AIX-MARSEILLE UNIVERSITE

FACULTE DE MEDECINE DE MARSEILLE

ECOLE DOCTORALE DES SCIENCES DE LA VIE ET DE LA SANTE

THESE DE DOCTORAT

Présentée par

DONG Xin

En vue de l'obtention du grade de Docteur de l'Université d'Aix-Marseille

Spécialité : Maladies Infectieuses et Pathologie Humaine

Comparative genomics of *Rickettsia* species

Soutenue le Lundi 10 décembre 2012

Devant le jury composé de

Pr. Jean-Louis Mège	**Président du jury**
Dr. Patricia Renesto	**Rapporteur**
Pr. Max Maurin	**Rapporteur**
Pr. Pierre-Edouard Founier	**Directeur de Thèse**

Unité de Recherche sur les Maladies Infectieuses et Tropicales Émergentes

UM63 - CNRS 7278 - IRD 198 – Inserm U1095

AVANT PROPOS

Le format de présentation de cette thèse correspond à une recommandation de la spécialité Maladies Infectieuses et Microbiologie, à l'intérieur du Master desSciences de la Vie et de la Santé qui dépend de l'Ecole Doctorale des Sciences de la Vie de Marseille.

Le candidat est amené à respecter des règles qui lui sont imposées et quicomportent un format de thèse utilisé dans le Nord de l'Europe et qui permet un meilleur rangement que les thèses traditionnelles. Par ailleurs**,** la partie introduction et bibliographie est remplacée par une revue envoyée dans un journal afin de permettre une évaluation extérieure de la qualité de la revue et de permettre à l'étudiant de commencer le plus tôt possible une bibliographie exhaustive sur le domaine de cette thèse. Par ailleurs, la thèse est présentée sur article publié, accepté ou soumis associé d'un bref commentaire donnant le sens général du travail. Cette forme de présentation a paru plus en adéquation avec les exigences dela compétition internationale et permet de se concentrer sur des travaux qui bénéficieront d'une diffusion internationale.

Prof. Didier Raoult

SOMMAIRE

Résumé

Le genre *Rickettsia*, sont des petites bactéries Gram-négatives et symbiotes intracellulaires obligatoires des eucaryotes. Les *Rickettsia* sont surtout connus pour leur pathogénicité et pour provoquer des maladies graves chez l'homme et les autres animaux. À ce jour, 26 espèces valides de Rickettsies ont été identifiées dans le monde entier, dont 20 sont des agents pathogènes éprouvées. Toutes les espèces de Rickettsies validées sont associées à des arthropodes. Les phylogénies basées sur divers marqueurs moléculaires ont présenté des topologies discordantes, avec seulement *R. bellii* et *R. canadensis* qui ne sont classées ni parmi la fièvre boutonneuse groupe rickettsies, ni parmi le typhus groupe rickettsies. En utilisant les méthodes avancées de séquençage de génomes entiers, nous avons obtenu et analysé quatre séquences génomiques de Rickettsies : *R. helvetica*, *R. honei*, *R. australis* et *R. japonica*. Via la phylogénomique qui constitue une nouvelle stratégie permettant de mieux comprendre leur évolution, l'on remarque que ces micro-organismes ont subi une évolution génomique réduite au cours de spécialisation en intracellulaire. Plusieurs caractéristiques évolutives, comme le réarrangement des gènes, la réduction génomique, le transfert horizontal de gènes et l'acquisition d'ADN égoïste, ont formé les génomes *Rickettsia* d'aujourd'hui. Ces processus peuvent jouer un rôle important pour équilibrer la taille du génome afin de l'adapter au mode de vie intracellulaire. En outre, la pathogénicité des rickettsies peut être associée à la réduction génomique.

Abstract

The *Rickettsia* genus is composed of small, Gram-negative, bacteria that are obligate intracellular eukaryotic symbionts. Members of the genus *Rickettsia* are best known for infecting and causing severe diseases in humans and other animals. To date, 26 valid *Rickettsia* species have been identified worldwide, including 20 that are proven pathogens. All validated *Rickettsia* species are associated to arthropods that act as vectors and/or reservoirs. The phylogenies based on various molecular markers have resulted in discrepant topologies, with *R. bellii* and *R. canadensis* being classified neither among spotted fever nor typhus group rickettsiae. In this thesis, using the advanced whole genomic sequencing methods, we have and analyzed the genomic sequences from four *Rickettsia* species, including *R. helvetica*, *R. honei*, *R. australis* and *R. japonica*. Phylogenomics constitute a new strategy to better understand their evolution. These microorganisms underwent a reductive genomic evolution during their specialization to their intracellular lifestyle. Several evolutive characteristics, such as gene rearrangement, reduction, horizontal gene transfer and aquisition of selfish DNA, have shaped *Rickettsia* genomes. These processes may play an important role in free-living bacteria for balancing the size of genome in order to adapt the intracellular life style. In addition, in contrast with the concept of bacteria becoming pathogens by acquisition of virulence factors, rickettsial pathogenecity may be linked to genomic reduction of metabolism and regulation pathways.

CHAPITRE I

Introduction générale

Le genre *Rickettsia* dans l'ordre des *Rickettsiales* (α-proteobacteria) est composé de petites bactéries à Gram-négatif, sont des parasites intracellulaires stricts des cellules eukaryotes (1). *Rickettsia prowazekii* a été l'une des premières bactéries intracellulaires à avoir bénéficié du séquençage du genome entier, dès l'année 1998 (2). A ce jour, le genre *Rickettsia* contient 26 espèces officiellement validées, dont 22 ont fait l'objet d'un séquençage génomique. Ces bactéries sont de petite taille, d'environ 1 µm x 0.5 µm. Leurs génomes varient en taille de 1,11 Mb pour *R. typhi* à 2,1 Mb pour une souche répertoriée comme *Rickettsia* endosymbiote d'*Ixodes scapularis*. Leur contenu en G+C varie de 29% pour *R. typhi* à 33% pour *Rickettsia* endosymbiote d'*Ixodes scapularis*.

Les bactéries du genre *Rickettsia* peuvent causer de nombreuses maladies chez l'homme. Tout au long de l'histoire, les espèces de *Rickettsia*, en particulier *R. prowazekii*, ont été responsables de millions de morts pendant les première et seconde guerres mondiales (3). Les symptômes les plus caractéristiques associent une éruption, une fièvre élevée et des céphalées. Les rickettsioses restent de maladies d'actualité. en 1995, une épidémie de typhus épidémique a touché plus de 50 000 personnes dans des camps de réfugiés au Burundi.

Les rickettsies sont une menace sérieuse pour l'homme, car elles sont le plus souvent associées avec des arthropodes se nourrissant de sang, comme les tiques, les puces, les poux et les acariens (4). L'homme peut entrer en contact avec ces hotes par l'intermédaire de leurs animaux de

compagnie ou des espèces sauvages, à domicile ou dans des zones de loisirs (figure 1). L'homme est infecté par les bactéries au travers de la piqûre d'arthropode dans le cas des tiques et des mites, ou par inoculation de matières fécales infectées sur la peau, comme c'est le cas avec les puces et les poux (5). A ce jour, parmi les 26 espèces de *Rickettsia* décrites, 20 ont fait preuve de pathogénicité chez l'homme (review tableau 1).

Selon le type de maladie causée chez l'homme et de réactivité antigénique croisée, le genre *Rickettsia* est divisé en deux biotypes: (i) le groupe des fièvres boutonneuses (SFG), qui comporte comprennent 22 espèces présentant des distributions géographiques restreintes, en dehors de *R. felis*. Des rickettsies du groupe SFG ont été trouvées sur tous les continents sauf en Antarctique. Ces bactéries ont une température de croissance optimale de 32°C et un contenu génomique en G+C entre 32 et 33%. Elles peuvent polymériser l'actine et donc être mobiles et entrer dans le noyau des cellules- hôtes (6); (ii) Le groupe du typhus (TG) inclut *R. prowazekii* et *R. typhi*. Les bactéries du groupe TG ont une température de croissance optimale de 35°C et un contenu génomique en G+C de 29%. De plus, elles ne se trouvent que dans le cytoplasme des cellules hôtes (6) ; (iii) les analyses phylogénétiques construites à partir des séquences des gènes d'ARN ribosomiques 16S et 23S indiquent que *R. bellii* et *R. canadensis* ne sont classées dans aucun des deux groupes (6).

Dans cette thèse, en guise d'introduction, nous avons fait une revue de la littérature scientifique sur les caractéristiques de pathogénicité des espèces de *Rickettsia* afin d'avoir une image générale de leur pathogénicité et voies de transmission. Le groupe SFG contient 18 espèces pathogènes, dont *R. conorii*, l'agent de la fièvre boutonneuse méditerranéenne, *R. rickettsii*, l'agent de la fièvre pourprée des montagnes rocheuses, *R. helvetica*, l'agent d'une fièvre anéruptive, *R. honei*, l'agent de la fièvre boutonneuse des iles Flinders, *R. australis*, l'agent du typhus à tiques du Qeensland, *R. japonica*, l'agent de la fièvre boutonneuse japonaise ou orientale. Les membres du groupe SFG sont distribués de l'Amérique du Nord à l'Asie du Sud-est, quasiment dans le monde entier (review tableau 1). Ces bactéries sont vectorisées principalement par des tiques dures, telles que *Dermacentor variabilis*, *Dermacentor adersoni*, *Rhipicephalus sanguineus* ou *Ixodes ricinus* (7,8). En revanche, *R. felis*, contrairement aux autres membres du groupe SFG, est principalement propagée par les puces de chat et de chien du genre *Ctenocephalides* (9). Les deux espèces du groupe TG sont également présentes dans le monde entier. Elles sont principalement vectorisées par des insectes, les puces et les poux (10, 11). Le typhus épidémique, causé par *R. prowazekii*, est la plus sévère rickettsiose, avec un taux de mortalité de près de 30% (10,12). Cette rickettsiose a actuellement une distribution géographique limitée à quelques zones restreintes en Europe de l'est, en Afrique et en Amérique du sud. Le typhus murin, causé par *R. typhi*, a une distribution mondiale et a été

impliqué dans une épidémie à Hawaii 2002 (13).

En raison de leur mode de vie intracellulaire, les génomes de rickettsies ont subi une évolution réductive. Leur proximité phylogénétique avec les mitochondries rend ces organismes encore plus intéressants (2). En plus de leur intérêt médical, les bactéries du genre *Rickettsia* sont également un modèle pour les études évolutives. Au cours des deux dernières décennies, un nombre croissant de nouvelles espèces de *Rickettsia* a été identifié dans différentes régions géographiques. Le développement du séquençage génomique complet à haut débit (appelé « next generation sequencing » par les anglo-saxons) nous a permis de révéler les caractéristiques génomiques des *Rickettsia*. Dans cette thèse, nous avons analysé 4 génomes de rickettsies ont séquencé, incluant *R. helvetica* strain C9P9, *R. honei* strain RB^T, *R. australis* strain $phillips^T$ et *R. japonica* strain YH^T. Nous avons également dans notre revue étudié les modes d'évolution des espèces de *Rickettsia*, dont les réarrangements génomiques, l'évolution réductive, l'expansion génomique et l'évolution des ADN égoïstes.

Au cours de cette thèse, nous nous sommes intéressés également aux facteurs de virulence des *Rickettsia*. Les éléments trouvés conservés dans les génomes bactériens appelés îlots de pathogénicité ne sont pas retrouvés dans les séquences des génomes de *Rickettsia* (14).De plus, dans les plasmides de des bactéries identifiés aujourd'hui, aucun facteur de virulence n'a pu être identifié non plus. Donc, dans notre travail, nous avons particulièrement étudié les éléments impliqués dans les processus

d'infection de l'hôte. Les antigènes principaux de *Rickettsia* sont les antigènes de surface cellulaire du groupe Sca. Les deux principaux membres en sont les protéines immunodominantes OmpA et OmpB. La famille Sca comporte 16 membres qui sont soit entiers, soit interrompus par des codons stop (pseudogènes), soit présents sous forme de fragments, soit totalement absents dans certaines espèces de *Rickettsia*. Seules Sca 1, OmpB et Sca4 sont présentes chez toutes les rickettsies (15, 16, 17). Ces protéines sont impliquées dans la reconnaissance et l'invasion dans la cellule hôte (18). Après phagocytose par les cellules eucaryotes, la membrane du phagosome contenant les *Rickettsia* est détruite par l'activité phospholipase D, et les bactéries sont libérées dans le cytoplasme des cellules hôtes (5). Le mécanisme d'infection de cellules est différent entre les deux groupes de rickettsies. Les bactéries du groupe SFG diffusent de cellule à cellule par polymérisation de l'actine cellulaire (protéine RickA, 19,20), tandis que les bactéries du groupe TG se multiplient dans les cellules-hôtes jusqu'à ce qu'elles explosent puis diffusent à d'autres cellules (Figure 3, 4, 5). De plus, la comparaison génomique entre *R. africae* et *R. rickettsii* a révélé que la pathogénicité de ces bactéries est influencée par la réduction de leur génome au travers d'une perte des gènes de régulation (21). Nous avons également comparé ces génomes avec les autres génomes de *Rickettsia*, et d'autres études seront menées pour étudier leur pathogénicité.

CHAPITRE II

Review

Review of characteristics of rickettsial genomes

DONG Xin, Pierre-Edouard Fournier

Aix-Marseille Université, URMITE, UM63, CNRS 7278,
IRD 198, Inserm U1095, Faculté de Médecine
27 Boulevard Jean Moulin, 13385 Marseille cedex 05, France
Tel : +33 491 324 375, Fax : + 33 491 387 772

1. Introduction

The genus *Rickettsia* in the order *Rickettsiales* of the *α-proteobacteria*, is composed of small Gram-negative bacteria, obligate intracellular parasites of eukaryotic cells (1). *Rickettsia* species have been found in many ecological niches and are most commonly associated with blood-feeding arthropod hosts, including lice, fleas, ticks and mites (2, 3, 15, 21, 22). To date, 26 *Rickettsia* species are described (table 1). Of these, 20 are proven human pathogen. Rickettsioses are found worldwide and include severe diseases such as epidemic typhus (*R. prowazekii*), Rocky Mountain spotted fever (*R. rickettsii*), Mediterranean spotted fever (*R. conorii*), and North Asian tick typhus (*R. sibirica*) (7, 11, 12, 14).

Arthropods are both vectors and reservoirs of *rickettsiae* (24, 25). Animals and humans may be infected through arthropod bites or contaminations of skin wounds by flea or louse feces (28, 31). Once injected into the skin, *rickettsiae* invade endothelial cells, multiply intracellularly and spreed from cell to cell.

The genus *Rickettsia* has traditionally been divided into 2 groups on the basis of phenotypic and genotypic traits (37,38, 39, 40): (i) the spotted fever group (SFG) include 18 species exhibiting restricted territorial distributions, most of which are transmitted by hard ticks (16, 17, 18, 20, 23). SFG rickettsiae have been found on every continent except Antarctica (27, 29, 30, 33). SFG rickettsiae have an optimal growth temperature of 32 °C and a genomic G+C content between 32% and 33%. They can polymerize actin

and hence enter the host cell nucleus (34); (ii) the typhus group (TG), including *R. prowazekii* and *R. typhi*, transmitted by body lice and rat fleas, respectively. TG rickettsiae have an optimal growth temperature of 35°C, a genomic G+C content of 29%, and are only found in the cytoplasm of host cells (34, 44). However this classification does not apply to all species. For example, the SFG *R. helvetica* exhibits no actin polymerization and only appears in the cytoplasm of host cells (35); the SFG *R. akari* is and *R. felis* transmitted by mites and cat fleas, respectively; *R. bellii* and *R. canadensis* cannot be classified within neither group (36, 41, 42, 43).

In evolutionary terms, *Rickettsia* species is close to microbial relatives of mitochondria (4). Rickettsiae and mitochondria might have shared a common ancestor and a similar evolutionary process to colonize the cytoplasm of eukaryotic cells 2 billion years ago (8). Some gene remnants and split genes are degraded protein-encoding genes (9), suggesting that during evolution, and adaptation to intracellular lifestyle, rickettsial genomes may have lost genes that were mandatory by progressive degradation or mutation. *R.bellii* is the earliest diverging species of known rickettsiae and can survive in phagocytic ameobae. Its genome exhibits many genes highly similar to orthologs in intracellular bacteria from ameobae. This suggests that rickettsial ancestors have exchanged genes with amoeba-like protozoa, eventually leading to their adaptation to the intracellular lifestyle within eukaryotic cells.

Rickettsia species was among the first intracellular bacteria that benefited from whole genome sequencing, as early as 1998 (4). Currently, the *Rickettsia* genus contains 26 officially validated species. Twenty-two genomes have been sequenced to date (table 2). *Rickettsia* genomes range in size from 1,11Mb for *R. typhi* to 2,1Mb for *Rickettsia endosymbiont of Ixodes scapularis*. Their G+C content ranges from 29% (*R. typhi*) to 33% (*Rickettsia endosymbiont of Ixodes scapularis)*. Compared to other free living bacteria, these genomes are smaller, and contain large numbers of degraded genes and high levels of non coding sequences (5, 52, 53). It is hypothesized that during their specialization to their intracellular lifestyle, they were eliminated some biosynthetic pathways producing the metabolite, and now provided by eucaryotic cells (5, 54). Some biosynthetic pathways are replaced by transport systems. Rickettsial genomes contain many genes encoding transposases and motifs involved in protein-protein interaction such as ankyrin repeats and tetratricopeptide repeats (6, 56, 57), which suggests a role in pathogenicity, possibly in interacting with host cell proteins.

In this review, we discuss the mechanisms of genomic evolution identified in several *Rickettsia* species.

2. Plasmids in *Rickettsia* species.

In bacteria, plasmids are known to play an important role in horizontal gene transfer, acquisition of virulence determinants and antibiotic resistance, and environmental adaptation (61, 62, 63, 64). Initially, rickettsiologists assumed that rickettsiae were devoid of plasmids (4, 51). This was supported by the sequencing of the first 3 rickettsial genomes of *R. prowazekii*, *R. conorii*, and *R. typhi* which were devoid of plasmids. However, in 2005, the sequencing of the *R. felis* genome enabled the discovery of a polymorphic plasmid, present under 2 forms, pRF and pRFδ, with size of 62,829-bp and 39,263-bp, respectively. pRF possessed 68 open reading frames (ORFs), 53 of which having orthologs in public databases. pRFδ was identical to pRF with the exception of a 23,566-bp segment that contains 24 ORFs (65). First, *R. felis* plasmids were considered to be unique among *rickettsiae*, because pRF-like sequence failed to be detected in other *Rickettsia* species (66). However, pRF being likely to be conjugative, and its phylogenetic analysis suggesting an origin in ancestral *rickettsiae*, it was suspected that plasmids existed in other *Rickettsia* species (59). Subsequently, a 23,486-bp plasmid, pRM, was identified in *R. monacensis* by PFGE and Southern blot methods (72, 73). The sequence analysis identified 23 ORFs, 22 of which encoded proteins that were highly similar to hypothetical proteins present in other bacteria, including 11 of which matched those encoded by the pRF plasmid. Both pRF and pRFδ plasmid contained transposases that are associated to horizontal gene transfer in

other bacteria (68). *R. massiliae* was the second fully sequenced rickettsial genome that harbored a plasmid. This 15,286-bp plasmid pRMA had 12 protein-coding genes and a pseudogene, 7 of them being present in pRF, including that 2 ORFs related to transposable elements (67). *R. peacokii* exhibited a 26,406-bp plasmid, pRPR. This plasmid contained 20 putative genes and 5 pseudogenes. 1 ParA protein, 2 hypothetical proteins and 1 transposase had a best BLAST match with pRF ORFs; 6 ORFs were similar to those from pRMA including 1 DnaA-like protein, 2 Hsp molecular chaperons and 3 transposases (69). In *R. africae*, one small plasmid pRA was detected, pRA exhibiting a size of 12,377-bp. pRA contained 11 genes, including 6 genes that were highly similar to a DnaA-like protein, 1 site-specific recombinase, 2 transposases, 1 autotransporter protein, Sca12, that may play a role in interaction with host cells, and 1 ParA-like plasmid stability protein. In *R. amblyommii*, 3 plasmids were detected (45): pRAM18 had a size of 18,344-bp, pRAM23 had a size of 22,852-bp and pRAM32 had a size of 31,972-bp. These two plasmids both encoded the DnaA-like replication initiator and ParA-like plasmid stability protein orthologous to genes from the pRM plasmids. Only pRAM23 encoded 1 Hsp protein that was also present in pRM. This plasmid also carried a PhrB UV damage repair enzyme that was only found in *R. bellii* and *Orientia tsutsugamushi*, as well as 2 *sca12* gene copies (71). pRAM32 encoded 22 genes, including 17 coding proteins. The shotgun sequencing project of *Rickettsia endosymbiont of Ixodes scapularis* detected three plasmids.

pREIS1 and pREIS3 had sizes of 55,147-bp, 66,811-bp, respectively. The pREIS3 plasmid was incompletely sequenced. The ParA protein was detected in all three plasmids, together with Tra family conjugal transfer proteins. However, no DnaA-like replication initiator protein was identified in these plasmids. In *R. helvetica*, we identified a 47,188-bp plasmid, pRhe (142). pRhe had 49 genes, including 22 plasmidic genes involved in maintenance, transmissibility and DNA processing machinery, and 14 genes possibly involved in adaptation and defense of *R. helvetica*. The genes encoding DnaA-like protein, ParA encoding plasmid stability protein and helicase recD/traA-like protein were present in two copies each. On the basis of PFGE and southern blot results, *R. akari* had one plasmid and *R. rhipicephali* had several plasmids (71). In *R. australis*, we identified a 26,608-bp conjugative plasmid, pRau (19, 143). pRau had 25 ORFs including 11 genes: Tra, Mob family proteins, DnaA-like protein, ParA stability protein, TldD/PmbA protein, transposases and hypothetical proteins.

The identification of plasmids in 9 *Rickettsia* species revealed an unsuspected genetic diversity in *Rickettsia* genomes. In addititon, the conjugative ability and transposon elements in plasmids suggest that rickettsial genomes may exchange genes through lateral gene transfer. Phylogenetic studies of the ParA plasmid stability protein shows that all the rickettsial plasmid-encoded ParA proteins, except that encoded by pREIS4 from rickettsia endosymbiont of I. scapularis, were highly diverse from the chromosomal ParA proteins and majority clustered in to three groups

(Figure1). These results suggested that *Rickettsia* plasmids possibly have mutiple origins (71).

3. Phylogenomic study and genomic evolution in the *Rickettsia* genus

3.1. Mechanisms of evolution

The obligate intracellular lifestyle of *Rickettsia* species drove their genomes to adapt to this specific environment, resulting not only in differences in the genome size, but also in genomic architecture (76, 77). Their phylogenetic relationship with mitochondria also make interesting in evolutionary studies (8).

3.1.1. Genomic rearrangements of Rickettsia genomes

By comparison of 7 *Rickettsia* genomes (5), except for *R. bellii*, all genomes express long range colinearities and only a few genomic rearrangements. Three genomes of *R. prowazekii*, *R. typhi* and *R. conorii* share 775 genes. Although most of these genes are colinear, in *R. typhi*, compared to *R. prowazekii* and *R. conorii*, there is a 35-kb inversion in gene order, which is close to the replication terminus. In addition, there is a *R. typhi*-specific inversion nearby the origin of replication, compared to *R. prowazekii* and *R. conorii* (10). Such inversions are also seen in the sequences of *R. helvetica*, R. *honei* and *R. australis* (142, 143, 144), indicating that this region is a hot spot for genomic rearrangements.

The major mechanism responsible for genomic rearrangement in

Rickettsia is homologous intra-chromosomal recombination which occurred between inversely oriented repetitive sequences and duplicated genes (79, 80). Homologous intra-chromosomal recombination results in a deletion or an invasin of one or several repetitive sequences.

Such events have been observed in *Rickettsia*, as shown by the case of the so-called super-ribosomal protein gene operon (81). This operon is highly conserved in a broad range of bacteria and archaea and consists of about 40 genes locating in seven operons with a conserved order (77). The conserved order of these 40 genes is considered as the reflection of the ancient origin of super ribosomal protein gene operons. This expected order is shared by the common ancestor of bacteria and archaea, but differences in the orientation of genes within the operons, as well as the numbers of operons and their locations in genomes vary from species to species (81).

Bacterial rRNA genes of bacteria are normally organized into operons with a conserved order 16S-23S-5S, and tRNA genes are often found in the spacer between the 16S and the 23S rRNA genes (32). However, it has been observed that *Rickettsia* species have an unusual arrangement of rRNA genes, as the 16S rRNA gene is separated from the 23S and 5S rRNA gene cluster (26, 78). This rearranged divided operon structure was observed in all rickettsial genomes so far, including *R. conorii, R. parkeri, R. sibirica, R. rickettsii, R. amblyomii, R. montana, R. rhipicephali, R. australis, R. akari, R. felis, R. canada, R. typhi* and *R. bellii*. The upstream spacer of the rearranged 23S rRNA gene in some *Rickettsia* species

contains short repetitive sequences that have been eliminated in other related species. This suggests that the rearrangement of rRNA genes occurred by intra-chromosomal recombination several times independently during the evolution of the *Rickettsia* lineage.

3.1.2. The reductive evolution of Rickettsia genomes.

Rickettsia species has small genome sizes ranging from 1.1 Mb (*R. typhi*) to 2.1 Mb (*rickettsia endosymbiont of Ixodes scapularis*). It has been suggested that obligate intracellular bacteria like *Rickettsia* have evolved from a common ancestor who had a larger genome size. The reconstruction of the genome from the last common ancestor genome of *Rickettsia* species provided a chromosome of 1,252 protein coding genes and 39 RNA-coding genes (5). *R. bellii, R. felis, R. massiliae, R. africae, R. conorii, R. prowazekii* and *R. typhi* retained 85%, 81%, 91%, 92%, 95%, 98% and 98% of these genes, respectively. TG rickettsiae only evolved by genome reduction from the common ancestor. The gene loss process is an important factor for the diversity of modern *Rickettsia* species Of the 1,252 genes from the common ancestor, 39% were lost in at least one species. *R. bellii* and *R. felis* have larger genomes, and underwent less gene loss. The highest degree of gene loss was observed in TG rickettsiae genomes. *R. massiliae*, *R. conorii*, and *R. africae* underwent an accelerated rate of gene loss after the separation of SFG and TG. The different gene loss degrees suggest that the genome reduction occurred through several modes of reductive evolution, and the

gene loss process is still ongoing today (9, 74, 75). Genome reduction may result from the elimination of large segments of DNA by homologous recombination, and progressive degradation and subsequent elimination of small segments of DNA.

Genome recombination

The study of rickettsial genomes highlighted many genes, such as rRNA (*rpsL* and *rpsG*) and elongation factor proteins (*tuf* and *fus*) present in more than one copy in *Rickettsia* genomes (58). The deletion or inactivation of these copies does not seem lethal for the bacterial life. These genes can serve as repeat sequences, and initiate a rapid gene loss through intra-chromosomal recombination (80). When compared to other bacteria genomes, *Rickettsia* genomes have a high percentage of non-coding DNA sequences which also contain many repeat sequences (83, 84). In addition, sequence comparison of the identified ORFs located in the intergenic regions of different *Rickettsia* species has provided several examples of homologous recombination as the responsible mechanism for gene deterioration which may also play a role in rapid gene loss. For example, a putative ORF has two pairs of short internal repeat elements at two sides, locating in the intergenic region of the S-adenosylmethionine tRNA gene (*gueA*), and ABC-transporter protein coding gene (*abcT3*). These repeat sequences provided the chance of occurrence of a rapid gene loss by homologous recombination. This putative ORF is absent in TG rickettsiae,

but inactive or degraded in SFG, and the size of this inactivated ORF varies between different members of the SFG rickettsiae (81). It suggests that the size differences in the SFG rickettsiae are a result of independent homologous recombination between the internal repeat elements in this region. The degree and positions of deletions caused by intra-chromosomal recombination in *Rickettsia* is different among the species, which suggests that the homologous recombination is an ongoing process that may result in an ongoing loss of genes under weak or no selection pressure (81).

Genome degradation

Rickettsia is characterized by high rate of accumulation of slightly harmful deletions, mutations and insertions (82). Alternatively, gene loss can also result from the accumulations of small mutations. Comparison of the *R. conorii* and *R. prowazekii* genomes enabled the different steps of gene degradation to be identified, from complete gene to functional split gene, to split gene not functional, to gene remnant (9). The formation of internal stop codon inside of the intact gene can occur through creation of a frameshift inside of genes by a single base mutation, insertion or deletion. The intact gene has been fragmented into several split genes. The overall images of the split gene still have high similarity to the full length ortholog in relative species. The ORFs could preserve its function and still can be transcribed, if the stop codon is generated in the region of polycistronic sequence which allows the conservation some of the enzymatic activities. By genomic repair

system, these alternations still can be restored. However, if the stop codon is formed in the region of high A+T%, the mutation in polyA is difficult to be rescued, because of its infidelity of the polymerase. As intracellular parasites, *Rickettsia* genomes have an evolutionary trend of genomic reduction. The genomic degradation is detrimental for G+C%, and enrichment for A+T%. The continue degradation would lead the split gene inactivation and become a split gene not functional rather than the genomic recovery. Then, the complete gene is interrupted by several stop codons, and performs no more functional. Parallel sequence comparisons of certain genes demonstrate the process from full functional gene to split gene not functional. As an example, the gene encoding for AmpG is widely present in *Rickettsia* genomes (9). AmpG protein cause resistance to antibiotics of the β-lactam family. The *ampG4* sequence was studied in five SFG and TG rickettsiae species. This gene was found intact in the SFG rickettsiae *R. conorii*, *R. rickettsii* and *R. montanensis*, but degraded in the TG rickettsiae. In the TG rickettsiae members, the *ampG4* gene contains many frameshifts and stop codons. The filamentation protein-encoding gene *fic* is also found intact in the SFG rickettsiae but present as a split gene in the TG rickettsiae. This gene is interrupted by 13 stop codons and contains a 279 bp-deletion in *R. prowazekii*, and contains 2 stop codons with a 579 bp-deletion in *R. typhi*. Furthermore, this degradation process may continue, split gene evolving into gene remnants or even being completely eliminated from *Rickettsia* genomes. The genomic sequence comparison of *R. prowazekii* and

R. conorii identified 229 ORFs, in the former genome which have significant similarity with the intergenic non coding regions of the later genome (9). Thus, intergenic spacers may contain the degraded gene remnants that are no longer recognized as functional genes but retain patial gene signatures. The *gabD* gene encodes succinate semialdehyde dehydronenase oxidoreductase (SSDH), playing a role in the formation of succinate from succinate semialdehyde by reducing NAD+/NAD(P)+ to NADH/NADPH (85). This gene is found to be complete in the SFG rickettsiae, but highly degraded in the TG rickettsiae, where only a 3'-end remnant of the gene remains (81). If the degradation inactivates a gene that plays a crucial part within a complex pathway or network, according to the domino theory of gene death (86), the gene loss may eventually cause a massive gene extinction of the dependent genes. As evidence, the entire cytochrome c oxidase pathway is missing from the *R. typhi* genome (10). By comparison with the reconstruction of the *Rickettsia* ancestor genome, modern rickettsial genomes are seriously reduced, including deletion of important biosynthesis pathways. In addition, comparative studies of *Rickettsia* genomes suggested that the virulence differences observed among *Rickettsia* species may be linked to the genomic reduction (70).

3.1.3. Genomic expansions of Rickettsia genomes

Although the reductive trend is dominant in *Rickettsia* genomes, the genetic elements replication, proliferation and gene integration still can occur. Genome size expansions may occur through of foreign DNA acquisition by gene transfer, intragenomic duplication or receive of reception of transposons or repetitive sequences.

Horizontal gene transfer (HGT)

In bacteria, gene transfer usually occurs via mobile genetic element such as plasmid and requires a set of conjugation genes such as *tra* family genes (87, 88). The *R. bellii* genome possesses a set of complete conjugation genes, and pilli like-filaments were observed on the bacterial surface (89). As described above, the pRF plasmid from *R. felis* contains the *tra* family conjugation genes. Among 13 tested *Rickettsia* collection strains, 11 got positive conjugation gene detection. This suggests that the conjugation elements are widely present among *Rickettsia* spp (88), and that horizontal gene transfer (HGT) occured at a high rate (88). To detect HGT events, rickettsial COGs missed from at least one genome were analysed by phylogeny. Six genes were indicated to be likely acquired by HGT. These genes encoding for enzymes with synthase or proteolytic activities, originate from *γ-proteobacteria* or eukaryotes (90). In some cases, a HGT can be followed by a gene loss, as the case for the *metK* gene. This gene, encoding the S-adenosylmethionine synthase, is considered to have been acquired

from *γ-proteobacteria* before the separation of SFG and TG. Later, the *metK* was lost by *R. prowazekii*. Conversly, *Rickettsia* species can also be gene donors. For example, the gene encoding the nucleotidyltransferase substrate binding protein from SFG rickettsiae is distributed to *Firmicutes* species and *Bacteroidetes* species In addition, a putative permease was provided to*γ-proteobacteria* species by a *Rickettsia* ancestor (90).

Ameobae have been proposed to be an evolutionary "training ground" for many ameobae-associated bacteria. Therefore, it is assumed that the *Rickettsia* ancestor also used ameobae as hosts (91). The phylogenetic reconstructions inferred the possibility of gene transfer between the *Rickettsia* ancestor and other ameobae parasite. The *R. bellii* genome presents many ORFs exhibiting a high degree of similarity with genes from two intracellular bacteria of ameobae: *Legionella pneumophila* and *Protochlamydia ameobophila*. By phylogenetic analysis, 15 genes between *R. bellii* and *L. pneumophila*, and 7 genes between *R. bellii* and *P. ameobophila* are considered as possible HGT elements. Within ameobae, HGT might have given the *Rickettsia* ancestor the access to novel gene pools, with possibility to acquire foreign DNA from other intracellular bacteria, thus, in capability of adaptation environment (89).

Gene duplication

Rickettsia genomes may also have gained genes via duplication (9, 65). Gene duplication is an important source of bacterial adaptation to variable

environmental conditions and new niches (93). But gene duplication is a rare event in *Rickettsia* genomes (92). Duplicated genes may still conserve the full function from the original source, and may play a role in increasing the effect of encoded function. Gene duplication may prevent the dysfunction of genes caused by deletion or intra-chromosomal homologous recombination or inactivation by SNP mutations. On the other hand, with the accumulation the mutations, a duplicated gene may become non functional or accept diverse mutation and finally develop into a new functional gene, providing a putative new selective advantage in a new environment (94). Even during the course of evolution, the *Rickettsia* genomes have a trend of gene loss rather than gene duplication, the strong selective effect co-existing with functional duplication required for survival.

Various gene families are duplicated in *Rickettsia* genomes, such as duplication of Type IV secretion systems components. Type IV secretion systems (T4SSs) is functional as multiple component membrane spanning transporters and is widely found in many *Rickettsia* genomes. The *R. prowazekii* and *R. conorii* genomes have encoded for six Vir components (*virB4*, *virB8*-*virB11*, *virD4*), and the *virB4*, *virB8* and *virB9* have two copies of each (95). These copies have been scattered in separated locales of the genome. The copies exhibit differences in nonsynonymous substitution. The *R. felis* genome contains 14 *spoT* transcribed genes involved in transcription and cellular metabolism (65). These duplicate *spoT* genes may help the bacteria to persist in the host and adapt to its intracellular environment. The

R. bellii and several SFG rickettsiae genomes contain many duplication toxin/antitoxin (TA) systems which involved in the stringent response pathway cascade. These genes may help the bacteria in global regulation of metabolism and the stress response (89). The *R. conorii* genome has multiple copies of *ampG* agent encoding β-lactamase, which may explain the resistance of these bacteria to β-lactam antibiotics (9). R. prowazekii, the most degraded *Rickettsia* genome, which lacks the genes encoding the biosynthesis of purines and pyrimidines (4), has five tlc1 gene paralogs (97). This gene encodes on ATP/ADP translocase responsible for the capture of ATP from the host cell (96). Similar sequences are found in *R. typhi*, *R. montanensis* and *R. rickettsii* without any stop codon or frameshift. It seems that these five paralogs are functional (81).

Selfish DNA

In addition to gene duplication, genome expansion can also occur through the duplication of selfish DNA (99, 100, 102). Due to their self-propagating characteristic these mobile genetic elements are found in most organisms (101). Selfish DNA related to the repeat elements has no obvious cellular function, but some evidences suggest that self-replicating elements have physiological functions such as stress response (103). In some cases, they can increase the fitness of their host significantly. Selfish DNA considered being an important factor for the genome evolution as well as transposaes, responsible for the genomic redistributions and expression

of regulation. The sequence analysis of the *R. conorii* genome uncovered the presence of the repeat mobile elements, the repeat palindromic element (RPE) (55, 98).

These RPEs have a length between 100-150 bp in *Rickettsia*. In the genome of *R. conorii*, 656 RPE sequences are identified. On the basis of sequence similarity, RPE sequences were classified into 8 families (RPE-1 to RPE-8). Their copy numbers ranged from 5 to 223 (99). The RPE-like repetitive element with palindromic sequence presents in both non coding sequences and genes in *Rickettsia*, unlike the presence only in the non coding sequences in *Enterobacteria* (45, 104, 105). 44 RPE-1 sequences are discovered in the *R. conorii* genome and 10 copies in the *R. prowazekii* genome. Surprisingly, *in R. conorii*, 19 RPE-1 targeted different genes, in *R. prowazekii*, 9 RPE-1 were inserted in the regions encoding proteins (98). We highlighted the same process of evolution of genes by inserting palindromic sequence in the genomes of *Wolbachia* genus bacteria (141, 145). But the existent of this phenomenon in other organisms remains unclear. Several observations suggest that the RPE is an element that has recently proliferated in the genomes of *Rickettsia* without destroying the expressions and functions of genes. RPE sequences are slightly hydrophilic. Most of the RPEs are capable to form a stable hairpin like mRNA secondary structures. Most genes with the insertion of RPE have homologues in bacteria distant. Some of these genes correspond to essential proteins for the metabolism, such as the glutamyl-tRNA synthetase and

tRNA(Ile)-lysidine synthetase. The positions of the RPE in protein sequences appear random, but their positions seem compatible with the 3-dimensional structures of proteins. By comparison with the homologous proteins, it demonstrated that the insertion sites of RPE are always located on the surface of the targeted proteins, far from the functional or catalytic sites (106).

By insertion into an existing gene, these repeats form a new ORF, giving an additional peptide segment in the final gene product. Under the evolution stress, these insertions followed by a long term of selection, the genes may appear the evolutionary advantage compared with the original genes and contribute the emergence of new protein sequences, domains and functions. This suggests that the RPEs participate directly in the generation of new gene sequences (98). The RPE elements in different proteins may have a possible functional role for the host. The *R. prowazekii* genome has a high evolutionary degradation rate. Indeed, the RPEs in *R. prowazekii* are more difficult to identified, only two of them are predicted to be able to form stable hairpin structures (106), most of them may already lost their functions. The loss of RPE may not deeply affect the protein function. RPE possible had a functional role in the early *Rickettsia* genomes but became weekly due to the rapid deterioration process in the current days.

By genomic sequence analysis, many other repeats have been found, in *Rickettsia* genomes including ankyrin repeats (ANK) and tetratricopeptide

repeats (TPR) (107). In eukaryotes, ANK containing proteins have a cell membrane connection function, including membrane of endosynmbiont to the cytoskeleton (108). But in bacteria, its function is poorly known. In *Pseudomonas aeruginosa*, the ANK like protein AnkB is essential for the periplasmic catalase activity. It probably plays a role of protective scaffold in the periplasm (109). In *Anaplasma phagocytophilum*, the ANK repeat protein AnkA was shown to be associated with the chromatin of infected cells. These suggest that ANK is involved in the alteration of the host cell gene transcription by binding to host chromatin (110). The protein with ANK domains expressed in *O. tsutsugamushi* can modify the stability of the host protein synthesis to control the cellular functions after infection, suggesting that the ANK protein plays a role in pathogenesis (113). TPRs identified as 34-amino acid degenerate sequences, are involved widely in many functions, but mainly in protein-protein interaction (111, 112). The TPR-containing proteins of *Legionella pneumophila* have been involved in interfering with the host cell trafficking events (114, 115). The ANK and TPR repeats have also been found in pathogenic bacteria and in viruses (108).

3.2. Virulence factors in *Rickettsia*

The study of the virulence of *Rickettsia* species has been initiated long ago. This research has mainly focused, through *in vitro* studies, on *Rickettsia*-host interactions. The pathogenicity of *Rickettsia* species involves several steps including recognition, invasion, phagosome escape,

intracellular multiplication and cell lysis (6). Unlike other Gram-negative bacteria with lipopolysaccharide outer membrane, after heating or trypsin treatment, *Rickettsia* species loses their adherence ability to host cells (116). In additional, *Rickettsia* species possesses a paralogous family of outer membrane protein, the Sca family proteins, among which rOmpA and rOmpB have been described as major rickettsial antigens (13, 49, 50, 60). It was described that Anti-rOmpA antibodies can prevent the attachment of *R. rickettsii* to host cells (117). This suggested that the invasion of *Rickettsia* into their targeted host cells requires a surface adherence by a protein ligand. rOmpA and rOmpB are globally present in SFG rickettsiae, the TG rickettsiae only having rOmpB (48). rOmpB is highly conserved in the *Rickettsia* genus (46, 47). It is involved in recognition and facilities in the invasion step to the host cell (119). Immunization against these two proteins provided a partial protection in mice against the SFG rickettsiae (120). rOmpB can interact with Ku70, a subunit of a nuclear DNA-dependent protein kinase present in the nucleus, cytoplasm and plasma membrane. This interaction is used by *R. rickettsii* to invade Vero and HeLa cells (121). Recent evidence demonstrated that the rOmpB-Ku70 ligand can mediate *Rickettsia* adhesion to nonphagocytic host cells (122). In additional, the overlay assays of 2D-PAGE and mass spectrometry revealed two additional putative adherence factors of *Rickettsia*, Adr1 and Adr2, encoded by the genes RC1281, RC1282 in *R. conorii* and the gene RP827, RP828 in *R. prowazekii* (118), respectively. These two proteins are present

ubiquitously among *Rickettsia* species. Furthermore, following phagocytosis, *Rickettsia* is able to escape into the cytoplasm before the phagolysosomal fusion (123). In bacteria, the escape from phagosomes and vesicle lysis is likely dependant on a phospholipase A2 (PLA2) activity (124). However, there is no PLA2-encoding gene in Rickettsia. But in *R. conorii* genome, the first rickettsial phospholipase-D (PLD) was found (125, 126). Salmonella transformed with the *R. prowazekii* PLD gene was able to escape from phagosomal vacuoles, which demonstrates that PLD was a major factor for *Rickettsia* phagosomal escape (127). Then, once *Rickettsia* penetrates into the cytoplasm of host cells, they are able to proliferate. By interaction with the host structure components, such as actin cytoskeleton (128, 129), these bacteria can search the nutrient environment, spread between cells and avoid the host immune response (130, 131). Actin motility is considered to be an important virulence factor in *Listeria* species and *Shigella* species (136, 137). One of the characteristics of SFG rickettsia is their actin-based motility, which has been identified in *R. conorii*, *R. rickettsii*, *R. monacensis*, *R. parkeri* and *R. australis* (132, 133, 134). By genomic comparison of SFG and TG rickettsiae, the protein RickA, who retains a Wlskott-Aldrich syndrome protein homolog domain, is considered as the responsible factor of actin polymerization (9). RickA can recognize and recruit the Arp2/3 complex which is an actin nucleator which can initiate the polymerization of new actin filaments and organize filaments into Y-branched arrays, then generate γ-branched networks in order to control the actin cytoskeleton

(135). RickA acts as a nucleator of actin polymerization. The *R. prowazekii* genus has no presence of *rickA* ortholog, and this species lacks actin motility (140). To realize the bacteria biology function *Rickettsia* needs to export the toxin proteins into the host cell or into the extracellular matrix. *Rickettsial* genomes contain genes encoding the type IV secretion system components (138). The homologies with the *virB*/*virD4* secretion system of *Agrobacterium tumefaciens* suggest that the type IV secretion system in *Rickettsia* genomes belongs to the T4SS type IV subgroup (139). The regulation of the type IV in *Rickettsia* species is poorly understood. Genes encoding for VirB-type T4SS have been found in all sequenced genomes of *Rickettsial* spp (66). The traditional definition of bacterial virulence factors is that any component produced by a pathogen contributes to a disease in a host (136). Proteins are thought to be the main responsible for the pathogenicity. The elements found conserved in bacteria genomes called pathogenicity islands are easily recognized in a distinct genome. But such structures cannot be found in *Rickettsia* genome sequences (61). However, in *Rickettsia* plasmids identified todate, no virulence factors could be identified. The genomic comparison between *R. africae* and *R. rickettsii* revealed that pathogenicity may be influenced by genome reduction, possibly by the loss of regulatory genes. In *R. rickettsii*, the loss of regulatory genes was associated with development of virulence under selective pressure (70).

4. Conclusion

In the last decades, more and more new *Rickettsia* species are continued discovered in different geographic regions. The next generation sequencing methods and advanced molecular analysis tools arising the enormous amount data and the innovation in the science research of this genus. The influence of obligate intracellular lifestyle reshaped their genomic architectures. The evolutionary stress drove *Rickettsia* species to a extreme genomic reduction. Different evolutionary processes such as genome rearrangements, loss of genomic information and gene duplications have been discused. Gene loss seems to be far more extensive than gene acquisition.

Rickettsia genomes are highly rearranged which is suggested to be the result of intrachromosomal recombination at duplicated sequences, such as the super ribosomal protein operons and the 16S-23S-5S rRNA operons appears different figures to the other bacteria and archaea. The evidence suggested that it is a consequence of intrachromosomal recombination at the inversely oriented repeat sequences of genes or spacers in ancestor at an early stage of the transition to an obligate intracellular life style.

Reductive evolution is the major evolution process that has influenced the genomes of *Rickettsia*. Primarily, the genome reduced through by homologous recombination between oriented repetitive sequences at the early stage. By reducing mass of repeated sequences, the rate of homologous recombination becomes minor. However, *Rickettsia* species

also exhibit genomic expansion event through the HGT, the duplication of genes or selfish DNA, such as RPE, ANK repeat sequences and the integration mobile insertion elements. The homologous recombination may still ongoing by the competition with genomic expansion. Sequence and phylogenetic analysis suggest that the deletion event occur more often than the insertion in *Rickettsia*. After the consumption of repeated sequences, *Rickettsia* genomes will continue genomic degradation by the short mutations. This will give rise to the formation of split genes, pseudo genes and junk DNA. Eventually, they are expected to slowly deteriorate from the genome. Phylogenomic analysis of the TG and SFG rickettsiae indicate that many fragment, split or completely disappeared genes of one species may remain active in another species, such as *ampG4*, *fic* and *gabD* genes present intact in the SFG but highly degraded or split in the TG rickettsiae. This process may play an important role in free-living bacteria for balancing the size of genome in order to adapt the intracellular life style. In addition, opposite with the concept of bacteria becoming pathogens by acquisition of virulence factors, *Rickettsia* pathogenecity may related with the genomic reduction of the metabolism and the regulation pathway.

5. Prospectives

The traditional phylogeny research based on various molecular markers have resulted a disordered phylogeny tree topologies. Today, the genomic studies diverted the focalization at genes and protein-coding sequences to

the comparison of the entire genomes. This whole-genome-based phylogenomic tool provides a new strategy to describe the phylogenetic position of *Rickettsia* species.

6. Acknowledgments

I would like to thank Khalid El Karkouri for his constructive comments.

7. References

1. The emerging diversity of Rickettsia, Steve J Perlman, Martha S Hunter, Einat Zchori-Fein, Proc Biol Sci. 2006 September 7; 273 (1598):2097-2106. Published online 2006 April 21. doi: 10.1098/rspb.2006.3541, PMCID: PMC1635513
2. Tick-and flea-borne rickettsial emerging zoonoses. Parola P, Davoust B, Raoult D. Vet Res. 2005 May-Jun;36(3):469-92.
3. Emerging and re-emerging rickettsioses: endothelial cell infection and early disease events, David H. Walker & Nahed Ismail, Nature Reviews Microbiology 6, 375-386 (May 2008) doi:10.1038/nrmicro1866
4. The genome sequence of Rickettsia prowazekii and the origin of mitochondria. Andersson SG, Zomorodipour A, Andersson JO, Sicheritz-Pontén T, Alsmark UC, Podowski RM, Näslund AK, Eriksson AS, Winkler HH, Kurland CG. Nature. 1998 Nov 12;396(6707):133-40.
5. Reductive Genome Evolution from the Mother of Rickettsia, Blanc G, Ogata H, Robert C, Audic S, Suhre K, et al. (2007) Reductive Genome Evolution from the Mother of Rickettsia. PLoS Genet 3(1): e14.doi:10.1371/journal.pgen.0030014
6. Some lessons from Rickettsia genomics. Renesto P, Ogata H, Audic S, Claverie JM, Raoult D. FEMS Microbiol Rev. 2005 Jan;29(1):99-117.
7. Pathogenic rickettsiae as bioterrorism agents. Azad AF, Radulovic S. Ann N Y Acad. Sci.2003 Jun;990:734-8.
8. Phylogenetic analysis based on groE shows the closest evolutionary relationship between mitochondria and Rickettsia. Emel'ianov VV, Sinitsyn BV. Genetika. 1999 Jun;35(6):733-43.
9. Mechanisms of evolution in Rickettsia conorii and R. prowazekii. Ogata H, Audic S, Renesto-Audiffren P, Fournier PE, Barbe V, Samson D, Roux V, Cossart P, Weissenbach J, Claverie JM, Raoult D. Science. 2001 Sep 14;293(5537):2093-8.

10. Complete genome sequence of Rickettsia typhi and comparison with sequences of other rickettsiae. McLeod MP, Qin X, Karpathy SE, Gioia J, Highlander SK, Fox GE, McNeill TZ, Jiang H, Muzny D, Jacob LS, Hawes AC, Sodergren E, Gill R, Hume J, Morgan M,Fan G, Amin AG, Gibbs RA, Hong C, Yu XJ, Walker DH, Weinstock GM. J Bacteriol. 2004 Sep;186(17):5842-55.

11. Material safety data sheet - *rickettsia prowazekii*. January 2001 Canadian Laboratory Centre for Disease Control. 4 Dec 2002.

12. Rickettsioses as paradigms of new or emerging infectious diseases. Raoult D, Roux V. Clin Microbiol Rev. 1997 Oct;10(4):694-719.

13. Genetic differentiation of Chinese isolates of Rickettsia sibirica by partial ompA gene sequencing and multispacer typing. Zhang L, Jin J, Fu X, Raoult D, Fournier PE. J Clin Microbiol. 2006 Jul;44(7):2465-7.

14. Typhus and other rickettsioses: emerging infections in Germany. Dobler G, Wölfel R. Dtsch Arztebl Int. 2009 May;106(20):348-54. Epub 2009 May

15. Ticks and tickborne bacterial diseases in humans: an emerging infectious threat. Parola P, Raoult D. Clin Infect Dis. 2001 Mar 15;32(6):897-928. Epub 2001 Mar

16. The relationship between spotted fever group Rickettsiae and ixodid ticks.Socolovschi C, Mediannikov O, Raoult D, Parola P. Vet Res. 2009 Mar-Apr;40(2):34. Epub 2009 Apr 10.

17. Spotless rickettsiosis caused by Rickettsia slovaca and associated with Dermacentor ticks. Raoult D, Lakos A, Fenollar F, Beytout J, Brouqui P, Fournier PE. Clin Infect Dis. 2002 May 15;34(10):1331-6. Epub 2002 Apr 1

18. A systematic study of the Australian species of the genusIxodes (Acarina: Ixodidae). F.H.S Roberts, Journal Article. 1960

19. Rickettsioses in Australia. Graves S, Stenos J. Ann N Y Acad Sci. 2009 May;1166:151-5.

20. Japanese spotted fever: report of 31 cases and review of the literature. F. Mahara Emerg Infect Dis. 1997 Apr-Jun; 3(2): 105–111. PMCID: PMC2627607

21. Rickettsioses in Japan and the far East. Mahara F. Ann N Y Acad Sci. 2006 Oct;1078:60-73.

22. Isolation of a spotted fever group rickettsia from a patient and related ecologic investigations in Xinjiang Uygur Autonomous Region of China. M Y Fan, J G Wang, Y X Jiang, D G Zong, B Lenz, D H Walker, J Clin Microbiol. 1987 April; 25(4): 628–632. PMCID: PMC266048

23. Causative agent of spotted fever group rickettsiosis in Japan. T Okada, Y Tange, and Y Kobayashi Infect Immun. 1990 April; 58(4): 887–892. PMCID: PMC258556

24. Genetic identification of rickettsiae isolated from ticks in Japan. Fournier PE, Fujita H, Takada N, Raoult D. J Clin Microbiol. 2002 Jun;40(6):2176-81.

25. Spotted fever group rickettsiosis and vectors in Kanagawa prefecture. Katayama T, Furuya Y, Yoshida Y, Kaiho I. Kansenshogaku Zasshi. 1996 Jun;70(6):561-8.

26. Molecular phylogeny and rearrangement of rRNA genes in Rickettsia species. Andersson SG, Stothard DR, Fuerst P, Kurland CG. Mol Biol Evol. 1999 Jul;16(7):987-95.

27. Bustamante ME, Varela G. Estudios de fiebre manchada en Mexico. Hallazgo del Amblyomma cajennense naturalmente infectado, en Veracruz. Rev Inst Salubr Enferm Trop. 1946;7:75–8.

28. Natural Infection of the Tick, Amblyomma Cajennense, with Rickettsia Rickettsii in Panama. Enid C. de Rodaniche Am J Trop Med Hyg July 1953 :696-699

29. Rocky Mountain spotted fever in Brazil. del Sá DelFiol F, Junqueira FM, da Rocha MC, de Toledo MI, Filho SB.

30. A focus of Rocky Mountain spotted fever within New York City. Salgo MP, Telzak EE, Currie B, Perlman DC, Litman N, Levi M, Nathenson G, Benach JL, Al-Hafidh R, Casey J. N Engl J Med. 1988 May 26;318(21):1345-8.

31. Rickettsial infection in animals and Brazilian spotted fever endemicity. Sangioni LA, Horta MC, Vianna MC, Gennari SM, Soares RM, Galvão MA, Schumaker TT, Ferreira F, Vidotto O, Labruna MB. Emerg Infect Dis. 2005 Feb;11(2):265-70.

32. Organization of the bacterial chromosome. S Krawiec and M Riley Microbiol Rev. 1990 December; 54(4): 502–539.PMCID: PMC372791

33. Tick-Borne Rickettsioses around the World: Emerging Diseases Challenging Old Concepts. Philippe Parola, Christopher D. Paddock, and Didier Raoult. Clin Microbiol Rev. 2005 October; 18(4): 719-756. doi: 10.1128/CMR.18.4.719-756. 2005 PMCID: PMC1265907

34. Res Microbiol. 1992 Nov-Dec;143(9):821-9. Intracellular movements of Rickettsia conorii and R. typhi based on actin polymerization. Teysseire N, Chiche-Portiche C, Raoult D.

35. Ultrastructure of a Japanese Rickettsial strain genetically identified as Rickettsia helvetica which was originally found in Europe. Yano Y, Fujita H, Takada N. Microbiol Immunol. 2004;48(7):535-9.

36. Ancestral Divergence of Rickettsia beZlii from the Spotted Fever and Typhus Groups of Rickettsia and Antiquity of the Genus Rickettsia. Diane R. Stothard, Jonathan B. Clark, Paul A. Fuerst. International Journal of Systematic Bacteriologyo,C t.1994, p. 798-804

37. Report of the Ad Hoc Committee on Reconciliation of Approaches to Bacterial Systematics. L. G. Wayne,I, D. J. Brenner, R. R. ColwelL, P. A. D. Grimont, Kandler, M. I. Krichevsky, L. H. Moore,W. E. C. Moore, R. G. E. murray, E. Stackebrandt, M. P. Starr,I and H. G.Truper

38. The species concept for prokaryotes. Rosselló-Mora R, Amann R. FEMS Microbiol Rev. 2001 Jan;25(1):39-67.

39. Kingdom protozoa and its 18 phyla. Cavalier-Smith T. Microbiol Rev. 1993 Dec;57(4):953-94.

40. Reassessment of the Taxonomic Position of Rickettsiella V. Roux, M. Bergoin, N. Lamaze, and D. Raoult.International Journal oF Systematic Bacteriologoyc,t . 1997, p. 1255-1257

41. Classification of Rickettsia tsutsugamushi in a new genus, Orientia gen. nov., as Orientia tsutsugamushi comb. nov. Tamura A, Ohashi N, Urakami H, Miyamura S. Int J Syst Bacteriol. 1995 Jul;45(3):589-91.

42. Phylogenetic analysis of the genus Rickettsia by 16S rDNA sequencing. Roux V, Raoult D.Res Microbiol. 1995 Jun;146 (5):385-96.

43. Serologic typing of rickettsiae of the spotted fever group by microimmunofluorescence. Philip RN, Casper EA, Burgdorfer W, Gerloff RK, Hughes LE, Bell EJ. J Immunol. 1978 Nov;121(5):1961-8.

44. Genotypic identification and phylogenetic analysis of the spotted fevergroup rickettsiae by pulsed-field gel electrophoresis. Roux V, Raoult D. J Bacteriol. 1993 Aug;175(15):4895-904.

45. Complete genome sequence of candidatus rickettsia amblyomma strain GAT-30v. Submitted (FEB-2012) to the EMBL/GenBank/DDBJ databases.

46. Phylogenetic analysis of members of the genus Rickettsia using the gene encoding the outer-membrane protein rOmpB (ompB). Roux V; Raoult D. International Journal of Systematic and Evolutionary Microbiology 1449-1455 Part 4 JUL 2000

47. The Rickettsial OmpB β-Peptide of Rickettsia conorii Is Sufficient To Facilitate Factor H-Mediated Serum Resistance. Riley SP, Patterson JL, Martinez JJ. Infect Immun. 2012 Aug;80(8):2735-43. Epub 2012 May 21.

48. Phylogeny of Rickettsia spp. inferred by comparing sequences of ‘gene D’, which encodes an intracytoplasmic protein. Z Sekeyova, V Roux and D Raoult. Int J Syst Evol Microbiol.2001 jul;51(Pt 4):1353-60

49. Sca1,a previously undescribed paralog from autotransporter protein-encoding genes in Rickettsiaspecies Maxime Ngwamidiba, Guillaume Blanc, Didier Raoult and Pierre Edouard Fournier BMC Microbiology 2006, 6:12 doi:10.1186/1471 -2180-6-12

50. Molecular Evolution of Rickettsia Surface Antigens: Evidence of Positive Selection. Guillaume Blanc, Maxime Ngwamidiba, Hiroyuki Ogata, Pierre-Edouard Fournier, Jean-Michel Claverie and Didier Raoult. Mol Biol Evol (October 2005) 2(10): 2073-2083.doi: 10.1093/molbev/msi199First published online: June 22, 2005

51. On the evolutionary origin of the plant mitochondrion and its genome. Gray, M. W., Cedergren, R., Abel, Y. & Sankoff, D. Proc. Natl. Acad. Sci. USA 86, 2267±2271 (1989).

52. Rickettsia identification by polymerase chain reaction amplification of DNA and restriction endonuclease fragment length polymorphism analysis. In: Kazar J, Raoulte D. Rickettsiae and Rickettsial Diseases: Proceedings of the IVth International Symposium. Bratislava, 1991:175-82

53. The winds of (evolutionary) change: breathing new life into microbiology. Olsen, G. J.,Woese, C. R. & Overbeek, J. Bacteriol. 176, 1±6 (1994).

54. S-adenosylmethionine transport in Rickettsia prowazekii. Tucker AM, Winkler HH, Driskell LO, Wood DO (2003) J Bacteriol 185: 3031–3035.

55. Proliferation and deterioration of Rickettsia palindromic elements. Amiri H, Alsmark CM, Andersson SG.Mol Biol Evol. 2002 Aug;19(8):1234-43.

56. TolC-Dependent Secretion of an Ankyrin Repeat-Containing Protein of Rickettsia typhi. Kaur SJ, Rahman MS, Ammerman NC, Beier-Sexton M, Ceraul SM, Gillespie JJ, Azad AF. J Bacteriol. 2012 Sep;194(18):4920-32. Epub 2012 Jul 6.

57. Rickettsia Phylogenomics: Unwinding the Intricacies of Obligate Intracellular Life. Gillespie JJ, Williams K, Shukla M, Snyder EE, Nordberg EK, et al. (2008) PLoS ONE 3(4): e2018. doi:10.1371/journal.pone.0002018

58. A chimeric disposition of the elongation factor genes in Rickettsia prowazekii. A C Syvänen, H Amiri, A Jamal, S G Andersson, and C G Kurland. J Bacteriol. 1996 November; 178(21): 6192–6199. PMCID: PMC178489

59. Plasmids and rickettsial evolution insight from rickettsia felis. Gillespie, J. J. et al. 2007. PLoS ONE 2: e266

60. Phylogenetic study of Rickettsia species using sequences of the autotransporter protein-encoding gene sca2. Ngwamidiba M, Blanc G, Ogata H, Raoult D, Fournier PE. Ann N Y Acad Sci. 2005 Dec;1063:94-9.

61. Pathogenicity islands and the evolution of bacterial pathogens.Infect Agents Dis. 1996 Jan;5(1):1-7. Lee CA. 62. Nature 405, 299-304 (18 May 2000)

62. Lateral gene transfer and the nature of bacterial innovation Howard Ochman, Jeffrey G. Lawrence & Eduardo A. Groisman. doi:10.1038/35012500

63. The evolution of bacterial toxins. Stine, O. C., J. P. Nataro. 2006. p. 167-188. In H. S. Siefert and V. J. Di Rita (ed.), Evolution of microbial pathogens. ASM Press, Washington, DC.

64. Effect of Rickettsial Toxin VapC on Its Eukaryotic Host. .Audoly G, Vincentelli R, Edouard S, Georgiades K, Mediannikov O, et al. (2011) PLoS ONE 6(10): e26528. doi:10.1371/journal.pone.0026528

65. The genome sequence of Rickettsia felis identifies the first putative conjugative plasmid in an obligate intracellular parasite. Ogata H, Renesto P, Audic S, Robert C, Blanc G, Fournier PE, Parinello H, Claverie JM, Raoult D. PLoS Biol. 2005 Aug;3(8):e248. Epub 2005 Jul 5.

66. Proteome analysis of Rickettsia felis highlights the expression profile of intracellular bacteria. Ogawa M, Renesto P, Azza S, Moinier D, Fourquet P, Gorvel JP, Raoult D. Proteomics. 2007 Apr;7(8):1232-48.

67. 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 November;17(11): 1657–1664. doi: 10.1101/gr.6742107 PMCID: PMC2045148

68. Transposon Insertion Reveals pRM, a Plasmid of Rickettsia monacensis Gerald D. Baldridge, Nicole Y. Burkhardt, Roderick F. Felsheim, Timothy J. Kurtti and Ulrike G. Munderloh. Published ahead of print 15 June 2007, doi: 10.1128/AEM. 00988-07 Appl. Environ. Microbiol. August 2007 vol. 73 no. 15 4984-4995

69. Genome Sequence of the Endosymbiont Rickettsia peacockii and Comparison with Virulent Rickettsia rickettsii: Identification of Virulence Factors Roderick F. Felsheim, Timothy J. Kurtti, Ulrike G. Munderloh PLoS One. 2009; 4(12): e8361. Published online 2009 December 21. doi: 10.1371/journal.pone.0008361

70. Analysis of the Rickettsia africae genome reveals that virulence acquisition in Rickettsia speciesmay be explained by genome reduction. Fournier PE, El Karkouri K, Leroy Q, Robert C, Giumelli B, Renesto P, Socolovschi C, Parola P, Audic S, Raoult D. BMC Genomics. 2009 Apr 20;10:166.

71. Wide dispersal and possible multiple origins of low-copy-number plasmids in rickettsia speciesassociated with blood-feeding arthropods. Baldridge GD, Burkhardt NY, Labruna MB, Pacheco RC, Paddock CD, Williamson PC, Billingsley PM, Felsheim RF, Kurtti TJ, Munderloh UG. Appl Environ Microbiol. 2010 Mar;76(6):1718-31. Epub 2010 Jan 22.

72. Plasmids of the pRM/pRF family occur in diverse Rickettsia species. Baldridge GD, Burkhardt NY, Felsheim RF, Kurtti TJ, Munderloh UG. Appl Environ Microbiol. 2008 Feb;74(3):645-52. Epub 2007 Dec 7.

73. Transposon insertion reveals pRM, a plasmid of Rickettsia monacensis. Baldridge GD, Burkhardt NY, Felsheim RF, Kurtti TJ, Munderloh UG. Appl Environ Microbiol. 2007 Aug;73(15):4984-95. Epub 2007 Jun 15.

74. Genomic, proteomic, and transcriptomic analysis of virulent and avirulent Rickettsia prowazekii reveals its adaptive mutation capabilities. Bechah Y, El Karkouri K, Mediannikov O, Leroy Q, Pelletier N, Robert C, Médigue C, Mege JL, Raoult D. Genome Res. 2010 May;20(5):655-63. Epub 2010 Apr 5.

75. The genome of Rickettsia prowazekii and some thoughts on the origin of mitochondria and hydrogenosomes. Müller M, Martin W. Bioessays. 1999 May;21(5):377-81.

76. Archaebacterial genomes: eubacterial form and eukaryotic content. Keeling PJ, Charlebois RL, Doolittle WF. Curr Opin Genet Dev. 1994 Dec;4(6):816-22.

77. GRS: a graphic tool for genome retrieval and segment analysis. Sicheritz-Pontén T, Andersson SG. Microb Comp Genomics. 1997;2(2):123-39.

78. Buchnera aphidicola (a prokaryotic endosymbiont of aphids) contains a putative 16S rRNA operon unlinked to the 23S rRNA-encoding gene: sequence determination, and promoter and terminator analysis. Munson MA, Baumann L, Baumann P. Gene. 1993 Dec 31;137(2):171-8.

79. Reductive evolution of resident genomes. Andersson SG, Kurland CG. Trends Microbiol. 1998 Jul;6(7):263-8.

80. Organization of the bacterial chromosome. S Krawiec and M Riley Microbiol Rev. 1990 December; 54(4): 502–539. PMCID: PMC372791

81. Patterns and processes of molecular evolution in rickettsia. Haleh Amiri. university dissertation from uppsala: Acta universitatis upsaliensis,2002

82. Evolutionary rates for tuf genes in endosymbionts of aphids. Brynnel EU, Kurland CG, Moran NA, Andersson SG. Mol Biol Evol. 1998 May;15(5):574-82.

83. Are noncoding sequences of Rickettsia prowazekii remnants of "neutralized" genes? Holste D, Weiss O, Grosse I, Herzel H. J Mol Evol. 2000 Oct;51(4):353-62.

84. Congruent evolution of different classes of non-coding DNA in prokaryotic genomes. Rogozin IB, Makarova KS, Natale DA, Spiridonov AN, Tatusov RL, Wolf YI, Yin J, Koonin EV. Nucleic Acids Res. 2002 Oct 1;30(19):4264-71.

85. KEG Database, http://www.genome.ad.jp/kegg/kegg2.html

86. The "domino theory" of gene death: gradual and mass gene extinction events in three lineages of obligate symbiotic bacterial pathogens. Dagan T, Blekhman R, Graur D. Mol Biol Evol. 2006 Feb;23(2):310-6. Epub 2005 Oct 19.

87. A Rickettsia genome overrun by mobile genetic elements provides insight into the acquisition of genes characteristic of an obligate intracellular lifestyle. Gillespie, J. J., Joardar, V., Williams, K. P., Driscoll, T., Hostetler, J. B., Nordberg, E., Shukla, M., Walenz, B., Hill, C. A., Nene, V. M., Azad, A. F., Sobral, B. W., Caler, E., Gillespie JJ, Joardar V, Williams KP, Driscoll T, Hostetler JB, Nordberg E, Shukla M, Walenz B, Hill CA, Nene VM, Azad AF, Sobral BW, Caler E. Journal of bacteriology. 2012 Jan 01; 194(2): 376-94.

88. Conjugation genes are common throughout the genus Rickettsia and are transmitted horizontally. Weinert LA, Welch JJ, Jiggins FM. Proc Biol Sci. 2009 Oct 22;276(1673):3619-27. Epub 2009 Jul 16.

89. Genome sequence of Rickettsia bellii illuminates the role of amoebae in gene exchanges between intracellular pathogens. Ogata H, La Scola B, Audic S, Renesto P, Blanc G, Robert C, Fournier PE, Claverie JM, Raoult D. PLoS Genet. 2006 May;2(5):e76. Epub 2006 May 12.

90. Gene gain and loss events in Rickettsia and Orientia species. Kalliopi Georgiades, Vicky Merhej, Khalid El Karkouri, Didier Raoult, and Pierre Pontarotti Biol Direct. 2011; 6: 6. Published online 2011 February 8. doi: 10.1186/1745-6150-6-6 PMCID: PMC3055210

91. Amoebae as training grounds for intracellular bacterial pathogens. Molmeret M, Horn M, Wagner M, Santic M, Abu Kwaik Y. Appl Environ Microbiol. 2005 Jan;71(1):20-8.

92. Bioinformatics Sheds Light On Evolutionary Origin Of Rickettsia Virulence Genes. Joseph J. Gillespie, Nicole C. Ammerman, Sheila M. Dreher-Lesnick, M. sayeedur, PHYSorg.com. 12 Mar 2009.

93. On the nature of gene innovation: duplication patterns in microbial genomes. Hooper SD, Berg OG. Mol Biol Evol. 2003 Jun;20(6):945-54. Epub 2003 Apr 25.

94. The Causes of Evolution by J. B. S. Haldane with a new afterword by Egbert G. Leigh, Jr. Princeton Science Library. Princeton University Press, 1990

95. An anomalous type IV secretion system in Rickettsia is evolutionarily conserved. Gillespie JJ, Ammerman NC, Dreher-Lesnick SM, Rahman MS, Worley MJ, Setubal JC, Sobral BS, Azad AF. PLoS One. 2009;4(3):e4833. Epub 2009 Mar 12.

96. Nucleotide sequence of the Rickettsia prowazekii ATP/ADP translocase-encoding gene. Williamson LR, Plano GV, Winkler HH, Krause DC, Wood DO. Gene. 1989 Aug 15;80(2):269-78.

97. Transcriptional regulation in the obligate intracytoplasmic bacterium Rickettsia prowazekii. Cai J, Winkler HH. J Bacteriol.1996 Sep;178(18):5543-5.

98. Selfish DNA in protein-coding genes of Rickettsia. Ogata H, Audic S, Barbe V, Artiguenave F, Fournier PE, Raoult D, Claverie JM. Science. 2000 Oct 13;290(5490):347-50.

99. Selfish DNA: the ultimate parasite. L. E. ORGEL & F. H. C. CRICK. Nature 284, 604- 607 (17 April 1980); doi:10.1038/284604a0

100. Selfish genes, the phenotype paradigm and genome evolution. Doolittle WF, Sapienza C. Nature. 1980 Apr 17;284(5757):601-3.

101. The impact of short interspersed elements (SINEs) on the host genome. Makalowski, W. 1995.

102. Selfish DNA: a sexually-transmitted nuclear parasite. Hickey DA. Genetics. 1982 Jul-Aug;101(3-4):519-31.

103. Perspective: transposable elements, parasitic DNA, and genome evolution. Kidwell MG, Lisch DR. Evolution. 2001 Jan;55(1):1-24.

104. Short-sequence DNA repeats in prokaryotic genomes. van Belkum A, Scherer S, van Alphen L, Verbrugh H. Microbiol Mol Biol Rev. 1998 Jun;62(2):275-93.

105. The evolution of bacterial genome. Herdman, M. 1985. In:T. Cavalier-Smith, Editor, the evolution of genome size, John Wiley & Sons pp. 37-68.

106. Direct demonstration of duplicate tuf genes in enteric bacteria. A V Furano Proc Natl Acad Sci U S A. 1978 July; 75(7): 3104–3108. PMCID: PMC392722

107. Rickettsia felis, from culture to genome sequencing. Ogata H, Robert C, Audic S, Robineau S, Blanc G, Fournier PE, Renesto P, Claverie JM, Raoult D.Ann N Y Acad Sci. 2005 Dec;1063:26-34.

108. Rubtsov AM, Lopina OD (2000) Ankyrins. FEBS Lett 482: 1-5.

109. AnkB, a periplasmic ankyrin-like protein in Pseudomonas aeruginosa, is required for optimal catalase B (KatB) activity and resistance to hydrogen peroxide. Howell ML, Alsabbagh E, Ma JF, Ochsner UA, Klotz MG, Beveridge TJ, Blumenthal KM, Niederhoffer EC, Morris RE, Needham D, Dean GE, Wani MA, Hassett DJ. J Bacteriol. 2000 Aug;182(16):4545-56.

110. AnkA: an Ehrlichia phagocytophila group gene encoding a cytoplasmic protein antigen with ankyrin repeats. Caturegli P, Asanovich KM, Walls JJ, Bakken JS, Madigan JE, Popov VL, Dumler JS. Infect Immun. 2000 Sep;68(9):5277-83.

111. The tetratricopeptide repeat: a structural motif mediating protein-protein interactions. Blatch GL, Lässle M. Bioessays. 1999 Nov;21(11):932-9.

112. Snap helix with knob and hole: essential repeats in S. pombe nuclear protein nuc2+. Hirano T, Kinoshita N, Morikawa K, Yanagida M. Cell. 1990 Jan 26;60(2):319-28.

113. Molecular characterization of a group of proteins containing ankyrin repeats in Orientia tsutsugamushi. Cho NH, Kim JM, Kwon EK, Kim SY, Han SH, Chu H, Lee JH, Choi MS, Kim IS. Ann N Y Acad Sci. 2005 Dec;1063:100-1.

114. Evidence in the Legionella pneumophila genome for exploitation of host cell functions and high genome plasticity. Cazalet C, Rusniok C, Brüggemann H, Zidane N, Magnier A, Ma L, Tichit M, Jarraud S, Bouchier C, Vandenesch F, Kunst F, Etienne J, Glaser P, Buchrieser C. Nat Genet. 2004 Nov;36(11):1165-73. Epub 2004 Oct 3.

115. Implication of Proteins Containing Tetratricopeptide Repeats in Conditional Virulence Phenotypes of Legionella pneumophila. Purnima Bandyopadhyay, Eren U. Sumer, Deepak Jayakumar, Shuqing Liu, Huifang Xiao and Howard M. Steinman. Published ahead of print 4 May 2012, doi: 10.1128/JB.00399-12. J. Bacteriol. July 2012 vol. 194 no. 14 3579-3588

116. Characterization of rickettsial attachment to host cells by flow cytometry. H Li and D H Walker Infect Immun. 1992 May; 60(5): 2030–2035. PMCID: PMC257111

117. rOmpA is a critical protein for the adhesion of Rickettsia rickettsii to host cells. Li H, Walker DH. Microb Pathog. 1998 May;24(5):289-98.

118. Identification of two putative rickettsial adhesins by proteomic analysis. Renesto P, Samson L, Ogata H, Azza S, Fourquet P, Gorvel JP, Heinzen RA, Raoult D. Res Microbiol. 2006 Sep;157(7):605-12. Epub 2006 Feb 28.

119. Adherence to and invasion of Vero cells by recombinant Escherichia coli expressing the outer membrane protein rOmpB of Rickettsia japonica. Uchiyama T. Ann N Y Acad Sci. 2003 Jun;990:585-90.

120. Identification of protective components of two major outer membrane proteins of spotted fever group Rickettsiae. Díaz Montero CM, Feng HM, Crocquet-Valdes PA, Walker DH. Am J Trop Med Hyg. 2001 Oct;65(4):371-8.

121. Ku70, a component of DNA-dependent protein kinase, is a mammalian receptor for Rickettsia conorii. Martinez JJ, Seveau S, Veiga E, Matsuyama S, Cossart P. Cell. 2005 Dec 16;123(6):1013-23.

122. Rickettsial outer-membrane protein B (rOmpB) mediates bacterial invasion through Ku70 in an actin, c-Cbl, clathrin and caveolin 2-dependent manner. Chan YG, Cardwell MM, Hermanas TM, Uchiyama T, Martinez JJ. Cell Microbiol. 2009 Apr;11(4):629-44. Epub 2009 Jan 7.

123. Rickettsia conorii entry into Vero cells. Teysseire N, Boudier JA, Raoult D. Infect Immun. 1995 Jan;63(1):366-74.

124. Expression of the Rickettsia prowazekii pld or tlyC gene in Salmonella enterica serovar Typhimurium mediates phagosomal escape. Whitworth T, Popov VL, Yu XJ, Walker DH, Bouyer DH. Infect Immun. 2005 Oct;73(10):6668-73.

125. Identification and characterization of a phospholipase D-superfamily gene in rickettsiae. Renesto P, Dehoux P, Gouin E, Touqui L, Cossart P, Raoult D. J Infect Dis. 2003 Nov 1;188(9):1276-83. Epub 2003 Oct 14.

126. A novel family of phospholipase D homologues that includes phospholipid synthases and putative endonucleases: identification of duplicated repeats and potential active site residues. Ponting CP, Kerr ID. Protein Sci. 1996 May;5(5):914-22.

127. Directed mutagenesis of the Rickettsia prowazekii pld gene encoding phospholipase D. Driskell LO, Yu XJ, Zhang L, Liu Y, Popov VL, Walker DH, Tucker AM, Wood DO. Infect Immun. 2009 Aug;77(8):3244-8. Epub 2009 Jun 8.

128. A comparative study of the actin-based motilities of the pathogenic bacteria Listeria monocytogenes, Shigella flexneri and Rickettsia conorii. Gouin E, Gantelet H, Egile C, Lasa I, Ohayon H, Villiers V, Gounon P, Sansonetti PJ, Cossart P. J Cell Sci. 1999 Jun;112 (Pt 11):1697-708.

129. Rickettsial actin-based motility: behavior and involvement of cytoskeletal regulators. Heinzen RA. Ann N Y Acad Sci. 2003 Jun;990:535-47.

130. Actin-dependent movement of bacterial pathogens. Stevens JM, Galyov EE, Stevens MP. Nat Rev Microbiol. 2006 Feb;4(2):91-101.

131. Actin-based motility of intracellular bacteria, and polarized surface distribution of the bacterial effector molecules. Carlsson F, Brown EJ. J Cell Physiol. 2006 Nov;209(2):288-96.

132. Comparison of Western immunoblotting and microimmunofluorescence for diagnosis of Mediterranean spotted fever. Teysseire N, Raoult D.J Clin Microbiol. 1992 Feb;30(2):455-60.

133. Directional actin polymerization associated with spotted fever group Rickettsia infection of Vero cells. R A Heinzen, S F Hayes, M G Peacock and T Hackstadt. Infect. Immun. May 1993 vol. 61 no. 5 1926-1935

134. Analysis of fluorescent protein expression in transformants of Rickettsia monacensis, an obligate intracellular tick symbiont. Baldridge GD, Burkhardt N, Herron MJ, Kurtti TJ, Munderloh UG. Appl Environ Microbiol. 2005 Apr;71(4):2095-105.

135. A Rickettsia WASP-like protein activates the Arp2/3 complex and mediates actin-based motility. Jeng RL, Goley ED, D'Alessio JA, Chaga OY, Svitkina TM, Borisy GG, Heinzen RA, Welch MD. Cell Microbiol. 2004 Aug;6(8):761-9.

136. Common themes in microbial pathogenicity revisited. Finlay BB, Falkow S. Microbiol Mol Biol Rev. 1997 Jun;61(2):136-69.

137. Surfing pathogens and the lessons learned for actin polymerization. Frischknecht F, Way M. Trends Cell Biol. 2001 Jan;11(1):30-38.

138. Bacterial type IV secretion: conjugation systems adapted to deliver effector molecules to host cells. Peter J Christiea and Joseph P Vogelb Trends in Microbiology, Volume 8, Issue 8, 354-360, 1 August 2000. doi:10.1016/S0966-842X(00)01792-3

139. Advances in rickettsia pathogenicity. Balraj P, Renesto P, Raoult D. Ann N Y Acad Sci. 2009 May;1166:94-105.

140. RickA Expression Is Not Sufficient to Promote Actin-Based Motility of Rickettsia raoultii. Premanand Balraj, Khalid El Karkouri, Guy Vestris, Leon Espinosa, Didier Raoult, and Patricia Renesto. PLoS ONE. 2008; 3(7): e2582. Published online 2008 July 9. doi: 10.1371/journal.pone.0002582. PMCID: PMC2440523

141. Discovery of protein-coding palindromic repeats in Wolbachia. Ogata H, Suhre K, Claverie JM. Trends Microbiol. 2005 Jun;13(6):253-5.

142. Genomic comparison of Rickettsia helvetica and other Rickettsia species. Dong X, El Karkouri K, Robert C, Gavory F, Raoult D, Fournier PE. J. Bacteriol. 2012 May;194(10):2751.

143. Genome Sequence of Rickettsia australis, the Agent of Queensland Tick Typhus. Dong X, El Karkouri K, Robert C, Raoult D, Fournier PE. J Bacteriol. 2012 Sep;194(18):5129.

144. Genomic Comparison of Rickettsia honei Strain RBT and Other Rickettsia Species. Xin Dong, El Karkouri K, Robert C, Raoult D, Fournier PE.J Bacteriol. 2012 Aug;194(15):4145.

145. Tandem repeat markers as novel diagnostic tools for high resolution fingerprinting of Wolbachia. Markus Riegler1, Iñaki Iturbe-Ormaetxe, Megan Woolfit, Wolfgang J Miller and Scott L O'Neill. BMC Microbiology 2012, 12(Suppl 1):S12 doi:10.1186/1471-2180-12-S1-S12

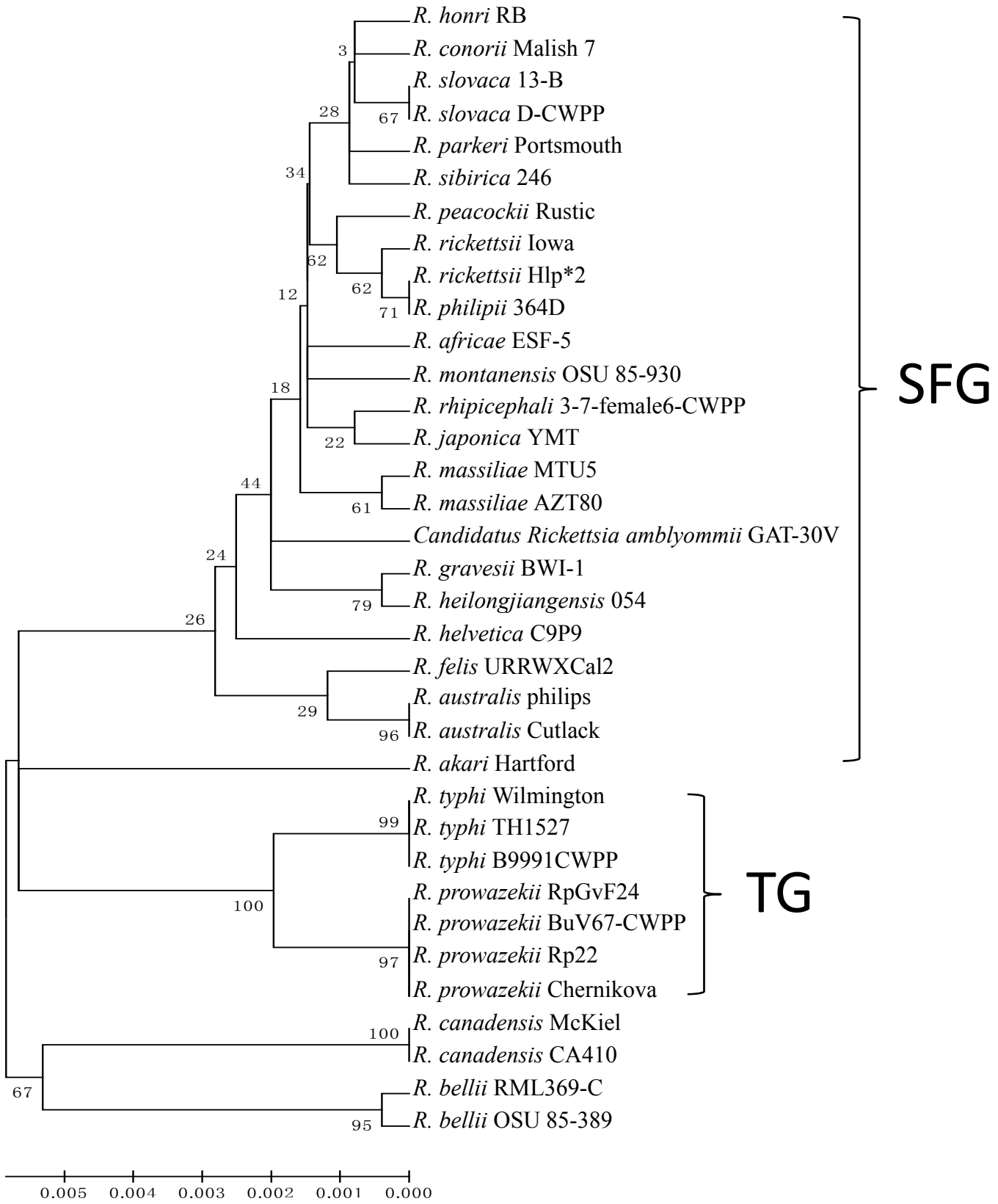

Fig.1. Phylogenetic tree of *Rickettsia* species besed on 16S rRNA. Phylogenetic tree is generated by MEGA4 neighbour-joining of 24 *Rickettsia* species.

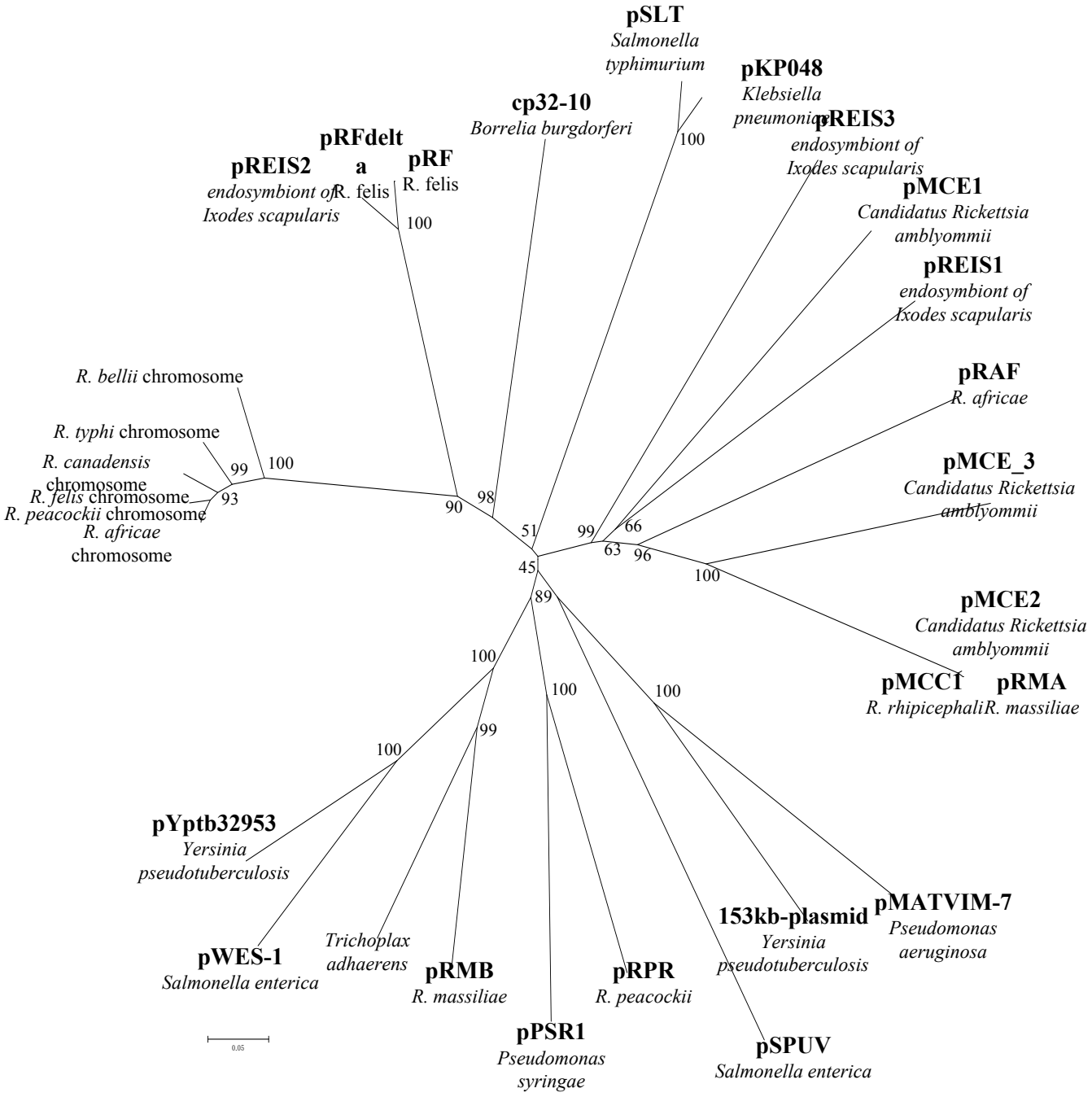

Fig.2. Maximum parsimony phylogenetic tree of rickettsial ParA proteins. GenBank accession numbers for the sequences used are as follows. Plasmid-encoded ParA sequences were from pMCE1, 2, 3 (NC_017020.1, NC_017029.1, NC_017021.1), pRAF (NC_012634.1), pREIS1, 2, 3 (NZ_CM000771.1, NZ_CM000772.1, NZ_CM000773.1), pRF (NC_007110.1), pRFdelta (NC_007111.1), pRMA (NC_009897.1), pRMB (NC_016939.1), pRPR (NC_012732.1), pMCC1 (NC_017055.1),cp32-10 (NC_017424.1), pPSR1 (NP_940697), pSLT (AE006471.1), pWes-1 (NC_011604.1), pSPUV (JN885081), pYptb32953 (NC_006154.1), pMATVIM7 (NC_009739.1), and a 153-kb plasmid (NC_009705.1), pKP048 (FJ628167), *Trichoplax adhaerens* (XP_002118475.1). Chromosome-encoded ParA sequences were from *R. africae* (YP_002844799), *R.bellii*(YP_538520), *R.canadensis* (YP_001491792), *R.felis* (YP_246133), *R. peacockii* (ACR47329), and *R. typhi* (YP_067042).

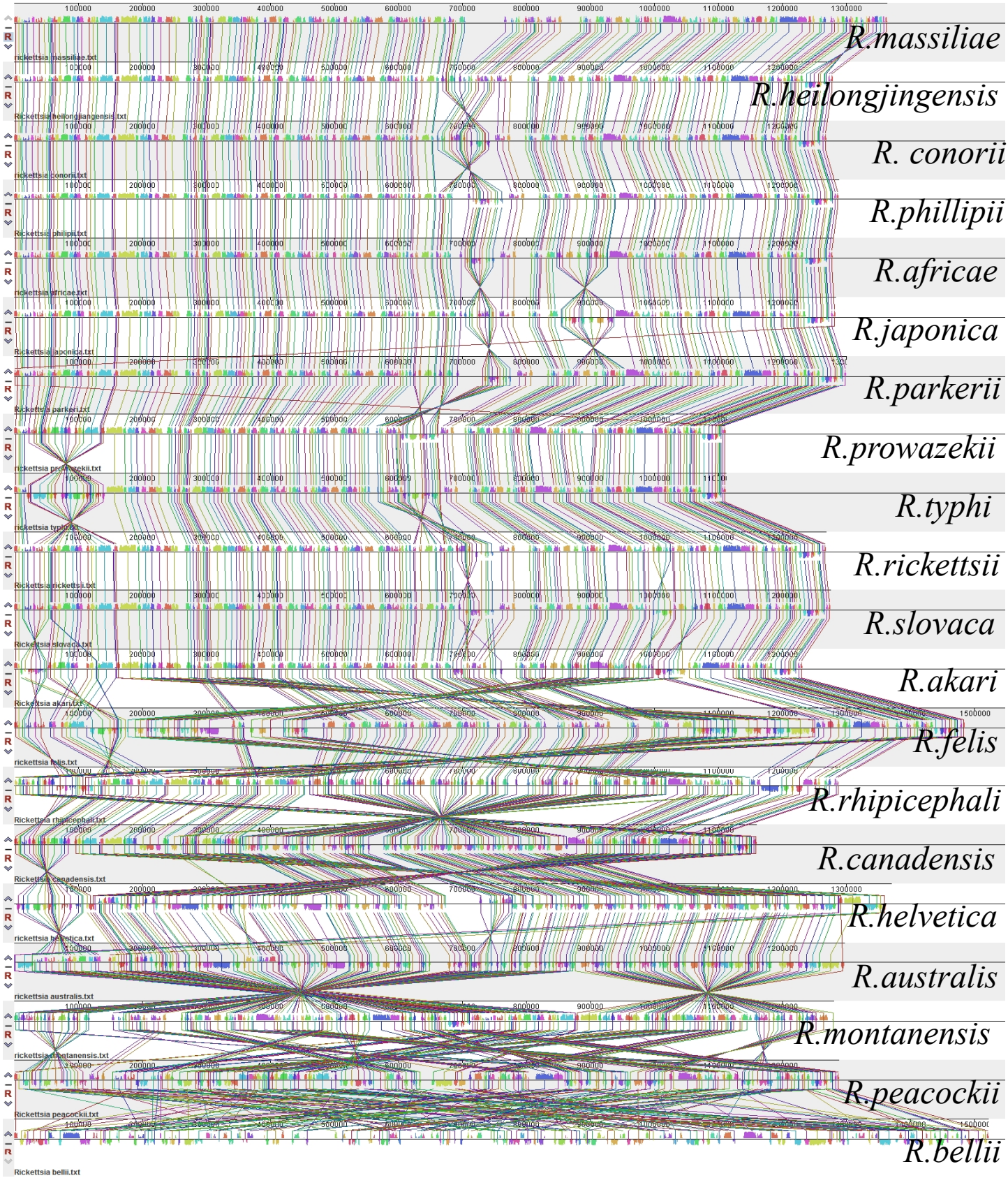

Fig.3. Whole genome alignment of *Rickettsia* species. The figure was generated by Mauve viewer (Darling *et al.*, 2004). This figure includes 20 Rickettsia genomes by order: *R. massiliae*, *R. heilongjiangensis*, *R. conorii*, *R. phillipii*, *R. africae*, *R. japonica*, *R. parkerii*, *R. prowazekii*, *R. typhi*, *R. rickettsii*, *R. slovaca*, *R. akari*, *R. felis*, *R. rhipicephali*, *R. canadensis*, *R. helvetica*, *R. australis*, *R. montanensis*, *R. peacockii* et *R. bellii*.

Table 1. Transmisson and diseases caused in humans by valided rickettsiae with their known pathogenic effects and distributions.

Pathogenic rickettsiae	Vector	Disease	Geographic Distribution
R. aeshlimannii	*-Hyalomma m. Marginatum* *-Hyalomma m.rufipes* *-Rhipicephalus appendiculatus* *-Haemaphysalis punctata*	Tick-transmitted disease	South Africa, Morocco, Mediterranean littoral
R. africae	*-Amblyomma hebraeum* *-Amblyomma variegatum* *-Rhipicephalus (Boophilus) decoloratus*	African tick bite fever	Sub-Saharan Africa, West Indies
R. akari	*-Allodermanyssus sanguineus* *-Liponyssoides sanguineus*	Rickettsial pox	Countries of the former Soviet Union, South Africa, Korea, Turkey, Balkan countries, North and South America
R. australis	*-Ixodes holocyclus* *-Ixodes tasmani* *-Ixodes cornuatus*	Queensland tick typhus	Australia, Tasmania
R. conorii subsp. *caspia*	*-Rhipicephalus pumilio* *-Rhipicephalus sanguineus*	Astrakham fever	North Caspian region of Russia
R. conorii subsp. *conorii*	*-Rhipicephalus sanguineus* *-Haemaphysalis leachii*	Mediterranean spotted fever	Southern Europe and Africa
R. conorii subsp. *indica*	*-Rhipicephalus sanguineus*	Indian tick typhus	South Asia
R. conorii subsp. *israelensis*	*-Rhipicephalus sanguineus* *-Amblyomma maculatum*	Israel tick typhus	Southern Europe and Middle East
R. felis	*-Ctenocephalides felis* *-Ctenocephalides canis* *-Archeopsylla erinacei* *-Pulex irritans* *-Xenopsylla cheopis* *-Anomiopsyllus nudata*	Flea-borne spotted fever	Europe, North and South America, Africa, Asia
R. heilongjiangensis	*-Dermacentor silvarum* *-Haemophysalis japonica douglasi* *-Haemophysalis concinna*	Far-eastern tick-borne rickettsiosis	Far East of Russia, Northern China, eastern Asia
R. helvetica	*-Ixodes ricinus* *-Ixodes ovatus* *-Ixodes persulcatus* *-Ixodes monospinosus*	Aneruptive fever	Central and northern Europe, Asia

Table 1. (*Cont.*)

Pathogenic rickettsiae	Vector	Disease	Geographic Distribution
R. honei	-*Aponomma hydrosauri* -*Ixodes granulatus*	Flinders Island spotted fever	Australia, Thailand
R. hoogstraalii	-*Haemaphysalis sulcata* -*Carios capensis*	Unknown pathogenicity	Croatia, Spain and Georgia, USA
R. japonica	-*Haemaphysalis flava* -*Haemaphysalis longicornis* -*Dermacentor taiwanensis* -*Ixodes ovatus*	Japanese or Oriental spotted fever	Japan
R. marmionii	-*Haemaphysalis novaeguinae* -*Ixodes holocytus*	Australian spotted fever	Australia
R. massiliae	-*Rhipicephalus sanguineus* -*Rhipicephalus turanicus* -*Rhipicephalus mushamae* -*Rhipicephalus lunulatus* -*Rhipicephalus sulcatus*	Spotted fever	France, Greece, Spain, Portugal, Switzerland, Siciliy, central Africa, and Mali
R. parkeri	-*Amblyomma maculatum* -*Amblyomma americanum*	Tick-transmitted disease skin lesions and lymphadenitis	North and South America
R. prowazekii	-*Pediculus humanus* -*Orchopeas howardii* -*Amblyomma cajennense*	Epidemic typhus, Brill-zinsser disease	Africa, Mexico, Central America, South America, Eastern Europe, Afghanistan, India, and China
R. raoultii	-*Dermacentor nutallii* -*Dermacentor silvarum* -*Dermacentor reticulatus* -*Dermacentor marginatus* -*Dermacentor niveus*	R. slovaca-like infection	France, Spain, Croatia, Russia and Kazakhstan
R. rickettsii	-*Dermacentor andersoni* -*Dermacentor variabilis* -*Rhipicephalus sanguineus* -*Amblyomma cajennense* -*Amblyomma aureolatum*	Rocky Mountain spotted fever, American spotted fever, or tick typhus	North, Central, and South America
R. sibirica subsp. *mongolitimonae*	-*Hyalomma asiaticum* -*Hyalomma truncatum*	Lymphangitis-associated rickettsioses	southern France, Greece, Spain, Portugal, South Africa, Egypt

Table 1. (*Cont.*)

Pathogenic rickettsiae	Vector	Disease	Geographic Distribution
R. sibirica subsp. *sibirica*	-*Dermacentor nutallii* -*Dermacentor marginatus* -*Dermacentor silvaum* -*Dermacentor pictus* -*Dermacentor sinicus* -*Dermacentor auratus* -*Haemaphysalis concinna*	Siberian tick typhus	Siberia and Far East, Asiatic part of Russia
R. slovaca	-*Dermacentor marginatus* -*Dermacentor reticulatus*	Tick-borne lymphadenopathy (Tibola) and Dermacentor-borne-necrosiserythema-lymphadenopathy (DEBONEL)	Southern and eastern Europe, Asia
R. typhi	-*Xenopsylla cheopis* -*Ctenocephalides felis* -*Leptopsylla segnis*	Endemic murine typhus	USA, Mediterranean area, Asia, Africa

Table 2. General features of available *Rickettsia* genomes.

Organism	BioProject	Chrs	plasmids	Size(Mb)	GC%	Gene	Protein
Candidatus Rickettsia amblyommii str. GAT-30V	PRJNA156845,PRJNA75043	1	3	1.48	32.4	1427	1390
Rickettsia africae ESF-5	PRJNA58799, PRJNA18269	1	1	1.29	32.4	1167	1041
Rickettsia akari str. Hartford	PRJNA58161, PRJNA12953	1	-	1.23	32.3	1293	1258
Rickettsia australis str. Cutlack	PRJNA158039, PRJNA75037	1	1	1.32	32.3	1297	1261
Rickettsia australis str. Phillips	PRJNA167984	1	1	1.32	32.3	-	-
Rickettsia bellii OSU 85-389	PRJNA58681, PRJNA17237	1	-	1.53	31.6	1511	1475
Rickettsia bellii RML369-C	PRJNA58405, PRJNA13996	1	-	1.52	31.6	1469	1429
Rickettsia canadensis str. CA410	PRJNA88063, PRJNA75029	1	-	1.15	31	1052	1016
Rickettsia canadensis str. McKiel	PRJNA58159, PRJNA12952	1	-	1.16	31.1	1126	1090
Rickettsia conorii str. Malish 7	PRJNA57633, PRJNA42	1	-	1.27	32.4	1414	1374
Rickettsia conorii subsp. caspia A-167	PRJNA156941	-	-	1.26	32.5	-	-
Rickettsia conorii subsp. indica ITTR	PRJNA89511	-	-	1.25	32.4	-	-
Rickettsia conorii subsp. israelensis ISTT CDC1	PRJNA156943	-	-	1.25	32.5	-	-
Rickettsia endosymbiont of Ixodes scapularis	PRJNA55851,PRJNA33979	1	3	2.1	33.3	2283	2117
Rickettsia felis URRWXCal2	PRJNA58331, PRJNA13884	1	2	1.59	32.6	1551	1512
Rickettsia heilongjiangensis 054	PRJNA70839, PRJNA66907	1	-	1.28	32.2	1338	1297
Rickettsia helvetica C9P9	PRJNA82855	1	1	1.37	32.2	-	-
Rickettsia honei RB	PRJNA158665	-	-	1.27	32.4	-	-
Rickettsia japonica YH	PRJNA73963, PRJDA38487	1	-	1.28	32.4	1010	971
Rickettsia massiliae MTU5	PRJNA58801, PRJNA18271	1	1	1	1.38	1436	980
Rickettsia massiliae str. AZT80	PRJNA86751, PRJNA75039	1	1	1	1.28	1243	1207

Table 2. (*Cont.*)

Organism	BioProject	Chrs	plasmids	Size(Mb)	GC%	Gene	Protein
Rickettsia montanensis str. OSU 85-930	PRJNA158043, PRJNA75045	1		1.28	32.6	1254	1217
Rickettsia parkeri str. Portsmouth	PRJNA158045, PRJNA75031	1	-	1.3	32.4	1355	1318
Rickettsia peacockii str. Rustic	PRJNA59301, PRJNA31309	1	1	1.31	32.6	984	947
Rickettsia philipii str. 364D	PRJNA89383, PRJNA75027	1	-	1.29	32.5	1380	1344
Rickettsia prowazekii Rp22	PRJNA161945, PRJNA19813	1	-	1.1	29	999	950
Rickettsia prowazekii str. BuV67-CWPP	PRJNA158063	1	-	1.1	29	879	843
Rickettsia prowazekii str. Chernikova	PRJNA158053	1	-	1.1	29	881	845
Rickettsia prowazekii str. Madrid E	PRJNA61565, PRJNA43	1	-	1.1	29	888	835
Rickettsia prowazekii str. RpGvF24	PRJNA158065	1	-	1.1	29	870	834
Rickettsia rhipicephali str. 3-7-female6-CWPP	PRJNA156977, PRJNA75041	1	1	1.31	32.4	1302	1266
Rickettsia rickettsii str. Hlp#2	PRJNA88067, PRJNA75015	1	-	1.27	32.5	1345	1308
Rickettsia rickettsii str. Iowa	PRJNA58961, PRJNA19943	1	-	1.27	32.4	1493	1384
Rickettsia sibirica 246	PRJNA54113,PRJNA1414	-	-	-	-	-	-
Rickettsia sibirica subsp. mongolitimonae HA-91	PRJNA82787	-	-	1.25	32.4	-	-
Rickettsia sibirica subsp. sibirica BJ-90	PRJNA80725	-	-	1.25	32.5	-	-
Rickettsia slovaca 13-B	PRJNA82369, PRJNA15712	1	-	1.28	32.5	1323	1112
Rickettsia slovaca str. D-CWPP	PRJNA158159, PRJNA75033	1	-	1.28	32.5	1386	1347
Rickettsia typhi str. B9991CWPP	PRJNA158357, PRJNA65221	1	-	1.11	28.9	875	839
Rickettsia typhi str. TH1527	PRJNA158161, PRJNA65219	1	-	1.11	28.9	874	838
Rickettsia typhi str. Wilmington	PRJNA58063, PRJNA10679	1	-	1.11	28.9	918	837

CHAPITRE III

Article I

Genomic Comparison of *Rickettsia helvetica* and Other *Rickettsia* Species

DONG Xin, Khalid El Karkouri, Catherine Robert, Frédérick Gavory, Didier Raoult, Pierre-Edouard Fournier

Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes,

CNRS UMR 6236, Faculté de médecine, Université de la Méditerranée

27 Boulevard Jean Moulin, 13385 Marseille cedex 05, France

Tel : (33).04.91.38.55.17, Fax : (33).04.91.83.03.90

E-mail: didier.raoult@gmail.com

Running head

Study of the *Rickettsia helvetica* genome

Word count: Abstract 56 words, text 569 words

Keywords: *Rickettsia helvetica*, genome

Abstract

The present study reports the complete and annotated genome sequence of *R. helvetica* strain C9P9 that was first isolated in 1979 from *Ixodes ricinus* ticks in Switzerland and is considered as a human pathogen. The comparison of this 1,36 Mb genome with other rickettsia genomes provides additional insight into the mechanisms of evolution in *Rickettsia* species.

Journal of Bacteriology

Genomic Comparison of Rickettsia helvetica and Other Rickettsia Species

Xin Dong, Khalid El Karkouri, Catherine Robert, Frédérick Gavory, Didier Raoult and Pierre-Edouard Fournier

J. Bacteriol. **2012, 194(10):2751. DOI: 10.1128/JB.00299-12.**

JB Journals.ASM.org

GENOME ANNOUNCEMENT

Genomic Comparison of *Rickettsia helvetica* and Other Rickettsia Species

Xin Dong,[a] Khalid El Karkouri,[a] Catherine Robert,[a] Frédérick Gavory,[b] Didier Raoult,[a] and Pierre-Edouard Fournier[a]

Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes, UMR CNRS 6236, IRD 198, Faculté de Médecine, Aix-Marseille Université, Marseille, France,[a] and Genoscope, Centre National de Séquençage, Evry, France[b]

We report the complete and annotated genome sequence of *Rickettsia helvetica* strain C9P9, which was first isolated in 1979 from *Ixodes ricinus* ticks in Switzerland and is considered a human pathogen.

Rickettsia helvetica was first isolated from *Ixodes ricinus* ticks in Switzerland in 1979. It was subsequently detected in many European countries, Russia, Japan, and Thailand. *R. helvetica* has been suggested to play a role in various human diseases, but its exact pathogenesis remains uncertain.

Genomic DNA extracted from *R. helvetica* strain C9P9 grown in Vero cells was pyrosequenced using the 454 GS FLX Titanium platform (Roche, Meylan, France) (8) and assembled using the Newbler software (Roche). Potential coding sequences (CDSs) were predicted using AMIGene (5), and split genes or nonpredicted genes were detected using Artemis (http://www.sanger.ac.uk/software/Artemis/) and BLASTN (1). Assignment of protein functions was performed by searching against the RickBase, GenBank, and Pfam databases using BLASTP (1, 4, 9), while rRNAs, tRNAs, and other RNAs were identified using BLASTN or tRNAscan-SE (7). Gene orthologs were identified using OrthoMCL (6) with a BLASTP E value cutoff of 1×10^{-5} and a Markov Clustering inflation parameter default of 1.5.

Thirty-two contigs were assembled into two scaffolds. The gaps were closed using Phusion DNA polymerase and specific primers. The *R. helvetica* genome consists of a 1,369,827-bp chromosome with a G+C content of 32.2%, in the range of other rickettsial genomes, and a 47,188-bp plasmid (pRhel) with a G+C content of 32.6%. The predicted total complement of 1,135 genes (1,515 open reading frames) includes 858 complete genes, 168 split genes, and 178 genes present only as fragments. Of these, 881 were assigned putative functions and 254 encode hypothetical proteins and proteins of unknown function. *R. helvetica* also contains 33 tRNAs, a single rRNA operon with noncontiguous 16S and 23S rRNAs, and 3 other RNAs. In addition, the pRhel plasmid encodes 49 genes (59 CDSs) but no tRNA.

The *R. helvetica* chromosome exhibits a high level of synteny with its closest phylogenetic neighbor, *R. massiliae* (3), with the exception of three inversions of 36,436 bp, 48,196 bp, and 17,375 bp, respectively. *R. helvetica* has 121 genes that are absent from *R. massiliae*, and the latter has 161 genes missing from *R. helvetica*. Most of these differentially present genes encode ankyrin repeat-, leucine-rich repeat-, or tetratricopeptide repeat-containing proteins, transposases, proteins of unknown function, and Tra family proteins. The *R. helvetica* genome appears more degraded than that of *R. massiliae* (168 versus 99 split genes and 178 versus 85 fragment genes, respectively). In comparison with the *R. prowazekii* genome (2), *R. helvetica* lacks only 36 genes, including the genes for 16 proteins of unknown function, three transposases, two ankyrin repeats, four transferases, two cell surface antigens, and one each for the integration host factor β subunit, a large extracellular α-helical protein, a monovalent cation/proton antiporter, peptide chain release factor RF-2, a prolyl endopeptidase precursor, pyrroloquinoline quinine biosynthesis protein C, signal peptide peptidase SppA, and a DNA invertase Pin-like protein. Thus, none of these genes is likely to be linked to rickettsial virulence. In contrast, many genes involved in the biosynthesis and regulation of biosynthesis of amino acids and nucleotides present in *R. helvetica* were absent from *R. prowazekii*. Further studies will be conducted to investigate the pathogenesis of *R. helvetica*.

Nucleotide sequence accession number. The *R. helvetica* strain C9P9 genome sequence and the pRhel plasmid sequence have been deposited in the DDBJ/EMBL/GenBank databases under accession no. AICO00000000.

ACKNOWLEDGMENT

This work was funded by the EuroPathoGenomics Network of Excellence.

REFERENCES

1. **Altschul SF, et al.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25**:3389–3402.
2. **Andersson SGE, et al.** 1998. The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria. Nature **396**:133–140.
3. **Blanc G, et al.** 2007. Lateral gene transfer between obligate intracellular bacteria: evidence from the Rickettsia massiliae genome. Genome Res. **17**:1657–1664.
4. **Blanc G, et al.** 2007. Reductive genome evolution from the mother of *Rickettsia*. PLoS Genet. **3**:e14.
5. **Bocs S, Cruveiller S, Vallenet D, Nuel G, Medigue C.** 2003. AMIGene: annotation of microbial genes. Nucleic Acids Res. **31**:3723–3726.
6. **Li L, Stoeckert CJ, Jr, Roos DS.** 2003. OrthoMCL: identification of ortholog groups for eukaryotic genomes. Genome Res. **13**:2178–2189.
7. **Lowe TM, Eddy SR.** 1997. t-RNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. **25**:955–964.
8. **Margulies M, et al.** 2005. Genome sequencing in microfabricated high-density picolitre reactors. Nature **437**:376–380.
9. **Punta M, et al.** 2012. The Pfam protein families database. Nucleic Acids Res. **40**:D290–D301.

Received 23 February 2012 **Accepted** 28 February 2012

Address correspondence to Pierre-Edouard Fournier, pierre-edouard.fournier@univmed.fr.

doi:10.1128/JB.00299-12

CHAPITRE IV

Article II

Genomic Comparison of *Rickettsia honei* Strain RB^T and Other *Rickettsia* Species

DONG Xin, Khalid El Karkouri, Catherine Robert, Didier Raoult,

Pierre-Edouard Fournier

Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes,

CNRS UMR 6236, Faculté de médecine, Université de la Méditerranée

27 Boulevard Jean Moulin, 13385 Marseille cedex 05, France

tel : (33).04.91.38.55.17, Fax : (33).04.91.83.03.90

E-mail: didier.raoult@gmail.com

Running head

Study of the *Rickettsia honei* genome

Word count: Abstract 64 words, text 504 words

Keywords: *Rickettsia honei*, genome

Abstract

R. honei strain RB^T has been isolated from a febrile patient on Flinders Island, Australia in 1991, and has been proposed for Flinders Island spotted fever (FISF) agent which is pathogenic for humans and is transmitted to human by tick bites. The comparison of this about 1,27 Mb genome with other Rickettsia genomes provides additional insight into the mechanisms of evolution in *Rickettsia* species.

Journal of Bacteriology

Genomic Comparison of Rickettsia honei Strain RB^T and Other Rickettsia Species

Dong Xin, Khalid El Karkouri, Catherine Robert, Didier Raoult and Pierre-Edouard Fournier

J. Bacteriol. **2012, 194(15):4145. DOI: 10.1128/JB.00802-12.**

JB Journals.ASM.org

GENOME ANNOUNCEMENT

Genomic Comparison of *Rickettsia honei* Strain RB^T and Other *Rickettsia* Species

Dong Xin, Khalid El Karkouri, Catherine Robert, Didier Raoult, and Pierre-Edouard Fournier

Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes, UMR CNRS 6236, IRD 198, Faculté de Médecine, Aix-Marseille Université, Marseille, France

***Rickettsia honei* strain RB^T was isolated from a febrile patient on Flinders Island, Australia, in 1991 and has been demonstrated to be the agent of Flinders Island spotted fever, a disease transmitted to humans by ticks. The comparison of this 1.27-Mb genome with other *Rickettsia* genomes provides additional insight into the mechanisms of evolution in *Rickettsia* species.**

The *Rickettsia* genus is composed of small, Gram-negative, obligate intracellular bacteria (1). These microorganisms underwent a reductive genomic evolution during their specialization to their intracellular lifestyle (2). Recent studies demonstrated that genome reduction was associated with virulence in rickettsiae (3). *R. honei* strain RB^T was isolated from a febrile patient on Flinders Island, Australia, in 1991 and has been described as a Flinders Island spotted fever (FISF) agent which is pathogenic for humans (4). FISF is characterized by fever, headache, myalgia, transient arthralgia, maculopapular rash, and in some cases cough (6). *R. honei* strain RB^T is transmitted to human mainly by *Bothriocroton hydrosauri* (formerly *Aponomma hydrosauri*) tick bites (5). Most cases occur in summer.

The genome sequencing of *R. honei* strain RB^T was performed by 454 shotgun sequencing. Briefly, the shotgun sequencing was performed using a GS-FLX Titanium sequencer (Roche, Meylan, France) and assembled into 11 contigs. All 11 contigs were part of the chromosome, with a size of 1,268,758 bp and a G+C content of 32.4%, which is similar to other rickettsial genomes. No plasmid was detected. The chromosome was predicted to encode 1,284 genes (1,595 open reading frames [ORFs]). Among these genes, 1,046 (81%) were complete, 158 (12%) were split into two or more ORFs, and 80 (6%) were present as fragments. Of the 1,284 genes, 751 (58%) encoded proteins with putative functions, and 533 (42%) encoded hypothetical proteins and proteins of unknown function. The *R. honei* genome had 3 noncontiguous rRNA genes (5S, 16S, and 23S rRNA), 33 tRNAs, and 3 other RNAs.

Phylogenically, *R. honei* is closely related to *R. rickettsii*, *R. conorii*, and *R. slovaca* (4). The 11 *R. honei* contigs exhibited an almost perfect colinearity with these 3 genomes, with the exception of an inversion of 34,133 bp, 81,501bp and 32,367 bp, respectively. By comparison with the closest genome, that of *R. conorii*, *R. honei* missed 56 genes but had an additional 89 specific genes. Most of these differentially present genes encoded ankyrin repeat-containing proteins, leucine-rich repeats (LRRs), tetratricopeptide repeat-containing proteins, transposases, and proteins of unknown function. By comparison with *R. prowazekii*, the agent of epidemic typhus, the most pathogenic *Rickettsia* species with the smallest genome (834 genes only), *R. honei* missed only 37 genes. These genes encoded proteins of unknown function ($n = 17$), transposases ($n = 3$), ankyrin repeat-containing proteins ($n = 2$), transferases ($n = 6$), and synthetases ($n = 3$) and one from each of the following categories: cell surface antigens, BioY family proteins, multisubunit Na^+/H^+ antiporters, protein kinase C inhibitors, Sec7 domain-containing proteins, site-specific recombinases, and VirD4 proteins. Thus, none of these genes is likely to play a role in rickettsial virulence. In contrast, many genes involved in the biosynthesis and regulation of biosynthesis of amino acids and nucleotides present in *R. honei* were absent from *R. prowazekii*. Further studies will be conducted to investigate whether a reduced metabolism of amino acids and/or nucleotides plays a role in the pathogenesis of rickettsiae.

Nucleotide sequence accession number. The genome sequence has been deposited in the GenBank database under accession number AJTT00000000.

ACKNOWLEDGMENT

This work did not benefit from any external funding.

REFERENCES

1. **Blanc G, et al.** 2007. Lateral gene transfer between obligate intracellular bacteria: evidence from the *Rickettsia massiliae* genome. Genome Res. **17**: 1657–1664.
2. **Blanc G, et al.** 2007. Reductive genome evolution from the mother of *Rickettsia.* PLoS Genet. **3**:e14. doi:10.1371/journal.pgen.0030014.
3. **Fournier PE, et al.** 2009. Analysis of the *Rickettsia africae* genome reveals that virulence acquisition in *Rickettsia* species may be explained by genome reduction. BMC Genomics **10**:166. doi:10.1186/1471-2164-10-166.
4. **Stenos J, Roux V, Walker D, Raoult D.** 1998. *Rickettsia honei* sp. nov., the aetiological agent of Flinders Island spotted fever in Australia. Int. J. Syst. Bacteriol. **48**:1399–1404.
5. **Stenos J, Graves SR, Popov VL, Walker DH.** 2003. *Aponomma hydrosauri*, the reptile-associated tick reservoir of *Rickettsia honei* on Flinders Island, Australia. Am. J. Trop. Med. Hyg. **69**:314–317.
6. **Stewart RS.** 1991. Flinders Island spotted fever: a newly recognized endemic focus of tick typhus in Bass Strait, part 1: clinical and epidemiological features. Med. J. Aust. **154**:94–99.

Received 7 May 2012 Accepted 21 May 2012

Address correspondence to Pierre-Edouard Fournier, pierre-edouard.fournier@univmed.fr.

doi:10.1128/JB.00802-12

CHAPITRE V

Article III

Genome Sequence of *Rickettsia australis*, the Agent of Queensland Tick Typhus

DONG Xin, Khalid El Karkouri, Catherine Robert, Didier Raoult, Pierre-Edouard Fournier

Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes, CNRS UMR 6236, Faculté de médecine, Université de la Méditerranée

27 Boulevard Jean Moulin, 13385 Marseille cedex 05, France

tel : (33).04.91.38.55.17, Fax : (33).04.91.83.03.90

E-mail: didier.raoult@gmail.com

Running head

Study of the *Rickettsia australis* genome

Word count: Abstract 57 words, text 521 words

Keywords: *Rickettsia australis*, genome

Abstract

R. australis strain phillips^{T} has been isolated in Queensland Australia 1950, and has been proposed for Queensland Tick Typhus (QTT) agent which is pathogenic for humans and is transmitted to human by tick bites. The comparison of this about 1,29 Mb genome with other Rickettsia genomes provides additional insight into the mechanisms of evolution in *Rickettsia* species.

Journal of Bacteriology

Genome Sequence of Rickettsia australis, the Agent of Queensland Tick Typhus

Xin Dong, Khalid El Karkouri, Catherine Robert, Didier Raoult and Pierre-Edouard Fournier

J. Bacteriol. **2012, 194(18):5129. DOI: 10.1128/JB.01117-12.**

JB Journals.ASM.org

GENOME ANNOUNCEMENT

Genome Sequence of *Rickettsia australis*, the Agent of Queensland Tick Typhus

Xin Dong, Khalid El Karkouri, Catherine Robert, Didier Raoult, and Pierre-Edouard Fournier

Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes, Institut Hospitalo-Universitaire Méditerranée-Infection, UMR CNRS 7278-IRD 198, Faculté de Médecine, Aix-Marseille Université, Marseille, France

***Rickettsia australis* strain Phillips^T was isolated in Queensland, Australia, in 1950. It is the tick-borne agent of Queensland tick typhus, a disease endemic in Australia. The 1.29-Mb genome sequence of this bacterium is highly similar to that of *Rickettsia akari* but contains two plasmids.**

The genus *Rickettsia* is composed of small Gram-negative, obligate intracellular alphaproteobacteria (1) that underwent progressive genomic reduction (2). However, paradoxically, recent genomic studies have suggested that genome reduction was associated with increased virulence in rickettsiae (3). *Rickettsia australis* was first identified in residents of Northern Queensland, Australia, in 1950 (4). This bacterium is a spotted fever group rickettsia and causes Queensland tick typhus (QTT). QTT is endemic in the eastern part of Australia, where it is transmitted to humans through the bites of *Ixodes holocyclus* or *Ixodes tasmani* ticks (6). It is considered to be mostly mild and is characterized by fever, headache, and myalgia followed by the development of a maculopapular or vesicular rash, an inoculation eschar (65% of cases), and lymphadenopathy (71%) (5).

The genome sequencing of *R. australis* strain Phillips^T was performed by 454 shotgun and 454 paired-end sequencing. Briefly, the shotgun sequencing was performed using a GS-FLX Titanium sequencer (Roche, Meylan, France) with assembly into 1 complete circularized chromosome, 1 complete circularized plasmid, and 1 putative incomplete plasmid. The chromosome has a predicted size of 1,297,390 bp and a G+C content of 31.7%, which is similar to that of other rickettsial genomes. The plasmid (pRau01) has a size of 26,608 bp with a G+C content of 33.7%. The second plasmid is made of 13 contigs, for a total size of 30,176 bp, and has a G+C content of 21.7%. The chromosome contains 1,110 genes (1,426 open reading frames [ORFs]), including 820 complete genes (74%), 139 split genes (12%), and 151 gene fragments (14%). Among these 1,110 genes, 855 genes (77%) encode proteins with putative functions and 255 (23%) encode hypothetical proteins and proteins of unknown function. The *R. australis* genome contains 3 noncontiguous rRNAs (5S, 16S, and 23S rRNA), 33 tRNAs, and 3 other RNAs. The pRau01 plasmid carries 16 genes (25 ORFs), whereas the second plasmid carries 20 genes (24 ORFs).

R. australis exhibits an almost perfect genomic synteny with *Rickettsia akari*, its closest phylogenetic neighbor, with the exception of two inversions of 78,519 bp and 36,776 bp. *R. australis* lacks only 46 genes that are present in *R. prowazekii*, the agent of epidemic typhus, the most severe human rickettsiosis (7). These genes encode proteins of unknown function (23 genes), transposases (3 genes), ankyrin repeat-containing proteins (2 genes), transferases (6 genes), cell surface antigens (1 gene), synthetases (3 genes), and 1 gene in each of the following categories: BioY family protein, ABC-type transporter related to toluene tolerance protein, dCTP deaminase, putative transcriptional regulator, Sco2 protein precursor, DNA invertase Pin-like protein, VirD4 protein, and beta-glucosidase. None of these genes appears to be related to rickettsial virulence. In contrast, many genes involved in the biosynthesis and regulation of biosynthesis of amino acids and nucleotides present in *R. australis* are absent from *R. prowazekii*. Further investigations will be conducted to study the genomic basis of rickettsial virulence and the potential role of plasmids.

Nucleotide sequence accession number. The whole-genome shotgun project has been deposited in the GenBank database under accession number AKVZ00000000.

ACKNOWLEDGMENT

This work did not benefit from any external funding.

REFERENCES

1. **Blanc G, et al.** 2007. Lateral gene transfer between obligate intracellular bacteria: evidence from the *Rickettsia massiliae* genome. Genome Res. **17**: 1657–1664.
2. **Blanc G, et al.** 2007. Reductive genome evolution from the mother of *Rickettsia.* PLoS Genet. **3**:e14. doi:10.1371/journal.pgen.0030014.
3. **Fournier PE, et al.** 2009. Analysis of the *Rickettsia africae* genome reveals that virulence acquisition in *Rickettsia* species may be explained by genome reduction. BMC Genomics **10**:166.
4. **Graves S, Stenos J.** 1999. Rickettsioses in Australia, p 244–246. *In* Raoult D, Brouqui P (ed), Rickettsiae and rickettsial diseases at the turn of the third millenium. Elsevier, Paris, France.
5. **McBride WJ, Hanson JP, Miller R, Wenck D.** 2007. Severe spotted fever group rickettsiosis, Australia. Emerg. Infect. Dis. **13**:1742–1744.
6. **Sexton DJ, Dwyer B, Kemp R, Graves S.** 1991. Spotted fever group rickettsial infections in Australia. Rev. Infect. Dis. **13**:876–886.
7. **Walker DH.** 1988. Pathology and pathogenesis of the vasculotropic rickettsioses, p 115–138. *In* Walker DH (ed), Biology of rickettsial disease. CRC Press Inc., Boca Raton, FL.

Received 22 June 2012 Accepted 29 June 2012

Address correspondence to Pierre-Edouard Fournier, pierre-edouard.fournier@univmed.fr.

doi:10.1128/JB.01117-12

CHAPITRE VI

Article IV

Genomic analysis of *Rickettsia japonica* strain YH^T

DONG Xin, Khalid El Karkouri, Catherine Robert, Didier Raoult, Pierre-Edouard Fournier

Aix Marseille Université, URMITE, UM63, CNRS 7278,

IRD 198, Inserm 1095, Faculté de Médecine

27 Boulevard Jean Moulin, 13385 Marseille cedex 05, France

tel : +33 (0)4 91 32 43 75, Fax : + 33 (0)4 91 38 77 72

E-mail: didier.raoult@gmail.com

Running head

Study of the *Rickettsia japonica* genome

Word count: Abstract 51 words, text 483 words

Keywords: *Rickettsia japonica*, genome

Abstract

Rickettsia japonica strain YH^T was reported in 1984 in Japan, and has been proposed for Japanese spotted fever (JSF) agent and is transmitted to human by tick bites. The comparison of this about 1,33 Mb genome with other *Rickettsia* genomes provides additional insight into the mechanisms of evolution in *Rickettsia* species.

The genus *Rickettsia* is composed of small Gram negative, obligate intracellular bacteria (1). These microorganisms underwent a reductive genomic evolution during their specialization to their intracellular lifestyle (2). The recent studies demonstrated that genome reduction was associated to virulence in rickettsiae (3). In Japan, *R. japonica*, classified within the spotted fever group, is the causative agent of Japanese spotted fever (JSF) (4). The first JSF patient was reported in 1984 in Tokushima Prefecture. According to the national surveillance system in Japan, JSF cases resulting in death have been gradually increasing (5, 7). The JSF is found mainly in the southwestern part of Japan (7). Recent epidemiologic evidences indicated that *Dermacentor taiwanensis* and *Haemaphysalis flava* can be both vector and reservoirs of *R. japonica* (6).

The genome sequencing of *R. japonica* strain YH^T was performed by 454 paired-end sequencing. Briefly, the sequencing result was performed by GS-FLX Titanium sequencer (Roche, Meylan, France). The genome of *R. japonica* strain YH^T was assembled to one chromosomic scaffold which contains 29 contigs (range from 951 bp to 197,977 bp), and one complete circularized plasmid, pRjap. All 29 chromosomic contigs are predicted with a size of 1,331,743 bp and G+C content of 32.7% which is similar to other rickettsial genomes. The plsmid, pRjap, has a size of 19,854 bp and a G+C content of 31.8%. The chromosome has been predicted of 1,239 genes (1,515 ORFs) and encodes 19 genes (21 ORFs). Among these genes, 889 genes (71,8%) are complete, 167 genes (13,5%) are present as fragment, and 183

genes (14,7%) are split genes. All over this genes, 894 genes (72,2%) encode proteins with putative functions, 345 genes (27,8%) encode hypothetical proteins or proteins of unknown function. We also identified 3 rRNA genes (5S, 16S and 23S rRNA), 33 tRNAs and 3 other RNAs.

Based on various molecular markers, the phylogenetic analsis indicates that *R. rhipicephili* is the closest species to *R. japonica*. Except for 4 different inversion regions (13,405 bp, 4,785 bp, 65,590 bp and 38,431 bp), the *R. japonica* genome exhibits a most perfect colinearity with *R. rhipicephili*. Furthermore, we compared the *R. japonica* genome with the *R. prowazekii* genome which causes the most serious epidemic typhus to human (8). This two genomes share 800 common genes. The *R. japonica* genome only missed 40 genes that were present in *R. prowazekii*. This genes encode proteins of unknown function (21), transposases (3), ankyrin repeats containing proteins (2), transferases (5), and one gene of each of the following categories: BioY family protein, Cell surface antigen Sca6 protein, LPS biosynthesis protein, Na+/proline symporter histidine kinase, Poly-beta-hydroxybutyrate polymerase, DNA invertase Pin-like protein, Twin-arginine translocation protein TatA, Universal stress protein UspA and VirD4 protein. But no rickettsial virulence factor have been identified. In contrast, many genes involved in the biosynthesis and regulation of biosynthesis of amino acids and nucleotides present in *R. japonica* were absent from *R. prowazekii*. To reveal the pathogenicity of *R. japonica*, more further studies will be conducted.

Nucleotide sequence accession number.

This Whole Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession AMRT00000000.

REFERENCES

1. Blanc, G., H. Ogata, et al. (2007). "Lateral gene transfer between obligate intracellular bacteria: evidence from the Rickettsia massiliae genome." Genome Res 17(11): 1657-1664.

2. Blanc, G., H. Ogata, et al. (2007). "Reductive genome evolution from the mother of Rickettsia." PLoS Genet 3(1): e14.

3. Fournier, P. E., K. K. El, Q. Leroy, C. Robert, B. Giumelli, P. Renesto, C. Socolovschi, P. Parola, S. Audic, and D. Raoult. 2009. "Analysis of the Rickettsia africae genome reveals that virulence acquisition in Rickettsia species may be explained by genome reduction." BMC. Genomics 10:166

4. Mahara, F., K. Koga, S. Sawaga, T. Taniguchi, F. Shigemi, T. Suto, Y. Tsuboi, A. Ooya, H. Koyama, T. Uchiyama, and T. Uchida. 1985. The first report of the rickettsial infection of spotted fever group in Japan: three clinical cases. Jpn. J. Assoc. Infect. Dis. 59:1165–1172.

5. Hanaoka N, Matsutani M, Kawabata H, Yamamoto S, Fujita H, Sakata A, Azuma Y, Ogawa M, Takano A, Watanabe H, Kishimoto T, Shirai M, Kurane I, Ando S. Diagnostic assay for Rickettsia japonica. Emerg Infect Dis. 2009 Dec;15(12):1994-7. National Institute of Infectious Diseases, Tokyo 162-8640, Japan.

6. Ishikura, M., H. Fujita, S. Ando, K. Matsuura, and M. Watanabe. 2002. Phylogenetic analysis of spotted fever group rickettsiae isolated from ticks in Japan. Microbiol. Immunol. 46:241–247.

7. Uchida, T., Y. Tsuboi, A. Oya, T. Funato, and Y. Kitamura. 1986. Serologic studies of spotted fever group rickettsiosis occurring in Muroto-City, Kochi Prefecture. Kansenshogaku Zasshi 60141-144. (In Japanese.)

8. Walker, D.H. (1988). Pathology and pathogenesis of the vasculotropic rickettsioses, *in:* D.H. Walker (Ed.), *Biology of Rickettsial Disease*, Florida: CRC Press Inc. 115-138.

CHAPITRE VII

Conclusion générale

Cette thèse est organisée en deux parties majeures. La première partie consiste en une revue de la littérature sur les génomes de bactéries du genre *Rickettsia*, notamment sur leurs caractéristiques générales et leurs mécanismes évolutifs. Le stress évolutif a conduit à une réduction génomique extrême des génomes de *Rickettsia*. Différents processus évolutifs tels que des réarrangements génomiques, la perte de l'information génomique et des duplications de gènes ont été identifiés. La perte de gènes est très supérieure à l'acquisition de gènes.

Les génomes de rickettsies ont été l'objet de réarrangements importants résultant de recombinaisons intra-chromosomiques au niveau de séquences répétées inversement orientées au cours de leur adaptation à un mode de vie intracellulaire obligatoire (Figure 4). Ainsi, les opérons 16S-23S-5S rRNA ont des organisations différentes de celles des autres genres bactériens et des archae.

L'évolution réductrive est le processus d'évolution majeure qu'ont subi les génomes de *Rickettsia*. Cette réduction a été le résultat de recombinaisons homologues entre les séquences répétées orientés à un stade précoce. En réduisant leur taille génomique, la masse des séquences répétées a diminué et taux de recombinaison homologue a également diminué. Cependant les espèces de *Rickettsia* présentent également des cas d'expansion génomique résultant de mécanismes de transfer latéral de gènes, de duplication de gènes ou de fragments d'ADN égoïste (rickettsial palindromic elements, ankyrin repeats, tetratricopeptide repeats, modules

toxine-antitoxine). L'analyse phylogénique suggère que les événements réductifs se produisent plus souvent que les processus expansifs chez Rickettsia. Après la réduction des séquences répétées, les génomes de *Rickettsia* continueraient la dégradation par des mutations ponctuelles qui donneraient lieu à la formation de pseudogènes et de fragments de gènes (Figure 2). L'analyse phylogénomique des rickettsies des groupes TG et SFG indique que de nombreux gènes dégradés présents chez certaines espèces sous forme de pseudogènes, de fragments ou ayant totalement disparu, peuvent toujours être fonctionnels dans d'autres espèces. C'est le cas des gènes *ampG4*, *fic* et *gadD* présents intacts dans le groupe SFG, mais pseudogènes ou très dégradés dans le groupe TG. Ce processus a pu jouer un rôle important chez les bactéries extra-cellulaires, leur permettant d'adapter la taille de leur génome au mode de vie intracellulaire. En outre, à l'inverse du concept classique que les bactéries deviennent pathogènes par l'acquisition des facteurs de virulence, la pathogénicité des *Rickettsia* est liée à la réduction de leurs génomes par perte de certains métabolismes et voies de régulation.

La seconde partie de notre thèse est consacrée à l'étude génomique de quatre espèces de *Rickettsia*. En utilisant les technologies de séquençage à haut débit : SOLiD, Roche 454 shotgun et paired-end, nous avons séquencé les génomes de *R. helvetica* strain C9P9, *R. honei* strain RB^T, *R. australis* strain $phillips^T$ et *R. japonica* strain YH^T. Ces 4 espèces de *Rickettsia* sont des pathogènes humains recoonus. Dans ces 4 génomes ont été identifiés

3 gènes ARN ribosomiques non contigus, 33 ARN de transfert et 3 autres ARN qui sont très similaires entre espèces.

R. helvetica a un génome composé d'un chromosome de 1,37 Mb et d'un plasmide pRhel de 47 Kb. Le chromosome comprend 1273 gènes (1619 ORFs), dont 923 (72%) sont complets, 203 (16%) sont interrompus par un ou plus codons stops, et 147 (12%) sont présents sous forme de fragments. pRhel comporte 49 gènes (59 ORFs). Ce travail a été publié dans Journal of Bacteriology (DONG X et al 2012 May).

Le génome de *R. honei* consiste en un scaffold chromosomique de 1,27 Mb (11 contigs). Aucun plasmide n'a été identifié. Le chromosome comporte 1284 gènes (1595 ORFs), dont 1046 (81%) sont complets, 158 (12%) sont des pseudogènes et 80 (6%) sont présents sous forme de fragments. Ce travail a publié dans Journal of Bacteriology (DONG X et al 2012 Aug).

Le génome de *R. australis* est formé d'un chromosome complet de 1,29 Mb, d'un plasmide pRau de 26 Kb et d'un plasmide putatif incomplet de 30 Kb (13 contigs). Le chromosome comporte 1110 gènes (1426 ORFs) dont 820 sont complets (74%), 151 sont présents sous forme de fragments (14%) et 139 sont des pseudogènes (12%). Le plasmide pRau comporte 16 gènes (25 ORFs). Le plasmide putatif comprend 20 gènes (24 ORFs). Ce travail a été publié dans Journal of Bacteriology (DONG X *et al* 2012 Sep).

Le génome de R. japonica consiste en un scaffold chromosomique de 1,33 Mb (29 contigs) et un plasmid pRjap de 20 Kb. Le chromosome

comporte 1,239 genes (1,515 ORFs) dont 889 genes sont completes (71,8%), 167 genes sont présentssous forme de fragments (13,5%), et 183 genes sont des pseudogènes (14,7%). Ce travail a été accepté dans Journal of Bacteriology récemment.

Reference

1. The emerging diversity of Rickettsia, Steve J Perlman, Martha S Hunter, Einat Zchori-Fein, Proc Biol Sci. 2006 September 7; 273(1598): 2097–2106. Published online 2006 April 21. doi: 10.1098/rspb.2006. 3541, PMCID: PMC1635513

2. The genome sequence of Rickettsia prowazekii and the origin of mitochondria. Andersson SG, Zomorodipour A, Andersson JO, Sicheritz-Pontén T, Alsmark UC, Podowski RM, Näslund AK, Eriksson AS, Winkler HH, Kurland CG. Nature. 1998 Nov 12;396(6707):133-40.

3. Durack, D.T., et al., Hellenic Holocaust: A historical clinico-pathologic conference. American Journal of Medicine, 2000. 109: p. 391-397.

4. Hackstadt, T., The biology of Rickettsiae. Infectious Agents and Disease, 1996. 5: p. 127-143.

5. Winkler, H.H., Rickettsia species (as organisms). Annual Reviews in Microbiology, 1990. 44: p. 131-153.

6. Res Microbiol. 1992 Nov-Dec;143(9):821-9. Intracellular movements of *Rickettsia conorii* and *R. typhi* based on actin polymerization. Teysseire N, Chiche-Portiche C, Raoult D.

7. Bior, A.D., R.C. Essenberg, and J.R. Sauer, Comparison of differentially expressed genes in the salivary glands of male ticks, Amblyomma americanum and Dermacentor andersoni. Insect Biochemistry and Molecular Biology, 2002. 32: p. 645-655.

8. Kollars, T.M., Jr., Interspecific differences between small mammals as hosts of immature Dermacentor variabilis (Acari: Ixodidae) and a model for detection of high risk areas of Rocky Mountain spotted fever. Journal of Parasitology, 1996. 82(5): p. 707-710.

9. Noden, B.H., et al., Molecular identification of Rickettsia typhi and R. felis in co-infected Ctenocephalides felis (Siphonaptera: Pulicidae). Journal of Medical Entomology, 1998. 35(4): p. 410-414.

10. Azad, A.F. and C.B. Beard, Rickettsial pathogens and their arthropod vectors. Emerging Infectious Diseases, 1998. 4(2): p. 179-186.

11. Ogata, H., et al., Mechanisms of evolution in *Rickettsia conorii* and *R. prowazekii*. Science, 2001. 293: p. 2093-2098.

12. Raoult, D. and V. Roux, Rickettsioses as paradigms of new or emerging infectious diseases. Clinical Microbiology Reviews, 1997. 10(4): p. 694-719.

13. Centers for Disease Control and Prevention, Summary of Notifiable Diseases - United States, 2004. Morbidity and Mortality Weekly Reports, 2004. 53(53).

14. Infect Agents Dis. 1996 Jan;5(1):1-7. Pathogenicity islands and the evolution of bacterial pathogens. Lee CA.

15. Roux, V. and D. Raoult, Phylogenetic analysis of members of the genus Rickettsia using the gene encoding the outer-membrane protein rOmpB (ompB). International Journal of Systematic and Evolutionary Microbiology, 2000. 50: p. 1449-1455.

16. Sekeyova, Z., V. Roux, and D. Raoult, Phylogeny of Rickettsia spp. inferred by comparing sequences of 'gene D', which encodes an intracytoplasmic protein. International Journal of Systematic and Evolutionary Microbiology, 2001. 51: p. 1353-1360.

17. Blanc, G., et al., Molecular evolution of Rickettsia surface antigens: evidence of positive selection. Molecular Biology and Evolution, 2005. 22(10): p. 2073-2083.

18. Ann N Y Acad Sci. 2003 Jun;990:585-90. Adherence to and invasion of Vero cells by recombinant Escherichia coli expressing the outer membrane protein rOmpB of *Rickettsia japonica*. Uchiyama T.

19. J Cell Sci. 1999 Jun;112 (Pt 11):1697-708. A comparative study of the actin-based motilities of the pathogenic bacteria Listeria monocytogenes, *Shigella flexneri* and *Rickettsia conorii*. Gouin E, Gantelet H, Egile C, Lasa I, Ohayon H, Villiers V, Gounon P, Sansonetti PJ, Cossart P.

20. Ann N Y Acad Sci. 2003 Jun;990:535-47. Rickettsial actin-based motility: behavior and involvement of cytoskeletal regulators. Heinzen RA.

21. BMC Genomics. 2009 Apr 20;10:166. Analysis of the *Rickettsia africae* genome reveals that virulence acquisition in Rickettsia species may be explained by genome reduction. Fournier PE, El Karkouri K, Leroy Q, Robert C, Giumelli B, Renesto P, Socolovschi C, Parola P, Audic S, Raoult D.

22. J Bacteriol. 2012 Aug;194(15):4145. Genomic Comparison of *Rickettsia honei* Strain RBT and Other Rickettsia Species. Xin D, El Karkouri K, Robert C, Raoult D, Fournier PE.

23. J Bacteriol. 2012 Sep;194(18):5129. Genome Sequence of *Rickettsia australis*, the Agent of Queensland Tick Typhus. Dong X, El Karkouri K, Robert C, Raoult D, Fournier PE.

24. J Bacteriol. 2012 May;194(10):2751. Genomic comparison of *Rickettsia helvetica* and other Rickettsia species. Dong X, El Karkouri K, Robert C, Gavory F, Raoult D, Fournier PE.

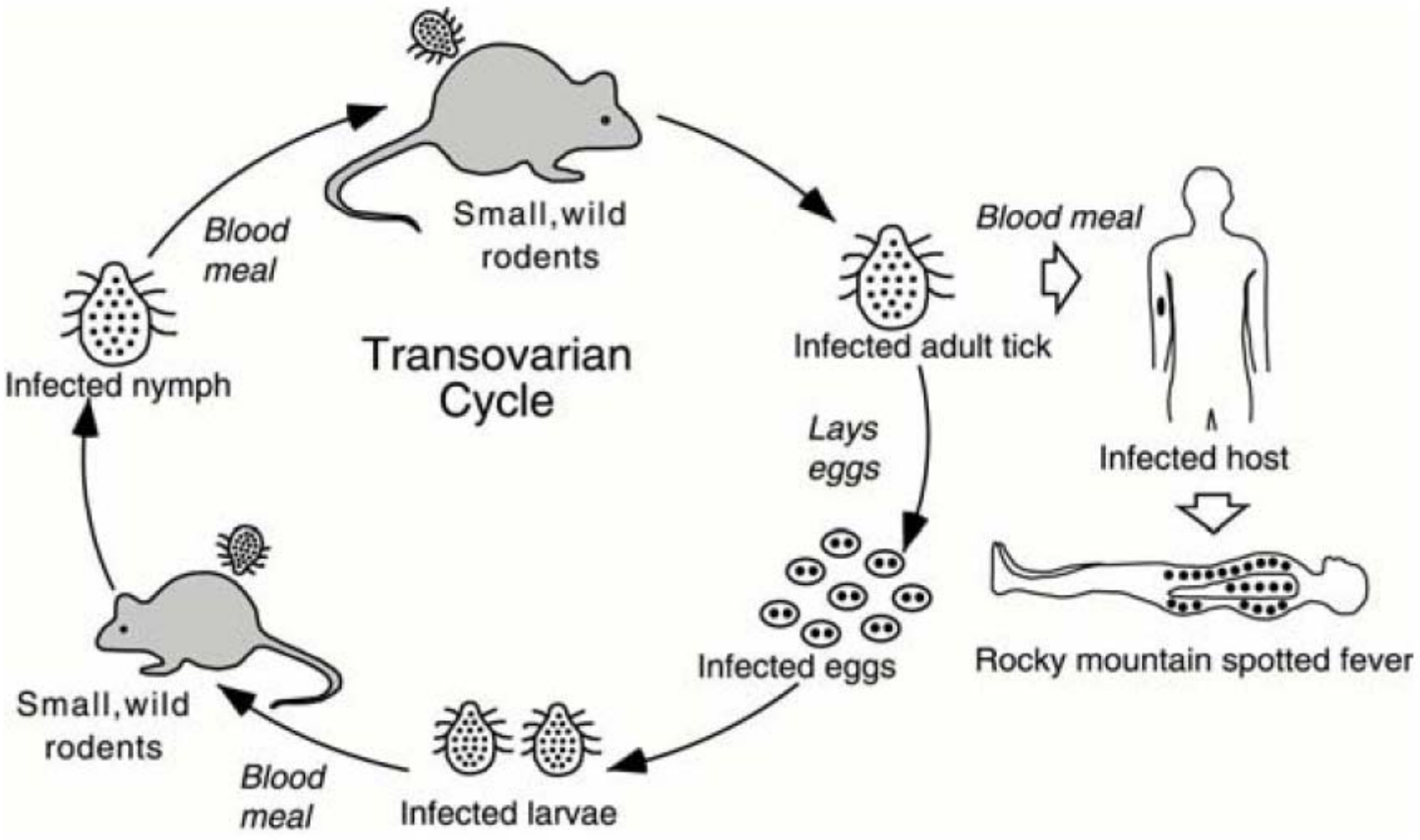

Figure 1. Le passage transmission joue un rôle important dans le maintien de l'infection à rickettsies dans la nature.

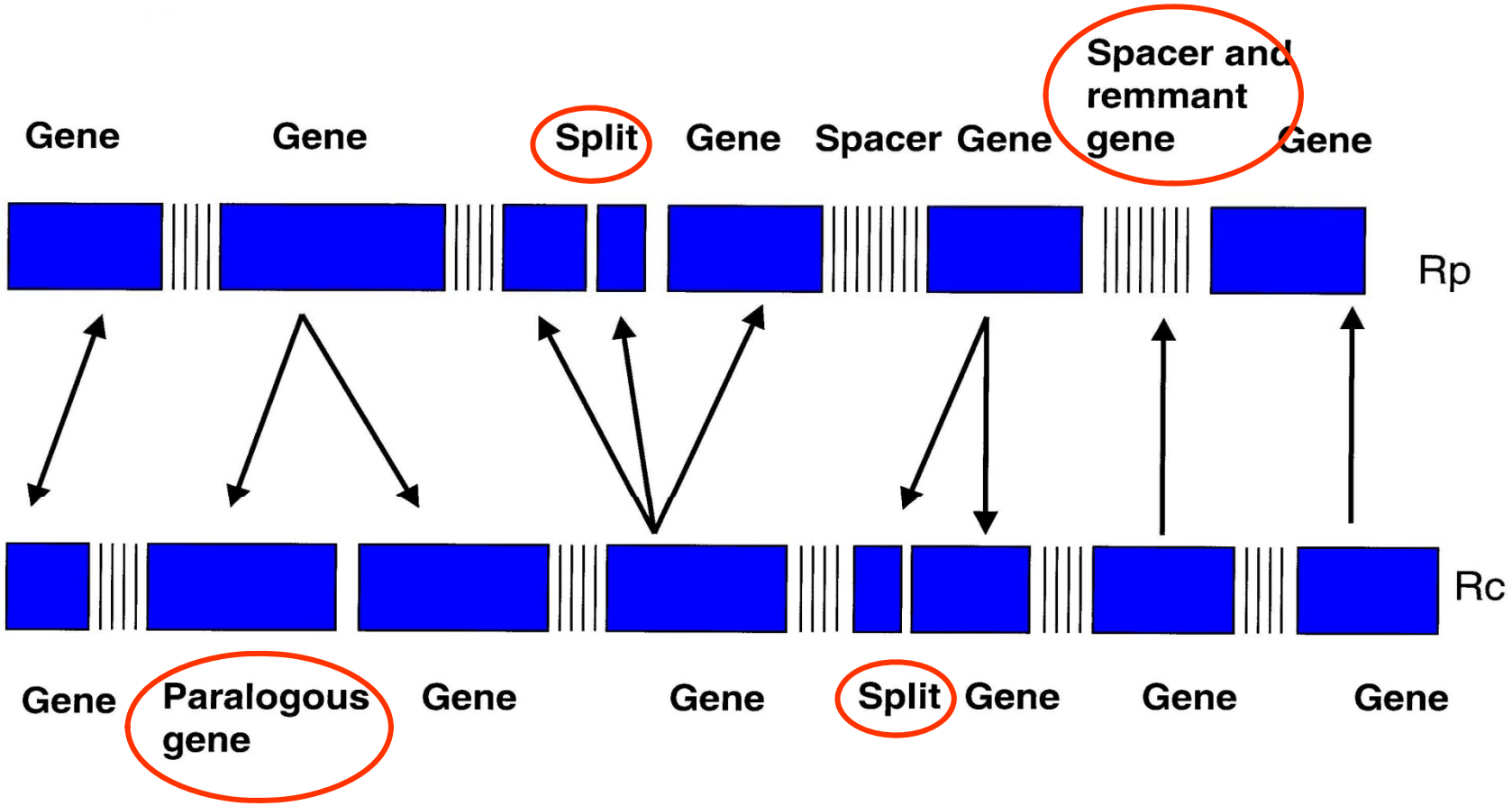

Figure 2. L'évolution genomique chez les Rickettsies.

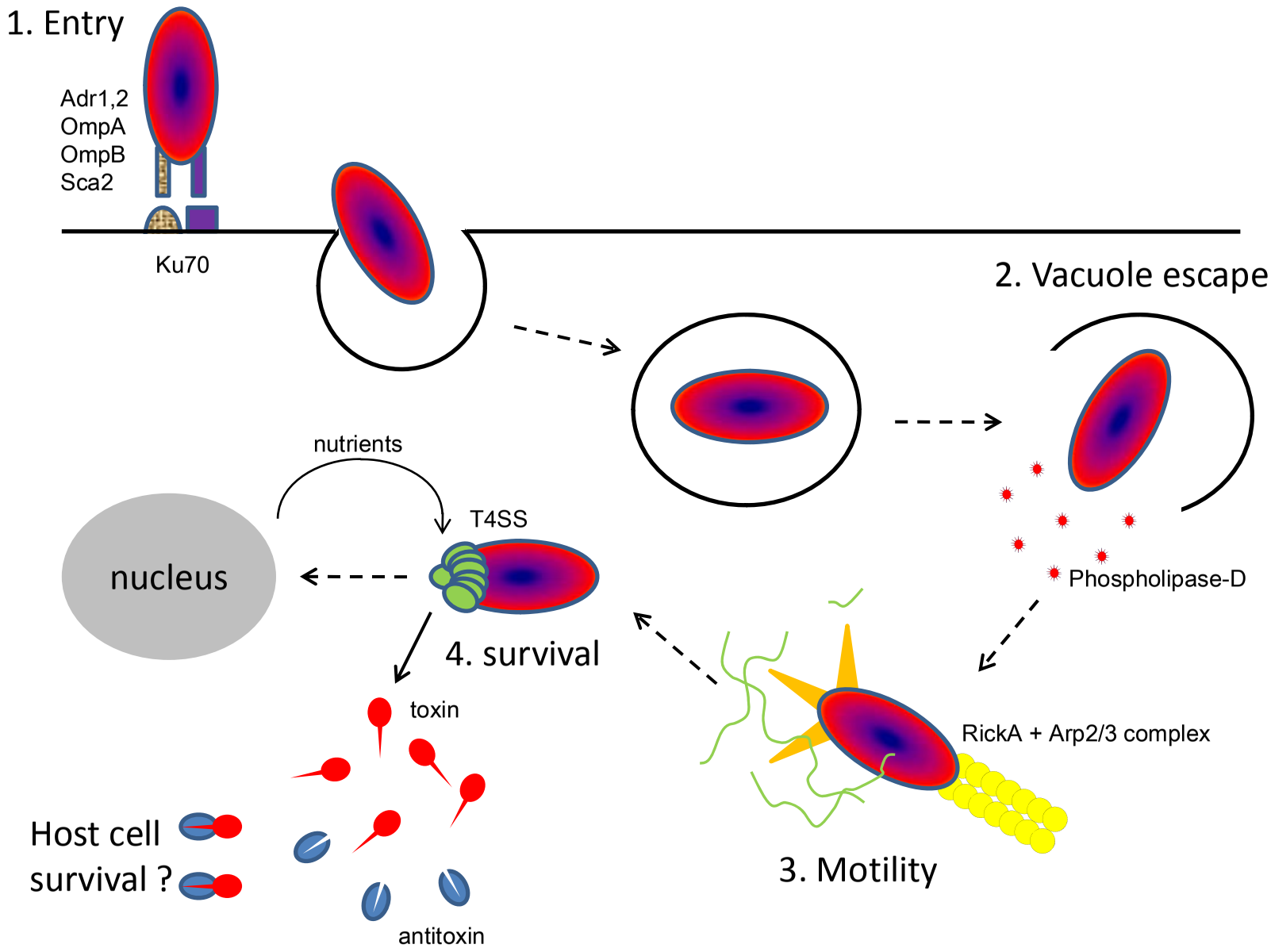

Figure 3. Intracellular behaviour of rickettsiae. Ce figure a modifié et prise de Dr. Patricia Renesto et Pr. Pierre edouard fournier.

Remerciements

Tout d’abord, je voudrais remercier le Professeur Didier Raoult pour m’avoir acceuilli au sein de l’Unité des Rickettsies.

Merci au Professeur Pierre-Edouard Fournier d’avoir dirigé mon travail de thèse, de votre supervision, de vos conseils et de votre guidage depuis le début de ma thèse. Merci de votre patience et enthousiasme pour reprendre la direction de ma thèse, votre esprit de la recherche scientifique qui sont des moteurs à l’avancement de mes travaux et aussi seront l’excitation dans ma carrière future. Ce fut un grand plaisir de passer ma thèse à vos cotés.

Ensuite je tiens à remercier les autres membres du jury pour avoir évalué soigneusement mes travaux de thèse. Merci au Pr Jean-Louis Mège d’avoir accepté de présider ce jury. Merci au Professeur Max Maurin et au Docteur Patricia Renesto d’avoir accepter de faire partie de mon jury de thèse en qualité de rapporteur.

Je voudrais adresser un grand merci à mon très précieux ami Khalid El Karkouri qui m'a donné de très nombreux conseils et corrections, et fait profite de son expérience inestimable dans le domaine de la bioinformatique depuis le tout début de ma thèse. Merci pour toute votre aide. Je n'aurais jamais pu mener à bien cette thèse sans vous.

Merci à Catherine Robert et toute l'équipe de séquençage, à tous mes amis parmi les bioinformaticiens, les chercheurs, les médecins, les secrétaires et les étudiants que j'ai cotoyés pendant ces 3 ans.

A la fin, je voudrais remerci mes parents et ma fiancée pour le support, l'amour et vos soins. Cette thèse est aussi la vôtre. Je vous aime.