolism, and poorly characterized function. The colors of the dots represent significant Blast match results: red for the SFG, yellow for the TG, green for *R. canadensis*, blue for *R. bellii*, purple for ORFs with no significant match with *Rickettsia* spp. other than *R. felis*, and gray for ORFans. HGT events through transposition are shown by disrupted, unfilled dots. Yellow dots with a red circle indicate ORFs that significantly match the SFG but seem to have been subject to recombination with the TG. ORFs were represented in the order of their position within the genome.

### Position of *R. felis* among *Rickettsia* spp.

We constructed single-gene phylogenies of the 701 chromosomal genes shared by the 11 currently available rickettsial genomes, using three construction methods (NJ, MP, and ML) and nucleotide sequences to resolve the phylogenetic position of *R. felis*. These analyses revealed extensive incongruence and produced more than 200 alternative topologies (supplementary file 2, tables S1–S3 and figs. S2–S5, Supplementary Material online). Concerning the placement of *R. felis*, robust phylogenies could be estimated from 494 genes (bootstrap values of the branch supporting *R. felis* clade  $\geq 75\%$ , with at least one construction method), with 476 trees (328 trees with ML method) placing *R. felis* within SFG. Out of these, 267 trees (220 ML trees) supported its presence within a sister group containing *R. akari*. In seven cases (out of which three ML trees), *R. felis* was placed in the TG, in three cases with TG and *R. akari*, in three cases (three ML trees) with *R. bellii*, in two cases (2 ML trees) with *R. canadensis*, in two cases (two ML trees) with *R. canadensis* and *R. bellii*, and in one ML tree with *R. canadensis* and *R. akari* (supplementary file 2, fig. S6, Supplementary Material online). These unconventional albeit robust phylogenies from core genes are highly suggestive of horizontal transfer events. The 207 core gene that did not generate reliable trees may be the result of hidden recombination, see subsequent analyses.

We then constructed phylogenies from the 598 *R. felis* genes common to some *Rickettsia* spp. using amino acid sequences extracted from BlastP matches. These genes placed *R. felis* reliably in 420 trees (corresponding to 326

robust ML trees) and 262 trees (184 ML trees) placed *R. felis* within the SFG, including 73 trees (44 ML trees) placing it in a sister group with *R. akari*. Interestingly, in 107 trees (96 ML trees), *R. felis* was placed in a clade with *R. bellii*; in three ML trees, with TG (supplementary file 2, fig. S7, Supplementary Material online); in four cases (3 ML trees), with *R. canadensis*; in two cases (1 ML tree), with *R. canadensis* and *R. bellii*; and in one NJ tree, with *R. canadensis* and *R. akari* (supplementary file 2, figs. S6 and S8, Supplementary Material online). These differentially distributed genes among *Rickettsia* spp. with unusual topologies may have been from horizontal transfer events; alternatively, they may have been from vertical transfer events but may have been lost from the other SFG species.

Finally, the identifiable gene ancestors of *R. felis* among the *Rickettsia* were mainly from the SFG in the *R. conorii* subgroup (43.5% of the total of 914 robust phylogenies obtained from *R. felis* genes common to all or some *Rickettsia* spp.), the *R. akari* subgroup (37%), the TG (1%), *R. bellii* (12%), and *R. canadensis* (0.1%). Based on this analysis, *R. felis* belongs primarily to the SFG of the *Rickettsia* genus, and the putative transitional group is inconsistent.

### Gene Recombination with the Sympatric Species *R. typhi*

Recombination occurs in sympatric organisms, contributing to their genetic diversity (Iredell et al. 2003; Santoyo and Romero 2005; Viezens and Arvand 2008; Vos and Didelot 2009; Touzain et al. 2010). The promiscuity of some *Rickettsia* spp. within the same host (e.g., *R. felis* and *R. typhi*

within flea (Eremeeva et al. 2008; Bitam et al. 2009) suggests the possibility of gene recombination.

When using the RDP3 program, putative recombination events were found between *Rickettsia* spp. in 420 out of 701 core genes (60%). Interestingly, *R. felis* presented the highest number of recombination events (supplementary file 3, Supplementary Material online). We investigated all the *R. felis* gene trees to verify the plausibility of these recombination events. Among the gene trees that show *R. felis* in a clade with *R. typhi*, we found 13 genes with pairwise regions of similarity between *R. typhi* and *R. felis*, and further statistical analyses identified these regions as potential hot spots for recombination (fig. 2) (supplementary file 3, Supplementary Material online). Four of these genes are involved in metabolic functions; five are involved in information storage and processing; one is involved in cellular processes and signaling, and three encode proteins with unknown functions. The recombined genes are evenly distributed around the genome showing genome-wide recombination. They are located in regions that are rich in repeats and split genes and exhibit a high level of colinearity with SFG species like *R. conorii* and *R. akari*.

Here, we limited our study to recombination between *R. felis* and *R. typhi* due to their known common habitat and found that recombination has played an important role in the genome evolution of these sympatric *Rickettsiae*. Other recombination events seem to have taken place as indicated by the RDP method, but we did not further verify this hypothesis owing to the lack of documentation on microorganisms that coinfect arthropods together with *R. felis*.

### Transferred Genes

Previous studies have provided evidence for HGT between the early ancestors of the *Rickettsiales* and amoebal parasites (Ogata et al. 2006) and for recent HGT events in the *R. massiliae* genome (Blanc et al. 2007). Our phylogenetic analysis of all *R. felis* genes revealed 73 individual genes that seem to have been acquired through HGT. Of these, two genes gave robust phylogenies that showed *R. felis* within the TG. The alignment of these gene sequences did not provide any evidence for gene recombination. Fifty genes displayed highly supported trees placing *R. felis* in a clade with *R. bellii*, including 12 genes that are present only in *R. bellii* and *R. felis*. *Rickettsia felis* was placed in a clade with *R. bellii* and *Francisella* spp. in two gene trees, with *R. bellii* and *Dictyostellium* spp. in one gene tree, with *R. bellii*, *Chthoniobacter* spp., and *Burkholderia* spp. in one gene tree, with *R. bellii* and *R. canadensis* in one gene tree, and with *R. canadensis* in one gene tree. Trees from 4 genes showed *R. felis* in a clade with *Alphaproteobacteria* spp. other than *rickettsia*. *Rickettsia felis* was placed in a clade with *Gammaproteobacteria* spp. (5 gene trees), with *Betaproteobacteria* spp. (1 gene trees), with *Deltaproteobacteria* spp. (1 gene tree), with bacteria from *Bacteroides/Chlorobi* spp. (2 gene trees), and with eukaryotes (2 gene trees) (supplementary files 1 and 2, fig. S9, and supplementary file 4, Supplementary Material online).

Furthermore, our single-gene phylogenetic approach that examined the whole genome, exploring it gene by gene

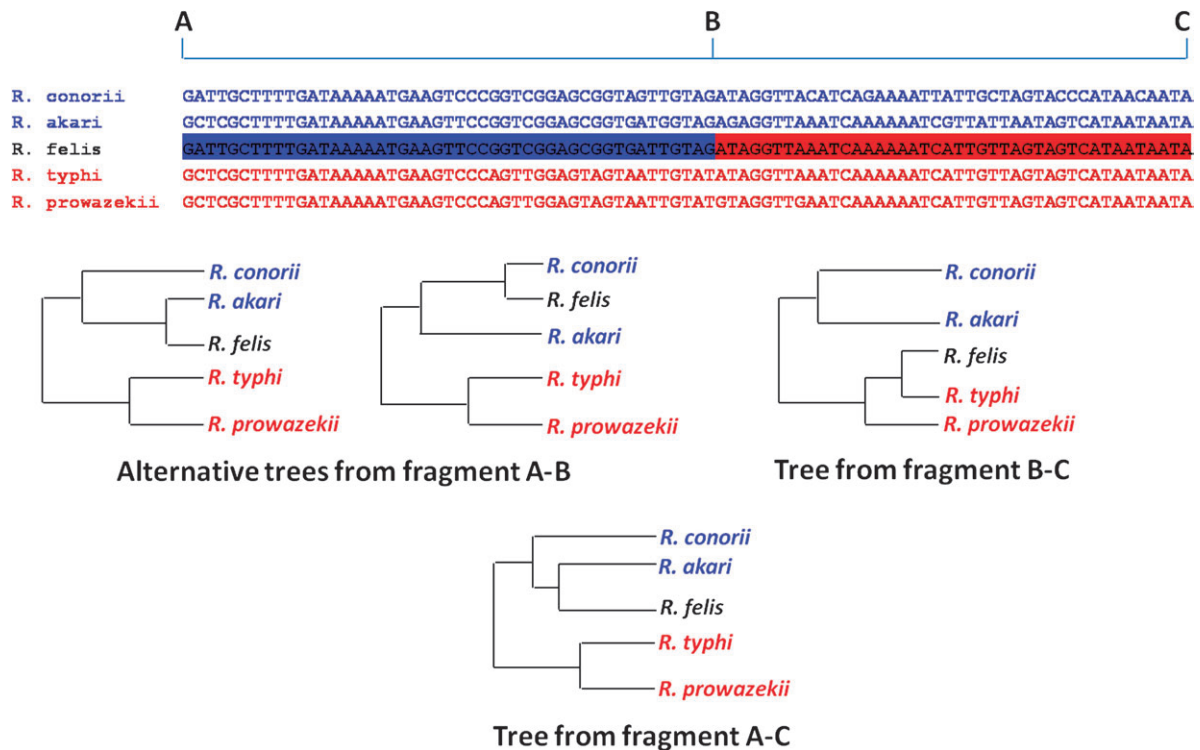
allowed us to locate 79 additional genes that are organized in 34 spatial clusters of 2 to 4 genes, giving similar topologies that were noncongruent with the dominant SFG clade (supplementary file 2, fig. S5, Supplementary Material online). These spatial clusters were distributed throughout the genome; the presence of many transposase sequences, especially in the flanking regions, suggests “en bloc” transmission. These spatial clusters might have been exchanged with other proteobacteria or with amoebae. Indeed, phylogenies from these genes showed *R. felis* in a clade with *R. bellii* (23 spatial clusters); with *R. bellii* and *Gammaproteobacteria* spp. such as *Legionella* spp. or *Francisella* spp. or *Burkholderia* spp. (7 spatial clusters); with *R. canadensis* (1 spatial cluster); with *Alphaproteobacteria* spp. other than *Rickettsia* such as *Wolbachia* spp. and *Orientia* spp. (2 spatial clusters); and with uncultured bacteria (1 spatial cluster) (figs. 3 and 4) (supplementary file 1, fig. S5, and supplementary file 4, Supplementary Material online). As some of these species share the amoebae as a common possible habitat, amoebae may have been the site of the horizontal transfer (Ogata et al. 2006; Thomas and Greub 2010).

### Evidence for Random Exchange

Genome analyses have suggested that the “informational” genes involved in replication, transcription, and translation are more refractory to HGT than “operational” genes involved in metabolism and cellular processes (Jain et al. 1999; Sorek et al. 2007).

Our analysis of the 420 recombined genes in *Rickettsia* spp. showed that 105 out of 190 in the core gene encode for information storage and processing functions, 117 out of 174 in the core gene encode for cellular processes and signaling 128 out of 198 encode for metabolism, and 80 out of 139 encode proteins with poorly characterized functions (9 genes are involved in 2 functions categories). These findings indicated that recombination in *Rickettsia* occurred independently of the gene function ( $P$  value = 0.064, chi-squared test for independence = 7.27, degrees of freedom [df] = 3). Similarly the functional analysis of the 132 recombined genes of *R. felis* showed that 47, 40, 39, and 30 recombination events concern genes encoding for information storage and processing functions, for cellular processes and signaling, for metabolism, and for proteins with poorly characterized functions, respectively. Hence, recombination in *R. felis* occurred independently of the gene function ( $P$  value = 0.679, chi-squared test for independence = 1.52, df = 3).

Among the 152 horizontally transferred genes in *R. felis*, 42 genes encode for information storage and processing functions (out of 315 in *R. felis* genome), 25 genes encode for cellular processes and signaling (out of 280 in *R. felis* genome), 9 genes encode for metabolism (out of 270 in *R. felis* genome), and 76 genes encode proteins with poorly characterized functions (out of 497 in *R. felis* genome). When we excluded the ORFs encoding transposases, our data showed that the transfers in *Rickettsia* did not occur independently of the gene function ( $P$  value =  $1.98 \times 10^{-06}$ , chi-squared test for independence = 29.25, df = 3). More precisely, genes encoding for metabolic functions and the genes encoding for poorly



**Fig. 2.** Phylogenetic tree revealing the presence of recombination and a sample of multiple sequence alignment from five nucleotide sequences from *R. conorii*, *R. akari*, *R. felis*, *R. typhi*, and *R. prowazekii*. These schematic phylogenetic trees depict the probable evolutionary histories of the five sequences as inferred from different portions of the nucleotide sequence alignment. The A-C tree corresponds to the topology obtained from the concatenation of the whole genomes. It shows two groups of *Rickettsia*: the SFG containing the *R. conorii* and the *R. akari* subgroups and the TG containing *R. typhi* and *R. prowazekii*. The recombinant sequence (*R. felis*) changed clades: it belonged to the *R. akari* or the *R. conorii* subgroup (in blue) in the A-B fragment tree and to the TG (in red) in the B-C fragment tree. The recombinant sequence inherited the A-B fragment from a sequence resembling an *R. conorii* sequence and the B-C fragment from a sequence resembling an *R. typhi* sequence.

characterized functions are less and more horizontally transferred than expected, respectively. Informational genes were not exempted from HGT; indeed, there was no significant preference between cellular or information functions ( $P$  value = 0.91, chi-squared test for independence = 0.012,  $df = 1$ ).

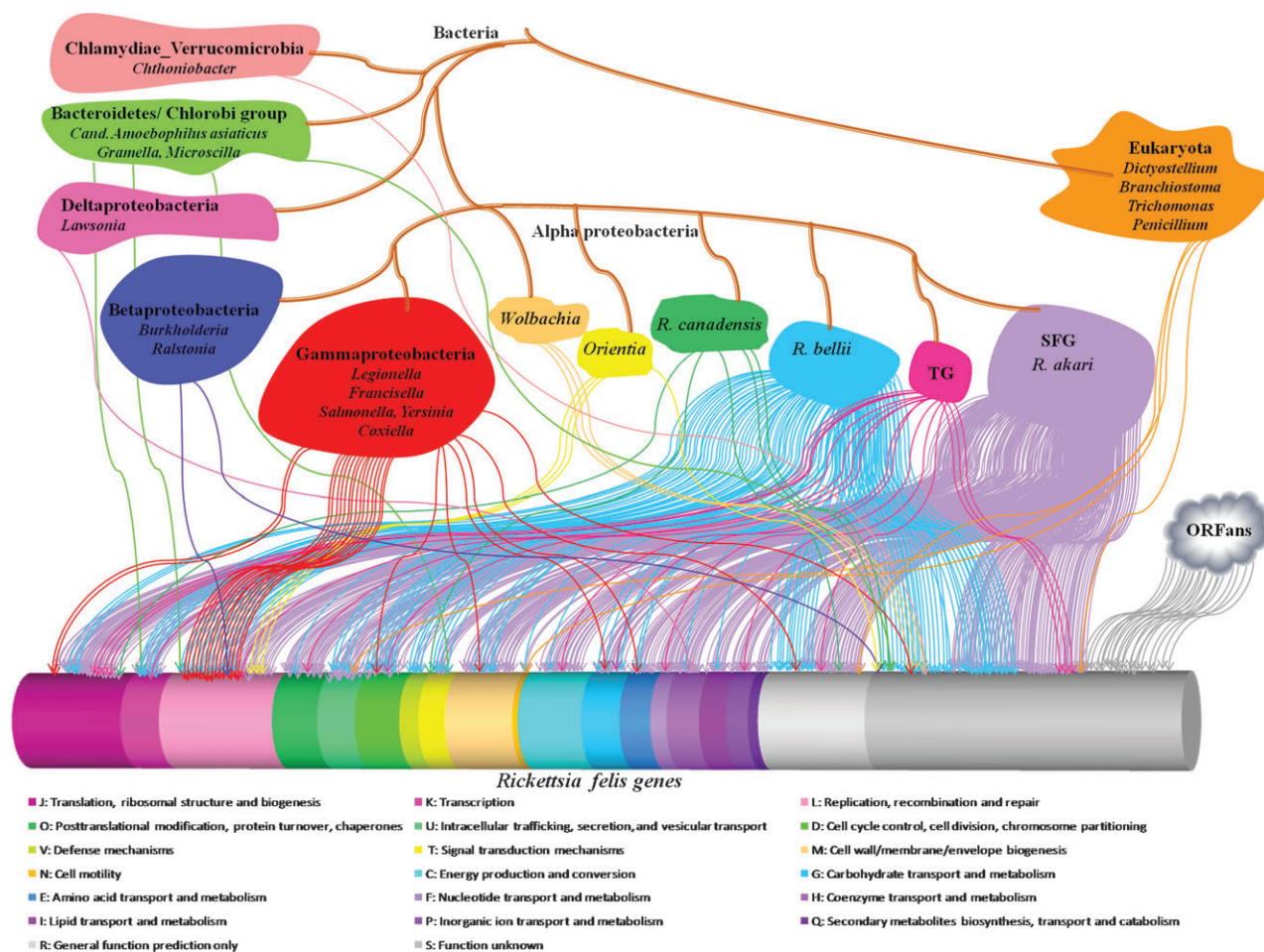
Moreover, in our analysis of sequence length, we found great variation in the length of the transferred sequences (fig. 4). The recombined sequences ranged from 12 to 83 nucleotides, with an average of 35 nucleotides. The average length of the genes where recombination has been detected was  $1,156.67 \pm 695.77$  bp versus  $789.76 \pm 528.19$  bp for the genes where no recombination event has been detected. Thus, recombined genes were significantly longer than genes that were not concerned by recombination ( $P$  value =  $9.55 \times 10^{-15}$ , Welch's  $t$ -test,  $t = 7.92$ ,  $df = 687.9$ ). The length of the horizontally transferred sequences ranged from 135 to 9,366 nucleotides, with an average of  $1081.99 \pm 1082.79$  bp. The length of the horizontally transferred genes slightly differs from that of the other genes in the genome (average length  $925.54 \pm 711.06$  bp), but the result is not relevant enough to conclude that HGT is length dependent ( $P$  value = 0.145, Welch's  $t$ -test,  $t = 1.466$ ,  $df = 114.5$ ).

These results reveal random transfers of DNA sequences in the *R. felis* genome that occur independently of gene function or sequence length. The DNA transfer seems to

depend on the only chance of meeting another organism in an intracellular community lifestyle.

## Discussion

Our phylogenetic analysis showed that the genome of *R. felis* is to some extent a collection of genes potentially having different evolutionary histories. The majority of the genes gave phylogenies in a general agreement with the phylogeny based on the core gene concatenation, showing *R. felis* in the SFG clade. However, some gene trees incongruence gave evidence for nonvertical inheritance. Although the percentage of horizontally transferred genes varies from 1.5% to 14.5% in bacterial genomes (Garcia-Vallvé and Palau 2000) or even 18% for *Escherichia coli* (Lawrence and Ochman 1998), our data showed that 12% of the genome of *R. felis* genome is of nonvertical inheritance. These results are consistent with previous work quantifying the amount of HGT in *Legionella* spp. (Coscollá et al. 2011); they emphasize the important role played by HGT in the composition of intracellular bacteria. The multiple origins of the *R. felis* gene repertoire make it impossible to represent its organismal evolution in the form of a tree, as illustrated in Darwin's book (Darwin 1859). This model of inheritance, which assigns a unique descent to every species, is inconsistent with data obtained in the



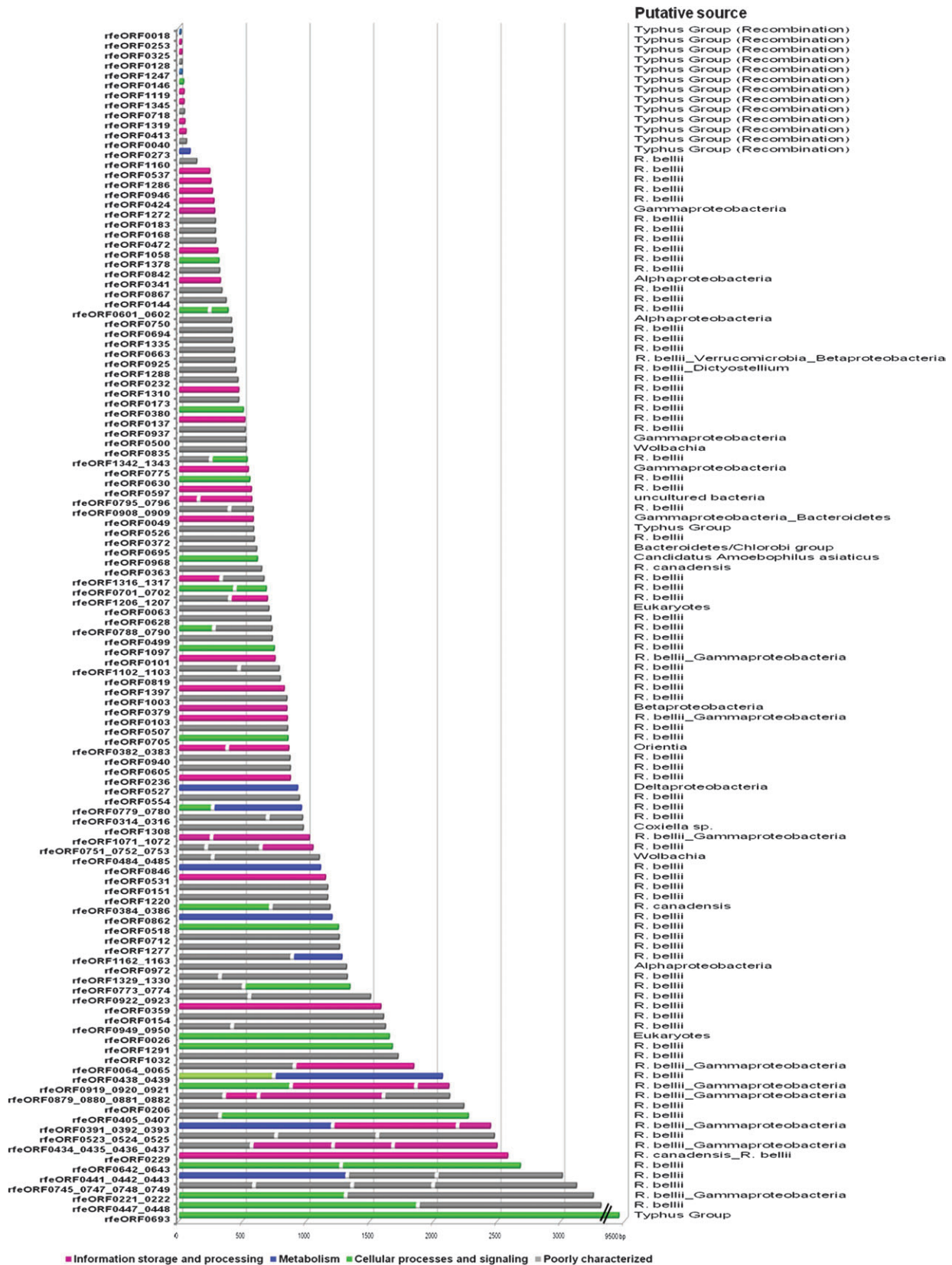
**FIG. 3.** The different origins of *Rickettsia felis* genes. The organismal phylogeny on the upper side of the figure was deduced from the concatenation of 31 universal protein families (Ciccarelli et al. 2006). The genome of *R. felis* was represented on the lower side of the figure with genes classified into functional categories of COGs. Arrows link *R. felis* genes to its closely phylogenetic species on the basis of the corresponding gene phylogeny. The number of arrows is proportional to the number of genes in the corresponding functional category.

genomic era (Doolittle 1999b). Indeed, the fact that spatial clusters of genes sometimes share the same incongruence suggest that, at least in some cases, these are not the consequence of tree reconstruction artifacts but rather the effect of genuine differences in the evolutionary histories of these genes brought about by horizontal transfer events (Boucher et al. 2003; Gogarten and Townsend 2005; Ochman et al. 2005). Horizontal transfer seems to challenge the traditional tree-based view of genome evolution described in terms of vertical descent in the organisms reproducing only sexually (Doolittle 1999a, 2000; Pennisi 1999; Koonin et al. 2001). Distinct from the concept of isolation, intracellular bacteria living sympatric lifestyles (e.g., in amoebae (Ogata et al. 2006)) seem to exchange genes and DNA, leading to large chimeric genome repertoires (Moliner et al. 2010). These chimeras constitute a formal refutation of the evolutionary vision derived from the tree of life. As a result, it is clear that a single gene tree, or even a tree from several highly conserved concatenated genes, is not informative enough to reflect the history of a whole genome (Doolittle 1999b; Dagan and Martin 2006). An adequate representation of the evolutionary relationships

and life histories of species might take the form of a forest, a synthesis of life, a network of genetic exchanges, or a rhizome of life showing the multiple roots and intertwining origins of currently living species, rather than a single tree (Baptiste et al. 2004; Doolittle and Baptiste 2007; Puigbo et al. 2009; Puigbo et al. 2010; Raoult 2010; Schliep et al. 2010). We applied a rhizome model showing the evolutionary relationships of *R. felis* as an alternative model to a single phylogenetic tree (fig. 3).

A phylogenetic reconstruction method combining the core of orthologous genes shared by members of a monophyletic group has been adopted by many studies (including previous analyses of our own [Merhej et al. 2009]) as the best way to infer organismal phylogeny (Lerat et al. 2003; Rokas et al. 2003). This assumption has been supported by arguments that core genes, which are largely comprised of essential genes, are inherited vertically. The single-gene phylogenies from the rickettsial core gene, however, exhibited a high degree of incongruence. The majority of the genes gave phylogenies in general agreement with the phylogeny based on core gene concatenation, showing *R. felis* in the SFG clade. To simplify the complex genealogy





Downloaded from https://academic.oup.com/mbe/article/28/1/1321/3/1/048891 by guest on 16 February 2022

**FIG. 4.** DNA transfer to *Rickettsia felis* from different sources. The horizontal transfer in *R. felis* concerns different length of DNA and various genes functions. The different ORFs at the left axis displayed original gene-single phylogeny showing *R. felis* in a cluster with the species at the right axis. The bars show the ORFs, and the different colors refer to their corresponding functions.

of *R. felis*, we have considered these genes as vertically inherited; nevertheless, some of these genes might result from recombination with *Rickettsia* spp. in SFG. Some gene trees had topologies that differed in only one or a few branches, for example, showing *R. felis* within the TG clade. The analysis of sequence alignments of these genes revealed cases of gene recombination likely due to the fact that *R. felis* and *R. typhi* can cohabit within the same flea (Eremeeva et al. 2008; Bitam et al. 2009). Coinfection limits bacterial allopatry and improves the ability of bacteria to undergo horizontal DNA transfer (Iredell et al. 2003; Chafee et al. 2009; Thomas and Greub 2010). Finally, some gene trees varied greatly from these topologies such that they placed *R. felis* in a clade with *R. canadensis* or *R. bellii*, often with low bootstrap support. These phylogenies might hide other instances of recombination that are more difficult to elucidate because unlike exchange with *R. typhi*, the places where these genetic exchanges occurred are more difficult to determine. Advances in molecular biology have revealed that these bacteria are quite ubiquitous: *R. bellii* can be found in both ticks and insects (Chen et al. 1996) and *R. felis* can live in many different hosts (Stevenson et al. 2005; Bitam et al. 2006; Venzal et al. 2006; Eremeeva et al. 2008; Reif and Macaluso 2009). Moreover, our analysis showed that nine genes in *R. felis* present high sequence similarity with amoebae and bacteria living in amoebae. Thus, it is possible that *R. felis* and these bacteria cohabited within an insect, tick, or protist that promotes horizontal DNA transfer. The frequency of recombination in *R. felis* seems to have been underestimated. Accordingly, the discovery of coinfection of related bacteria associated with atypical gene tree topologies should direct future research toward in-depth analysis of sequence alignments to identify eventual gene recombination events.

Previous studies have suggested that barriers might exist that select transferable genes according to their function. In fact, HGT occurs more often among operational genes (those involved in metabolic and cellular processes) than among informational genes (those involved in replication, transcription and translation) (Jain et al. 1999; Sorek et al. 2007; Wellner et al. 2007). Moreover, many studies have looked for the transfer of single genes, and the transfer of a block of genes has been investigated in the case of pathogenicity islands due to the importance of this transfer in pathogenicity (Hacker and Kaper 2000). Our phylogenetic analysis of the whole genome of *R. felis* allowed us to detect apparent horizontal transfer of DNA segments with different sizes corresponding to genes in various functional categories. These results indicate that any segment of DNA might have the opportunity to be horizontally transferred. Thus, the transfer of large DNA segments containing genes with diverse functions is possible; in fact, *Wolbachia* spp. can transfer their entire genomes to their eukaryotic host cells (Nikoh et al. 2008). Transferred genes are subsequently sorted and then kept or eliminated depending on whether or not they confer evolutionary advantages. Many of the transferred genes identified in this study have unknown functions. In some cases, the transferred genes were

present next to gene fragments, which may indicate random transfer and subsequent degradation. Our results refute the idea of bacteria with a thoughtful character in which intelligent donors and recipients select and promote the transfer of genes.

## Conclusion

Using a whole-genome approach, which examined the history of every gene, we found evidence that the *R. felis* genome not only evolves through vertical inheritance but also exhibits a mosaic structure of genes from different parental sources. These findings allowed us to reconstruct a partial rhizome of *R. felis* representing the different ascents of genes. This rhizome highlights the fact that contrary to predictions based on perfect colinearity of the rickettsiae and obligate intracellular isolation, a significant fraction of the *R. felis* genes have been subject to horizontal transfer. The amount and sources of horizontal transfer may be linked to the sympatric lifestyle of *R. felis* within arthropods. HGT events challenge the Darwinian vision of purely vertical inheritance from parents and stress the importance of considering pervasive random horizontal transfer in the representation of evolutionary relationships to more reliably map species history.

## Supplementary Material

Supplementary files 1–4 are available at *Molecular Biology and Evolution* online (<http://www.mbe.oxfordjournals.org/>).

## Acknowledgments

We would like to thank E. V. Koonin for providing us with the algorithm for COGs search, O. Chabrol for helping us in the implementation and the use of CONSEL program, and the reviewers for their valuable and constructive comments.

## References

- Andersson SG, Zomorodipour A, Andersson JO, Sicheritz-Ponten T, Alsmark UC, Podowski RM, Naslund AK, Eriksson AS, Winkler HH, Kurland CG. 1998. The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria. *Nature* 396:133–140.
- Bapteste E, Boucher Y, Leigh J, Doolittle WF. 2004. Phylogenetic reconstruction and lateral gene transfer. *Trends Microbiol.* 12:406–411.
- Bitam I, Baziz B, Kernif T, Harrat Z, Parola P, Raoult D. 2009. Molecular detection of *Rickettsia typhi* and *Rickettsia felis* in fleas from Algeria. *Clin Microbiol Infect.* 15:255–256.
- Bitam I, Parola P, De La Cruz KD, Matsumoto K, Baziz B, Rolain JM, Belkaid M, Raoult D. 2006. First molecular detection of *Rickettsia felis* in fleas from Algeria. *Am J Trop Med Hyg.* 74:532–535.
- Blanc G, Ogata H, Robert C, Audic S, Claverie JM, Raoult D. 2007. Lateral gene transfer between obligate intracellular bacteria: evidence from the *Rickettsia massiliae* genome. *Genome Res.* 17:1657–1664.
- 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.
- Bork P, Dandekar T, Diaz-Lazcoz Y, Eisenhaber F, Huynen M, Yuan Y. 1998. Predicting function: from genes to genomes and back. *J Mol Biol.* 283:707–725.

- Boucher Y, Douady CJ, Papke RT, Walsh DA, Boudreau ME, Nesbo CL, Case RJ, Doolittle WF. 2003. Lateral gene transfer and the origins of prokaryotic groups. *Annu Rev Genet.* 37: 283–328.
- Boyer M, Azza S, Barrassi L, et al. (15 co-authors). 2011. Mimivirus shows dramatic genome reduction after intraamoebal culture. *Proc Natl Acad Sci U S A.* 108:10296–10301.
- Boyer M, Gimenez G, Monti M-S, Raoult D. 2010. Classification and possible origin of orfan through analysis of large nucleocytoplasmic DNA viruses. *Intervirology* 53:310–320.
- Boyer M, Yutin N, Pagnier I, et al. (14 co-authors). 2009. Giant Marseillevirus highlights the role of amoebae as a melting pot in emergence of chimeric microorganisms. *Proc Natl Acad Sci U S A.* 106:21848–21853.
- Chafee ME, Funk DJ, Harrison RG, Bordenstein SR. 2009. Lateral phage transfer in obligate intracellular bacteria (*Wolbachia*): Verification from natural populations. *Mol Biol Evol.* 27:501–516.
- Chen DQ, Campbell BC, Purcell AH. 1996. A new rickettsia from a herbivorous insect, the pea aphid *Acyrtosiphon pisum* (Harris). *Curr Microbiol.* 33:123–128.
- Cochran WG. 1950. The comparison of percentages in matched samples. *Biometrika* 37:256–266.
- Coscollá M, Comas I, González-Candelas F. 2011. Quantifying nonvertical inheritance in the evolution of *Legionella pneumophila*. *Mol Biol Evol.* 28(2):985–1001.
- Ciccarelli FD, Doerks T, von Mering C, Creevey CJ, Snel B, Bork P. 2011. Toward automatic reconstruction of a highly resolved tree of life. *Science* 311:1283–1287.
- Dagan T, Martin W. 2006. The tree of one percent. *Genome Biol.* 7:118.
- Darwin CR. 1859. On the origin of species by means of natural selection. London: John Murray.
- Doolittle WF. 1999a. Lateral genomics. *Trends Cell Biol.* 9:M5–M8.
- Doolittle WF. 1999b. Phylogenetic classification and the universal tree. *Science* 284:2124–2129.
- Doolittle WF. 2000. Uprooting the tree of life. *Sci Am.* 282:90–95.
- Doolittle WF, Bapteste E. 2007. Pattern pluralism and the tree of life hypothesis. *Proc Natl Acad Sci U S A.* 104:2043–2049.
- Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32:1792–1797.
- Ellison DW, Clark TR, Sturdevant DE, Virtaneva K, Porcella SF, Hackstadt T. 2008. Genomic comparison of virulent *Rickettsia rickettsii* Sheila Smith and avirulent *Rickettsia rickettsii* Iowa. *Infect Immun.* 76:542–550.
- Eremeeva ME, Warashina WR, Sturgeon MM, Buchholz AE, Olmsted GK, Park SY, Effler PV, Karpathy SE. 2008. *Rickettsia typhi* and *Rickettsia felis* in rat fleas (*Xenopsylla cheopis*), Oahu, Hawaii. *Emerg Infect Dis.* 14:1613–1615.
- Felsenstein J. 1993. PHYLIP (Phylogeny Inference Package) version 3.5c. Distributed by the author. Seattle: Department of Genetics, University of Washington.
- Fournier PE, El Karkouri K, Leroy Q, Robert C, Giumelli B, Renesto P, Socolovschi C, Parola P, Audic S, Raoult D. 2009. Analysis of the *Rickettsia africae* genome reveals that virulence acquisition in *Rickettsia* species may be explained by genome reduction. *BMC Genomics.* 10:166.
- Frost LS, Leplae R, Summers AO, Toussaint A. 2005. Mobile genetic elements: the agents of open source evolution. *Nat Rev Microbiol.* 3:722–732.
- García-Vallvé S, Romeu A, Palau J. 2000. Horizontal gene transfer in bacterial and archaeal complete genomes. *Genome Res.* 10: 1719–1725.
- Gillespie JJ, Beier MS, Rahman MS, Ammerman NC, Shallom JM, Purkayastha A, Sobral BS, Azad AF. 2007. Plasmids and rickettsial evolution: insight from *Rickettsia felis*. *PLoS One.* 2:e266.
- Gogarten JP, Townsend JP. 2005. Horizontal gene transfer, genome innovation and evolution. *Nat Rev Microbiol.* 3:679–687.
- Guindon S, Gascuel O. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol.* 52:696–704.
- Hacker J, Kaper JB. 2000. Pathogenicity islands and the evolution of microbes. *Annu Rev Microbiol.* 54:641–679.
- Iredell J, Blanckenberg D, Arvand M, Grauling S, Feil EJ, Birtles RJ. 2003. Characterization of the natural population of *Bartonella henselae* by multilocus sequence typing. *J Clin Microbiol.* 41:5071–5079.
- Jain R, Rivera MC, Lake JA. 1999. Horizontal gene transfer among genomes: the complexity hypothesis. *Proc Natl Acad Sci U S A.* 96:3801–3806.
- Jain R, Rivera MC, Moore JE, Lake JA. 2003. Horizontal gene transfer accelerates genome innovation and evolution. *Mol Biol Evol.* 20:1598–1602.
- Johannesson K. 2001. Parallel speciation: a key to sympatric divergence. *Trends Ecol Evol.* 16(3):148–153.
- Koonin EV, Makarova KS, Aravind L. 2001. Horizontal gene transfer in prokaryotes: quantification and classification. *Annu Rev Microbiol.* 55:709–742.
- Lawrence JG, Ochman H. 1998. Molecular archaeology of the *Escherichia coli* genome. *Proc Natl Acad Sci U S A.* 95:9413–9417.
- Lerat E, Daubin V, Moran NA. 2003. From gene trees to organismal phylogeny in prokaryotes: the case of the gamma-Proteobacteria. *PLoS Biol.* 1:E19.
- Liddell FDK. 1983. Simplified exact analysis of case-referent studies: matched pairs; dichotomous exposure. *J Epidemiol Comm Health.* 37:82–84.
- Marchler-Bauer A, Anderson JB, DeWeese-Scott C, et al. (24 co-authors). 2003. CDD: a curated Entrez database of conserved domain alignments. *Nucleic Acids Res.* 31:383–387.
- Martin DP, Lemey P, Lott M, Moulton V, Posada D, Lefeuve P. 2010. RDP3: a flexible and fast computer program for analyzing recombination. *Bioinformatics* 26:2462–2463.
- Mayr E. 1942. Systematics and the origin of species from the viewpoint of a zoologist. New York: Columbia University Press. p. 334.
- McLeod MP, Qin X, Karpathy SE, et al (22 coauthors). 2004. Complete genome sequence of *Rickettsia typhi* and comparison with sequences of other rickettsiae. *J Bacteriol.* 186:5842–5855.
- Merhej V, El Karkouri K, Raoult D. 2009. Whole genome-based phylogenetic analysis of Rickettsiae. *Clin Microbiol Infect.* 15(Suppl 2): 336–337.
- Merhej V, Raoult D. 2010. Rickettsial evolution in the light of comparative genomics. *Biol Rev Camb Philos Soc.* 86:379–405.
- Moliner C, Fournier PE, Raoult D. 2010. Genome analysis of microorganisms living in amoebae reveals a melting pot of evolution. *FEMS Microbiol Rev.* 34:281–294.
- Nikoh N, Tanaka K, Shibata F, Kondo N, Hizume M, Shimada M, Fukatsu T. 2008. *Wolbachia* genome integrated in an insect chromosome: evolution and fate of laterally transferred endosymbiont genes. *Genome Res.* 18:272–280.
- Notredame C, Higgins DG, Heringa J. 2000. T-Coffee: a novel method for fast and accurate multiple sequence alignment. *J Mol Biol.* 302:205–217.
- Ochman H, Lawrence JG, Groisman EA. 2000. Lateral gene transfer and the nature of bacterial innovation. *Nature* 405:299–304.
- Ochman H, Lerat E, Daubin V. 2005. Examining bacterial species under the specter of gene transfer and exchange. *Proc Natl Acad Sci U S A.* 102(Suppl 1):6595–6599.
- Ogata H, Audic S, Barbe V, Artiguenave F, Fournier PE, Raoult D, Claverie JM. 2000. Selfish DNA in protein-coding genes of *Rickettsia* *Science.* 290:347–350.
- Ogata H, Audic S, Renesto-Audiffren P, et al (11 coauthors). 2001. Mechanisms of evolution in *Rickettsia conorii* and *Rickettsia prowazekii*. *Science* 293:2093–2098.
- Ogata H, La Scola B, Audic S, Renesto P, Blanc G, Robert C, Fournier PE, Claverie JM, Raoult D. 2006. Genome sequence of

- Rickettsia bellii* illuminates the role of *amoebae* in gene exchanges between intracellular pathogens. *PLoS Genet.* 2:e76.
- Ogata H, Renesto P, Audic S, Robert C, Blanc G, Fournier PE, Parinello H, Claverie JM, Raoult D. 2005. The genome sequence of *Rickettsia felis* identifies the first putative conjugative plasmid in an obligate intracellular parasite. *PLoS Biol.* 3:e248.
- Pennisi E. 1999. Is it time to uproot the tree of life? *Science* 284:1305–1307.
- Perlman SJ, Hunter MS, Zchori-Fein E. 2006. The emerging diversity of *Rickettsia*. *Proc Biol Sci.* 273:2097–2106.
- Popa O, Hazkani-Covo E, Landan G, Martin W, Dagan T. 2011. Directed networks reveal genomic barriers and DNA repair bypasses to lateral gene transfer among prokaryotes. *Genome Res.* 21:599–609.
- Puigbo P, Wolf YI, Koonin EV. 2010. The tree and net components of prokaryote evolution. *Genome Biol Evol.* 2:745–756.
- Puigbo P, Wolf YI, Koonin EV. 2009. Search for a 'Tree of Life' in the thicket of the phylogenetic forest. *J Biol.* 8:59.
- Raoult D. 2010. The post-Darwinist rhizome of life. *Lancet* 375:104–105.
- Raoult D, Roux V. 1997. Rickettsioses as paradigms of new or emerging infectious diseases. *Clin Microbiol Rev.* 10:694–719.
- Reif KE, Macaluso KR. 2009. Ecology of *Rickettsia felis*: a review. *J Med Entomol.* 46:723–736.
- Robinson D, Foulds L. 1981. Comparison of phylogenetic trees. *Math Biosci.* 53:131–147.
- Rokas A, Williams BL, King N, Carroll SB. 2003. Genome-scale approaches to resolving incongruence in molecular phylogenies. *Nature* 425:798–804.
- Roux V, Raoult D. 1999. Phylogenetic analysis and taxonomic relationships among the genus *Rickettsia*. In: Raoult D, Brouqui P, editors. *Rickettsiae and rickettsial diseases at the turn of the third millenium*. Marseilles (France): Elsevier production. p. 52–66.
- Saitou N, Nei M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol.* 4:406–425.
- Santoyo G, Romero D. 2005. Gene conversion and concerted evolution in bacterial genomes. *FEMS Microbiol Rev.* 29:169–183.
- Schliep K, Lopez P, Lapointe FJ, Baptiste E. 2010. Harvesting evolutionary signals in a forest of prokaryotic gene trees. *Mol Biol Evol.* 28:1393–1405.
- Shimodaira H. 2002. An approximately unbiased test of phylogenetic tree selection. *Syst Biol.* 51:492–508.
- Shimodaira H, Hasegawa M. 2001. CONSEL: for assessing the confidence of phylogenetic tree selection. *Bioinformatics* 17:1246–1247.
- Smets BF, Barkay T. 2005. Horizontal gene transfer: perspectives at a crossroads of scientific disciplines. *Nat Rev Microbiol.* 3:675–678.
- Sorek R, Zhu Y, Creevey CJ, Francino MP, Bork P, Rubin EM. 2007. Genome-wide experimental determination of barriers to horizontal gene transfer. *Science* 318:1449–1452.
- Stevenson HL, Labruna MB, Monteneri JA, Kosoy MY, Gage KL, Walker DH. 2005. Detection of *Rickettsia felis* in a New World flea species, *Anomiopsyllus nudata* (Siphonaptera: Ctenophthalmidae). *J Med Entomol.* 42:163–167.
- Stothard DR, Clark JB, Fuerst PA. 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.
- Stothard DR, Fuerst PA. 1995. Evolutionary analysis of the spotted fever and typhus group of *Rickettsia* using 16S rRNA gene sequences. *Syst Appl Microbiol.* 18:52–61.
- Tatusov RL, Fedorova ND, Jackson JD, et al (17 couthors). 2003. The COG database: an updated version includes eukaryotes. *BMC Bioinformatics.* 4:41.
- Tatusov RL, Koonin EV, Lipman DJ. 1997. A genomic perspective on protein families. *Science* 278:631–637.
- Thomas V, Greub G. 2010. *Amoebae/amoebal* symbionts genetic transfers: lessons from giant viruses neighbours. *Intervirology* 53:254–267.
- Touzain F, Denamur E, Medigue C, Barbe V, El Karkouri M, Petit MA. 2010. Small variable segments constitute a major type of diversity of bacterial genomes at the species level. *Genome Biol.* 11:R45.
- Venzal JM, Perez-Martinez L, Felix ML, Portillo A, Blanco JR, Oteo JA. 2006. Prevalence of *Rickettsia felis* in *Ctenocephalides felis* and *Ctenocephalides canis* from Uruguay. *Ann N Y Acad Sci.* 1078:305–308.
- Viezens J, Arvand M. 2008. Simultaneous presence of two different copies of the 16S rRNA gene in *Bartonella henselae*. *Microbiology* 154:2881–2886.
- Vos M, Didelot X. 2009. A comparison of homologous recombination rates in bacteria and archaea. *ISME* 3:199–208.
- Weinert LA, Werren JH, Aebi A, Stone GN, Jiggins FM. 2009. Evolution and diversity of *Rickettsia* bacteria. *BMC Biol.* 7:6.
- Wellner A, Lurie MN, Gophna U. 2007. Complexity, connectivity, and duplicability as barriers to lateral gene transfer. *Genome Biol.* 8:R156.
- Woese CR. 1987. Bacterial evolution. *Microbiol Rev.* 51:221–271.