AIX-MARSEILLE UNIVERSITE
FACULTE DE MEDECINE DE MARSEILLE
ECOLE DOCTORALE: SCIENCES DE LA VIE ET DE LA SANTE

THESE

Présentée et publiquement soutenue devant
LA FACULTE DE MEDECINE DE MARSEILLE

Le 8 octobre 2015
par

Mlle **FELICETTA D'AMATO**

Date et lieu de naissance 26/12/1986 Bitonto, Italie

TITRE DE LA THESE:

PANGENOME DE *COXIELLA BURNETII*

Etude pangénomique de *C.burnetii* : relations entre profil génétique et pathogénicité

Pour obtenir le grade de DOCTORAT d'AIX-MARSEILLE UNIVERSITE
SPECIALITE: **Pathologies humaines-MALADIES INFECTIEUSES**

Membres du Jury de la Thèse:

M le Professeur Jean-Louis MEGE	**Président du jury**
M le Docteur Olivier DURON	**Rapporteur**
M le Professeur Max MAURIN	**Rapporteur**
M le Professeur Didier RAOULT	**Directeur de thèse**

Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes «URMITE »
UM63, CNRS 7278, IRD 198, Inserm 1095, Faculté de médecine

Directeur : Pr. Didier Raoult

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 de Sciences 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 qui comportent un format de thèse utilisé dans le Nord de l'Europe permettant 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 de la compétition internationale et permet de se concentrer sur des travaux qui bénéficieront d'une diffusion internationale.

Professeur **Didier RAOULT**

"Per aspera sic itur ad astra"

Ai miei genitori

A chi non smette mai di credere in me

TABLE DES MATIERES

RÉSUMÉ

Coxiella burnetii est l'agent pathogène responsable de la fièvre Q. Ce coccobacille gram négatif appartient à la famille des *γ*-protéobactéries. Grâce à de nombreuses études effectuées, nos connaissances sur sa taxonomie, son mode de vie et sa résistance aux antibiotiques se sont approfondies. Cependant, plusieurs aspects nécessitent encore plus d'investigations afin d'améliorer notre compréhension notamment des liens entre les différentes souches ainsi que les relations entre le génome et la virulence.

Dans le cadre de cette thèse nous nous sommes intéressés à l'étude de souches de *C.burnetii* responsables d'événements épidémiques. Nous avons séquencé une souche de génotype MST33 (Z3055), proche de la souche responsable de l'épidémie de fièvre Q aux Pays-Bas, et une souche de génotype MST17 (Cb175) clone provoquant l'une des formes les plus virulentes de fièvre Q aiguë jamais décrite auparavant et retrouvée à ce jour uniquement en Guyane Française. Ensuite, nous avons effectué une analyse génomique comparative de ces souches avec d'autres souches disponibles dans GenBank. Nous avons effectué un génotypage en utilisant la méthode de « multispacer typing (MST) » et, afin d'explorer les niveaux de conservation et les différences génétiques entre les souches étudiées, nous avons utilisé une approche basée sur l'algorithme de Blast Ratio Score (BSR).

Les résultats de ces analyses montrent que le génome de la souche Z3055 était très similaire à celui de la souche de référence Nine Mile I. Les différences observées sont liées à la présence de mutations non synonymes dans le génome de Z3055. Ces mutations sont plus fréquentes dans les gènes codant pour des protéines membranaires, des protéines contenant des domaines d'ankyrine, des facteurs de transcription et de traduction. Cette grande proportion de protéines membranaires mutées pourrait expliquer l'ampleur de cette épidémie en

Hollande. En effet, le changement de profil antigénique pourrait être à l'origine de la formation d'un nouveau sérotype capable d'échapper à la réponse immunitaire de l'hôte et de diffuser facilement dans une population au système immunitaire naïf.

Nous avons d'ailleurs montré que la souche responsable de la fièvre Q en Guyane (Cb175) présente des différences chromosomiques importantes par rapport à NMI, bien au contraire de la souche Z3055. Ces différences se manifestent principalement par la présence d'une délétion d'une région de 6105pb contenant l'opéron hlyCABD du système de sécrétion de type 1 (T1SS). Ce résultat est cohérent avec ce qui a été observé chez les bactéries épidémiques les plus dangereuses comparées à leurs espèces non-épidémiques plus proches qui ont un génome réduit et contiennent moins de protéines du système de sécrétion.

D'autre part, le génotypage des souches de notre laboratoire révèle la corrélation entre le génome et leur répartition géographique, confirmant la notion de « geotyping ». En effet, un seul génotype MST21, a été détecté au Canada au cours des récentes épidémies de fièvre Q rencontrées dans ce pays. Nous avons séquencé la souche canadienne épidémique DOG UTAD (MST21) et comparé le « draft génome » correspondant avec une souche de même génotype. Les résultats indiquent que la souche DOG UTAD est très similaire à CbuG_Q212 en partageant le même génotype MST21 et la même origine géographique (Canada), ce qui suggère une évolution clonale récente. Nous avons, aussi, séquencé la souche Cb196 de génotype MST51 (uniquement décrit en Arabie Saoudite) et avons observé une forte similarité de séquence avec CbuK_Q154.

Enfin, l'analyse par BSR indique un ratio core génome/pangenome de 96%, ce qui suggère que le pangénome de *C. burnetii* est fermé. Une analyse

pangénomique globale impliquant toutes les souches collectées dans notre laboratoire permettra de confirmer ce résultat préliminaire et de clarifier les liens entre l'origine géographique, le génotype et la virulence des souches.

ABSTRACT

Coxiella burnetii is a human pathogen that causes the zoonotic disease Q fever. This coccobacillus gram-negative belongs to the family of γ-proteobacteria. Several studies on this pathogen have improved our understanding of its taxonomy, lifestyle, virulence and the resistance to antibiotics. However, until nowadays, our knowledge of many aspects is still weak and need more highlights, such as the relationship between strains and the link between genome and virulence. In order to answer these questions, we focused on the study of strains responsible for epidemic events. Particularly, we sequenced the clone of the strain responsible for Netherlands outbreak having genotype MST33 (Z3055), and strain having MST17 (Cb175) responsible for one of the most severe form of acute Q fever never reported in literature and uniquely described in French Guiana. We performed a comparative genomic analysis of *C. burnetii* strains sequenced in our laboratory with other strains whose genomes are available in GenBank database. The genotyping of *C. burnetii* was performed using multispacer sequence typing (MST) based on the most variable intergenic spacers of *C. burnetii* genome. In the aim to explore the conservation levels and genetic differences among the strains studied in this thesis, we performed the Blast Score Ratio (BSR).

Our findings showed that the Netherlands outbreak responsible strain (clone Z3055) was highly similar to the reference strain Nine Mile I. Only slight differences were observed, which were related to non-synonymous mutations in Z3055 genome. These mutations were significantly increased in genes encoding membrane proteins, ankyrin repeat domains containing proteins, transcription factors and translation proteins. The high proportion of mutated membrane proteins could explain this large-scale outbreak. Change of antigenic profile may have led to a new serotype, conferring to the novel clone the capacity to escape

the host immune response and to disseminate easily in a immunologically naïve population.

On the contrary, the type strain responsible for Q fever in Guiana (Cb175) showed an important difference in its chromosome sequence compared to the reference NMI because of the deletion of a sequence of 6105bp containing the Type 1 secretion systems (T1SS) hlyCABD operon. This result appear consistent with previous findings that showed the most dangerous epidemic bacteria compared with their closest non-epidemic species are characterized by reduced genomes accompanied by significant decrease in ORF content and contain less secretion system proteins.

Moreover, the genotyping of our strains has shown a link between genome and geographic distribution of *C.burnetii*, confirming the concept of "geotyping". In fact, a single genotype MST21 has been detected in Canada during the multiple cases of Q fever. We sequenced the epidemic strain DOG UTAD (MST21) and we compared the corresponding draft genome to a strain of the same genotype. The results indicated a high similarity between DOG UTAD and CbuG_Q212. These two strains share the same genotype MST21 and were both isolated in Canada, suggesting a recent clonal radiation.

Recently, a novel and unique genotype MST51 has been described in Saudi Arabia and we sequenced the corresponding strain Cb196 of the same genotype. The results showed a high sequence similarity with strain CbuK_Q154 with a phylogenetically close genotype (MST8).

Finally, results of the analysis based on the Blast Score Ratio considering 7 whole genomes showed an estimated ratio core genome/pangenome of 96%, suggesting that the pangenome of *C. burnetii* is closed. A larger pangenomic study involving all strains (more than 40) collected in our laboratory will allow confirming this result. It would also highlight the pathophysiology of Q fever by

investigating the link between geographical origin, genotype, antibiotic susceptibility and other genetic specificities with clinical issues.

INTRODUCTION

Avant-propos

La première partie de notre travail est une synthèse bibliographique sous forme de revue qui a été soumise au journal ***Future Microbiology***. Ce travail vise à montrer le rôle important que la génomique a joué dans l'étude de la fièvre Q. La naissance de la génomique, puis le développement des méthodes de séquençage de nouvelle génération (NGS), ont permis d'explorer le répertoire génétique complet de *Coxiella burnetii* . En 2003, Le séquençage du premier génome de *C.burnetii* Nine Mile I RSA493, une souche isolée en 1935 aux Etats-Unis sur une tique, a été effectué. Depuis, le séquençage de plusieurs autres souches a été réalisé et la génomique a apporté des informations utiles pour le développement de méthodes de culture d'outils de diagnostic, de vaccins, et l'étude de la sensibilité aux antibiotiques.

Plusieurs techniques de génotypage ont été développées pour distinguer les souches de *C. burnetii* notamment la méthode de « multispacer typing (MST) » caractérisée par sa sensibilité et reproductibilité et un bon pouvoir discriminant. Cette méthode a permis l'introduction du concept de géotypage ou « geotyping ». Le geotyping est l'étude de la répartition géographique des différents génotypes de *C. burnetii* dans le monde. Il permet de déterminer si un pays donné est caractérisé par une large distribution de génotypes ou au contraire par un seul clone hyper virulent qui circule avec d'autres microvariants. Ce dernier cas a été observé aux Pays- Bas où le génotype MST33 était prédominant et responsable de la plus grande épidémie de fièvre Q jamais décrite . Egalement en Guyane française, un seul génotype circulant, MST17 a été décrit comme responsable d'une des formes de fièvre Q les plus sévères avec un taux de prévalence très élevé .

De plus, un seul génotype MST21 a été détecté au Canada lors des différentes épidémies de fièvre Q et un nouveau génotype MST51 a été décrit uniquement en Arabie Saoudite récemment .

L'objectif principal de cette thèse a été d'étudier les liens entre le profil génétique des souches épidémiques et leur pathogénicité grâce au séquençage de deux souches épidémiques de notre laboratoire : Z3055 et Cb175. De plus, nous avons effectué le séquençage de deux souches représentant des clones ayant une répartition géographique spécifique, à savoir la souche Canadienne DOG UTAD et la souche Cb196 d'Arabie Saoudite.
Comme illustré dans la Figure 1 expliquant les différentes étapes d'analyses effectuées, chaque souche a été séquencée et le génome a été fermé par méthode de « finishing ». Le génotypage a été fait aussi bien *en silico* que par biologie moléculaire. Dans le cas de la souche Z3055, l'étude de Multiple Locus Variable number of tandem repeats (VNTR) Analysis (MLVA) a été réalisée également pour la comparer aux autres souches isolées aux Pays-Bas. Une fois la séquence complète obtenue, nous avons fait une analyse comparative basée sur l'alignement multiple, la phylogénie ainsi que la recherche des mutations ponctuelles, tout en se focalisant sur les mutations non-synonymes dans la séquence codante. L'étape suivante a consisté en la prédiction de gènes et à l'annotation fonctionnelle par COG et par KEGG. Nous avons ensuite estimé le pourcentage du core génome en appliquant l'algorithme de Blast Score Ratio (BSR) sur un échantillon des cinq génomes complets disponibles dans GenBank et deux génomes complets obtenus dans par ces travaux de thèse.

Dans un premier temps , nous nous sommes intéressés à l'étude du clone responsable d'une des plus grandes épidémies de fièvre Q aux Pays-Bas. Entre 2007 et 2010, un total de 4000 cas d'infection a été reporté . Au moment de notre étude, le génome de la souche isolée pendant l'épidémie n'était pas

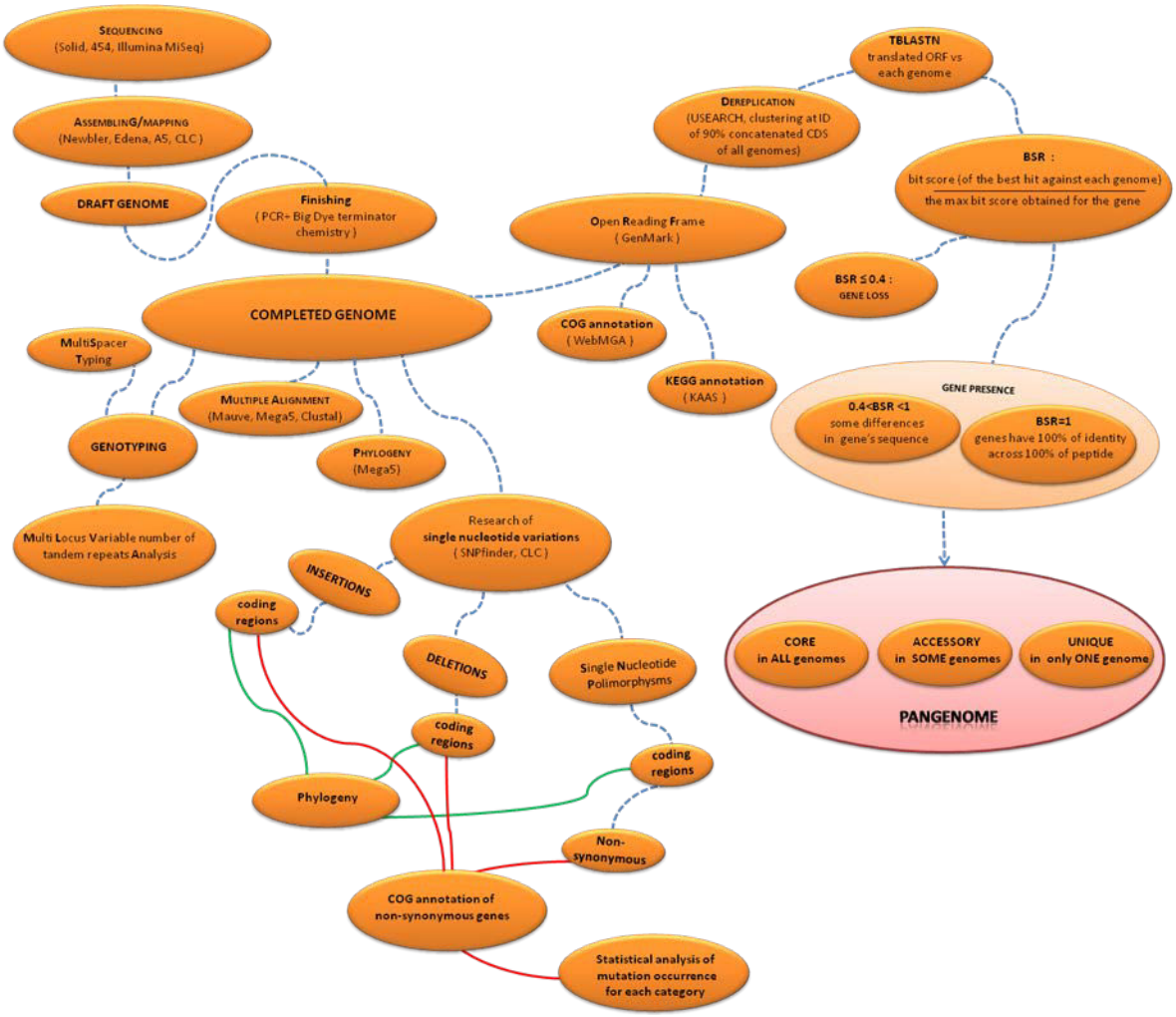
disponible, nous avons alors analysé la souche Z3055. C'est un clone de la souche épidémique ayant le même génotype 33 et le même profil MLVA. L'analyse génomique comparative de cette souche et la référence Nine Mile nous a permis de proposer une possible explication de cette épidémie.

En suite, nous avons étudié la souche Cb175 isolé d'un patient ayant la fièvre Q provenant de Guyane française. Dans la région de Cayenne, la fièvre Q devient de plus en plus un problème de santé publique. L'incidence annuelle est passée de 37 cas pour 100.000 habitants en 1996 à 150 cas pour 100.000 habitants en 2005 . La fièvre Q est la cause de 24% des pneumonies d'origine communautaire, la prévalence la plus forte jamais décrite dans la littérature . Nous avons analysé le profil génétique de la souche Cb175 pour connaitre les causes de cette forte virulence. Pendant l'étape de finishing, nous avons révélé la particularité de la souche : une large délétion dans le chromosome comparé à la souche de référence NMI a été mise en évidence.

Grâce à la naissance du séquençage haut débit, en 2005 le pangénome a été défini pour la première fois . Ses applications sont multiples, il a été utilisé dans l'étude des clones responsables de la pandémie *Vibrio cholerae* en 2009 ainsi que dans l'étude de l'épidémie *d'Escherichia coli* qui a eu lieu en 2011 en Allemagne . Le pangénome d'une espèce bactérienne donnée peut être soit ouvert ou fermé selon sa capacité à acquérir de l'ADN exogène. En effet, les espèces allopatriques qui vivent isolées dans une niche étroite possèdent la plupart du temps un petit génome ainsi qu'un pangenome fermé comme dans le cas de *Bacillus anthracis* qui possède un ratio core génome/pangenome de 99%. Au contraire, les espèces sympatriques vivant en communauté de bactéries pouvant s'échanger des gènes ont tendance à avoir un génome plus grand et un pangenome ouvert comme *Campylobacter jejuni, Legionella pneumophila*, *Acinetobacter baumannii* etc. avec un ratio inférieur à 87% .

Dans le cadre de cette thèse, l'analyse du pangenome a permis la recherche de gènes d'antivirulence dont la perte (associée à la réduction de la taille du génome) est un phénomène à l'origine de l'hyperpathogenicité de certaines espèces comparées aux espèces proches et moins pathogènes.

Figure 1 Pipeline de notre travail. Workflow illustrant les différentes étapes d'analyse, de séquençage de la souche jusqu'à l'analyse du pangenome.



REVUE

The contribution of genomics to the study of Q fever

Felicetta D'AMATO[1], Carole ELDIN[1], Didier RAOULT[1*]

[1]Aix-Marseille Université, Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes, UM63, CNRS 7278, IRD 198, INSERM 1095, Marseille, France

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

The contribution of genomics to the study of Q fever

Abstract

Introduction

1. The genome

1.1. Comparative genomics and the pan-genome study

1.2. Phylogenetic analysis

1.3. Genotyping

1.3.1. VNTR-MLVA genotyping

1.3.2. MST and the concept of 'geotyping'

1.4. Diagnostic tools

1.4.1. PCR and qPCR for direct diagnosis of Q fever

1.4.2. Genome-based design of culture medium

1.4.3. Serology: potential immunodominant antigens

1.5. Pathophysiology/lifestyle

1.5.1. Secretion systems in Legionellaceae and *C. burnetii*

1.5.2. Response to stress conditions

1.5.3. Metabolism

1.6. Pathogenicity and genomic characteristics of *C.burnetii* strains

1.6.1. Gene loss and strain virulence

1.6.2. French Guiana

1.6.3. Canada

1.6.4. Holland

1.7. Therapeutic tools

1.7.1. Antibiotic susceptibility

1.7.2. Putative antigens: surfaceome and vaccinology

2. Conclusion: Pangenome of all strains sequenced in our laboratory

The contribution of genomics to the study of Q fever

Felicetta D'AMATO[1], Carole ELDIN[1], Didier RAOULT[1*]

[1]Aix-Marseille Université, Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes, UM63, CNRS 7278, IRD 198, INSERM 1095, Marseille, France

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

Phone: (33) 491 38 55 17

Fax: (33) 491 83 03 90

Journal: **FUTURE MICROBIOLOGY**

Abstract

Coxiella burnetii is the etiological agent of Q fever, a worldwide zoonosis which can result in large outbreaks. The birth of genomics and sequencing of *C.burnetii* strains has revolutionised many fields of study of this infection. Accurate genotyping methods and comparative genomic analysis have enabled description of the diversity of strains around the world and their link with pathogenicity. Genomics has also permitted the development of qPCR tools and axenic culture medium, facilitating the diagnosis of Q fever. Moreover, several pathophysiological mechanisms can now be predicted and therapeutic strategies can be determined thanks to *in silico* genome analysis. An extensive pangenomic analysis will allow for a comprehensive view of the clonal diversity of *C.burnetii* and its link with virulence.

Keywords: Q fever, *Coxiella burnetii*, genomics, virulence, pangenome

Introduction

C.burnetii is the agent of Q fever, a worldwide zoonosis. It is a small, pleomorphic, obligate intracellular Gram-negative bacterium [1]. The primary reservoirs of this zoonosis are cattle, goats or sheep which shed microorganisms in urine, faeces, milk and birth products. Inhalation of contaminated aerosols represents the principal route of transmission to humans [1]. Genomics has emerged in Q fever studies since the first genome sequence of a *C.burnetii* strain was conducted in 2003 [2]. This breakthrough had major implications in all fields of study of this fastidious microorganism. Crucial domains such as phylogeny, epidemiology, diagnostic, culture methods, pathophysiology, virulence factors and antibiotic resistance analysis have benefited from this development. Until recently, the strict intracellular nature and high infectivity of *C.burnetii,* requiring a biosafety level 3 laboratory, has considerably complicated the study of its pathophysiology and mechanisms of virulence. The use of classical molecular-genetic tools and genetic transformations was fastidious and time-consuming [3]. The elaboration of an axenic medium to cultivate *C.burnetii* was a major issue that has been significantly accelerated by the availability of a genome sequence. Omsland *et al.* performed *in silico* the prediction of the main metabolic pathways and the identification of the growth requirements of the bacterium from genome analysis [4]. This was the key step forward, following multiple combinations of culture conditions, culminating in an efficient axenic medium called ACCM2 (Acidified Cystein Citrate Medium 2[5]. This major advance for all researchers working on *C.burnetii* is a first illustration of the contribution of genomics to the study of Q fever.

Q fever is characterised by a wide clinical polymorphism. The most commonly described symptomatic forms of Q fever primo infections, or ‘acute’ infections are flu-like illness, pneumonia and hepatitis [1]. However, during the phase of primo infection, the infection can range from an asymptomatic sero-conversion to severe pneumonia causing death

by respiratory distress syndrome [6].When the infection persists, the two main reported presentations are endocarditis and vascular infections [7, 8]. These infections were initially described as 'chronic Q fever', a nomenclature that appears obsolete because it covers very different foci of infection that can be diagnosed today thanks to new, efficient diagnostic tools [9, 10]. It appears that as in the case of tuberculosis, *C.burnetii* causes a primo infection, sometimes followed by an active, focalised persistent infection.

Historically, the principal factors identified to explain the clinical variability and outcome of Q fever were patient susceptibilities. Before the genomics era, the age and sex of patients were the first variables identified [11] as influencing the clinical expression of the acute form, with men aged >45 being at greater risk of more severe manifestations. In terms of persistent diseases, the presence of an immunosuppression, a valvulopathy and vascular aneurysms were soon associated with evolution to endocarditis of vascular infections [11]. The particular severity of Q fever in pregnant women and its implication in obstetrical complications was also described [12]. However, it was also hypothesised that clinical presentation could vary depending on the strain of *C.burnetii* involved. Initially, because of the established nomenclature, a difference between strains causing 'acute Q fever' and strains causing 'chronic Q fever' was actively sought after. Initially, a correlation between plasmid types and clinical features was put forward [13]. Then, the availability of a genomic sequence enabled the development of efficient genotyping tools to describe the genetic diversity of *C.burnetii* strains [14], demonstrating that no systematic link existed between a given genotype and the acute or chronic form of the disease. For example almost all genotypes could be represented in 'chronic' isolates of *C.burnetii* [14]. Thus, the evolution to a persistent form of Q fever appears to be largely dependent on host factors. Conversely, clinical expression of the primo infection or 'acute form' has been demonstrated to be dependent on the strain involved. For example in French Guiana, the epidemiological profile

of the disease is atypical, with Q fever accounting for 24% of hospitalised community-acquired pneumonia, the highest prevalence ever recorded [15]. The incidence of the disease rose to 150 cases/100,000 persons in 2005 [16]. These pneumonia are characterised by a more severe initial presentation than community-acquired pneumonia caused by other microorganisms in this region [15] in contrast to hepatitis and endocarditis which are less frequent than in metropolitan France [17]. Genotyping methods revealed that the epidemic strain causing Q fever in French Guiana was clonal and had only been reported in this territory [18, 19]. This example is a good illustration that clinical presentation of Q fever primo-infection and its epidemiology is dependent upon the strain involved [20].

The global advances provided by genomics helped to establish that the generic term of 'Q fever' covered a large diversity of clinical diseases and epidemiological profiles, depending largely upon the genetic properties of *C.burnetii* strains. We are beginning to describe this genetic diversity and its mechanisms. In this review we focus on what genomics has brought to all fields of the study of Q fever and detail the different stages of these advances (Figure 1).

1. **The genome**

1.1. **Comparative genomics and the pan-genome study**

Comparative genomics consists of studying the similarities and differences in genomes structures and functions between different microorganisms at the intra- or inter-specific level [21]. The first *C.burnetii* genome was sequenced in 2003; the Nine Mile phase I RSA493 strain, isolated from a tick in 1935 [2]. It was composed of a 1,995,275 bp chromosome and a 37,393 bp QpH1 plasmid. This genome sequence enabled the *C.burnetii* genome to be compared with those of other intracellular bacteria. Several elements considerably differentiating *C.burnetii* from other sequenced intracellular organisms (*Rickettsia* and *Chlamydia*) and its closest related gamma proteobacteria (*Legionella*) were described [2].

This genome was predicted to encode 2,134 coding sequences, of which 719 were hypothetical with no equivalent in other gamma proteobacterial sequenced genomes. Moreover, a much higher percentage of the genome was coding, in comparison with other species known to have undergone genome reduction such as *Rickettsia prowazekii* or *Mycobacterium leprae*. Also, a high number of pseudogenes (83) were found, suggesting that genome reduction is still in progress. An unusually high number of 29 IS elements was reported [2], suggesting high genomic plasticity. Transportation, biosynthesis and metabolic capabilities were also greater than in other intracellular bacteria such as *Rickettsia* and *Chlamydia*. Finally, a new family of ankyrin-repeat containing proteins was described. The *C.burnetii* RSA493 strain also encoded some genes very similar to eukaryotic genes which didn't match any prokaryotic genes sequences, reflecting its intracellular nature. As with *Legionella pneumophila, C.burnetii* encodes multiple eucaryotic-like proteins, probably to modulate its host cell functions. The genomic sequence of this first strain allowed for the specific features of *C.burnetii* to be deciphered within other intracellular bacteria, with ongoing genome reduction probably due to a recent adaptation to intracellular parasitism [2].

Since this first genome publication, seven whole *C.burnetii* strain genome sequences have been published and are available in GenBank: Dugway 5J108-111 [22] ,CbuG_Q212 [2, 22], CbuK_Q154 [22], RSA 331 [2, 22], Nine Mile phase I (NMI) [2, 22] ,Z3055 [23] and Cb175_Cayenne_FrenchGuiana [19]. Sixteen incomplete genomes are also available: Cb109 [24], Q321, MSU Goat Q177, Cb_C2 [25], Cb_B1 [25], EV-Cb_BK10 [25], Cb_O184 [25], EV-Cb_C13 [25], Cb_B18 [25], Cb171_QLYMPHOMA, DOG UTAD [26], Cb196_Saudi_Arabia [27], Namibia [28], Cb185 [29], AuQ01 [30] and NL-Limburg. These new data enable a deeper comparative genomic analysis to be performed at an intraspecific level. To date, the largest *C.burnetii* genome reported is that of the Dugway strain, isolated from a rodent, with a 2,158,758 bp chromosome and a 54,179 bp QpDG plasmid. The

smallest is that of Cb175, isolated from the blood of a patient from French Guiana, with a 1,989,565 bp chromosome and a 37,398 bp QpH1 plasmid [19]. All sequenced strains have an autonomously replicating plasmid (QpH1, QpDV, QpRS, QpDV) or plasmid sequences integrated into their chromosome [19]. The availability of whole genome sequences has enabled a pangenomic analysis to be conducted of 7 strains of *C.burnetii*. D'Amato *et al.* analysed entire genomes to obtain an estimated percentage of core genes (genes present in all strains), accessory genes (genes present in some strains) and unique genes (present in only one strain), [19, 23] using a blast score ratio algorithm [31, 32]. The core genome/pangenome ratio was 96%, with a total of 13,542 core genes, 498 accessory genes and 88 unique genes [19] (Figure 2). Among the unique genes, 74 belonged to the Dugway strain, 13 to strain Q212 and one single gene to RSA331 [19]. This result showed a strong similarity between all considered *C.burnetii* strains and suggested that the pangenome is closed; however, given the limited number of strains in the study, a more comprehensive analysis is needed to confirm this hypothesis. In addition, a detailed comparative analysis of some epidemic strains of *C.burnetii* has revealed some particularities which are discussed below in the 'pathogenicity' chapter.

1.2. Phylogenetic analysis

Initially, in the absence of genomic phylogenetic information and basing on ecologic similarities (such as culture methods), *C. burnetii* was assigned to the order of the alphaproteobacteria such as *Rickettsiales* and *Chlamydiales* [33]. Before whole genome sequencing, the sequencing of 16 rRNA definitely excluded any relationship between *C.burnetii* and members of this order [34], rather showing a common line of descendent with *Legionella pneumophila* and *Rickettsiella grylli*, belonging to the family of gammaproteobacteria [34, 35]. Members of the *Legionella* genus can survive in the environment within free-living amoeba thanks to a mechanism of resistance to phagocytosis

166 [36] and *C.burnetii* can also survive in a spore-like form in vacuoles of *Acanthamoeba*
167 *castellani* [37].

168 Several studies have shown that horizontal gene transfers between intra-amoebal
169 microorganisms take place in free-living amoebae [38]. On the genomic level, the presence of
170 a high number of ankyrin repeat domains and a similar sterol delta-reductase gene in
171 *Legionella* and *C.burnetii* genomes suggest that they exchange genes within the amoeba [38].
172 Similarly, some mycobacterial genes encoding for pyr-redox end up phylogenetically close to
173 *C. burnetii* and other amoeba-resistant bacteria, suggesting abundant horizontal gene transfer
174 in the amoeba (Figure 3A) [38]. In this context, the similarities of genome sequence is a good
175 illustration that *C.burnetii* shares a common ecosystem with other intra-amoebal pathogens.

176 For years, *C.burnetii* has been the only representative of its genus, preventing accurate
177 phylogenetic studies about a putative ancestor. In 2000, *Coxiella cheraxi,* a new member of
178 the genus showing 95.6% similarities with *C.burnetii* in 16S rRNA sequences was isolated
179 from crayfish in Australia [39]. However, the genome sequence of this bacterium is not
180 available to date. More recently, several *Coxiella*-like organisms, isolated from non-vertebrate
181 species (mainly ticks) were identified using 16SrRNA sequencing [40]. Some of these
182 *Coxiella*-like organisms play the role of tick endosymbionts [40]. Duron *et al.* performed a
183 phylogenetic study among *C.burnetii* and *Coxiella*-like tick endosymbionts, combining a new
184 multilocus typing method and whole genome sequencing. This work revealed the presence of
185 high diversity among *Coxiella*-like organisms, with four clades described (A to D) [41].
186 Interestingly, in this phylogenetic network, *C.burnetii* belonged to a unique subclade within
187 the clade of *Coxiella* associated with soft ticks (clade A) (Figure 3B). This finding suggests
188 that *C.burnetii* originated from a tick-borne *Coxiella* ancestor. There was also a wide
189 distribution of *Coxiella*-like microorganism across tick species, and a large variety of
190 genotypes, suggesting a long evolutionary history of the association between *Coxiella* and

ticks. In contrast, the low genetic variability of *C.burnetii* strains, suggests that a vertebrate pathogen recently emerged from a tick-associated ancestor [41]. The mechanism of its evolution may be spontaneous mutations or transfers and acquisition of genes from pathogens that co-infected the same tick or vertebrate host [40, 41]. This study is a major step forward in the conception of *C.burnetii* phylogeny, recounting the story of an evolutionary transformation from a tick endosymbiont to a highly virulent vertebrate pathogen. By comparing the gene content of *Coxiella-like* endosymbionts with *C.burnetii* we observed a pattern of gene gain in *C.burnetii*. The same trend is observed in the example of *Buchnera*, an endosymbiont [42] of aphid species which is very close to the pathogenic *E.coli*. [42-46]

BOX 1.

BOX 1. *Coxiella burnetii* and *Coxiella*- endosymbiont

We compared the gene content of Coxiella endosymbiont of Amblyomma americanum (NZ_CP007541) and *Coxiella burnetii* by using Proteinortho [43] at parameters of 50% identity and 70% coverage.
In *Coxiella*-endosymbiont, 523 genes (i.e. 94% of its total genes) are shared with *C.burnetii*, while 1,319 genes of *C.burnetii* are absent. These 1,319 belong mostly to COG categories of transcription (K), cell wall/membrane/envelope biogenesis (M), replication, recombination and repair (L) and defence mechanisms (V). These findings are consistent with what has been previously reported in the case of Buchnera endosymbiont, compared to its closely related nonsymbiotic species *E.coli* [44-46] which showed a limited capacity for DNA recombination and repair (indicating a vulnerability to DNA damage), a low number of transcriptional regulators, of genes involved in the envelope biogenesis (suggesting a weak surface structure) and few genes encoding ABC transporters [42] (Figure 4).
Moreover, no loss was described for *Buchnera* in genes encoding Sec proteins, which are conserved with *E.coli* and other closely related species [42]. In *Coxiella*-like endosymbiont, these proteins are also conserved and we found the genes encoding for subunit SecE, SecB, YidC, SecA, SecY, YajC, SecG, SecD, and SecF of a preprotein translocase involved in secretion systems of type II, IV and V.

1.3. Genotyping

Several methodologies have been developed for differentiating *Coxiella burnetii* isolates according to their environment, hosts and pathogenicity. It was soon demonstrated that

223 classical genotyping methods such as 16S ribosomal RNA sequencing, 16S-23S ribosomal
224 DNA (rDNA) sequencing, internal transcribed spacer (ITS) sequencing, and RNA
225 polymerase's β-subunit (rpoB) sequencing were not relevant, having no sufficient
226 discriminant power in *C.burnetii* [47-50]. Thus, many attempts were made to find a good
227 differentiating method which could be used routinely. Initially, before the availability of a
228 genomic sequence, Restriction Fragment Length Polymorphism (RFLP) analysis of genomic
229 DNA and PCR-RFLP [51-53] was used and did reveal genetic diversity, but these methods
230 were shown to have limited sensitivity and reproducibility. Differentiation of strains based on
231 sequences or PCR-RLFP analysis of several genes such as *Com1*, *DjlA (MucZ)* and *icd* was
232 also proposed [54-56]. Subsequently, thanks to the availability of a genome sequence, a more
233 global method using Comparative Genomic Hybridisation (CGH) using a whole-genome
234 microarray was performed , allowing for polymorphisms in chromosomal and plasmid ORF
235 sequences to be detected, determining genomotypes [20, 57]. Some of these genomotypes
236 have been associated with host specificity and virulence, but this fastidious method cannot be
237 performed routinely in a clinical setting. Currently, two discriminant sequence-based
238 genotype methods are most frequently used: Multispacer Sequence Typing (MST) [14], and
239 Multiple Locus Variable Number of Tandem Repeats (VNTR) Analysis (MLVA) [58, 59].

240 **1.3.1. VNTR-MLVA genotyping**

241 Svraka *et al.* searched for tandem repeats in the RSA 493 strain sequence. They identified
242 seven VNTR sequences which were amplified from 21 *C.burnetii* strains [58], resulting in
243 five major clusters and nine MLVA types. In a subsequent study, Arricau-Bouvery *et al.* [59]
244 identified 36 MLVA genotypes from 42 *C.burnetii* isolates and proposed to use two panels of
245 markers to enhance the discriminatory power of this method. This method has been applied in
246 recent epidemiological studies in Germany and Poland [60, 61] and shows a good
247 discriminant power but lacks interlaboratory reproducibility [59].

1.3.2. MST and the concept of 'geotyping'

MST genotyping was developed in 2005. Glazunova *et al.* investigated the variability of sequences of spacers located between two ORF, which were supposed to present less selection pressure than adjacent genes. The authors identified ten highly variable spacers that allow three monophyletic groups [14] and 30 different sequences types to be defined from 173 *C.burnetii* isolates. Among these monophyletic groups, the second group contained only strains presenting a QpH1plasmid. Conversely, in the first group only QpDV and QpRS were represented while the third group showed only QpH1 or plasmidless strains. This result illustrates that plasmid types are correlated with genotypes. Most of the genotypes are detected both in animals and humans [14, 62-64] while others show host specificity [65]. This genotyping method is highly discriminant and has been used in several other studies to characterise strains from all over the world so that databases are available allowing for easy interlaboratory comparison [14]. The MST genotyping of a high number of strains from different geographical areas has helped to reconstitute the epidemiology of *C.burnetii* from one region to another and to identify epidemic clones. Currently, some MST are spread across the five continents [14, 62, 63, 66], while others are very specific to one geographical area. This is the case of MST 17 in French Guiana (discussed later in the text)[18] and MST 51 [67], a novel genotype detected in Saudi Arabia (strain Cb196) [27]. Thus, given the geographical distribution of strains, this method can be considered as a "geotyping" tool, as illustrated in Figure 5. For example, genotype MST33 was found to be the predominant clone causing the historically most significant outbreak in the Netherlands [62, 68, 69] and a phylogenetic analysis revealed that this clone had probably spread from Germany to the Netherlands via France. MST 20 is found in Europe and in the USA, suggesting an historical spread of the disease by infected animals or humans.

As suggested by Figure 5, this 'geotyping' is still incomplete and should be improved by further studies using this technique to provide a complete cartography of the genetic diversity of *C.burnetii*.

1.4. Diagnostic tools

Because *C.burnetii* is an intracellular bacterium with a fastidious culture, indirect diagnosis from clinical samples using serology (indirect immunofluorescence assay) has long been the gold standard for diagnosis in humans [70]. However, serology has the principal disadvantage of being positive after a significant delay of one to two weeks from the beginning of the infection [70]. Thanks to the analysis of the *C.burnetii* genome sequence, new tools have been developed for direct diagnosis. The identification of stable repetitive sequences has allowed efficient primers and probes for qPCR systems to be defined. *In silico* prediction of *C.burnetii* metabolic requirements from genomic sequence has accelerated the design of an axenic medium [4, 5]. Also, potential specific antigenic proteins can be deduced from the genomic sequence to try to develop more sensitive serological assays.

1.4.1. PCR and qPCR for direct diagnosis of Q fever

The first PCR systems detecting *C.burnetii* were developed before the availability of the entire genome sequence of the bacterium. Primers and probes for standard PCR systems were designed targeting the Superoxide Dismutase gene, the *com1* gene, 16S-23S RNA and IS1111 repetitive elements, with detection limits ranging from 10 to 10^2 bacteria [49, 71-74] in human and animal samples. Nested-PCR systems targeting IS111, *com1* and plasmid sequences have also been proposed [70, 75-78].

RT-PCR or quantitative PCR (qPCR) has the advantages of both detecting and quantifying the amount of bacteria from a given sample. Consequently, this method, which is also less time-consuming, has superseded the standard PCR systems for routine diagnosis.

Currently, the most sensitive RT-PCR system targets the IS1111 repetitive elements, present in about 20 copies in the *C.burnetii* genome [79-82]. This IS1111 qPCR allow for early detection of the bacterium in the serum of patients in the first two weeks of the infection, when serology is still negative, and in patients with Q fever endocarditis or vascular infections[83-86]. Another qPCR system targeting IS30A repetitive elements is also available, but is much less sensitive [82].

1.4.2. Genome-based design of culture medium

Recently, an axenic medium has been developed for *C.burnetii,* allowing host-cell free growth [4]. The first crucial step of the design of this medium has been to deduce metabolic pathway deficiencies from an *in silico* genomic analysis to design a medium that sustained a minimal metabolic activity for the bacterium [87]. Results suggested that in order to replicate in an axenic medium, the bacterium needed the presence of L-cysteine which, being a source of sulphur and a precursor in protein synthesis, provided a sufficient expression of ribosomal genes. Finding genes encoding cytochrome b which has a high affinity for oxygen also explained the optimal replication in 2.5% oxygen [4]. In a second step, transcript profiles and phenotype microarrays were analysed, leading to the formulation of the final medium known as ACCM2 (Acidified Citrate Cysteine Medium 2). Boden *et al*. recently achieved the first isolation of *C.burnetii* in ACCM2 from a clinical sample (heart valve). Thus, this medium has the potential to simplify the isolation of *C.burnetii* from clinical samples. However, to date, the sensitivity of ACCM2 for different pathogenic strains of *C.burnetii* is not known.

1.4.3. Serology: potential immunodominant antigens

The gold standard method for *C.burnetii* serology is immunofluorescence assay based on the serum antibody response to *C.burnetii* Nine Mile phase I and phase II. This technique requires culture and purification of the microorganism so that only specialised laboratories

320 can carry out this method [70]. The development of new serological markers is an interesting
321 challenge to improving the sensitivity and specificity of serological methods. In this context,
322 the availability of a genome sequence has enabled the prediction of proteins produced by
323 *C.burnetii* and their functions. This approach, combined with proteomic studies has been
324 applied to search for antigenic candidates for serodiagnosis of Q fever. Potential antigens
325 belonging to COG categories of membrane proteins, translation and post-translational
326 modification, DNA replication recombination and repair, proteins preventing DNA
327 denaturation in stress conditions, and heat shock proteins, have been selected [88]. Among
328 these, OmpH, YbgF, Com1, OmpH, Mip, GroEL, DnaK, RplL, hsp60, Com-1, RecA,
329 elongation factor Tu, OmpA-like transmembrane domain and FtsZ have been tested as
330 potential Q fever markers [89]. However, these markers showed a moderate specificity when
331 testing the microarray with sera infected by *Rickettsia*, *Legionella* and other species close to
332 *C.burnetii* [88-91]. Some authors have also tried to identify serological markers which are
333 specific for a given clinical form of the disease. An immunodominant polypeptide of 28-
334 Kilodalton, the outer membrane protein *adaA* (acute disease antigen A) has been identified
335 and proposed as a marker of acute Q fever [92]. It has been also detected as a potential
336 specific marker for Q fever abortion in goats [93]. Differential protein expression involved in
337 LPS biosynthesis between phase I and phase II of Nine Mile strain has also been studied [94].
338 However, to date, all these data need to be confirmed by further investigations to be applied in
339 routine practice.

340 **1.5. Pathophysiology/lifestyle**

341 *C.burnetii* presents a very particular intracellular lifestyle. The bacterium has evolved to
342 survive and multiply in the harshest of intracellular compartments: the phagolysosome [95].
343 The parasitophorous vacuole (PV) in which *C.burnetii* enters its host cell progressively
344 develops lysosomal characteristics with acidic pH, low oxygen tension, hydrolases and

cationic peptides. Genomics has enabled an understanding of some of the mechanisms of this particular pathophysiology.

1.5.1. Secretion systems in *Legionellaceae* and *C.burnetii*

C.burnetii is closely related in evolutionary terms to *L.pneumophila* (Figure 3A). Bioinformatics approaches investigating the mechanism of infection have revealed a common virulence strategy for these bacteria [96]. Like *Legionella, C. burnetii* encodes a type IV secretion system (T4SS), allowing for the delivery of effector proteins into the host cytosol during infection [95]. Twenty-three of the 26 Dot/Icm T4SS proteins described in *L. pneumophila* were found to be homologous in *C.burnetii* [2].

Several studies have been performed to identify the substrates of *C.burnetii* Dot/Icm machinery. Initially, because of the impossibility of obtaining *C.burnetii* mutants, the transformation of *L. pneumophila* was used. The candidate substrates of T4SS are characterised by eukaryotic-like domains able to mimic or inhibit host cellular processes. These proteins are involved in phagocytosis, cytotoxicity, inhibition of phagosome and lysosome fusion, association of the phagosome with the endoplasmic reticulum, and apoptosis and exit from the phagosome [96, 98]. Four *C.burnetii* Anks (AnkA, B, F and G) have been identified as the result of transformation of *L.pneumophila* [97]. The addition of AnkG to *L.pneumophila* enabled its replication in host cells by preventing apoptosis, mimicking the way *C.burnetii* persists in its host cells [98]. This result shows that, although both *C.burnetii* and *L.pneumophila* virulence strategies are based on T4SS, they have some differences in effector proteins [98]. Also, using *L. pneumophila* as a secretion model, six plasmid effectors were found to be translocated into the host cell cytosol by the Dot/Icm T4SS during infection [99].

Recently, the development of axenic media has removed the constraints of genetic transformation of *C.burnetii* and allows transformants to be obtained in only 16 days [5] instead of the two-three months previously required [100]. Thanks to the expression of recombinant proteins in *C.burnetii* using a shuttle plasmid, *Chen et al.* demonstrated the functionality of T4SS, confirming previous findings based on *L.pneumophila* transformation [101]. Moreover, *C.burnetii* mutants with deletion of genes dotA and dotB, which are components of type IVB secretion systems, appeared unable to secrete T4BSS substrates to infect the host cells and to produce the replication vacuole typical of *C.burnetii* [102].

Legionella and *Coxiella* also encode components of Type 1 Secretion System (T1SS). In *L. pneumophila*, it has been recently demonstrated that T1SS is functional and involved in the internalisation of the bacterium into its host cells [103]. In *C.burnetii* Nine Mile, some effectors proteins of this system are mutated but this is possibly functional [19]. In contrast, in the Cb175 strain from French Guiana, T1SS operon is characterised by a large 6105 pb deletion and the same gene loss is observed in *L. longbeachae*. Further experimental studies are needed to determine the exact role of T1SS in *C.burnetii* and to decipher the impact of such deletions in this region.

1.5.2. Response to stress conditions

Several studies have been performed in order to elucidate the *C.burnetii* protection strategy under the exceptionally stressful conditions of the phagolysosome. *C.burnetii* encodes an unusually high number of basic proteins, comparable to that which is found in *Helicobacter pylori*, which resides in the extremely acidic environment of the gastric mucosa [2]. These proteins are most probably involved in buffering the acidic environment of the phagolysosome. In the same vein, four predicted sodium ion/proton exchangers are present in *C.burnetii* genome. Also, genes involved in vacuole detoxification and transporters for

osmoprotectants have been identified, enabling to face osmotic and oxidative stress conditions [2]. When it is outside of its cell host, *C.burnetii* is in a spore-like form resisting desiccation, chemical products, disinfectants and UV radiation [1]. It is also highly resistant to temperature stress, as the need for pasteurisation has been established in order to eliminate *C.burnetii* [104]. A micro-array transcriptional analysis was performed by Leroy *et al.* to study genes involved in the response to temperature stresses. Under cold shock and heat shock stress conditions, genes that are differentially expressed are organised into regulation clusters [105]. The same genes are differently expressed in these two different conditions. These genes are associated with membrane biogenesis, particularly lipopolysaccharide and peptidoglycan synthesis, bacterial division, and bacterial sporulation [105].

1.5.3. Metabolism

C.burnetii's genome contains a high proportion of transporters for organic nutrients when compared to *Chlamydia* and *Rickettsia*. The biosynthetic pathway is found for 11 amino acids, although none of them are complete. The presence of 15 amino acids and three peptide transporters suggests that the bacterium acquires these substrates from the host to compensate for this deficiency [2]. Enzymes for glucose metabolism, the electron transport chain, pentose phosphate pathway and tricarboxylic pathway are found. Pathways of lipids and phospholipids, purine and pyrimidine, coenzymes and cofactors are also intact. Pathways for the use of glucose, galactose, glycerol and xylose are present, in contrast to enzymes involved in cholesterol synthesis, with the exception of two sterol reductases (CBU1206 and CBU1158), suggesting the uptake and conversion of metabolites from the host, and not a *de novo* biosynthesis [2].

1.6. Pathogenicity and genomic characteristics of *C.burnetii* strains

Q fever clinical polymorphism and its link with *C. burnetii* strains was investigated at a very early stage in Q fever studies [106]. In the absence of an available genomic sequence, initial studies on this issue were based on the comparison of plasmid types [13, 47, 51, 107-111], in an attempt to find a link between a given plasmid and the 'acute' or 'chronic' forms of the disease. Animal models have also been used to compare strains pathogenicity [112-116] (**BOX 2-3**) but extrapolation of these results to human beings seems hazardous to date. Globally, these methods have failed to precisely describe virulence determinants in *C.burnetii* and this failure illustrates that the traditional dichotomy between 'acute' and 'chronic' Q fever is without biological foundation.

BOX 2

Plasmid types and pathogenicity

First, a correlation of plasmid type with clinical issues was proposed [13]. Strains harbouring QpH1 plasmid were associated with the acute form, while those harbouring QpRS plasmid or QpRS chromosomal integrated sequence were associated with endocarditis, abortion and the other most 'chronic' forms of Q fever [51]. To investigate this correlation more thoroughly, it has been suggested that unique specific sequences for each plasmid were correlated with virulence. In this context, the gene *cbbE'*, encoding for a surface protein, was found to be specific to plasmid QpRS [107] and the gene *cbhE'*, encoding for a hydrophilic protein was specific to QpH1 plasmid [108]. However, some years later, the finding of a novel QpDV plasmid [109], detected in both 'acute' and 'chronic'isolates of *C.burnetii* suggested that the classification in plasmid groups was not sufficiently comprehensive [47]. Finally, Stein *et al.* demonstrated that the *CbhE'* gene specific to QpH1 plasmid could be detected by PCR in isolates from 'chronic' Q fever and was not systematically detected in isolates of 'acute' Q fever. [47]. Further studies confirmed that neither type of plasmid nor gene unique to a type of plasmid could be used as reliable marker for distinguishing different pathotypes [110]. However, a recent correlation has been found between abortion in pregnant women and strains harbouring a QpDV plasmid [111], although further analyses are needed to determine whether this result is correlated with specific plasmidic sequences or particular associated genomic virulence properties of these strains.

BOX 3

Comparison of strain pathogenicity in animal models

The pathogenicity of several strains belonging to four different genomic groups (based on MST genotyping) has been compared in immunocompetent and immunocompromised mice and guinea pigs [112]. The authors found that development of the disease after infection with Priscilla strain and Q212 strain was slower and milder than infection with Nine Mile strain for example. Globally, strains belonging to genomic group I were responsible for more rapidly progressing and more severe disease and determined a stronger immune response than other strains [112, 113]. Nine Mile strain appeared to be the most pathogenic, being the only one to cause an inflammatory reaction in infected mice [114]. However, these results should be interpreted with caution. Currently, comparing observations in experimental models with naturally infected humans could sometimes be erroneous in certain aspects and the research for a perfect animal model that truly mimics Q fever in humans is still a hurdle to overcome. *C. burnetii* is pathogenic but has low virulence in murine rodents which show an inflammatory response with lower febrile disease and minimal signs of infection compared to humans. Large doses of phase I *C. burnetii* are necessary to induce endocarditis in mice [112]. In contrast, in guinea pigs Q fever may represent a life threatening disease [115]. It has been hypothesised that these differences are due to the ability of mice to restore affected cardiac valves [116]. By means of example, in the study by Samuel *et al.*, the Canadian strain Q212 [114] was much less pathogenic than Nine Mile Strain in a murine model. Nevertheless, this clone (MST 21) was the only one to be detected in Canada and to have triggered several outbreaks of both acute infections and endocarditis. Unfortunately, this is evidence that the human pathogenicity of a given *C.burnetii* strain cannot be deduced from an animal model.

1.6.1. Gene loss and strain virulence

The development of sequencing methods and massive comparative genomic analyses have revolutionised the paradigm of virulence in microbiology. During the pre-genomic era, the virulence of bacteria was thought to be linked to the acquisition of 'virulence genes'. However, several studies have shown that some genes supposed to code for 'virulence factors' were also present in non-pathogenic species [117]. It has now been demonstrated that specialisation of bacteria to eukaryotic hosts and evolution to hyper pathogenicity is rather driven by genome reduction with massive gene loss [117]. This is illustrated by a recent study comparing the 12 most dangerous bacterial species of all times for humans ('bad bugs') to their phylogenetically closest, but non-epidemic species. Ten functional COG categories contained significantly fewer genes in 'bad bugs': transcription, regulation, mechanisms of signal transduction, categories of metabolism, cell motility and energy production [118]. Five 'bad bugs' had a higher large number of poly (A) tails compared to their non-epidemic relatives. The pseudogenisation mechanism begins with a shift to a higher AT nucleotide composition and consequently to an excess homopolymer (poly (A) tails). This leads to accumulation of stop codons, which create cut or damaged genes that are subsequently removed through large deletions [119]. Another example of genomic reduction has been found in *Mycobacteria*. Leprosy bacillus lost 2000 genes after its divergence from the last common mycobacterial ancestor [120-125]. This phenomenon has also been observed between *Shigella spp* and *E.coli*. These two bacteria have always been considered to be very close and have even been classified as a single species [126]. However, *Shigella spp* causes entero-invasive disease in humans, whereas *E. coli*, with the exception of some pathogenic clones, is a human commensal [127]. It has been demonstrated that loss of the lysine enzyme decarboxylase (LDC) gene is linked to the increased virulence of *S. dysenteriae* by suppressing the inhibition of enterotoxins [127].

Rickettsiales have represented the best model for studying reductive evolution [128, 129]. Sequencing of the *Rickettsia prowazekii* genome, the most dangerous epidemic member of the genus Rickettsia, has not shown the presence of virulence factors [129]. Genome comparison of *R. prowazekii* with the less virulent *R. conorii* revealed that *R. prowazekii* is a subset of *R. conorii*, with only 834 open reading frames (ORFs) compared to the 1,374 ORFs of *R. conorii* [128]. Similarly, a comparison of *R. africae* with highly pathogenic *R. rickettsii* showed the loss of genes in *R. rickettsii* [130]. In general, pathogenic Rickettsia species lack what have been defined as 'pathogenicity islands' [131]. During the passage from extracellular to intracellular lifestyle, Rickettsiaceae lost 2135 genes and, in particular, they lost the entire mismatch excision repair and a large part of the recombinational repair machinery. *R. prowazekii* and *R. typhi* have lost the greatest number of ancestral genes. Taken together, these findings favour the hypothesis that genome reduction, due to genetic isolation in a specific host, consists of the loss of genes mainly involved in metabolic regulation pathways, repair machinery and secretion systems, leading to a deregulated multiplication in the host which increases pathogenicity [118]. In the case of *C.burnetii*, genome sequencing of strains from exceptional epidemic situations has enabled the description of links between virulence and strains' genetic characteristics, in some cases involving gene loss.

1.6.2. French Guiana

Sporadic cases of Q fever have been described in French Guiana until 1998, when a study showed an increase in the seropositivity rate of Q fever from 1.9% in 1992 to 23.9% in 1996, in serum samples of febrile patients. [132].The annual incidence then increased from 37 cases per 100,000 inhabitants in 1996 to 150 cases per 100,000 inhabitants in 2005 [133]. A recent study confirmed that *Coxiella burnetii* was found in 24% of community-acquired pneumonia cases in French Guiana, which is the highest prevalence ever described in the

521 literature [15]. The majority of cases diagnosed since 1996 occurred in Cayenne and its
522 suburbs, indicating an unusual epidemiology of Q fever in this region [6], with no classical
523 reservoir having been identified in previous environmental studies [6]. However we recently
524 detected *C.burnetii* by qPCR in stools, ticks and spleen samples of a dead three-toed sloth
525 [134]. This animal is present in the rainforest hills surrounding Cayenne. We found a
526 correlation between its reproductive season (the rainy season), and the number of cases of Q
527 fever in Cayenne, with a time lag of between one and two months, suggesting that three toed
528 sloths could be the reservoir for the disease [82]. A clinical and serological comparison
529 between 115 patients with Q fever from French Guyana and 182 from metropolitan France
530 showed that acute Q fever has a higher incidence in Cayenne than in Marseilles, with
531 pneumonia as the main clinical presentation and a uniquely strong serological response with
532 high levels of phase I IgG, suggesting the circulation of a hypervirulent clone [17]. After
533 isolation of five *Coxiella burnetii* from biological samples taken from patients from Cayenne
534 (two with endocarditis and three with pneumonia), MST genotyping was performed and all
535 samples showed a genotype MST17 [49]. This genotype seemed to be unique to French
536 Guiana, which is very different from the situation in metropolitan France where 21 different
537 circulating genotypes have been identified [12]. Genome sequencing of the Cb175 strain,
538 (isolated from a Guianese patient with endocarditis) has been performed [19]. Contrary to
539 Z3055, Cb175 showed a significant difference compared to the reference Nine Mile strain: a
540 large deletion consisting of 6105 bp, resulting in a genome reduction. This deletion has only
541 been detected using a specific qPCR in eight other strains isolated from this territory and in
542 none of 298 *C.burnetii* isolates from other areas, having different genotypes (8/8 vs 0/298,
543 Fisher's exact test, $p < 0.0000001$) [19]. Genome reduction is consistent with previous
544 findings that the most dangerous epidemic bacteria, compared with their closest non-epidemic
545 species, are characterised by reduced genomes [121, 130, 135-139] and contain fewer

546 secretion system proteins [118]. Interestingly, the missing region contained the type I
547 secretion system (T1SS) hlyCABD operon and the same deletion is observed in *Legionella*
548 *longbeachae*. Fuche *et al.* recently showed that T1SS is functional in *L. pneumophila* and is
549 involved in the mechanism of internalisation into host cells [140]. Further studies are needed
550 to understand the exact mechanism by which this deletion in T1SS increases the virulence of
551 MST 17 strains.

552 **1.6.3. Canada**

553 In spring 1982, the 'Jordan Bay' Q fever outbreak occurred in a village in Nova Scotia
554 among thirteen members of a family and their friends after a cat parturition [141]. Fever,
555 headache, myalgia and malaise were the principal symptoms accompanied by a four-fold rise
556 in antibody titers for phase II antigens; atypical pneumonia and urinary and gastrointestinal
557 pains were also reported [141]. Between May and August 1985, an outbreak of Q fever
558 involved 24 confirmed cases in the village of Baddeck, Nova Scotia [142]. The source was
559 also a parturient cat which had bled per vaginum for three weeks prior to delivery and had
560 antibodies to *Coxiella burnetii* [142]. Another outbreak was described in 1987 in a truck
561 repair plant where one of the employees, his wife and 16 of the 32 employees developed Q
562 fever [143]. He owned a cat which turned out to have high antibody titers to phase I and II
563 *Coxiella burnetii* antigens [143]. Later, in 1989, another cat-associated outbreak was
564 described in eastern Maine, contiguous to maritime Canada: acute Q fever developed among
565 15 members of a family following a reunion during which they had been exposed to a
566 parturient cat [144]. Q fever was also caused by a parturient dog in 1994 in Nova Scotia in
567 three members of the same family [145]. We sequenced the strain responsible for this last
568 outbreak; DOG UTAD which harbours the MST21 genotype, the only genotype detected in
569 Canadian isolates to date [22, 26]. The genome of this strain is very similar to Q212, another
570 strain isolated from a Q fever endocarditis in Canada, harbouring the same genotype. The

571 DOG UTAD strain is plasmidless and shows 70 mutations in genes of which 47 are non-
572 synonymous compared to the Q212 strain. This low number of mutations suggests a very
573 short genetic distance between these two strains, indicating a recent clonal radiation of
574 MST21 in Canada [26].

575 **1.6.4. Holland**

576 The largest Q fever outbreak was reported in Netherlands between 2007 and 2010
577 [146]. The rapid intensification of goat farming contributed to the epidemics, with a
578 serological attack rate of 92% [147]. A uniquely predominant genotype, MST 33, was
579 detected in dairy goatherds, in one sheep in the southern Netherlands and on a farm in the
580 eastern part of the country [148]. It represented 91% of identified genotypes while nine other
581 genotypes accounted for only 0.8% of all circulating genotypes. The most probable
582 explanation is the emergence of a new genotype, responsible for abortion in dairy goats,
583 which could spread more easily over the dense goat population in the southeast [148]. This
584 genotype was found in samples from humans and from goats and sheep, confirming the
585 hypothesis that goats and sheep are the source of the human Dutch Q fever outbreak [62].
586 MST33 probably spread to the Netherlands from Germany via France [62, 69], because it was
587 isolated in sporadic human cases in France in 1996, 1998, and 1999 and from the placenta of
588 an asymptomatic ewe in Germany in 1992 [62].

589 Since the genome of the Netherlands outbreak strain is not available, a genomic
590 analysis of strain Z3055 isolated from a ewe placenta in Germany in 1992 belonging to
591 genotype MST33 was performed. A comparative genomic analysis using five other whole
592 genomes available on GenBank at the time of the study was conducted [23]. Genome analysis
593 showed an absence of gene loss or gain, with only slight differences to the reference Nine
594 Mile consisting in point mutations, consistent with a clinical spectrum which is not that

595 different from that which has previously been described in the literature [149, 150].
596 Interestingly, a high proportion of mutated proteins were found among membrane proteins,
597 ankyrin repeat domains containing proteins and proteins involved in transcription and
598 translation processes. Similarly to the influenza virus [151], a possible explanation for the
599 outbreak is that mutation in genes encoding for membrane proteins determined changes in
600 surface antigens, allowing the clone to escape the host immune response of goats and spread
601 rapidly [23].

602 **1.7. Therapeutic tools**

603 **1.7.1. Antibiotic susceptibility**

604 Currently, the treatment for Q fever is based on doxycycline, which belongs to the
605 tetracycline family and is able to inhibit bacterial protein synthesis by targeting the 30S
606 ribosomal subunit [1]. It has also been demonstrated that new generation antibiotics, such as
607 telithromycin [152] and tigecycline, are effective against *C. burnetii* [153].
608 To date, some isolates have been shown to be resistant to doxycycline, such as the German
609 strain Cb109 [24]. It was isolated from the cardiac valve of a patient with endocarditis who
610 died during treatment. Analysis of the draft genome of this strain did not reveal any genetic
611 differences with other strains which could explain its resistance; further and more detailed
612 analysis could be useful to speculate on the reason for this important feature [24].

613 Thanks to genomics, some mechanisms of antibiotic resistance have been studied. For
614 instance, genomics allowed for speculation on the mechanism of a high level of resistance to
615 quinolones, which consisted of mutations in DNA gyrase (gyrA) genes resulting in amino
616 acid substitution of Gly in place of Glu at position 87 and causing alterations to the GyrA
617 protein, necessary for replication [154]. The fluoroquinolone-resistant *C.burnetii* Nine Mile I
618 developed *in vitro* showed the up-regulation of glutathione S-transferase and fabZ. The

susceptible strain was characterised by up-regulation of 13 proteins, of which one enhanced entry protein, reducing the ability to penetrate the host cell in the presence of quinolones [155]. The *C.burnetii* proteome has also been studied to investigate the mechanism of resistance to tetracycline. Overexpressed proteins include: a signal peptidase I, involved in cell growth and division; a transaldolase, that provides NADPH cells responding to a higher energy requirement to excrete the antibiotic; the DNA polymerase III alpha subunit, which may induce resistance to antibiotics by mutagenicity; and the 3-methyl-2-oxobutanoate hydroxymethyltransferase, which reduces affinity between antibiotics and ribosome. Moreover, genes encoding the iron transport protein B, reducing activation of cytokine and factor pY, increasing synthesis of proteins to respond to protein synthesis inhibition determined by tetracycline are down-regulated [156].

1.7.2. Putative antigens: surfaceome and vaccinology

Since Q fever in humans is often an occupational hazard, vaccination is considered in exposed populations and in immunocompromised patients or those with cardiac valvular dysfunctions [1]. To date, three types of vaccine have been proposed for Q fever: the attenuated live vaccine, produced and tested in Russia but discarded because of possible consequences for human health; the chloroform-methanol residue extracted vaccine or other extracted vaccines, which have been tested in animals but not in humans; and the whole-cell formalin-inactivated vaccine (Q-Vax), considered to be quite safe for humans and with 98% efficacy [157]. Zhang *et al*. combined 2-D gel and genomics to study possible specific proteins of *C.burnetii* that could be used as an antigen for the development of vaccines. After screening a genomic DNA library with convalescent-phase sera from mice, 20 different immunoreactive proteins were cloned and identified [55]. Further studies are needed to confirm the possible use of these proteins as vaccine candidates.

2. Conclusion

The introduction of genomics has led to important changes in the management of Q fever and the study of *Coxiella burnetii*. It has been possible to speculate on genetic diversity and evolution by performing comparative genomic analysis and accurate phylogeny. The epidemiology of Q fever around the world is better known thanks to the description of 'geotypes'. Genomics also enabled the development of new diagnostic tools, based on qPCR, genotyping of strains and axenic culture. The detection of putative mechanisms of resistance to antibiotics, as well as the immunodominant antigens which are potential targets for vaccines are also important developing fields.

In particular, the sequencing and analysis of the *C.burnetii* genome of epidemic strains made it possible to study the potential links between pathogenicity and genetic properties. The discovery of a unique genome reduction in the hypervirulent strain from French Guiana is the best illustration of this phenomenon. This is a crucial breakthrough in the study of Q fever epidemiology, showing that *C.burnetii* causes more or less severe clinical presentation during primo infection, depending mainly on genomic determinants. On the contrary, to date, no genetic determinants appear to be linked to a 'chronic' or 'acute' form of the disease, confirming that this dichotomy is outdated. The persistence of the disease in an infectious focus seems to be largely determined by predisposing host factors (age, sex, valvulopathy, vascular aneurysms, joint prosthesis and pregnancy).

We are only at the beginning of the description of *C.burnetii* genetic diversity and its links with pathogenicity and epidemiology.

Future perspective

A more comprehensive pangenome analysis including more than 40 strains collected in our laboratory is currently underway to provide an overview of *C. burnetii* clonal diversity and its relation to virulence. These advances will significantly help clinicians and epidemiologists with the diagnosis and management of Q fever around the world.

Executive summary

- Genomics has been crucial for the rapid progress in the knowledge of *Coxiella burnetii* and to improve the management of Q fever
- Genomic phylogenetic information allowed understanding the correct taxonomy of *C.burnetii* that belongs to the family of gammaproteobacteria in the order of Legionellales.
- Multispacer Sequence Typing (MST) can be considered as a "geotyping" tool that provides a complete cartography of the genetic diversity of *C.burnetii*.
- Thanks to genomics, IS1111 qPCR has been developed for early detection of the bacterium in the serum of patients when serology is still negative
- Genomics helped to develop the ACCM axenic medium that allows host-cell free growth of *C.burnetii* and enabled the genetic manipulation of this bacterium
- The dichotomy between 'acute' and 'chronic' Q fever is without biological foundation; the genomic determinants of each strain are mainly responsible for more or less severe clinical presentation during primo infection
- Genome sequencing of strains from exceptional epidemic situations has enabled the description of links between strains' genetic profile and virulence. In some cases this implicates gene loss and genomic reduction
- Preliminary pangenomic study suggested that *C.burnetii* has a closed pangenome

References

1. Maurin M, Raoult D. Q fever. *Clinical microbiology reviews* 12(4), 518-553 (1999).
2. Seshadri R, Paulsen IT, Eisen JA *et al*. Complete genome sequence of the q-fever pathogen coxiella burnetii. *Proceedings of the National Academy of Sciences of the United States of America* 100(9), 5455-5460 (2003).
3. Beare PA, Howe D, Cockrell DC, Omsland A, Hansen B, Heinzen RA. Characterization of a coxiella burnetii ftsz mutant generated by himar1 transposon mutagenesis. *Journal of bacteriology* 191(5), 1369-1381 (2009).
4. Omsland A, Cockrell DC, Howe D *et al*. Host cell-free growth of the q fever bacterium coxiella burnetii. *Proceedings of the National Academy of Sciences of the United States of America* 106(11), 4430-4434 (2009).
5. Omsland A, Beare PA, Hill J *et al*. Isolation from animal tissue and genetic transformation of coxiella burnetii are facilitated by an improved axenic growth medium. *Applied and environmental microbiology* 77(11), 3720-3725 (2011).
6. Gardon J, Heraud JM, Laventure S *et al*. Suburban transmission of q fever in french guiana: Evidence of a wild reservoir. *The Journal of infectious diseases* 184(3), 278-284 (2001).
7. Million M, Walter G, Thuny F, Habib G, Raoult D. Evolution from acute q fever to endocarditis is associated with underlying valvulopathy and age and can be prevented by prolonged antibiotic treatment. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* 57(6), 836-844 (2013).
8. Botelho-Nevers E, Fournier PE, Richet H *et al*. Coxiella burnetii infection of aortic aneurysms or vascular grafts: Report of 30 new cases and evaluation of outcome. *European journal of clinical microbiology & infectious diseases : official publication of the European Society of Clinical Microbiology* 26(9), 635-640 (2007).
9. Million M, Raoult D. Recent advances in the study of q fever epidemiology, diagnosis and management. *The Journal of infection* 71 Suppl 1, S2-9 (2015).
10. Raoult D. Chronic q fever: Expert opinion versus literature analysis and consensus. *The Journal of infection* 65(2), 102-108 (2012).
11. Tissot-Dupont H, Vaillant V, Rey S, Raoult D. Role of sex, age, previous valve lesion, and pregnancy in the clinical expression and outcome of q fever after a large outbreak. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* 44(2), 232-237 (2007).
12. Carcopino X, Raoult D, Bretelle F, Boubli L, Stein A. Q fever during pregnancy: A cause of poor fetal and maternal outcome. *Annals of the New York Academy of Sciences* 1166, 79-89 (2009).
13. Samuel JE, Frazier ME, Mallavia LP. Correlation of plasmid type and disease caused by coxiella burnetii. *Infection and immunity* 49(3), 775-779 (1985).
14. Glazunova O, Roux V, Freylikman O *et al*. Coxiella burnetii genotyping. *Emerging infectious diseases* 11(8), 1211-1217 (2005).
15. Epelboin L, Chesnais C, Boulle C *et al*. Q fever pneumonia in french guiana: Prevalence, risk factors, and prognostic score. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* 55(1), 67-74 (2012).
16. Eldin C, Mahamat A, Demar M, Abboud P, Djossou F, Raoult D. Q fever in french guiana. *The American journal of tropical medicine and hygiene* 91(4), 771-776 (2014).
17. Edouard S, Mahamat A, Demar M, Abboud P, Djossou F, Raoult D. Comparison between emerging q fever in french guiana and endemic q fever in marseille, france. *The American journal of tropical medicine and hygiene* 90(5), 915-919 (2014).
18. Mahamat A, Edouard S, Demar M *et al*. Unique clone of coxiella burnetii causing severe q fever, french guiana. *Emerging infectious diseases* 19(7), 1102-1104 (2013).

19. D'amato F, Eldin C, Georgiades K *et al*. Loss of tss1 in hypervirulent coxiella burnetii 175, the causative agent of q fever in french guiana. *Comparative immunology, microbiology and infectious diseases*, (2015).
20. Leroy Q, Armougom F, Barbry P, Raoult D. Genomotyping of coxiella burnetii using microarrays reveals a conserved genomotype for hard tick isolates. *PloS one* 6(10), e25781 (2011).
21. Fournier PE, Raoult D. Prospects for the future using genomics and proteomics in clinical microbiology. *Annual review of microbiology* 65, 169-188 (2011).
22. Beare PA, Unsworth N, Andoh M *et al*. Comparative genomics reveal extensive transposon-mediated genomic plasticity and diversity among potential effector proteins within the genus coxiella. *Infection and immunity* 77(2), 642-656 (2009).
23. D'amato F, Rouli L, Edouard S *et al*. The genome of coxiella burnetii z3055, a clone linked to the netherlands q fever outbreaks, provides evidence for the role of drift in the emergence of epidemic clones. *Comparative immunology, microbiology and infectious diseases* 37(5-6), 281-288 (2014).
24. Rouli L, Rolain JM, El Filali A, Robert C, Raoult D. Genome sequence of coxiella burnetii 109, a doxycycline-resistant clinical isolate. *Journal of bacteriology* 194(24), 6939 (2012).
25. Sidi-Boumedine K, Ellis RJ, Adam G *et al*. Draft genome sequences of six ruminant coxiella burnetii isolates of european origin. *Genome announcements* 2(3), (2014).
26. D'amato F, Million M, Edouard S *et al*. Draft genome sequence of coxiella burnetii dog utad, a strain isolated from a dog-related outbreak of q fever. *New microbes and new infections* 2(4), 136-137 (2014).
27. D'amato F, Robert C, Azhar EI, Fournier PE, Raoult D. Draft genome sequence of coxiella burnetii strain cb196, an agent of endocarditis in saudi arabia. *Genome announcements* 2(6), (2014).
28. Walter MC, Ohrman C, Myrtennas K *et al*. Genome sequence of coxiella burnetii strain namibia. *Standards in genomic sciences* 9, 22 (2014).
29. Million M, Roblot F, Carles D *et al*. Reevaluation of the risk of fetal death and malformation after q fever. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* 59(2), 256-260 (2014).
30. Walter MC, Vincent GA, Stenos J, Graves S, Frangoulidis D. Genome sequence of coxiella burnetii strain auq01 (arandale) from an australian patient with acute q fever. *Genome announcements* 2(5), (2014).
31. Rasko DA, Myers GS, Ravel J. Visualization of comparative genomic analyses by blast score ratio. *BMC bioinformatics* 6, 2 (2005).
32. Sahl JW, Gillece JD, Schupp JM *et al*. Evolution of a pathogen: A comparative genomics analysis identifies a genetic pathway to pathogenesis in acinetobacter. *PloS one* 8(1), e54287 (2013).
33. Weiss E MJ. Order i rickettsiales, gieszczkiewicz 1939, 25 al. *In: Krieg NR, Holt JG, eds. Bergey's Manual of Systematic Bacteriology.* Vol. 1. The William & Wilkins Co. Baltimore, p.687-704. (1984).
34. Weisburg WG, Dobson ME, Samuel JE *et al*. Phylogenetic diversity of the rickettsiae. *Journal of bacteriology* 171(8), 4202-4206 (1989).
35. Roux V, Bergoin M, Lamaze N, Raoult D. Reassessment of the taxonomic position of rickettsiella grylli. *International journal of systematic bacteriology* 47(4), 1255-1257 (1997).
36. Greub G, Raoult D. Microorganisms resistant to free-living amoebae. *Clinical microbiology reviews* 17(2), 413-433 (2004).
37. La Scola B, Raoult D. Survival of coxiella burnetii within free-living amoeba acanthamoeba castellanii. *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases* 7(2), 75-79 (2001).

38. Lamrabet O, Merhej V, Pontarotti P, Raoult D, Drancourt M. The genealogic tree of mycobacteria reveals a long-standing sympatric life into free-living protozoa. *PloS one* 7(4), e34754 (2012).
39. Tan CK, Owens L. Infectivity, transmission and 16s rrna sequencing of a rickettsia, coxiella cheraxi sp. Nov., from the freshwater crayfish cherax quadricarinatus. *Diseases of aquatic organisms* 41(2), 115-122 (2000).
40. Smith TA, Driscoll T, Gillespie JJ, Raghavan R. A coxiella-like endosymbiont is a potential vitamin source for the lone star tick. *Genome biology and evolution* 7(3), 831-838 (2015).
41. Duron O, Noel V, Mccoy KD *et al*. The recent evolution of a maternally-inherited endosymbiont of ticks led to the emergence of the q fever pathogen, coxiella burnetii. *PLoS pathogens* 11(5), e1004892 (2015).
42. Shigenobu S, Watanabe H, Hattori M, Sakaki Y, Ishikawa H. Genome sequence of the endocellular bacterial symbiont of aphids buchnera sp. Aps. *Nature* 407(6800), 81-86 (2000).
43. Lechner M, Findeiss S, Steiner L, Marz M, Stadler PF, Prohaska SJ. Proteinortho: Detection of (co-)orthologs in large-scale analysis. *BMC bioinformatics* 12, 124 (2011).
44. Munson MA, Baumann P, Clark MA *et al*. Evidence for the establishment of aphid-eubacterium endosymbiosis in an ancestor of four aphid families. *Journal of bacteriology* 173(20), 6321-6324 (1991).
45. Moran NA. Microbial minimalism: Genome reduction in bacterial pathogens. *Cell* 108(5), 583-586 (2002).
46. Moran NA, Plague GR, Sandstrom JP, Wilcox JL. A genomic perspective on nutrient provisioning by bacterial symbionts of insects. *Proceedings of the National Academy of Sciences of the United States of America* 100 Suppl 2, 14543-14548 (2003).
47. Stein A, Raoult D. Lack of pathotype specific gene in human coxiella burnetii isolates. *Microbial pathogenesis* 15(3), 177-185 (1993).
48. Masuzawa T, Sawaki K, Nagaoka H, Akiyama M, Hirai K, Yanagihara Y. Identification of rickettsiae isolated in japan as coxiella burnetii by 16s rrna sequencing. *International journal of systematic bacteriology* 47(3), 883-884 (1997).
49. Stein A, Kruszewska D, Gouvernet J, Raoult D. Study of the 16s-23s ribosomal DNA internal spacer of coxiella burnetii. *European journal of epidemiology* 13(4), 471-475 (1997).
50. Mollet C, Drancourt M, Raoult D. Determination of coxiella burnetii rpob sequence and its use for phylogenetic analysis. *Gene* 207(1), 97-103 (1998).
51. Hendrix LR, Samuel JE, Mallavia LP. Differentiation of coxiella burnetii isolates by analysis of restriction-endonuclease-digested DNA separated by sds-page. *Journal of general microbiology* 137(2), 269-276 (1991).
52. Heinzen R, Stiegler GL, Whiting LL, Schmitt SA, Mallavia LP, Frazier ME. Use of pulsed field gel electrophoresis to differentiate coxiella burnetii strains. *Annals of the New York Academy of Sciences* 590, 504-513 (1990).
53. Jager C, Willems H, Thiele D, Baljer G. Molecular characterization of coxiella burnetii isolates. *Epidemiology and infection* 120(2), 157-164 (1998).
54. Nguyen SV, Hirai K. Differentiation of coxiella burnetii isolates by sequence determination and pcr-restriction fragment length polymorphism analysis of isocitrate dehydrogenase gene. *FEMS microbiology letters* 180(2), 249-254 (1999).
55. Zhang G, Kiss K, Seshadri R, Hendrix LR, Samuel JE. Identification and cloning of immunodominant antigens of coxiella burnetii. *Infection and immunity* 72(2), 844-852 (2004).
56. Sekeyova Z, Roux V, Raoult D. Intraspecies diversity of coxiella burnetii as revealed by com1 and mucz sequence comparison. *FEMS microbiology letters* 180(1), 61-67 (1999).
57. Beare PA, Samuel JE, Howe D, Virtaneva K, Porcella SF, Heinzen RA. Genetic diversity of the q fever agent, coxiella burnetii, assessed by microarray-based whole-genome comparisons. *Journal of bacteriology* 188(7), 2309-2324 (2006).

58. Svraka S, Toman R, Skultety L, Slaba K, Homan WL. Establishment of a genotyping scheme for coxiella burnetii. *FEMS microbiology letters* 254(2), 268-274 (2006).
59. Arricau-Bouvery N, Hauck Y, Bejaoui A *et al*. Molecular characterization of coxiella burnetii isolates by infrequent restriction site-pcr and mlva typing. *BMC microbiology* 6, 38 (2006).
60. Frangoulidis D, Walter MC, Antwerpen M *et al*. Molecular analysis of coxiella burnetii in germany reveals evolution of unique clonal clusters. *International journal of medical microbiology : IJMM*, (2014).
61. Chmielewski T, Sidi-Boumedine K, Duquesne V, Podsiadly E, Thiery R, Tylewska-Wierzbanowska S. Molecular epidemiology of q fever in poland. *Polish journal of microbiology / Polskie Towarzystwo Mikrobiologow = The Polish Society of Microbiologists* 58(1), 9-13 (2009).
62. Tilburg JJ, Roest HJ, Buffet S *et al*. Epidemic genotype of coxiella burnetii among goats, sheep, and humans in the netherlands. *Emerging infectious diseases* 18(5), 887-889 (2012).
63. Reichel R, Mearns R, Brunton L *et al*. Description of a coxiella burnetii abortion outbreak in a dairy goat herd, and associated serology, pcr and genotyping results. *Research in veterinary science* 93(3), 1217-1224 (2012).
64. Pearson T, Hornstra HM, Hilsabeck R *et al*. High prevalence and two dominant host-specific genotypes of coxiella burnetii in u.S. Milk. *BMC microbiology* 14, 41 (2014).
65. Sulyok KM, Kreizinger Z, Hornstra HM *et al*. Genotyping of coxiella burnetii from domestic ruminants and human in hungary: Indication of various genotypes. *BMC veterinary research* 10, 107 (2014).
66. Santos AS, Tilburg JJ, Botelho A *et al*. Genotypic diversity of clinical coxiella burnetii isolates from portugal based on mst and mlva typing. *International journal of medical microbiology : IJMM* 302(6), 253-256 (2012).
67. Angelakis E, Johani S, Ahsan A, Memish Z, Raoult D. Q fever endocarditis and new coxiella burnetii genotype, saudi arabia. *Emerging infectious diseases* 20(4), 726-728 (2014).
68. Tilburg JJ, Roest HJ, Nabuurs-Franssen MH, Horrevorts AM, Klaassen CH. Genotyping reveals the presence of a predominant genotype of coxiella burnetii in consumer milk products. *Journal of clinical microbiology* 50(6), 2156-2158 (2012).
69. Tilburg JJ, Rossen JW, Van Hannen EJ *et al*. Genotypic diversity of coxiella burnetii in the 2007-2010 q fever outbreak episodes in the netherlands. *Journal of clinical microbiology* 50(3), 1076-1078 (2012).
70. Fournier PE, Marrie TJ, Raoult D. Diagnosis of q fever. *Journal of clinical microbiology* 36(7), 1823-1834 (1998).
71. Ibrahim A, Norlander L, Macellaro A, Sjostedt A. Specific detection of coxiella burnetii through partial amplification of 23s rdna. *European journal of epidemiology* 13(3), 329-334 (1997).
72. Willems H, Thiele D, Frolich-Ritter R, Krauss H. Detection of coxiella burnetii in cow's milk using the polymerase chain reaction (pcr). *Zentralblatt fur Veterinarmedizin. Reihe B. Journal of veterinary medicine. Series B* 41(9), 580-587 (1994).
73. Lorenz H, Jager C, Willems H, Baljer G. Pcr detection of coxiella burnetii from different clinical specimens, especially bovine milk, on the basis of DNA preparation with a silica matrix. *Applied and environmental microbiology* 64(11), 4234-4237 (1998).
74. Vaidya VM, Malik SV, Bhilegaonkar KN, Rathore RS, Kaur S, Barbuddhe SB. Prevalence of q fever in domestic animals with reproductive disorders. *Comparative immunology, microbiology and infectious diseases* 33(4), 307-321 (2010).
75. Fenollar F, Raoult D. Molecular genetic methods for the diagnosis of fastidious microorganisms. *APMIS : acta pathologica, microbiologica, et immunologica Scandinavica* 112(11-12), 785-807 (2004).
76. Fenollar F, Fournier PE, Raoult D. Molecular detection of coxiella burnetii in the sera of patients with q fever endocarditis or vascular infection. *Journal of clinical microbiology* 42(11), 4919-4924 (2004).

77. Kato K, Arashima Y, Asai S *et al*. Detection of coxiella burnetii specific DNA in blood samples from japanese patients with chronic nonspecific symptoms by nested polymerase chain reaction. *FEMS immunology and medical microbiology* 21(2), 139-144 (1998).
78. Zhang GQ, Hotta A, Mizutani M *et al*. Direct identification of coxiella burnetii plasmids in human sera by nested pcr. *Journal of clinical microbiology* 36(8), 2210-2213 (1998).
79. Denison AM, Thompson HA, Massung RF. Is1111 insertion sequences of coxiella burnetii: Characterization and use for repetitive element pcr-based differentiation of coxiella burnetii isolates. *BMC microbiology* 7, 91 (2007).
80. Klee SR, Tyczka J, Ellerbrok H *et al*. Highly sensitive real-time pcr for specific detection and quantification of coxiella burnetii. *BMC microbiology* 6, 2 (2006).
81. Klee SR, Ellerbrok H, Tyczka J, Franz T, Appel B. Evaluation of a real-time pcr assay to detect coxiella burnetii. *Annals of the New York Academy of Sciences* 1078, 563-565 (2006).
82. Eldin C, Angelakis E, Renvoise A, Raoult D. Coxiella burnetii DNA, but not viable bacteria, in dairy products in france. *The American journal of tropical medicine and hygiene* 88(4), 765-769 (2013).
83. Angelakis E, Mediannikov O, Socolovschi C *et al*. Coxiella burnetii-positive pcr in febrile patients in rural and urban africa. *International journal of infectious diseases : IJID : official publication of the International Society for Infectious Diseases* 28, 107-110 (2014).
84. Mediannikov O, Fenollar F, Socolovschi C *et al*. Coxiella burnetii in humans and ticks in rural senegal. *PLoS neglected tropical diseases* 4(4), e654 (2010).
85. Grisoli D, Million M, Edouard S *et al*. Latent q fever endocarditis in patients undergoing routine valve surgery. *The Journal of heart valve disease* 23(6), 735-743 (2014).
86. Rolain JM, Raoult D. Molecular detection of coxiella burnetii in blood and sera during q fever. *QJM : monthly journal of the Association of Physicians* 98(8), 615-617; author reply 617-620 (2005).
87. Lok C. Mining the microbial dark matter. *Nature* 522(7556), 270-273 (2015).
88. Hernychova L, Toman R, Ciampor F *et al*. Detection and identification of coxiella burnetii based on the mass spectrometric analyses of the extracted proteins. *Analytical chemistry* 80(18), 7097-7104 (2008).
89. Sekeyova Z, Kowalczewska M, Decloquement P, Pelletier N, Spitalska E, Raoult D. Identification of protein candidates for the serodiagnosis of q fever endocarditis by an immunoproteomic approach. *European journal of clinical microbiology & infectious diseases : official publication of the European Society of Clinical Microbiology* 28(3), 287-295 (2009).
90. Chao CC, Chen HW, Li X, Xu WB, Hanson B, Ching WM. Identification, cloning, and expression of potential diagnostic markers for q fever. *Annals of the New York Academy of Sciences* 1063, 76-78 (2005).
91. Xiong X, Wang X, Wen B, Graves S, Stenos J. Potential serodiagnostic markers for q fever identified in coxiella burnetii by immunoproteomic and protein microarray approaches. *BMC microbiology* 12, 35 (2012).
92. Zhang G, To H, Russell KE *et al*. Identification and characterization of an immunodominant 28-kilodalton coxiella burnetii outer membrane protein specific to isolates associated with acute disease. *Infection and immunity* 73(3), 1561-1567 (2005).
93. Fernandes I, Rousset E, Dufour P *et al*. Serological and molecular characterization of adaa: A potential marker of q-fever abortion in goats? *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases* 15 Suppl 2, 163-164 (2009).
94. Skultety L, Hajduch M, Flores-Ramirez G *et al*. Proteomic comparison of virulent phase i and avirulent phase ii of coxiella burnetii, the causative agent of q fever. *Journal of proteomics* 74(10), 1974-1984 (2011).
95. Voth DE, Heinzen RA. Coxiella type iv secretion and cellular microbiology. *Current opinion in microbiology* 12(1), 74-80 (2009).

96. Zusman T, Yerushalmi G, Segal G. Functional similarities between the icm/dot pathogenesis systems of coxiella burnetii and legionella pneumophila. *Infection and immunity* 71(7), 3714-3723 (2003).
97. Pan X, Luhrmann A, Satoh A, Laskowski-Arce MA, Roy CR. Ankyrin repeat proteins comprise a diverse family of bacterial type iv effectors. *Science* 320(5883), 1651-1654 (2008).
98. Luhrmann A, Nogueira CV, Carey KL, Roy CR. Inhibition of pathogen-induced apoptosis by a coxiella burnetii type iv effector protein. *Proceedings of the National Academy of Sciences of the United States of America* 107(44), 18997-19001 (2010).
99. Voth DE, Beare PA, Howe D *et al*. The coxiella burnetii cryptic plasmid is enriched in genes encoding type iv secretion system substrates. *Journal of bacteriology* 193(7), 1493-1503 (2011).
100. Beare PA, Sandoz KM, Omsland A, Rockey DD, Heinzen RA. Advances in genetic manipulation of obligate intracellular bacterial pathogens. *Frontiers in microbiology* 2, 97 (2011).
101. Chen C, Banga S, Mertens K *et al*. Large-scale identification and translocation of type iv secretion substrates by coxiella burnetii. *Proceedings of the National Academy of Sciences of the United States of America* 107(50), 21755-21760 (2010).
102. Beare PA, Larson CL, Gilk SD, Heinzen RA. Two systems for targeted gene deletion in coxiella burnetii. *Applied and environmental microbiology* 78(13), 4580-4589 (2012).
103. Takamatsu R, Teruya H, Takeshima E *et al*. Molecular characterization of legionella pneumophila-induced interleukin-8 expression in t cells. *BMC microbiology* 10, 1 (2010).
104. Cerf O, Condron R. Coxiella burnetii and milk pasteurization: An early application of the precautionary principle? *Epidemiology and infection* 134(5), 946-951 (2006).
105. Leroy Q, Lebrigand K, Armougom F, Barbry P, Thiery R, Raoult D. Coxiella burnetii transcriptional analysis reveals serendipity clusters of regulation in intracellular bacteria. *PloS one* 5(12), e15321 (2010).
106. Raoult D, Marrie T, Mege J. Natural history and pathophysiology of q fever. *The Lancet. Infectious diseases* 5(4), 219-226 (2005).
107. Minnick MF, Heinzen RA, Reschke DK, Frazier ME, Mallavia LP. A plasmid-encoded surface protein found in chronic-disease isolates of coxiella burnetti. *Infection and immunity* 59(12), 4735-4739 (1991).
108. Minnick MF, Small CL, Frazier ME, Mallavia LP. Analysis of the cbhe' plasmid gene from acute disease-causing isolates of coxiella burnetii. *Gene* 103(1), 113-118 (1991).
109. Valkova D, Kazar J. A new plasmid (qpdv) common to coxiella burnetii isolates associated with acute and chronic q fever. *FEMS microbiology letters* 125(2-3), 275-280 (1995).
110. Thiele D, Willems H. Is plasmid based differentiation of coxiella burnetii in 'acute' and 'chronic' isolates still valid? *European journal of epidemiology* 10(4), 427-434 (1994).
111. Angelakis E, Million M, D'amato F *et al*. Q fever and pregnancy: Disease, prevention, and strain specificity. *European journal of clinical microbiology & infectious diseases : official publication of the European Society of Clinical Microbiology* 32(3), 361-368 (2013).
112. Russell-Lodrigue KE, Andoh M, Poels MW *et al*. Coxiella burnetii isolates cause genogroup-specific virulence in mouse and guinea pig models of acute q fever. *Infection and immunity* 77(12), 5640-5650 (2009).
113. Russell-Lodrigue KE, Zhang GQ, Mcmurray DN, Samuel JE. Clinical and pathologic changes in a guinea pig aerosol challenge model of acute q fever. *Infection and immunity* 74(11), 6085-6091 (2006).
114. Stein A, Louveau C, Lepidi H *et al*. Q fever pneumonia: Virulence of coxiella burnetii pathovars in a murine model of aerosol infection. *Infection and immunity* 73(4), 2469-2477 (2005).
115. Bewley KR. Animal models of q fever (coxiella burnetii). *Comparative medicine* 63(6), 469-476 (2013).
116. La Scola B, Lepidi H, Maurin M, Raoult D. A guinea pig model for q fever endocarditis. *The Journal of infectious diseases* 178(1), 278-281 (1998).

117. Georgiades K. Genomics of epidemic pathogens. *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases* 18(3), 213-217 (2012).
118. Georgiades K, Raoult D. Genomes of the most dangerous epidemic bacteria have a virulence repertoire characterized by fewer genes but more toxin-antitoxin modules. *PloS one* 6(3), e17962 (2011).
119. Medina M, Sachs JL. Symbiont genomics, our new tangled bank. *Genomics* 95(3), 129-137 (2010).
120. Cole ST, Brosch R, Parkhill J *et al*. Deciphering the biology of mycobacterium tuberculosis from the complete genome sequence. *Nature* 393(6685), 537-544 (1998).
121. Cole ST, Eiglmeier K, Parkhill J *et al*. Massive gene decay in the leprosy bacillus. *Nature* 409(6823), 1007-1011 (2001).
122. Cole ST. Comparative and functional genomics of the mycobacterium tuberculosis complex. *Microbiology* 148(Pt 10), 2919-2928 (2002).
123. Stinear TP, Seemann T, Harrison PF *et al*. Insights from the complete genome sequence of mycobacterium marinum on the evolution of mycobacterium tuberculosis. *Genome research* 18(5), 729-741 (2008).
124. Demangel C, Stinear TP, Cole ST. Buruli ulcer: Reductive evolution enhances pathogenicity of mycobacterium ulcerans. *Nature reviews. Microbiology* 7(1), 50-60 (2009).
125. Monot M, Honore N, Garnier T *et al*. Comparative genomic and phylogeographic analysis of mycobacterium leprae. *Nature genetics* 41(12), 1282-1289 (2009).
126. Pupo GM, Lan R, Reeves PR. Multiple independent origins of shigella clones of escherichia coli and convergent evolution of many of their characteristics. *Proceedings of the National Academy of Sciences of the United States of America* 97(19), 10567-10572 (2000).
127. Maurelli AT, Fernandez RE, Bloch CA, Rode CK, Fasano A. "Black holes" and bacterial pathogenicity: A large genomic deletion that enhances the virulence of shigella spp. And enteroinvasive escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America* 95(7), 3943-3948 (1998).
128. Ogata H, Audic S, Renesto-Audiffren P *et al*. Mechanisms of evolution in rickettsia conorii and r. Prowazekii. *Science* 293(5537), 2093-2098 (2001).
129. Bechah Y, El Karkouri K, Mediannikov O *et al*. Genomic, proteomic, and transcriptomic analysis of virulent and avirulent rickettsia prowazekii reveals its adaptive mutation capabilities. *Genome research* 20(5), 655-663 (2010).
130. Fournier PE, El Karkouri K, Leroy Q *et al*. Analysis of the rickettsia africae genome reveals that virulence acquisition in rickettsia species may be explained by genome reduction. *BMC genomics* 10, 166 (2009).
131. Hacker J, Kaper JB. Pathogenicity islands and the evolution of microbes. *Annual review of microbiology* 54, 641-679 (2000).
132. Pfaff F, Francois A, Hommel D *et al*. Q fever in french guiana: New trends. *Emerging infectious diseases* 4(1), 131-132 (1998).
133. Grangier C DM, Ardillon V, Mahamat a, Fournier P, Simmonnet C, Et Al. . Epidemiologie de la fièvre q en guyanne, 1990–2006. . *Le Bulletin de Veille Sanitaire.CIRE Antilles Guyane* 10 (10), 2– 4. (2009).
134. Davoust B, Marie JL, Pommier De Santi V, Berenger JM, Edouard S, Raoult D. Three-toed sloth as putative reservoir of coxiella burnetii, cayenne, french guiana. *Emerging infectious diseases* 20(10), 1760-1761 (2014).
135. Dobrindt U, Hochhut B, Hentschel U, Hacker J. Genomic islands in pathogenic and environmental microorganisms. *Nature reviews. Microbiology* 2(5), 414-424 (2004).
136. Wixon J. Featured organism: Reductive evolution in bacteria: Buchnera sp., rickettsia prowazekii and mycobacterium leprae. *Comparative and functional genomics* 2(1), 44-48 (2001).

137. Sakharkar KR, Dhar PK, Chow VT. Genome reduction in prokaryotic obligatory intracellular parasites of humans: A comparative analysis. *International journal of systematic and evolutionary microbiology* 54(Pt 6), 1937-1941 (2004).
138. Andersson JO, Andersson SG. Insights into the evolutionary process of genome degradation. *Current opinion in genetics & development* 9(6), 664-671 (1999).
139. Merhej V, Raoult D. Rickettsial evolution in the light of comparative genomics. *Biological reviews of the Cambridge Philosophical Society* 86(2), 379-405 (2011).
140. Fuche F, Vianney A, Andrea C, Doublet P, Gilbert C. Functional type 1 secretion system involved in legionella pneumophila virulence. *Journal of bacteriology* 197(3), 563-571 (2015).
141. Kosatsky T. Household outbreak of q-fever pneumonia related to a parturient cat. *Lancet* 2(8417-8418), 1447-1449 (1984).
142. Marrie TJ, Macdonald A, Durant H, Yates L, Mccormick L. An outbreak of q fever probably due to contact with a parturient cat. *Chest* 93(1), 98-103 (1988).
143. Marrie TJ, Langille D, Papukna V, Yates L. Truckin' pneumonia--an outbreak of q fever in a truck repair plant probably due to aerosols from clothing contaminated by contact with newborn kittens. *Epidemiology and infection* 102(1), 119-127 (1989).
144. Pinsky RL, Fishbein DB, Greene CR, Gensheimer KF. An outbreak of cat-associated q fever in the united states. *The Journal of infectious diseases* 164(1), 202-204 (1991).
145. Buhariwalla F, Cann B, Marrie TJ. A dog-related outbreak of q fever. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* 23(4), 753-755 (1996).
146. Speelman P. The largest q fever outbreak ever reported. *The Netherlands journal of medicine* 68(12), 380-381 (2010).
147. Hackert VH, Van Der Hoek W, Dukers-Muijrers N *et al*. Q fever: Single-point source outbreak with high attack rates and massive numbers of undetected infections across an entire region. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* 55(12), 1591-1599 (2012).
148. Roest HI, Ruuls RC, Tilburg JJ *et al*. Molecular epidemiology of coxiella burnetii from ruminants in q fever outbreak, the netherlands. *Emerging infectious diseases* 17(4), 668-675 (2011).
149. Million M, Raoult D. Reply to kampschreur et al. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* 58(3), 447-448 (2014).
150. Munster JM, Leenders AC, Hamilton CJ, Hak E, Aarnoudse JG, Timmer A. Placental histopathology after coxiella burnetii infection during pregnancy. *Placenta* 33(2), 128-131 (2012).
151. Grodeland G, Mjaaland S, Roux KH, Fredriksen AB, Bogen B. DNA vaccine that targets hemagglutinin to mhc class ii molecules rapidly induces antibody-mediated protection against influenza. *Journal of immunology* 191(6), 3221-3231 (2013).
152. Rolain JM, Lambert F, Raoult D. Activity of telithromycin against thirteen new isolates of c. Burnetii including three resistant to doxycycline. *Annals of the New York Academy of Sciences* 1063, 252-256 (2005).
153. Spyridaki I, Psaroulaki A, Vranakis I, Tselentis Y, Gikas A. Bacteriostatic and bactericidal activities of tigecycline against coxiella burnetii and comparison with those of six other antibiotics. *Antimicrobial agents and chemotherapy* 53(6), 2690-2692 (2009).
154. Musso D, Drancourt M, Osscini S, Raoult D. Sequence of quinolone resistance-determining region of gyra gene for clinical isolates and for an in vitro-selected quinolone-resistant strain of coxiella burnetii. *Antimicrobial agents and chemotherapy* 40(4), 870-873 (1996).
155. Vranakis I, De Bock PJ, Papadioti A *et al*. Identification of potentially involved proteins in levofloxacin resistance mechanisms in coxiella burnetii. *Journal of proteome research* 10(2), 756-762 (2011).
156. Vranakis I, De Bock PJ, Papadioti A *et al*. Quantitative proteome profiling of c. Burnetii under tetracycline stress conditions. *PloS one* 7(3), e33599 (2012).

157. Chiu CK, Durrheim DN. A review of the efficacy of human q fever vaccine registered in australia. *New South Wales public health bulletin* 18(7-8), 133-136 (2007).

Figure 1 Application of genomics to the study of Q fever. The branches connect the genomic repertoire with the different aspects of Q fever whose investigation it enabled.

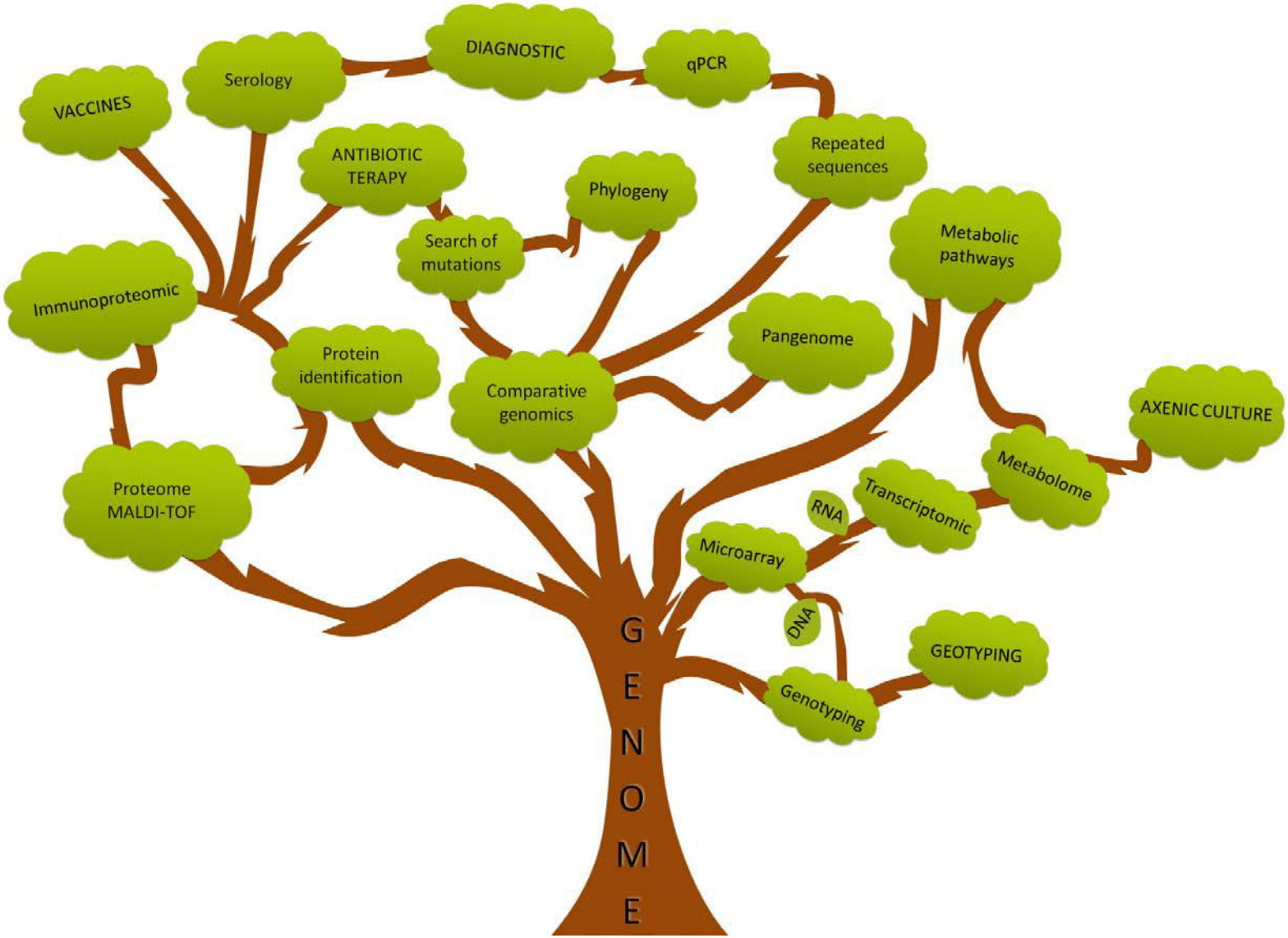


Figure 2 Pangenome of *Coxiella burnetii*. Pangenome representation: the middle of the flower represents the core genes, and the petals show number of genes which are unique for each strain.

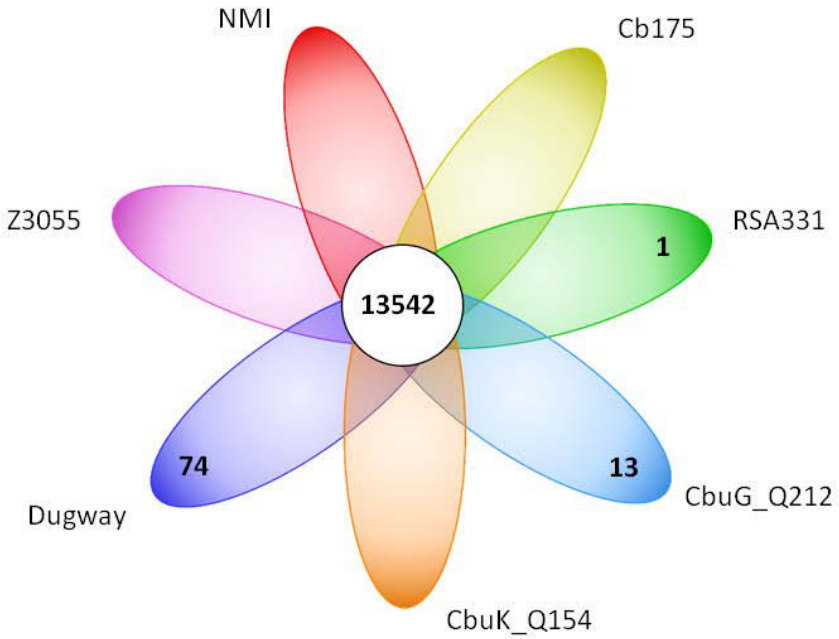


Figure 3A Phylogeny estimation of Proteobacteria 16S rRNA genes. Evolutionary analyses were conducted using the MEGA6 tool. Multiple alignment was performed using the Muscle method, and the tree was created with the NJ algorithm with 100 bootstrap iterations. α-proteobacteria are coloured yellow, β-proteobacteria in green, δ-proteobacteria in blue, ε-proteobacteria in black and γ-proteobacteria in red.. In the tree, we can observe *Coxiella burnetii* strains belonging to the γ-proteobacteria group.

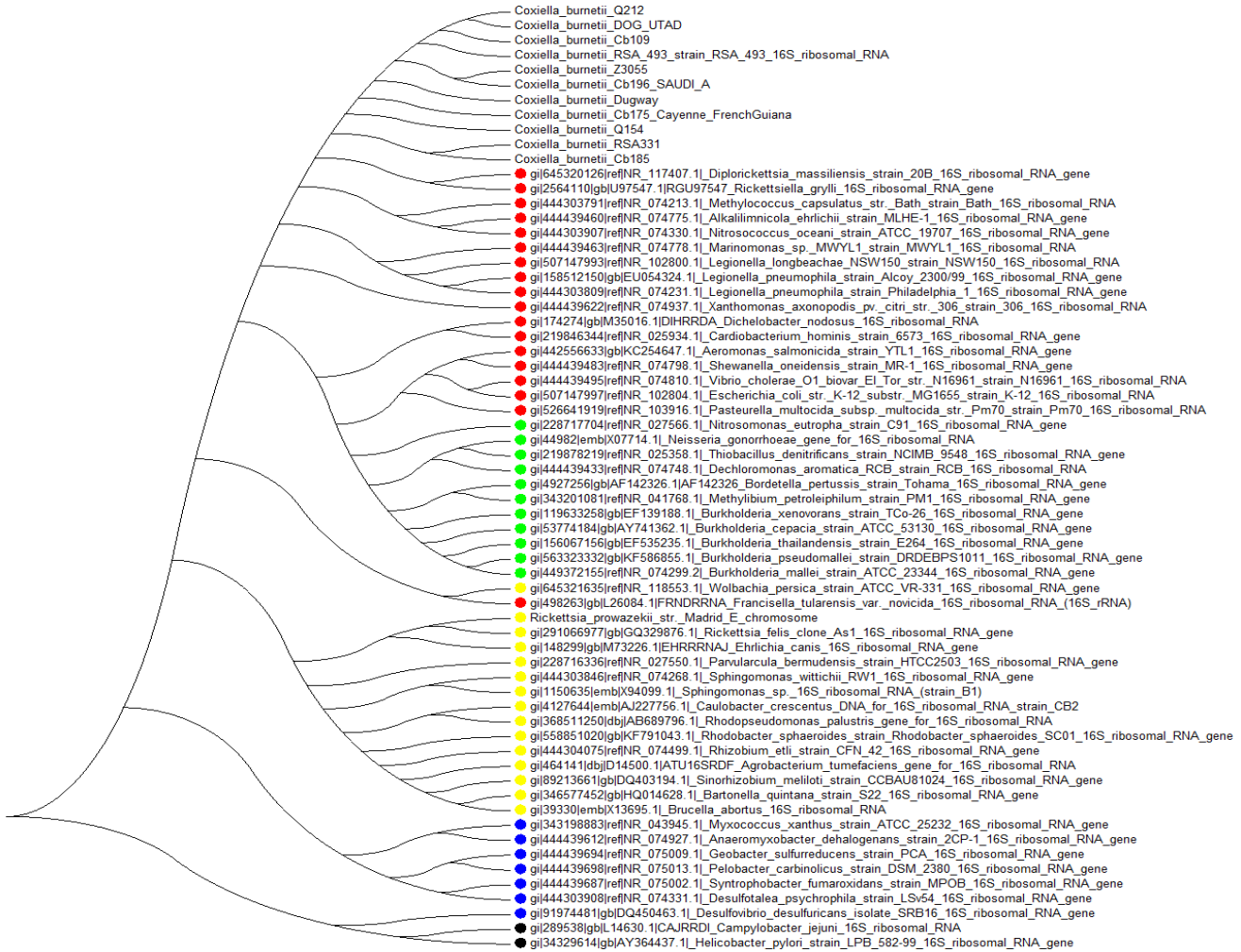
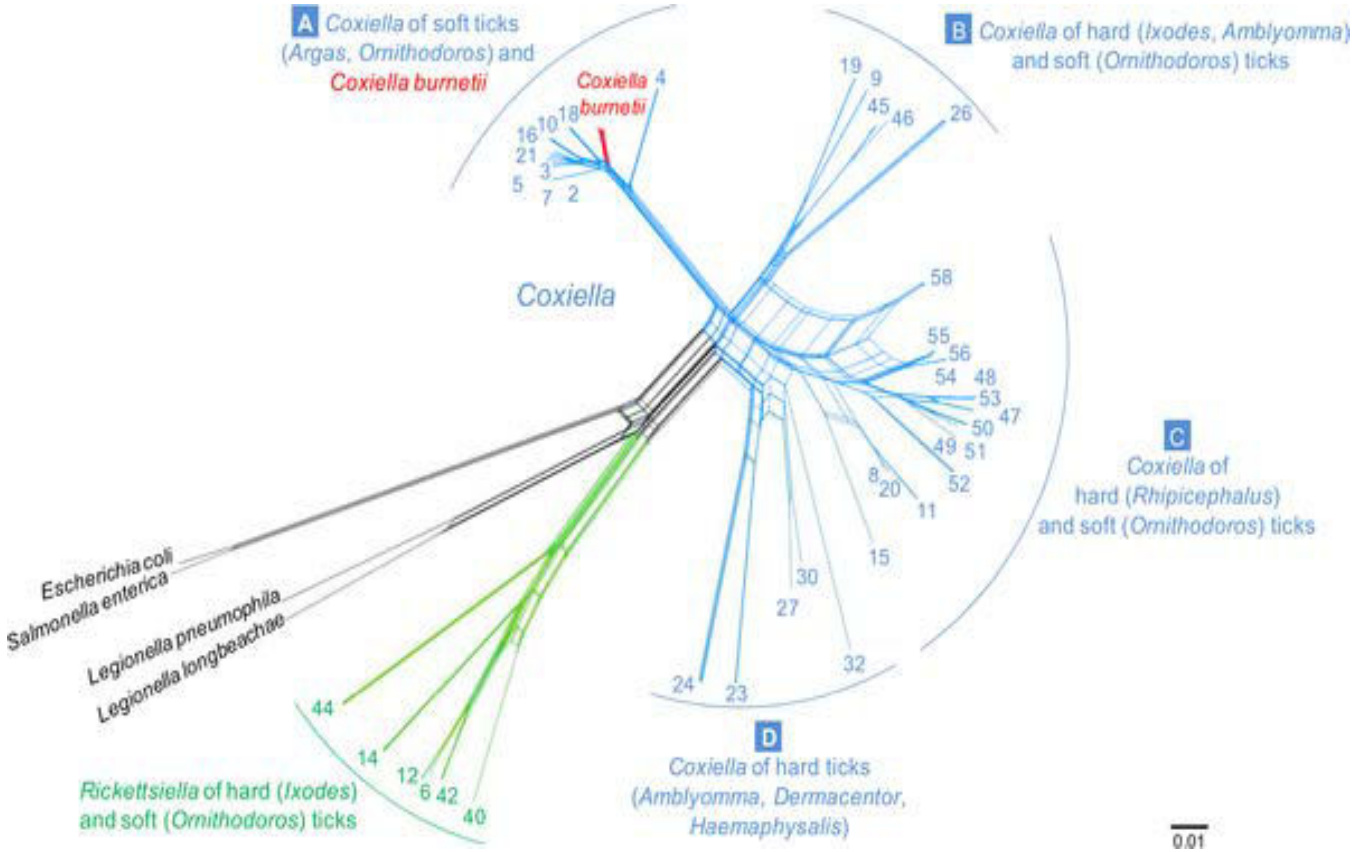


Figure 3B Phylogenetic network with concatenated 16S rRNA, 23S rRNA, GroEL, rpoB and dnaK sequences (3009 unambiguously aligned bp), including 71 Coxiella-like strains of ticks, 15 *C. burnetii* reference strains, and bacterial outgroups.



Duron O, Noël V, McCoy KD, Bonazzi M, Sidi-Boumedine K, *et al.* (2015) The Recent Evolution of a Maternally-Inherited Endosymbiont of Ticks Led to the Emergence of the Q Fever Pathogen, Coxiella burnetii. PLoS Pathog 11(5): e1004892. doi:10.1371/journal.ppat.1004892
http://127.0.0.1:8081/plospathogens/article?id=info:doi/10.1371/journal.ppat.1004892

Figure 4. ***Coxiella burnetii* and *Coxiella*-like vs *Buchnera* sp. APS and *E.coli***. On the left, the comparison of the gene content between *Coxiella burnetii* (core genes) and *Coxiella*-like. On the right, the comparison of gene content between references strains *Buchnera* sp. APS (NC_002528.1) and *E.coli* K-12 MG1655 (NC_000913.3). The same parameters of 50% identity and 70% coverage were used for both comparisons. Genes in the '"function unknown' category are not included. Classes with a high number of absent genes in endosymbionts include Transcription (K), Cell wall/membrane/envelope biogenesis (M), Replication, recombination and repair (L) and Defence mechanisms (V).

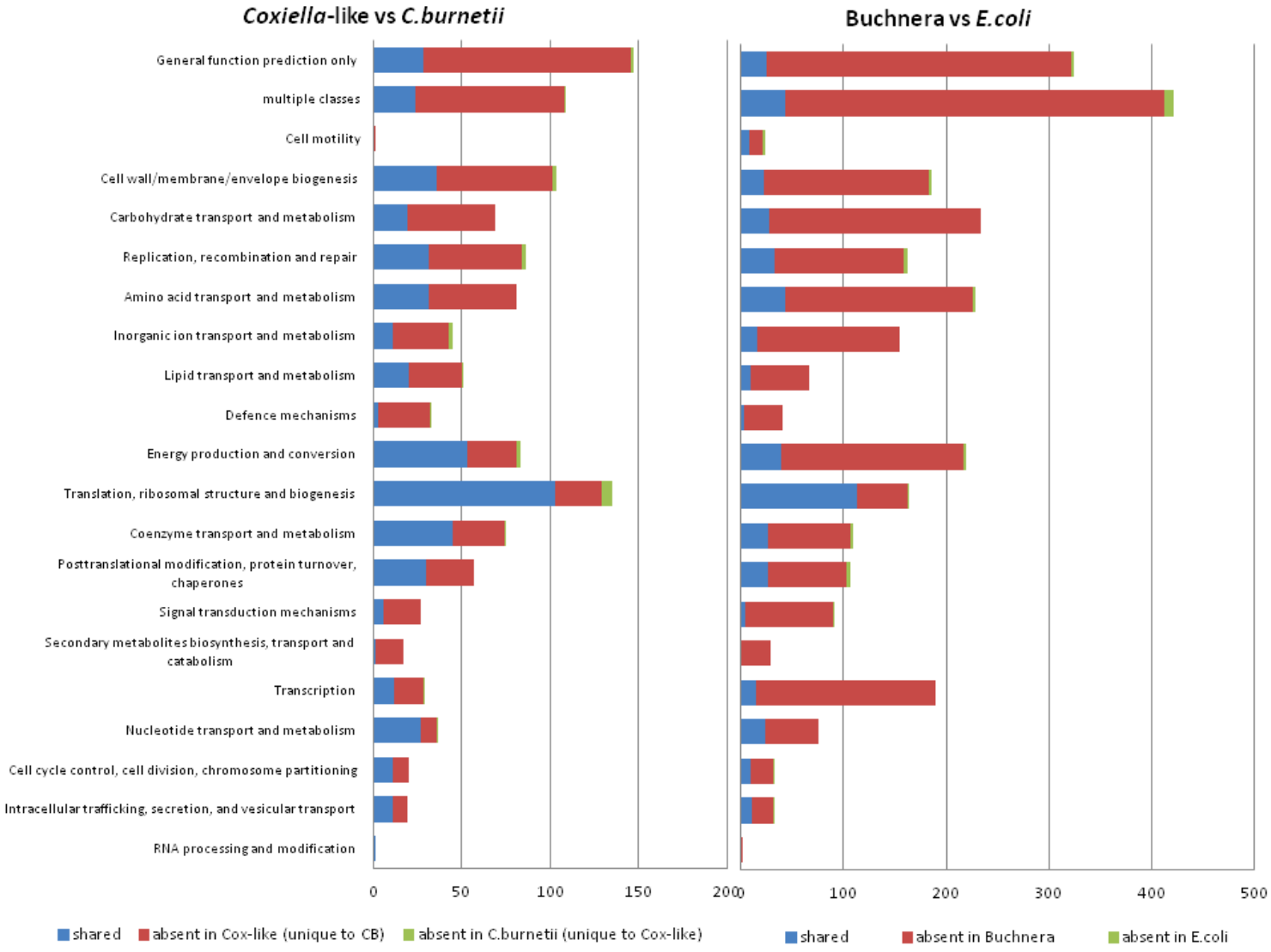
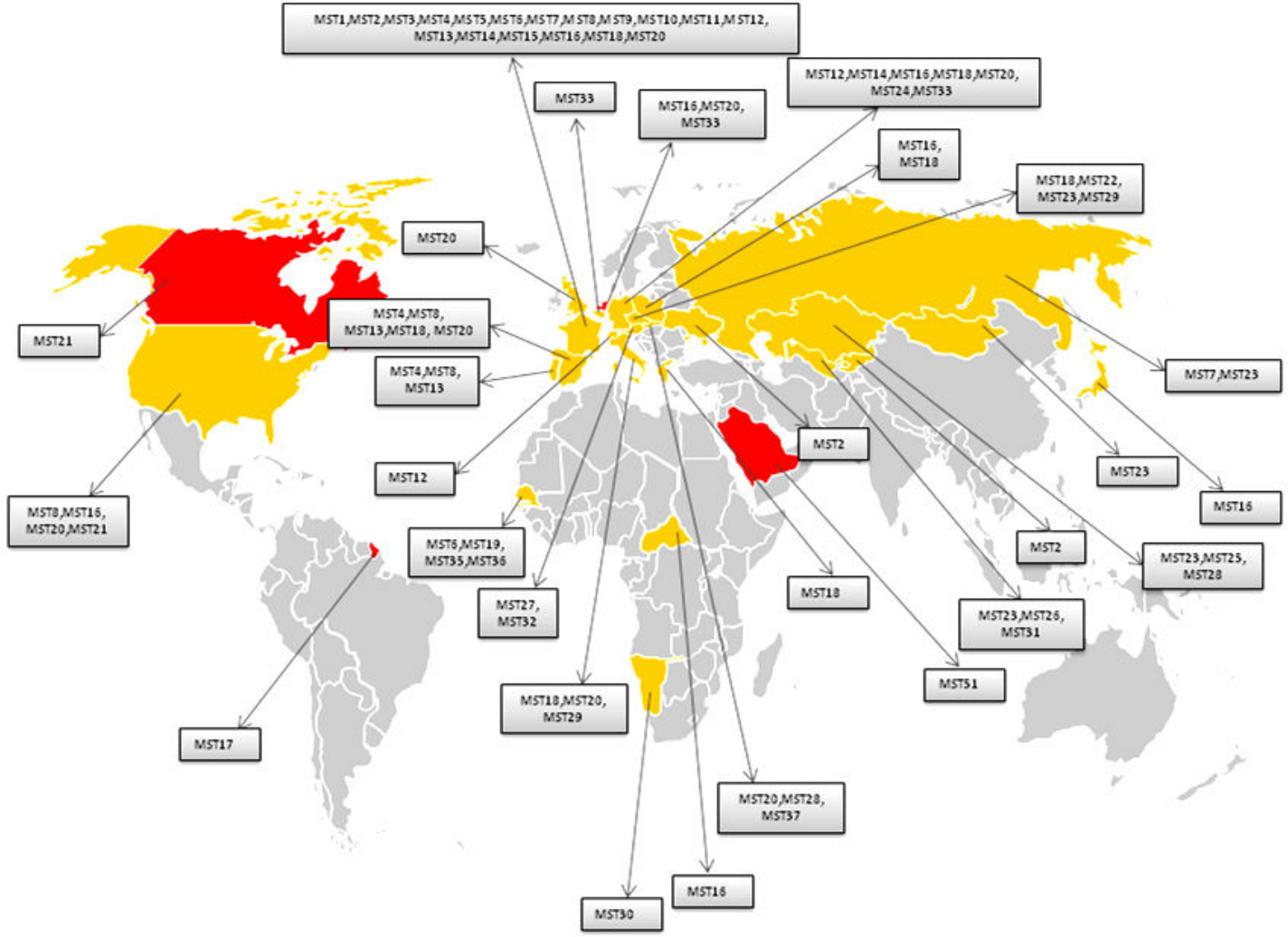


Figure 5 'Geotyping'. Geographical distribution of *Coxiella burnetii* detected genotypes. In red, countries characterised by a unique circulating clone. In yellow, countries where other MST genotypes have been described.



CHAPITRE 1: Etude du génome de *Coxiella burnetii* Z3055, un clone de la souche responsable de l'épidémie de fièvre Q aux Pays-Bas

Avant-propos

Une épidémie humaine de fièvre Q sans précédent a eu lieu aux Pays-Bas entre 2007 et 2010, entraînant au total jusqu'à 4000 cas . La proximité géographique de fermes avec de petits ruminants (principalement des chèvres) et l'intensification rapide de l'élevage dans le Sud-Est du pays (province du Brabant-du-Nord) ont probablement contribué à l'épidémie . Les cas humains sont survenus après des épisodes d'avortement d'animaux dans les fermes.

Lors de cette épidémie, plusieurs génotypes de *C. burnetii* ont été identifiés. La plupart d'entres eux ne possèdent que très peu de différence par rapport au génotype prédominant. Ainsi, les clones responsables de l'épidémie sont tous de génotype MST33, qui s'est propagé de l'Allemagne aux Pays-Bas, via la France .

Nous avons cultivé et séquencé la souche Z3055 (isolée du placenta d'une brebis en Allemagne en 1992), de génotype MST 33 donc proche de la souche responsable de l'épidémie aux Pays-Bas. Notre objectif était de comprendre quels étaient les déterminants génomiques qui ont permis à ce clone de se propager plus facilement parmi la population de chèvres, puis chez les humains, provocant l'épidémie. La comparaison de cette souche (Z3055) avec la souche de référence NMI n'a montré aucune perte ou acquisition de gènes. Cette souche possède un chromosome d'une taille similaire de 1, 995,463pb (1, 995,281pb pour NMI) et le même type de plasmide QpH1 que NMI. De plus, de légères différences sont remarquées entre Z3055 et NMI consistant principalement en des mutations ponctuelles.

Cependant nous avons constaté une augmentation significative de la fréquence de mutation pour les quatre catégories précédemment décrites comme ayant un rôle important dans la virulence: les protéines membranaires, les protéines

contenants des domaines de répétition d'ankyrines, les facteurs de transcription et les protéines ribosomiques.

Les gènes codant pour des protéines de transcription, de traduction et membranaires ont déjà été décrits comme gènes "critiques" inactivés et/ou perdus durant le processus de pathoadaptation et de spécialisation d'une niche spécifique. La perte des opérons ribosomiques dans *Bartonella birtlesii* et des protéines membranaires dans *E. coli* a été associée à une meilleure capacité de la bactérie à se multiplier .

Les protéines contenants des domaines de répétition d'ankyrines semblent être liées à la virulence, étant sécrétées dans la cellule hôte par le système de sécrétion de type IV (T4SS) . *Coxiella burnetii* partage sa stratégie de virulence avec *Legionella pneumophila* basée sur le système de sécrétion de type IV Icm/Dot impliqué dans la phagocytose, la cytotoxicité, l'inhibition de la fusion phagosome - lysosome, l'association du phagosome avec le réticulum endoplasmique rugueux, l'apoptose et la sortie du phagosome .

La forte similarité entre Z3055 et NMI semble être en accord avec le fait que les aspects cliniques de l'épidémie observée aux Pays-Bas sont assez similaires à ceux précédemment décrits dans la littérature. En ce qui concerne les mutations dans les protéines membranaires, il est possible qu'une nouvelle combinaison des antigènes de surface ait pu créer un nouveau clone, comme observé dans le cas de la grippe .

Notre hypothèse est que l'épidémie a été la conséquence d'une surpopulation dans les zones péri- urbaines de chèvres avec un système immunitaire partiellement ou totalement naïf contre ce nouveau clone qui a pu échapper à la réponse immunitaire de l'hôte.
Ce travail a été publié dans le journal ***Comparative Immunology, Microbiology & Infectious Diseases***.

The genome of *Coxiella burnetii* Z3055, a clone linked to the Netherlands Q fever outbreaks, provides evidence for the role of drift in the emergence of epidemic clones

Felicetta D'Amato[1], Laetitia Rouli[1], Sophie Edouard[1], Judith Tyczka[2], Matthieu Million[1], Catherine Robert[1], Ti Thien Ngyuen[1], Didier Raoult[1]*

[1]Aix Marseille Université, URMITE, UM63, CNRS 7278, IRD 198, Inserm 1095, 13005 Marseille, France.

[2] Chemisches und Veterinäruntersuchungsamt Karlsruhe, Weißenburger Straße 3,Abteilung 7, Karlsruhe,76187, Germany

Comparative Immunology, Microbiology and Infectious Diseases 37 (2014) 281–288

Contents lists available at ScienceDirect

Comparative Immunology, Microbiology and Infectious Diseases

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

The genome of *Coxiella burnetii* Z3055, a clone linked to the Netherlands Q fever outbreaks, provides evidence for the role of drift in the emergence of epidemic clones

Felicetta D'Amato[a], Laetitia Rouli[a], Sophie Edouard[a], Judith Tyczka[b], Matthieu Million[a], Catherine Robert[a], Thi Tien Nguyen[a], Didier Raoult[a,*]

[a] *Aix Marseille Université, URMITE, UM63, CNRS 7278, IRD 198, Inserm 1095, 13005 Marseille, France*
[b] *Chemisches und Veterinäruntersuchungsamt Karlsruhe, Weißenburger Straße 3, Abteilung 7, Karlsruhe 76187, Germany*

ARTICLE INFO

Article history:
Received 30 June 2014
Accepted 18 August 2014

Keywords:
Coxiella burnetii
Z3055
Epidemic
Netherlands

ABSTRACT

Coxiella burnetii is a pathogen causing Q fever. The aim of our work was to study Z3055, a strain that is genotypically related to the strain causing the Netherlands outbreak. We compared Z3055 to 5 other completed genomes available in GenBank. We calculated the blast score ratio (BSR) to analyze genetic differences among the strains. The ratio core genome/pangenome was 98% likely other bacteria with closed pangenomes. Differences between Z3055 and the reference NMI consisted only of point mutations and insertion/deletion (INDELs). Non-synonymous mutations significantly increased in genes coding for membrane proteins (16/156 vs 103/1757, bilateral Chi^2 test, $p < 0.05$), ankyrin repeat domains containing proteins (2/9 vs 117/1904, bilateral Chi^2 test, $p < 0.05$), transcription factors (7/53 vs 112/1860, bilateral Chi^2 test, $p < 0.05$) and translation proteins (15/144 vs 109/1655, bilateral Chi^2 test, $p < 0.05$). The evolution of this strain may have been driven by mutations in critical genes.

1. Introduction

Coxiella burnetii is a human pathogen that causes the zoonotic disease Q fever [1]. This Gram-negative bacterium, which belongs to the Gamma-Proteobacteria class, is an obligate intracellular bacterium that exhibits a complex life cycle in the host's macrophages [2]. Currently, the sequences for six *C. burnetii* genomes are available in GenBank: Dugway 5J108-111 [3], CbuG_Q212 [3,4], CbuK_Q154 [3], RSA 331, Nine Mile phase I (NMI) [3,4] and Cb109 [5]. The most common mechanism of transmission of *C. burnetii* from animals to humans is the inhalation of contaminated aerosols or direct contact with infected animals or infected animal products, such as milk, urine, feces or semen. Less common modes of transmission are infection through the digestive tract and injection through the skin [2]. Cattle, sheep and goats serve as the primary reservoirs for this bacterial species. Moreover, placenta and birth fluids from infected animals may contain a high concentration of bacteria and may contaminate the environment [1].

* Corresponding author at: URMITE, UMR CNRS 7278, IRD 198, INSERM U1095, Faculté de Médecine, 27 Bd Jean Moulin, 13385 Marseille cedex 5, France. Tel.: +33 04 91 32 43 75; fax: +33 04 91 38 77 72.
E-mail addresses: felicettadamato@gmail.com (F. D'Amato), laetitia.rouli@gmail.com (L. Rouli), soph.edouard@gmail.com (S. Edouard), judith.tyczka@cvuaka.bwl.de (J. Tyczka), matthieumillion@gmail.com (M. Million), catherine.robert@univmed.fr (C. Robert), evenguyen1@yahoo.fr (T.T. Nguyen), didier.raoult@gmail.com (D. Raoult).

http://dx.doi.org/10.1016/j.cimid.2014.08.003
0147-9571/

Very few Q fever infections were reported in the Netherlands before 2007. During the four years following 2007, one of the largest Q fever outbreaks in the world occurred in southeastern Netherlands [6], during which 3921 cases of human Q fever were recorded, and abortion commonly occurred on dairy farms. Genotypic analyses of bacterial isolates obtained from these cases were conducted to confirm the possible link between the human and animal cases of Q fever [7]. A specific genotype of *C. burnetii* (MST33) was responsible for the Q fever epidemic in the Netherlands [7]. Microbial genotyping is useful for studying population dynamics and for detecting epidemics, and it is currently the most sensitive method for identifying bacterial clones [8]. Isolates exhibiting similar genotypic profiles according to this method are considered to belong to the same clone. Using this method, it has been also determined that the genotype MST17 was responsible for an outbreak in French Guiana [9,10]. The presence of the MST33 genotype in clinical samples from humans and animals involved in the Dutch outbreak confirmed that goats and sheep were the source of this outbreak [7].

MLVA also represents a molecular typing tool with high resolution and low costs. MLVA has been used for epidemiological studies [11]. A 10-locus variable-number tandem repeat analysis (MLVA) panel revealed one predominant genotype among goats and sheep throughout the Netherlands. Because the strain responsible for the epidemic in Holland was not available for genomic analysis, we sequenced and analyzed another strain that has the same VNTR profile as the Netherlands' strains and that belongs to the same MST genotype: the Z3055 strain, isolated from a ewe placenta in Germany in 1992 and also identified as genotype MST33. Because the genotyping performed in this analysis enables the detection of clones, we considered the genome of the Z3055 strain to be genetically related to the epidemic strain circulating in the Netherlands.

In our study, we performed a comparative genomic analysis of strain Z3055, which is related to the strain responsible for the Q fever epidemic in the Netherlands, and the sequences of the 5 strains of *C. burnetii* that have available genomes using several bioinformatics tools. We analyzed genetic difference among the strains. We then compared only the Z3055 genome to the reference NMI genome and focused on point mutations and different proteins.

2. Methods

2.1. Strains culture

C. burnetii were cultured at 35 °C on L929 cells using MEM (GIBCO, Invitrogen, Cergy-Pontoise, France) supplemented with 4% SVF (GIBCO) and 1% L-glutamine (GIBCO). Monolayers of cells and the supernatants from three 175 cm^2 flasks were harvested and incubated with 1% trypsin (GIBCO) for 1 h at 37 °C. Released bacteria were purified from L929 cell debris into a discontinuous Gastrografine (Schering, Lys-Lez-Lannoy, France) gradient (45%, 36% and 28%) and ultracentrifuged at 5000 Tars/min for 1 h 10.

2.2. Basic genomic data

The genome of the culture-positive strain Z3055 was subjected to paired-end SoLiD sequencing. DNA was extracted using QIAamp® DNA Mini Kit protocol, by performing two subsequent elutions in a total volume of 30 μl. We used Quant-iT™ PicoGreen® dsDNA Assay Kit – Life Technologies for DNA quantitation. The paired end library was constructed from 1 μg of purified genomic DNA of strain (with an obtained concentration of 58 ngr/μl). The sequencing was carried out to 50 × 35 bp using SOLiD™ V4 chemistry on one full slide associate with 95 others projects on an Applied BiosystemsSOLiD4 machine. All of these 96 genomic DNA were barcoded with the module 1–96 barcodes provided by Life Technologies. The DNA was fragmented on the Covaris device and the concentration of the library was measured on the Qbitfluorometer at 20.3 nmol/l. Libraries were pooled in equimolar ratios and size selected on the E-Gel iBase system at 240–270 bp. The PCR was performed according to Life Technologies specification templated bead preparation kits on the EZ beads automates Emulsifier, Amplifier and Enricher E80 for full scale. For the run, a total amount of 708 million P2 positive beads was loaded onto the flowcell. The output read length was as expected 85 bp (50 × 35 bp). On a total of 39.8 Gb for the full slide, the *Coxiella* project obtained 4,488,534 barcoded reads which lead to 381.5 Mb. Mapping of reads obtained by SoliD against reference NMI was performed using CLC Genomics (Copyright CLC bio) at parameters of length fraction 0.7 and similarity fraction 0.8 to obtain a consensus. Multiple rounds of PCR and sequencing using BigDye terminator chemistry on an ABI3730 sequencing machine (Applied Biosciences) were performed to close gaps and to complete the genome.

The complete sequence that will replace the draft genome previously submitted was deposited in GenBank the project PRJEB1438. To perform genomic comparisons, we used 5 *C. burnetii* complete genomes that are available from NCBI: Dugway 5J108-111 (accession number: NC_009727.1), CbuG_Q212 (accession number: NC_011527.1), CbuK_Q154 (accession number: NC_011528.1), RSA 331 (accession number: NC_010117.1), and Nine Mile phase I (NMI, accession number: NC_002971.3). To avoid biases, we did not use strains Q321, CB109 and MSU Goat Q177 because the number of available contigs changed during the period in which we conducted the analysis.

2.3. Genotyping of the strain

Multi-spacer sequence typing (MST) genotyping was performed in silico. It consisted of 10 different spacers of the *C. burnetii* genome: Cox2, Cox 5, Cox 6, Cox 18, Cox 20, Cox 22, Cox 37, Cox 51, Cox 56, Cox 57 and Cox61 [7,8,12]. Using a web-based MST database (http://ifr48.timone.univ-mrs.fr/MST Coxiella/mst), we aligned each sequence of Cox with the Z3055 genome and determined its genotype. To verify that strain Z3055 is a clone of the epidemic strain, we also performed a multiple-locus variable-number tandem repeat analysis (MLVA). We used 6 variable loci described by de Bruin et al. [13] to define the

type of *C. burnetii* strains isolated from animal and environmental samples during the Q fever epidemics in the Netherlands.

Furthermore, we determined in silico the repeat number of each microsatellite marker by analyzing sequence regions targeted by each pair of primers, and we confirmed the results by experimental assay. Standard PCR was performed in a total volume of 20 μl containing 0.02 U of Phusion High-Fidelity DNA Polymerase (Finnzyme, Thermo Scientific), 200 μM dNTPs, 1× reaction buffer, 0.5 μM amplification primers, and 0.5 μl template DNA. Amplification products were sequenced using BigDye terminator chemistry on an ABI3130 sequencing machine (Applied Biosciences). Then, we blasted repeat markers against the sequenced fragments to calculate the number of tandem repeats.

Finally, we entered the VNTR allelic profile of Z3055 in a MLVA database (http://mlva.u-psud.fr/mlvav4/genotyping/) and compared it to the profile of other typed strains. We constructed a dendrogram based on the origin of strains and drew a Newick tree in MEGA5 [14].

2.4. Blast score ratio analysis

Coding regions (CDSs) were predicted and translated using GenMarK [15]. CDSs from each genome were concatenated and then de-replicated by clustering with USEARCH [16] at an identity of 0.9 to reduce the size of the dataset. The translated sequence was then aligned against each genome in the dataset with TBLASTN [17]. The query bit score for each genome alignment was tabulated and subsequently divided by the maximum bit score in all genomes to obtain the blast score ratio; we determined the conservation of peptides in each genome by calculating the blast score ratio [18,19], which can range from 1.0 (exact peptide match) to 0.0 (no significant alignment). Thus, genes with a blast score ratio of ≤0.4 were considered to be lost (equivalent to ≤40% identity over 100% of the peptide length). Calculated blast score ratio values across groups were visualized with the multi-experiment viewer [20,21] to analyze clusters of strains based on gene difference. The blast score ratio algorithm allows comparative analysis of multiple proteomes and nucleotide sequence at the same time [18]. Using this method, we avoided the limitation of BLAST E-values normalizing the BLAST raw scores. We determined the proportion of core, accessory and unique genes of the *C. burnetii* pangenome. Genes with a blast score ratio >0.4 in only one strain but ≤0.4 in others represent unique genes. Genes with a blast score ratio >0.4 in some genomes but not in all, i.e., genes that are found in some *C. burnetii* strains but not others, represent accessory genes. The core genome encompasses genes with a blast score ratio >0.4 in all strains and genes filtered from the dataset by dereplication.

2.5. Point mutations analysis

We analyzed 6 whole genomes with the SNPfinder analysis tool [22], which generated multiple sequence alignments based on single nucleotide polymorphisms (SNPs) and insertion and deletion of single base (INDEL) events. We subsequently focused on point mutations between Z3055 and NMI. We aligned protein sequences to verify whether point mutations were silent or result in a change in the amino acid (non-synonymous mutation). We confirmed the presence of mutations by mapping reads obtained by Z3055 genome sequencing against NMI. We performed COG (Cluster of Orthologous Groups) annotation for all NMI genes and for non-synonymous genes. We analyzed the proportion of different proteins for each COG category.

2.6. Comparison and phylogeny

We used MEGA5 software [14] to perform multiple sequence alignment and to generate trees. Multiple alignment was performed using the ClustalW method, and the tree was created with the Neighbor Joining (NJ) algorithm (part of the distance method). We performed phylogenetic analysis based on SNPs in ORFs, discarding intergenic regions to minimize phylogenetic error because these regions are not likely to be evolutionarily stable. We also performed a phylogenomic analysis based on all 6 complete genomes and a phylogenetic analysis based on the strains' MST genotype.

3. Results

3.1. Z3055 is closely related to Nine Mile

After 454 and SOliD sequencing of the Z3055 strain, we obtained one scaffold corresponding to the bacterial chromosome of 1,995,463 bp and a second scaffold corresponding to the QpH1 plasmid of 37,316 bp, with a GC content 42.6%, 45 tRNAs and the same number of proteins as NMI (i.e., 1913), as the BSR showed that there was neither gain nor loss of genes.

MST genotyping in silico showed 100% of identity for matching sequences of Cox2.7, Cox5.5, Cox18.1, Cox20.6, Cox22.5, Cox37.9, Cox51.9, Cox56.4, Cox57.3, and Cox61.2, corresponding to genotype 33, the genotype described from humans in Holland. The number of tandem repeats calculated in silico for each locus was confirmed by sequencing. We found 5 repetitions for Ms27, 3 repetitions for Ms28, 7 repetitions for Ms34, 7.5 repetitions for Ms 20b, 13 repetitions for Ms24 and 4 repetitions for Ms31. The comparison of the allelic profile of this strain with other typed *C. burnetii* strains showed that this strain is in the same cluster of Netherland strains (Fig. 1).

The analysis of the blast score ratio, when considering all strains, showed that the genetic variability between *C. burnetii* strains is low and that the strains are very similar (Fig. 2). The core genome/pangenome ratio was 98%, which was consistent with the high ratio value found for other bacteria having a closed pangenome, such as *Bacillus anthracis* [23]. According to the Pangenomic analysis, we found 11595 core genes, 183 accessory genes (1.5% of pangenome) and 88 unique genes. The unique genes of *C. burnetii* Dugway (74 unique genes) encode different proteins: 41 are hypothetical, 15 are involved in metabolic pathways, 5 are involved in transcription, 5 are involved in environmental information processes, 2 belong to the

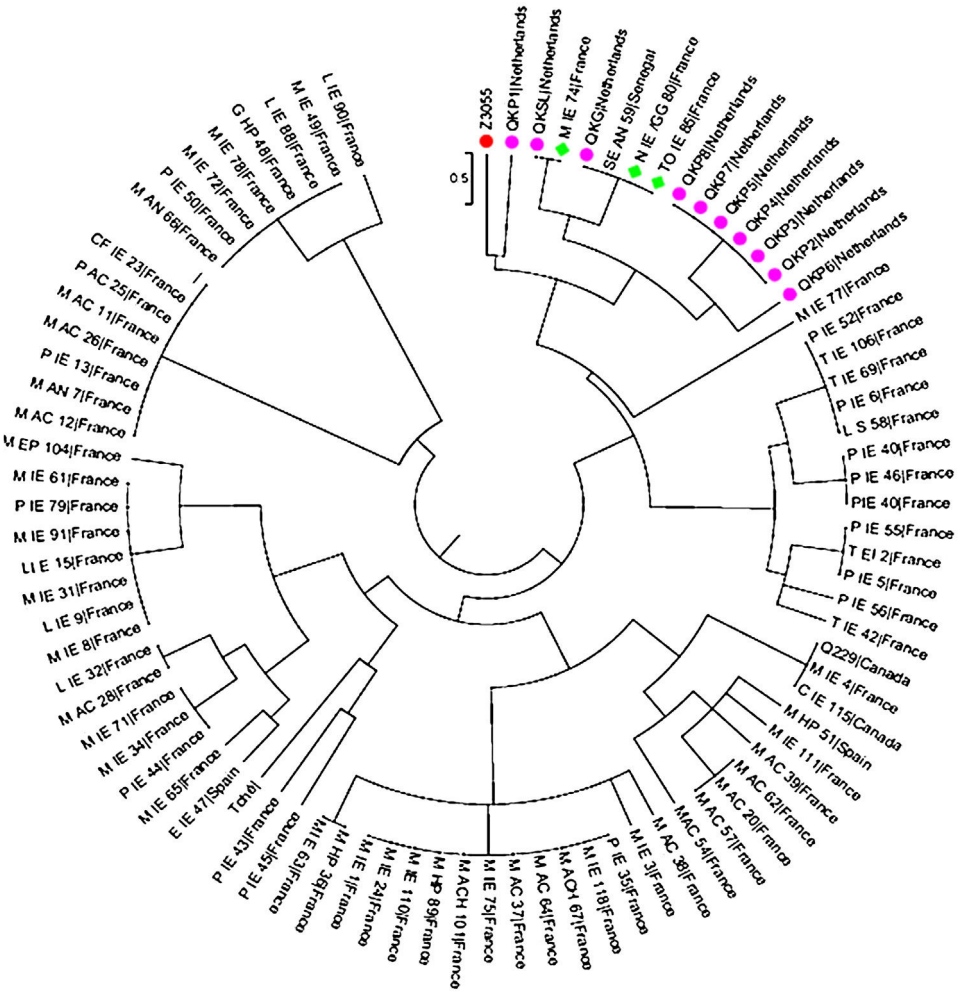


Fig. 1. VNTR-based tree. This phylogenetic tree, generated through MLST database show Z3055 clusterize with all the other Netherlands strains.

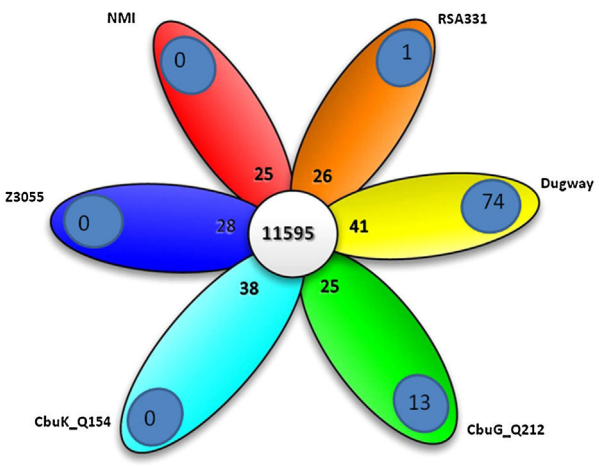


Fig. 2. Pangenome representation for 6 analyzed genomes. On the upper end of each petal, genes unique to each strain. At the bottom of the petals, accessory genes. On the center, of the core-genome.

rhs family, 2 belong to the tetratricopeptide repeat family, 1 is involved in cellular processes, 1 is involved in replication, 1 is an ankyrin repeat protein and 1 is a transposase. In addition, 13 unique genes were found in *C. burnetii* Q212 encoding 9 hypothetical proteins, two proteins of the phage integrase family, 1 DNA-binding protein and one phosphoesterase. However, only 1 unique gene (encoding a hypothetical protein) was identified in *C. burnetii* RSA331, while the following strains have no unique genes: *C. burnetii* Q154, *C. burnetii* Z3055 and *C. burnetii* NMI. Phylogenetic analysis based on the MST genotype showed that strains NMI and RSA331, which both contained the plasmid QpH1, were close to each other (same clade), while MST33 was found to not be very close to these strains (Fig. 3A). Moreover, the results of different analysis methods using phylogenomic analysis (Fig. 3B), phylogenetic analysis based on the core genome (Fig. 3C) and clustering with the blast score ratio algorithm (Fig. 3D), demonstrated a high level of similarity between the 6 strains, and among these, Z3055 and NMI were the closest.

Phylogenetic analysis based on concatenated SNP in coding regions (Fig. 3E) confirmed that NMI is the most

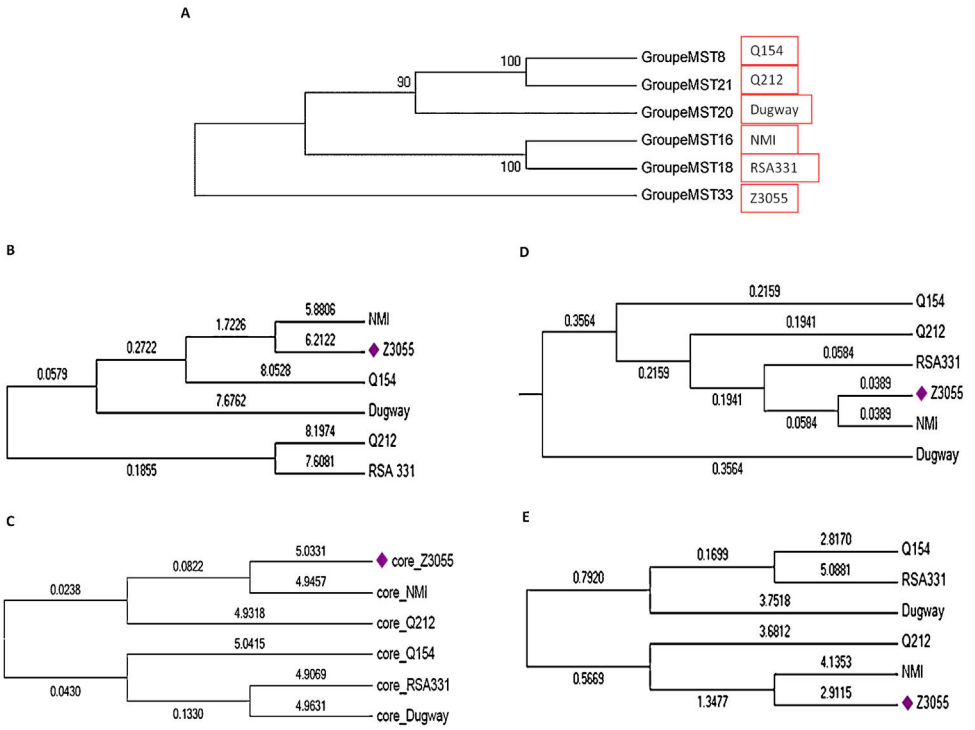


Fig. 3. Trees validating NMI-Z3055 cluster, based on different data. (A) MST-based tree. This phylogenetic tree was generated by NJ algorithm with 100 bootstrap iterations. It showed how our 6 genomes of interest clusterized based on their genotype. (B) Phylogenetic tree (NJ method) based on all genomic content. (C) Phylogenetic tree (NJ method) based on core genome. (D) Tree based on blast score ratio results. (E) Phylogenetic tree (NJ method) based on coding sequences' SNPs.

similar strain to Z3055 and that all genomes are very close, suggesting a clonal mode of genetic inheritance [24] with a low substitution rate.

3.2. Slight differences with NMI

We found a total of 2917 SNPs between Z3055 and NMI, including 2507 SNPs in intergenic regions, 410 in coding regions and 228 non-synonymous SNPs (Fig. 4, Supplementary Table 1), corresponding to 119 mutated genes. These mutated genes encode 16 membrane proteins, 2 ankyrin repeat domain-containing proteins, 7 transcription factors and 15 translation proteins. We found 569 INDELs in intergenic regions and 2 INDELS in two genes encoding ankyrin proteins, which were also mutated by SNPs. Statistical tests showed that for membrane proteins, the proportion of different proteins is higher than for other proteins (16/156 vs 103/1757, bilateral Chi^2 test, $p < 0.05$), as is true for ankyrin repeat domain-containing proteins (2/9 vs 117/1904, bilateral Chi^2 test, $p < 0.05$), transcription factors (7/53 vs 112/1860, bilateral Chi^2 test, $p < 0.05$) and translation proteins (15/144 vs 109/1655, bilateral Chi^2 test, $p < 0.05$).

Supplementary Table 1 related to this article can be found, in the online version, at http://dx.doi.org/10.1016/j.cimid.2014.08.003.

When we classified proteins by COG categories, we found that the proportion of mutated proteins (Supplementary Table 2) was significantly increased for category J (Translation, ribosomal structure and biogenesis) compared to other categories, i.e., 115/131 vs 104/1782, bilateral Chi^2 test, $p < 0.05$. We observed a trend of an increase in the proportion of different proteins for category M (cell wall/membrane/envelope biogenesis), 11/108 vs 108/1805, bilateral Chi^2 test $p = 0.08$; for category H (Coenzyme transport and metabolism), 9/77 vs 110/1836 bilateral Fischer test, $p = 0.09$; and for category I (Lipid transport and metabolism), 7/55 vs 112/1858 bilateral Fischer test, $p = 0.29$.

Supplementary Table 2 related to this article can be found, in the online version, at http://dx.doi.org/10.1016/j.cimid.2014.08.003.

Additionally, we found a significant decrease in the proportion of different proteins for category L (Replication, recombination and repair), i.e., 5/110 vs 114/1803, bilateral Chi^2 test, $p < 0.05$ and for category D (Cell cycle control, cell division, chromosome partitioning), i.e., 1/22 vs 118/1891 bilateral Fischer test, $p < 0.05$).

4. Discussion

Our analysis showed that the strain responsible for the Netherlands outbreak and Z3055 are clonal, as they both contain plasmid QpH1 and have the same MST33 genotype and the same VNTR profile. The first isolate identified as MST33 was isolated from a placenta from an ewe in Germany in 1992. This genotype was identified several times in ungulates and human clinical samples

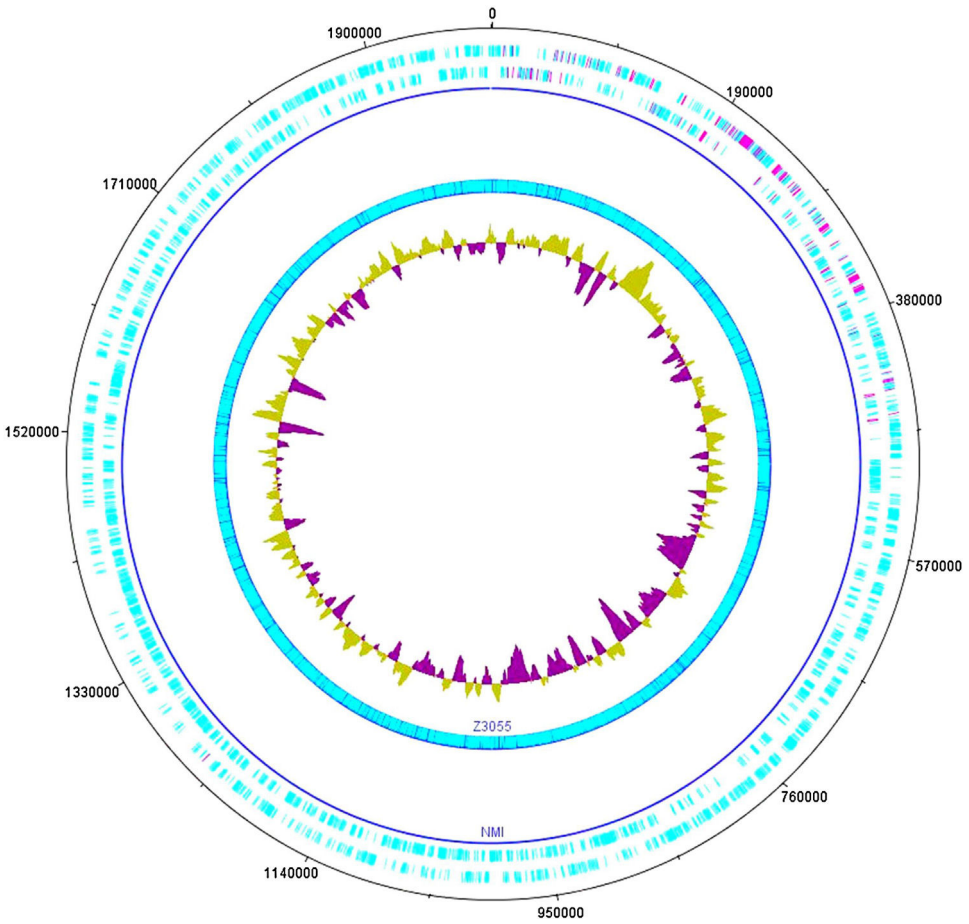


Fig. 4. Circular genomic comparison between Z3055 and NMI. On the circle non-synonymous genes are marked with fuchsia color. Violet: GC content below average. Green: GC content above average. (For interpretation of the references to color in figure legend, the reader is referred to the web version of the article.)

(Supplementary Table 3). The clone identified as responsible for this epidemic [7,12] most likely spread from Germany to the Netherlands via France (Supplementary Fig. 1) and evolved into an epidemic strain that infected both animals and humans. Moreover, MST33 has been isolated in non-outbreak situations in human clinical samples obtained in France in 1996, 1998, and 1999 and from the placenta of an asymptomatic ewe in Germany in 1992 [7]. A second genotype, MST20, was also identified in 2009 in the Netherlands but was only found in cattle and goat and cow's milk but not in humans [7,12]. Because the strain responsible for the Q fever pandemic in the Netherlands is not available for study, it is of great interest for the scientific community to know the characteristics of the most closely related isolated strain based on MST and MLVA. We cannot exclude some genomic differences between the two strains, but it is very unlikely given that very little variation exists among isolates of *C. burnetii*, as *C. burnetii* has a closed pan-genome.

Supplementary Table 3 and Figure 1 related to this article can be found, in the online version, at http://dx.doi.org/10.1016/j.cimid.2014.08.003.

SNP-based assays have been performed to elucidate the evolutionary history of bacteria such as *Salmonella enterica* [25], to study the diversification of *Streptococcus pneumonia* [26], and to analyze the evolution of the *Escherichia coli* population [27]. Patho-adaptation of *C. burnetii* appears to be based on the loss or modification of existing genes [28], as observed in other specialized parasites [29]. Based on our results, the strain of *C. burnetii* responsible for the outbreak in the Netherlands is highly similar to the reference strain Nine Mile I at both the genomic and metabolic levels and with respect to pathogenic features. We analyzed in detail the point mutations between these two strains (2917). Interestingly, we found a significant increase in non-synonymous gene mutations (Supplementary Table 3) for proteins belonging to the categories of membrane proteins, ankyrin repeat domain-containing proteins, transcription factors and translational proteins.

We hypothesize that the epidemic potential of this strain is based on changes in antigenic surface proteins and a reshaped growth rate. Membrane protein modification may have resulted in a new serotype, similar to what is observed in influenza virus, which is secondary to

stabilizing selection [30]. New combinations of membrane proteins could lead to a human population that is partially or completely naive to such influenza strains [31]. The disruption of genes encoding transcription factors, proteins involved in translation and outer membrane proteins are frequently implicated in pathoadaptation and specialization in a specific niche [32]. Genes belonging to these categories have been described as inactivated in *E. coli* cultured for thousands of passages [27]. Disruption of ribosomal operons allows *Bartonella birtlesii* to increase its multiplication ability [32]. Furthermore, ankyrin repeat domains appear to be linked with pathogenicity [33], as they are candidates for Dot/Icm substrates in *Legionella pneumophila* [34,35]. Finally, although the epidemic potential of the Dutch strain has been shown by the unprecedented magnitude of the epidemic, the pathogenicity of this epidemic strain is most likely not very different from other strains described to date. This can be confirmed by the fact that the clinical spectrum ultimately does not appear to be different from the published literature, including cardiovascular infections [36] and obstetrical involvement in humans [37].

5. Conclusion

Z3055 and NMI are very close strains. Genetic differences between the two strains are minimal, and Z3055 is not characterized by gene gain or loss. The difference between the two strains consisted only of point mutations. *C. burnetii* cannot exchange exogenous DNA, and pathoadaptation appears to be based on the loss or modification of existing genes, as in other specialized parasites [29]. The presence of non-synonymous mutations in membrane proteins suggests a possible explanation for the outbreak. Changes in the antigenic profile may have led to a new serotype, conferring to the novel clone the capacity to escape the host immune response and to disseminate easily in a immunologically naïve population.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by the Mediterranean Infection Foundation.

References

[1] Raoult D. Chronic Q fever: expert opinion versus literature analysis and consensus. J Infect 2012;65:102–8, http://dx.doi.org/10.1016/j.jinf.2012.04.006. pii:S0163-4453(12)00106-5.

[2] Raoult D, Marrie T, Mege J. Natural history and pathophysiology of Q fever. Lancet Infect Dis 2005;5:219–26, http://dx.doi.org/10.1016/S1473-3099(05)70052-9. pii:S1473-3099(05)70052-9.

[3] Beare PA, Unsworth N, Andoh M, Voth DE, Omsland A, Gilk SD, et al. Comparative genomics reveal extensive transposon-mediated genomic plasticity and diversity among potential effector proteins within the genus Coxiella. Infect Immun 2009;77:642–56, http://dx.doi.org/10.1128/IAI.01141-08, pii:IAI.01141-08.

[4] Seshadri R, Paulsen IT, Eisen JA, Read TD, Nelson KE, Nelson WC, et al. Complete genome sequence of the Q-fever pathogen *Coxiella burnetii*. Proc Natl Acad Sci USA 2003;100:5455–60, http://dx.doi.org/10.1073/pnas.0931379100, pii:0931379100.

[5] Rouli L, Rolain JM, El FA, Robert C, Raoult D. Genome sequence of *Coxiella burnetii* 109, a doxycycline-resistant clinical isolate. J Bacteriol 2012;194:6939, http://dx.doi.org/10.1128/JB.01856-12, pii:194/24/6939.

[6] Roest HI, Ruuls RC, Tilburg JJ, Nabuurs-Franssen MH, Klaassen CH, Vellema P, et al. Molecular epidemiology of *Coxiella burnetii* from ruminants in Q fever outbreak, the Netherlands. Emerg Infect Dis 2011;17:668–75, http://dx.doi.org/10.3201/eid1704.101562.

[7] Tilburg JJ, Roest HJ, Buffet S, Nabuurs-Franssen MH, Horrevorts AM, Raoult D, et al. Epidemic genotype of *Coxiella burnetii* among goats, sheep, and humans in the Netherlands. Emerg Infect Dis 2012;18:887–9, http://dx.doi.org/10.3201/eid1805.111907.

[8] Glazunova O, Roux V, Freylikman O, Sekeyova Z, Fournous G, Tyczka J, et al. *Coxiella burnetii* genotyping. Emerg Infect Dis 2005;11:1211–7, http://dx.doi.org/10.3201/eid1108.041354.

[9] Mahamat A, Edouard S, Demar M, Abboud P, Patrice JY, La SB, et al. Unique clone of *Coxiella burnetii* causing severe Q fever, French Guiana. Emerg Infect Dis 2013;19:1102–4, http://dx.doi.org/10.3201/eid1907.130044.

[10] Edouard S, Mahamat A, Demar M, Abboud P, Djossou F, Raoult D. Comparison between Emerging Q Fever in French Guiana and Endemic Q fever in Marseille, France. Am J Trop Med Hyg 2014, http://dx.doi.org/10.4269/ajtmh.13-0164, pii:ajtmh.13-0164.

[11] Arricau-Bouvery N, Hauck Y, Bejaoui A, Frangoulidis D, Bodier CC, Souriau A, et al. Molecular characterization of *Coxiella burnetii* isolates by infrequent restriction site-PCR and MLVA typing. BMC Microbiol 2006;6:38, http://dx.doi.org/10.1186/1471-2180-6-38, pii:1471-2180-6-38.

[12] Tilburg JJ, Rossen JW, van Hannen EJ, Melchers WJ, Hermans MH, van de Bovenkamp J, et al. Genotypic diversity of *Coxiella burnetii* in the 2007–2010 Q fever outbreak episodes in The Netherlands. J Clin Microbiol 2012;50:1076–8, http://dx.doi.org/10.1128/JCM.05497-11, pii:JCM.05497-11.

[13] de BA, van Alphen PT, van der Plaats RQ, de Heer LN, Reusken CB, van Rotterdam BJ, et al. Molecular typing of *Coxiella burnetii* from animal and environmental matrices during Q fever epidemics in the Netherlands. BMC Vet Res 2012;8:165, http://dx.doi.org/10.1186/1746-6148-8-165, pii:1746-6148-8-165.

[14] Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 2011;28:2731–9, http://dx.doi.org/10.1093/molbev/msr121, pii:msr121.

[15] Besemer J, Lomsadze A, Borodovsky M. GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions. Nucleic Acids Res 2001;29:2607–18.

[16] Edgar RC. Search and clustering orders of magnitude faster than BLAST. Bioinformatics 2010;26:2460–1, http://dx.doi.org/10.1093/bioinformatics/btq461, pii:btq461.

[17] Gertz EM, Yu YK, Agarwala R, Schaffer AA, Altschul SF. Composition-based statistics and translated nucleotide searches: improving the TBLASTN module of BLAST. BMC Biol 2006;4:41, http://dx.doi.org/10.1186/1741-7007-4-41, pii:1741-7007-4-41.

[18] Rasko DA, Myers GS, Ravel J. Visualization of comparative genomic analyses by BLAST score ratio. BMC Bioinform 2005;6:2, http://dx.doi.org/10.1186/1471-2105-6-2, pii:1471-2105-6-2.

[19] Sahl JW, Gillece JD, Schupp JM, Waddell VG, Driebe EM, Engelthaler DM, et al. Evolution of a pathogen: a comparative genomics analysis identifies a genetic pathway to pathogenesis in acinetobacter. PLOS ONE 2013;8:e54287, http://dx.doi.org/10.1371/journal.pone.0054287, pii:PONE-D-12-31404.

[20] Saeed AI, Sharov V, White J, Li J, Liang W, Bhagabati N, et al. TM4: a free, open-source system for microarray data management and analysis. Biotechniques 2003;34:374–8.

[21] Saeed AI, Bhagabati NK, Braisted JC, Liang W, Sharov V, Howe EA, et al. TM4 microarray software suite. Methods Enzymol 2006;411:134–93, http://dx.doi.org/10.1016/S0076-6879(06)11009-5. pii:S0076-6879(06)11009-5.

[22] Song J, Xu Y, White S, Miller KW, Wolinsky M. SNPsFinder – a web-based application for genome-wide discovery of single nucleotide polymorphisms in microbial genomes. Bioinformatics 2005;21:2083–4, http://dx.doi.org/10.1093/bioinformatics/bti176, pii:bti176.

[23] Read TD, Peterson SN, Tourasse N, Baillie LW, Paulsen IT, Nelson KE, et al. The genome sequence of *Bacillus anthracis* Ames and comparison to closely related bacteria. Nature 2003;423:81–6, http://dx.doi.org/10.1038/nature01586, pii:nature01586.

[24] Pearson T, Hornstra HM, Sahl JW, Schaack S, Schupp JM, Beckstrom-Sternberg SM, et al. When outgroups fail; phylogenomics of rooting the emerging pathogen, *Coxiella burnetii*. Syst Biol 2013;62:752–62, http://dx.doi.org/10.1093/sysbio/syt038, pii:syt038.

[25] Timme RE, Pettengill JB, Allard MW, Strain E, Barrangou R, Wehnes C, et al. Phylogenetic diversity of the enteric pathogen *Salmonella enterica* subsp. enterica inferred from genome-wide reference-free SNP characters. Genome Biol Evol 2013;5:2109–23, http://dx.doi.org/10.1093/gbe/evt159, pii:evt159.

[26] Croucher NJ, Mitchell AM, Gould KA, Inverarity D, Barquist L, Feltwell T, et al. Dominant role of nucleotide substitution in the diversification of serotype 3 pneumococci over decades and during a single infection. PLoS Genet 2013;9:e1003868, http://dx.doi.org/10.1371/journal.pgen.1003868, pii:PGENETICS-D-13-01204.

[27] Barrick JE, Lenski RE. Genome-wide mutational diversity in an evolving population of *Escherichia coli*. Cold Spring Harb Symp Quant Biol 2009;74:119–29, http://dx.doi.org/10.1101/sqb.2009.74.018, pii:sqb.2009.74.018.

[28] Merhej V, Georgiades K, Raoult D. Postgenomic analysis of bacterial pathogens repertoire reveals genome reduction rather than virulence factors. Brief Funct Genomics 2013;12:291–304, http://dx.doi.org/10.1093/bfgp/elt015, pii:elt015.

[29] Georgiades K, Raoult D. Genomes of the most dangerous epidemic bacteria have a virulence repertoire characterized by fewer genes but more toxin–antitoxin modules. PLoS ONE 2011;6:e17962, http://dx.doi.org/10.1371/journal.pone.0017962.

[30] Gatherer D. The 2009 H1N1 influenza outbreak in its historical context. J Clin Virol 2009;45:174–8, http://dx.doi.org/10.1016/j.jcv.2009.06.004. pii:S1386-6532(09)00253-4.

[31] Grodeland G, Mjaaland S, Roux KH, Fredriksen AB, Bogen B. DNA vaccine that targets hemagglutinin to MHC class II molecules rapidly induces antibody-mediated protection against influenza. J Immunol 2013;191:3221–31, http://dx.doi.org/10.4049/jimmunol.1300504, pii:jimmunol.1300504.

[32] Rolain JM, Vayssier-Taussat M, Saisongkorh W, Merhej V, Gimenez G, Robert C, et al. Partial disruption of translational and posttranslational machinery reshapes growth rates of *Bartonella birtlesii*. MBio 2013;4:e00115–213, http://dx.doi.org/10.1128/mBio.00115-13, pii:mBio.00115-13.

[33] Bruggemann H, Cazalet C, Buchrieser C. Adaptation of *Legionella pneumophila* to the host environment: role of protein secretion, effectors and eukaryotic-like proteins. Curr Opin Microbiol 2006;9:86–94, http://dx.doi.org/10.1016/j.mib.2005.12.009. pii:S1369-5274(05)00203-1.

[34] Voth DE, Howe D, Beare PA, Vogel JP, Unsworth N, Samuel JE, et al. The *Coxiella burnetii* ankyrin repeat domain-containing protein family is heterogeneous, with C-terminal truncations that influence Dot/Icm-mediated secretion. J Bacteriol 2009;191:4232–42, http://dx.doi.org/10.1128/JB.01656-08, pii:JB.01656-08.

[35] Zusman T, Yerushalmi G, Segal G. Functional similarities between the icm/dot pathogenesis systems of *Coxiella burnetii* and *Legionella pneumophila*. Infect Immun 2003;71:3714–23.

[36] Million M, Raoult D. Reply to kampschreur et Al. Clin Infect Dis 2014;58:447–8, http://dx.doi.org/10.1093/cid/cit720, pii:cit720.

[37] Munster JM, Leenders AC, Hamilton CJ, Hak E, Aarnoudse JG, Timmer A. Placental histopathology after *Coxiella burnetii* infection during pregnancy. Placenta 2012;33:128–31, http://dx.doi.org/10.1016/j.placenta.2011.11.012. pii:S0143-4004(11)00544-3.

Supplementary data

Supplementary Table 1. All non-synonymous mutations. The table shows SNiPs and INDELs determining change in protein sequence with position respectively in NMI and Z3055.

Genome coordinate	NMI	Z3055	Gene_ID	Product
10441	A	C	CBU_0009	D-alanyl-meso-diaminopimelate_endopeptidase
10516	A	G	CBU_0009	D-alanyl-meso-diaminopimelate_endopeptidase
13364	G	A	CBU_0014	low-affinity_inorganic_phosphate_transporter
14114	G	A	CBU_0014	low-affinity_inorganic_phosphate_transporter
14383	T	C	CBU_0014	low-affinity_inorganic_phosphate_transporter
16513	A	C	CBU_0017	deoxyribose-phosphate_aldolase
18808	C	G	CBU_0021	hypothetical_protein
18932	C	T	CBU_0021	hypothetical_protein
19322	T	C	CBU_0021	hypothetical_protein
19336	T	C	CBU_0021	hypothetical_protein
19472	T	A	CBU_0021	hypothetical_protein
19844	C	T	CBU_0021	hypothetical_protein
22020	G	A	CBU_0023	hypothetical_cytosolic_protein
22291	G	C	CBU_0023	hypothetical_cytosolic_protein
24340	C	G	CBU_0027	acyltransferase_family_protein
24393	G	A	CBU_0027	acyltransferase_family_protein
26856	G	A	CBU_0029	1-acyl-sn-glycerol-3-phosphate_acyltransferase
28606	C	T	CBU_0031	3-oxoacyl-[acyl-carrier_protein]_reductase
28729	G	T	CBU_0031	3-oxoacyl-[acyl-carrier_protein]_reductase
31274	A	T	CBU_0035	3-oxoacyl-[acyl-carrier-protein]_synthase
31874	G	C	CBU_0035	3-oxoacyl-[acyl-carrier-protein]_synthase
33185	G	A	CBU_0037	3-hydroxydecanoyl-[acyl-carrier-protein]_dehydratase
33413	T	G	CBU_0037a	hypothetical_protein
33528	C	T	CBU_0038	3-oxoacyl-[acyl-carrier-protein]_synthase_III
33835	A	G	CBU_0038	3-oxoacyl-[acyl-carrier-protein]_synthase_III
34416	A	G	CBU_0038	3-oxoacyl-[acyl-carrier-protein]_synthase_III
38353	G	A	CBU_0041	hypothetical_protein
39324	C	T	CBU_0041	hypothetical_protein
39339	C	G	CBU_0041	hypothetical_protein
39734	C	A	CBU_0041	hypothetical_protein
39963	A	G	CBU_0041	hypothetical_protein
47396	T	C	CBU_0049	hypothetical_protein
47523	C	A	CBU_0049	hypothetical_protein
47707	G	A	CBU_0049	hypothetical_protein
49203	G	T	CBU_0051	hypothetical_protein
49362	G	A	CBU_0051	hypothetical_protein
49536	G	A	CBU_0053	enhanced_entry_protein
50218	A	G	CBU_0053	enhanced_entry_protein
53071	G	T	CBU_0056	type_I_secretion_outer_membrane_protein

56251	C	G	CBU_0062	DnaJ_domain_protein
56440	C	T	CBU_0062	DnaJ_domain_protein
57430	C	T	CBU_0063	3-deoxy-D-manno-octulosonic-acid_transferase
57686	A	G	CBU_0063	3-deoxy-D-manno-octulosonic-acid_transferase
58207	C	T	CBU_0063	3-deoxy-D-manno-octulosonic-acid_transferase
61503	T	C	CBU_0065a	hypothetical_protein
63246	T	G	CBU_0068a	hypothetical_protein
65625	-	T	CBU_0072	ankyrin_repeat_protein
65727	T	G	CBU_0072	ankyrin_repeat_protein
65860	A	T	CBU_0072	ankyrin_repeat_protein
65871	C	T	CBU_0072	ankyrin_repeat_protein
65876	G	A	CBU_0072	ankyrin_repeat_protein
65887	A	C	CBU_0072	ankyrin_repeat_protein
65894	A	T	CBU_0072	ankyrin_repeat_protein
66630	C	G	CBU_0072	ankyrin_repeat_protein
66632	A	T	CBU_0072	ankyrin_repeat_protein
66634	A	G	CBU_0072	ankyrin_repeat_protein
66636	C	G	CBU_0072	ankyrin_repeat_protein
76520	C	T	CBU_0084	phosphoglycerol_transferase_MdoB_and_related_protein-like_protein,_alkaline_phosphatase_superfamily
78401	C	A	CBU_0086	putative_lactonase
84536	C	G	CBU_0092	tol_system_periplasmic_component
84695	G	A	CBU_0092	tol_system_periplasmic_component
100077	G	A	CBU_0107	methionine_transport_ATP-binding_protein
101819	C	T	CBU_0109	methionine-binding_protein
102000	G	A	CBU_0109	methionine-binding_protein
102294	G	T	CBU_0109	methionine-binding_protein
102523	G	A	CBU_0109	methionine-binding_protein
111904	T	C	CBU_0122	hypothetical_membrane_associated_protein
112005	A	G	CBU_0122	hypothetical_membrane_associated_protein
112416	T	C	CBU_0123	UDP-N-acetylmuramoylalanyl-D-glutamate--2,6-diaminopimelate_ligase
113066	G	C	CBU_0123	UDP-N-acetylmuramoylalanyl-D-glutamate--2,6-diaminopimelate_ligase
114873	C	T	CBU_0124	UDP-N-acetylmuramoyl-tripeptide--D-alanyl-D-alanine_ligase
129425	T	A	CBU_0141	cell_division_protein
130935	A	G	CBU_0143	hypothetical_protein
130940	G	T	CBU_0143	hypothetical_protein
136877	C	T	CBU_0150	hypothetical_protein
137014	C	T	CBU_0150	hypothetical_protein
137974	G	A	CBU_0152	dephospho-CoA_kinase
137982	C	T	CBU_0152	dephospho-CoA_kinase
138732	A	C	CBU_0153	type_4_prepilin_peptidase
139485	A	G	CBU_0154	type_4_pili_biogenesis_protein_(plasma_membrane_protein)
139663	T	G	CBU_0154	type_4_pili_biogenesis_protein_(plasma_membrane_protein)
139681	G	A	CBU_0154	type_4_pili_biogenesis_protein_(plasma_membrane_protein)
139754	G	T	CBU_0154	type_4_pili_biogenesis_protein_(plasma_membrane_protein)
140060	T	C	CBU_0154	type_4_pili_biogenesis_protein_(plasma_membrane_protein)

143981	G	A	CBU_0159	hypothetical_cytosolic_protein
156818	T	C	CBU_0175	serine/threonine_protein_kinase_domain_protein
157539	T	C	CBU_0175	serine/threonine_protein_kinase_domain_protein
157815	C	A	CBU_0176	endopeptidase
162655	C	T	CBU_0180	peptidoglycan-specific_endopeptidase,_M23_family
163370	C	T	CBU_0180	peptidoglycan-specific_endopeptidase,_M23_family
172742	T	C	CBU_0184	hypothetical_membrane_associated_protein
172793	C	T	CBU_0184	hypothetical_membrane_associated_protein
172796	A	G	CBU_0184	hypothetical_membrane_associated_protein
172799	A	C	CBU_0184	hypothetical_membrane_associated_protein
172801	A	C	CBU_0184	hypothetical_membrane_associated_protein
173013	C	T	CBU_0184	hypothetical_membrane_associated_protein
173462	C	T	CBU_0184	hypothetical_membrane_associated_protein
184188	C	T	CBU_0197	hypothetical_outer_membrane_protein
186346	G	A	CBU_0198	outer_membrane_protein
187019	A	G	CBU_0198	outer_membrane_protein
187778	C	G	CBU_0199	pantothenate_kinase
189378	C	T	CBU_0200	aldose_1-epimerase_family_protein
195233	A	G	CBU_0209	hypothetical_protein
195460	C	G	CBU_0209	hypothetical_protein
197986	G	A	CBU_0214	integral_membrane_protein
199667	C	T	CBU_0215	peptidase,_C40_family
201953	T	C	CBU_0217a	hypothetical_protein
203270	A	G	CBU_0218	cytochrome_d_ubiquinol_oxidase_subunit_I
203721	G	C	CBU_0218	cytochrome_d_ubiquinol_oxidase_subunit_I
205467	C	T	CBU_0221	4'-phosphopantetheinyl_transferase
207884	G	A	CBU_0224	protein_translocase_subunit
207980	A	G	CBU_0224	protein_translocase_subunit
208526	G	C	CBU_0225	transcription_antitermination_protein
213267	A	G	CBU_0231	DNA-directed_RNA_polymerase_beta_chain
216431	A	G	CBU_0232	DNA-directed_RNA_polymerase_beta'_chain
217483	G	A	CBU_0232	DNA-directed_RNA_polymerase_beta'_chain
217709	G	T	CBU_0232	DNA-directed_RNA_polymerase_beta'_chain
221268	A	G	CBU_0234	SSU_ribosomal_protein_S7P
226425	G	A	CBU_0240	LSU_ribosomal_protein_L23P
226792	G	A	CBU_0241	LSU_ribosomal_protein_L2P
228446	C	T	CBU_0244	SSU_ribosomal_protein_S3P
228459	T	C	CBU_0244	SSU_ribosomal_protein_S3P
228937	G	A	CBU_0245	LSU_ribosomal_protein_L16P
229649	G	T	CBU_0247	SSU_ribosomal_protein_S17P
230626	G	A	CBU_0250	LSU_ribosomal_protein_L5P
231894	C	T	CBU_0253	LSU_ribosomal_protein_L6P
233576	C	T	CBU_0257	LSU_ribosomal_protein_L15P
235881	G	A	CBU_0260	SSU_ribosomal_protein_S13P
242483	A	G	CBU_0271	single-strand_DNA_binding_protein
242676	G	A	CBU_0271	single-strand_DNA_binding_protein

243567	G	T	CBU_0272	transporter,_MFS_superfamily
243730	C	T	CBU_0272	transporter,_MFS_superfamily
246794	G	A	CBU_0274	excinuclease_ABC_subunit_A
247709	A	C	CBU_0274	excinuclease_ABC_subunit_A
247756	A	G	CBU_0274	excinuclease_ABC_subunit_A
249092	A	G	CBU_0275	uroporphyrinogen_decarboxylase
249940	A	G	CBU_0276	rubredoxin-NAD(+)_reductase
250554	A	G	CBU_0276	rubredoxin-NAD(+)_reductase
250581	G	A	CBU_0276	rubredoxin-NAD(+)_reductase
251928	G	A	CBU_0279	amidinotransferase_family_protein
252215	G	A	CBU_0279	amidinotransferase_family_protein
254296	C	A	CBU_0282	transcriptional_regulator
254305	T	C	CBU_0282	transcriptional_regulator
254314	A	G	CBU_0282	transcriptional_regulator
254316	C	G	CBU_0282	transcriptional_regulator
254332	T	A	CBU_0282	transcriptional_regulator
254541	C	T	CBU_0282	transcriptional_regulator
254961	G	A	CBU_0284	hypothetical_cytosolic_protein
254992	C	T	CBU_0284	hypothetical_cytosolic_protein
255316	C	T	CBU_0285	hypothetical_protein
255943	G	A	CBU_0286	poly(A)_polymerase
256552	C	T	CBU_0286	poly(A)_polymerase
257974	C	T	CBU_0288	phosphopantetheine_adenylyltransferase
262337	G	C	CBU_0295	hypothetical_protein
263004	C	T	CBU_0295	hypothetical_protein
264413	T	C	CBU_0297	exodeoxyribonuclease_III
267453	T	G	CBU_0300	stress-induced_protein,_putative
270570	T	C	CBU_0303	GTP_pyrophosphokinase
271002	C	T	CBU_0304	translation_initiation_inhibitor
271632	C	T	CBU_0305	ATP-dependent_DNA_helicase
281044	A	C	CBU_0314	hydroxyacylglutathione_hydrolase
281307	C	T	CBU_0314	hydroxyacylglutathione_hydrolase
281621	A	G	CBU_0314	hydroxyacylglutathione_hydrolase
281708	C	T	CBU_0314	hydroxyacylglutathione_hydrolase
282227	C	T	CBU_0315	methyltransferase
282552	G	A	CBU_0315	methyltransferase
290819	G	A	CBU_0324	acyl-CoA_synthetase
296946	G	A	CBU_0330	thiamine_biosynthesis_protein
301062	C	T	CBU_0334	phosphomethylpyrimidine_kinase
301138	G	A	CBU_0334	phosphomethylpyrimidine_kinase
301343	T	C	CBU_0334	phosphomethylpyrimidine_kinase
305808	A	G	CBU_0338	membrane_alanine_aminopeptidase
308270	T	C	CBU_0340	hypothetical_cytosolic_protein
308362	G	A	CBU_0340	hypothetical_cytosolic_protein
308384	C	T	CBU_0340	hypothetical_cytosolic_protein
309657	G	T	CBU_0341	6-phosphofructokinase

313747	C	T	CBU_0346	xylulose_kinase
313916	G	T	CBU_0346	xylulose_kinase
314697	G	A	CBU_0346	xylulose_kinase
320548	T	C	CBU_0354	amino_acid_permease
320666	G	T	CBU_0354	amino_acid_permease
320846	C	T	CBU_0354	amino_acid_permease
321089	A	G	CBU_0354	amino_acid_permease
324099	G	C	CBU_0357	hypothetical_protein
329469	T	G	CBU_0364	phosphate_transporter
329517	A	G	CBU_0364	phosphate_transporter
333350	G	A	CBU_0368	hypothetical_exported_protein
334798	C	T	CBU_0371	multidrug_resistance_transporter,_Bcr_family
334883	G	A	CBU_0371	multidrug_resistance_transporter,_Bcr_family
335092	G	A	CBU_0371	multidrug_resistance_transporter,_Bcr_family
336502	G	A	CBU_0372	Fic_family_protein
336657	G	A	CBU_0372	Fic_family_protein
340567	C	G	CBU_0378	hypothetical_membrane_associated_protein
343543	C	G	CBU_0382	farnesyl_pyrophosphate_synthetase
347889	C	T	CBU_0387	GTP-binding_protein_(probably_involved_in_DNA_repair)
349185	A	G	CBU_0388	hypothetical_protein
349890	G	T	CBU_0388	hypothetical_protein
350032	C	A	CBU_0388	hypothetical_protein
350636	G	T	CBU_0388	hypothetical_protein
351689	A	G	CBU_0388	hypothetical_protein
354773	C	T	CBU_0390	virulence_factor
358263	T	C	CBU_0395	hypothetical_exported_protein
359858	C	A	CBU_0396	isoleucyl-tRNA_synthetase
360204	A	T	CBU_0396	isoleucyl-tRNA_synthetase
361825	T	C	CBU_0397	lipoprotein_signal_peptidase
361950	T	C	CBU_0397	lipoprotein_signal_peptidase
362558	C	G	CBU_0398	hypothetical_membrane_spanning_protein
366457	T	A	CBU_0410	hypothetical_membrane_spanning_protein
366823	G	A	CBU_0410	hypothetical_membrane_spanning_protein
366824	G	A	CBU_0410	hypothetical_membrane_spanning_protein
367392	C	A	CBU_0410	hypothetical_membrane_spanning_protein
367473	C	T	CBU_0410	hypothetical_membrane_spanning_protein
367732	G	C	CBU_0410	hypothetical_membrane_spanning_protein
386395	C	A	CBU_0433	mechanosensitive_ion_channel
388137	C	T	CBU_0436	hypothetical_membrane_spanning_protein
391279	A	G	CBU_0443	tRNA_(Guanine-N(1)-)-methyltransferase
391484	C	A	CBU_0443	tRNA_(Guanine-N(1)-)-methyltransferase
391497	G	A	CBU_0443	tRNA_(Guanine-N(1)-)-methyltransferase
398780	C	T	CBU_0453	type_4_pili_biogenesis_protein_(prepilin-like_protein)
399217	C	T	CBU_0453	type_4_pili_biogenesis_protein_(prepilin-like_protein)
401744	G	A	CBU_0459	potassium/proton_antiporter
417415	C	A	CBU_0476	hypothetical_protein

417670	T	C	CBU_0476	hypothetical_protein
422628	C	T	CBU_0483	arginine_transport_system_permease_protein
428181	G	T	CBU_0487	ribosomal_large_subunit_pseudouridine_synthase_C
428774	C	A	CBU_0487	ribosomal_large_subunit_pseudouridine_synthase_C
433616	C	T	CBU_0492	fatty_acid/phospholipid_synthesis_protein
434538	G	A	CBU_0493	3-oxoacyl-[acyl-carrier-protein]_synthase_III
434828	A	G	CBU_0493	3-oxoacyl-[acyl-carrier-protein]_synthase_III
435669	A	G	CBU_0494	malonyl-CoA-[acyl-carrier-protein]_transacylase
437219	C	T	CBU_0496	acyl_carrier_protein
441711	C	G	CBU_0502	DNase,_TatD_family
442406	A	T	CBU_0503	glutamine_synthetase
444103	G	T	CBU_0504	di-/tripeptide_transporter
450928	A	G	CBU_0510a	hypothetical_cytosolic_protein
461364	C	T	CBU_0518	excinuclease_ABC_subunit_B
462286	G	A	CBU_0519	DedA_family_protein
463291	G	A	CBU_0520	hydroxymethylglutaryl-CoA_lyase
1247388	G	A	CBU_1292	ankyrin_repeat_protein

Supplementary Table 2. Proportion of different proteins classified in each COG category. The table shows proteins changed by SNiPs, proteins changed by INDELs and proteins changed by SNiPs or INDELs. The ratio different proteins/total proteins is calculated for proteins changed by SNiPs or INDELs. Statistical test are shown.

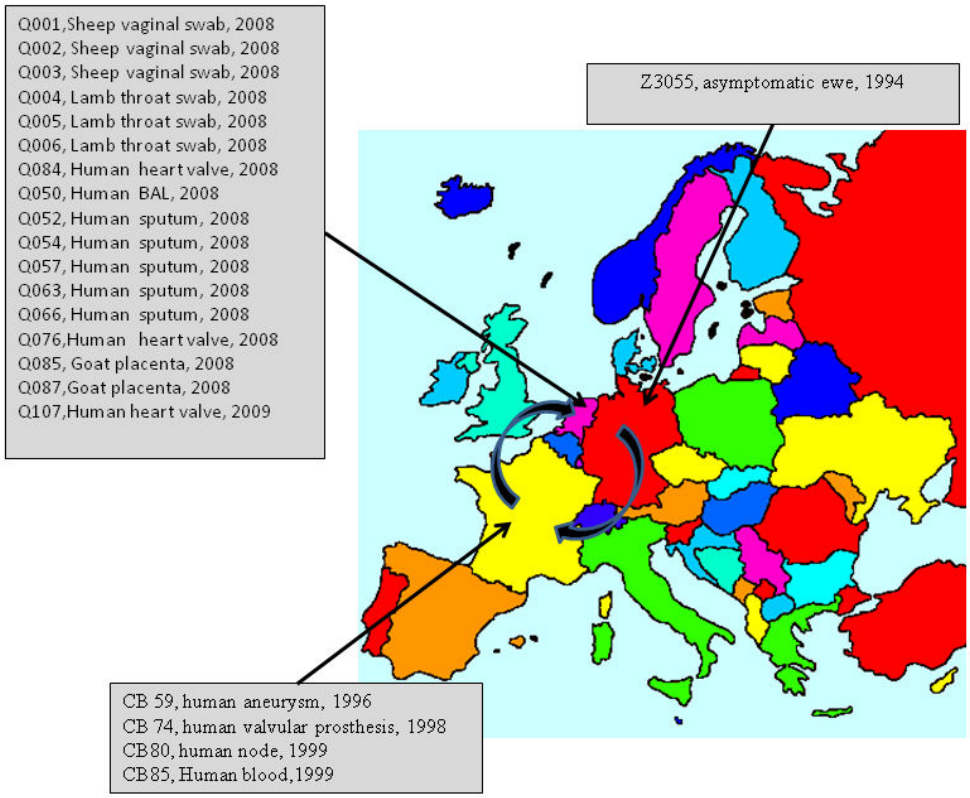
COG	Nb.different proteins encoded by genes mutated by SNPs	Nb.different proteins encoded by genes mutated by INDELs	Nb.different proteins	Total prot.	Nb.different proteins/ Total protein s%	p-value		
						khi2	Fisher	Mid-P
Other (not in COG)	19	0	19	484	3,9%	0,01558		
Translation, ribosomal structure and biogenesis	15	0	15	131	11,5%	0,01024		
RNA processing and modification	0	0	0	1	0,0%	-	-	-
Transcription	4	1	4	31	12,9%		0,24	0,16
Replication, recombination and repair	5	0	5	110	4,5%	0,003264		
Chromatin structure and dynamics	0	0	0	0	#	-	-	-
Cell cycle control, cell division, chromosome partitioning	1	0	1	22	4,5%		2,894E-06	1,599E-06
Defense mechanisms	0	0	0	33	0,0%	-	-	-
Signal transduction mechanisms	0	0	0	39	0,0%	-	-	-
Cell wall/membrane/envelope biogenesis	11	0	11	108	10,2%	0,08		
Cell motility	0	0	0	0	#	-	-	-
Cytoskeleton	0	0	0	0	#	-	-	-
Extracellular structures	0	0	0	0	#	-	-	-
Intracellular trafficking, secretion, and vesicular transport	1	0	1	20	5,0%		>0,9999999	0,92
Posttranslational modification, protein turnover, chaperones	2	0	2	58	3,4%		0,58	0,4
Energy production and conversion	2	0	2	82	2,4%	0,15		
Carbohydrate transport and metabolism	6	0	6	74	8,1%		0,62	0,48
Amino acid transport and metabolism	7	0	7	86	8,1%	0,45		
Nucleotide transport and metabolism	1	0	1	42	2,4%		0,5	0,32
Coenzyme transport and metabolism	9	0	9	77	11,7%		0,09	0,06
Lipid transport and metabolism	7	0	7	55	12,7%		0,1	0,07
Inorganic ion transport and metabolism	5	0	5	45	11,1%		0,29	0,20
Secondary metabolites biosynthesis, transport and catabolism	0	0	0	18	0,0%	-	-	-
General function prediction only	7	1	7	164	4,3%	0,2802		
Function unknown	6	0	6	120	5,0%		0,74	0,59
Multiple class	11	0	11	113	9,7%	<0,0000001		
Total	**119**	**2**	**119**	**1913**	**6,2%**			

Supplementary Table 3: Genotype MST33 history

Year		Isolate	Host	Origin	Disease, clinical findings	Geographic source	MST	Ref
1992	First isolation of *C. burnetii* MST33 from a ewe in Germany	Z3055	Ewe	Placenta	Asymptomatic	Germany	33	[7]
1996	Identification of *C. burnetii* MST33 in human clinical samples in France	CB59	Human	Aneurysm	Vascular infection	Saint Etienne, France	33	
1998		CB74	Human	Valvular prosthesis	Endocarditis	Toulouse, France	33	
1999		CB80	Human	Node	ND	Niort, France	33	
1999		CB85	Human	Blood	Endocarditis	Tour, France	33	
2002	Isolation of a doxycycline resistant clinical strain of *C. burnetii* in Germany presenting a high similarity with Z3055.	CB 109	Human	Heart valve	Endocarditis	Berlin, Germany	12	[5]
2008	The *C. burnetii* MST33 was identified among farm animals and human over the entire area affected with Q fever in Netherland implicating goat and sheep as possible source of the outbreak	Q001	Sheep	Vaginal swab	ND	Netherlands	33	[7]
2008		Q002	Sheep	Vaginal swab	ND	Netherlands	33	
2008		Q003	Sheep	Vaginal swab	ND	Netherlands	33	
2008		Q004	Lamb	Throat swab	ND	Netherlands	33	

2008		Q005	Lamb	Throat swab	ND	Netherlands	33
2008		Q006	Lamb	Throat swab	ND	Netherlands	33
2008		Q084	Human	Heart valve	ND	Netherlands	33
2009		Q050	Human	BAL	ND	Netherlands	33
2009		Q052	Human	Sputum	ND	Netherlands	33
2009		Q054	Human	Sputum	ND	Netherlands	33
2009		Q057	Human	Sputum	ND	Netherlands	33
2009		Q063	Human	Sputum	ND	Netherlands	33
2009		Q066	Human	Sputum	ND	Netherlands	33
2009		Q076	Human	Heart valve	ND	Netherlands	33
2009		Q085	Goat	Placenta	ND	Netherlands	33
2009		Q087	Goat	Placenta	ND	Netherlands	33
2009	A second genotype, *C. burnetii* MST 20, was identified from cattle, one goat and cow milk in the Netherlands but not in human that indicate that outbreak in the Netherlands is not linked with cattle.	Q086	Goat	Placenta	ND	Netherlands	20
2009		Q097	Cattle	Swab	ND	Netherlands	20
2010		Q107	Human	Heart valve	ND	Netherlands	33
2010		Q090	cattle	Milk	ND	Netherlands	20
2010		Q091	cattle	Milk	ND	Netherlands	20
2010		Q093	cattle	Milk	ND	Netherlands	20
2010		Q096	cattle	Milk	ND	Netherlands	20
2010		Q123	cattle	Milk	ND	Netherlands	20

Supplementary Figure 1. *C.burnetii* strains' distribution in Europe.



CHAPITRE 2: De l'hyper virulence du clone circulant en Guyane française

Avant-propos

En Guyane Française, la fièvre Q est un problème de santé publique. Ceci représente une exception en comparaison des autres pays d'Amérique du Sud où les données de prévalence sur cette pathologie sont rares. *C.burnetii* est responsable dans cette région de la plus forte prévalence de pneumonies communautaires (24,4%) jamais rapportée dans la littérature. L'incidence annuelle est passée de 37 cas pour 100.000 habitants en 1996 à 150 cas pour 100.000 habitants en 2005 . D'autre part, une comparaison clinique et sérologique entre 115 patients atteints de la fièvre Q en Guyane française et 182 patients en France métropolitaine, a montré que la pneumonie est la présentation clinique principale avec une forte réponse sérologique et une sévérité plus importante que les pneumonies non liées à la fièvre Q , ce qui a suggéré la circulation d'un clone hyper virulent . Le génotypage de souches provenant de Guyane Française a montré qu'un seul clone, de génotype MST17, circule spécifiquement dans cette zone géographique. En effet, il a été seulement détecté dans des échantillons prélevés sur des patients qui vivaient ou avaient vécu à Cayenne .

Nous avons séquencé la souche Cb175 isolée de la valve cardiaque d'un patient guyanais souffrant d'une endocardite à *C.burnetii*, afin d'étudier les déterminants génétiques de cette hypervirulence.
Après le « mapping » des reads contre la référence NMI, nous avons obtenu un « draft génome » contenant un gap d'une taille estimée d'environ 7kb. Lorsque nous avons essayé de compléter ce gap, nous avons obtenu une séquence de 767 pb. Après l'alignement par BLASTN contre le génome de NMI, nous avons observé sa position à 224 et 546 pb dans la région adjacente aux 6105 pb manquants dans la souche Guyane.

Nous avons, donc, testé les amorces dessinées pour le finishing sur les souches NMI et Z3055, ayant cette région, afin d'exclure une erreur dans le design des amorces et nous avons obtenu comme attendu une séquence de 7Kb. Ensuite, nous avons testé les amorces sur l'ADN de deux souches isolées directement à partir des valves cardiaques et sur des souches avec peu de passages de culture: nous avons obtenu une bande de 767 pb et non de 7 Kb. Ce dernier résultat a confirmé que la délétion n'a pas été une conséquence du passage en phase II. La délétion a été également trouvée dans deux échantillons d'une tique de paresseux et sur toutes les souches isolées chez des patients guyanais. Nous avons recherché la délétion dans les génomes disponibles sur GenBank à la fois *in silico* et par biologie moléculaire (avec un système de qPCR) dans les souches collectées dans notre laboratoire. Celle-ci était absente de toutes les souches testées. Cette délétion est donc spécifique des souches circulant en Guyane Française et entraîne donc une réduction significative de la taille du génome de la bactérie comparée aux souches d'autres origines géographiques. La réduction du génome a été précédemment décrite pour les bactéries épidémiques les plus dangereuses telles que les Mycobactéries et les Rickettsies ayant des tailles de génomes réduites par rapport à leurs espèces les plus proches non-épidémiques . Une étude réalisée dans notre laboratoire a comparé les 12 bactéries épidémiques les plus dangereuses avec leurs espèces proches non-épidémiques et a constaté que les espèces épidémiques sont caractérisées par des génomes réduits, une réduction significative du contenu d'ORF et une disparition progressive des gènes .

En outre, nous avons constaté que la région manquante contient l'opéron hlyCABD du système de sécrétion de type I (T1SS). Il a été démontré que les espèces bactériennes épidémiques les plus dangereuses contiennent beaucoup moins de protéines du système de sécrétion par rapport aux espèces proches moins pathogènes . Cependant, dans *Coxiella burnetii* l'opéron hlyCABD n'est

pas complet et la fonctionnalité du T1SS n'a pas encore été clarifiée. A rajouter à l'intérêt de cette délétion, le fait d'avoir trouvé une analogie supplémentaire avec Legionella : les mêmes gènes sont perdus par *L. longbeachae* comparée à *L.pneumophila* .

Enfin, nous avons également constaté une fréquence élevée de mutations dans les gènes codants pour les protéines du métabolisme. Il a été décrit que l'activité métabolique des bactéries intracellulaires les plus pathogènes est diminuée et qu'elles nécessitent moins de gènes du métabolisme . Toutefois, les effets des mutations au sein de ces gènes sur la fonctionnalité de leurs produits doivent être étudiés.

Ce travail a été publié dans le journal ***Comparative Immunology, Microbiology & Infectious Diseases***.

Loss of TSS1 in Hypervirulent *Coxiella burnetii* 175, the Causative Agent of Q Fever in French Guiana

Running title: **Genome reduction in *Coxiella burnetii* from French Guiana**

Felicetta D'Amato[1a], Carole Eldin[1a], Kalliopi Georgiades[2], Sophie Edouard[1], Jeremy Delerce[1], Noémie Labas[1], Didier Raoult[1]*

[1]Aix Marseille Université, URMITE, UM63, CNRS 7278, IRD 198, Inserm 1095, 13005 Marseille, France.

[2] University of Cyprus, Medical School, PO Box 20537, CY 1678, Nicosia, Cyprus

Comparative Immunology, Microbiology and Infectious Diseases xxx (2015) xxx–xxx

Contents lists available at ScienceDirect

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

Loss of TSS1 in hypervirulent *Coxiella burnetii* 175, the causative agent of Q fever in French Guiana

Felicetta D'Amato[a,1], Carole Eldin[a,1], Kalliopi Georgiades[b], Sophie Edouard[a], Jeremy Delerce[a], Noémie Labas[a], Didier Raoult[a,*]

[a] Aix Marseille Université, URMITE, UM63, CNRS 7278, IRD 198, INSERM 1095, 13005 Marseille, France
[b] University of Cyprus, Medical School, PO Box 20537, CY 1678 Nicosia, Cyprus

ARTICLE INFO

Article history:
Received 16 January 2015
Received in revised form 23 April 2015
Accepted 28 April 2015

Keywords:
Coxiella burnetii
Guiana
Cb175

ABSTRACT

In French Guiana, the unique *Coxiella burnetii* circulating genotype 17 causes 24% of community-acquired pneumonia, the highest prevalence ever described. To explain this unusual virulence, we performed a comparative genomic analysis of strain Cb175, which was isolated from a patient from French Guiana. Cb175 has a greater number of mutations in genes involved in metabolism compared with the Nine Mile I strain. We found a 6105 bp fragment missing in Cb175, which corresponds to the Type 1 secretion systems (T1SS) hlyCABD operon region. This deletion was detected by a specific qPCR in the 8 other strains available from this territory an in none of 298 *C.burnetii* strains from other areas and other genotypes (8/8 vs 0/298, Fisher's exact test, $p < 0.0000001$). Loss of genes implicated in secretion systems has been observed in other epidemic bacterial strains. Thus, the virulence of Cb175 may be linked to this genome reduction.

1. Introduction

Coxiella burnetii is an obligate intracellular bacterium that belongs to the Gamma-Proteobacteria family [1] and causes Q fever, a worldwide occurring zoonosis [2]. In the acute form of the disease, clinical presentation ranges from asymptomatic seroconversion to various clinical conditions, such as flu-like symptoms, hepatitis or pneumonia. The clinical presentation and severity of Q fever can vary depending on the strain of *C. burnetii* involved [3]. In metropolitan France, where Q fever is endemic, the annual incidence of acute infection is estimated to be 2.5 per 100,000 inhabitants [4], and fever and transaminitis are the most common clinical symptoms [5]. The epidemiology, clinical features and serological responses of Q fever reported in French Guiana are different from what has been described throughout the rest of the world [6].

Since its first report in 1955 [7], the incidence of acute Q fever in Cayenne increased to 37 per 100,000 inhabitants in 1996 [8]. That year, 3 patients were admitted to the intensive care unit in Cayenne Hospital for Q fever pneumonia, and one patient died as a result of distress respiratory syndrome [6]. Subsequently, the rate of incidence has continued to increase and peaked at 150 per 100,000 inhabitants in 2005 [9]. Q fever currently represents 24% of cases of community-acquired pneumonia in Cayenne [10]. Patients with Q fever pneumonia in Cayenne exhibit a more severe initial presentation with significantly more frequent chills, night sweats, headache and arthromyalgia than patients with other etiologies of pneumonia. These patients also have a more marked inflammatory response with higher CRP levels, but lower leukocyte counts [10]. Additionally, Q fever patients from Cayenne have a higher prevalence of fever (97%) and pneumonia (83%) than patients from metropolitan France (81% and 8%, respectively) [7]. Regarding the serological response, Guianan Q fever patients exhibit higher levels of phase I antibodies in the acute form of the disease [7].

In 2012, five *C. burnetii* isolates obtained from samples of five patients from Cayenne were cultured for the first time [10]. Genotypic analysis of these strains revealed that a single clone (MST17) circulates in Cayenne and is related to genotypes that harbor the QpH1 plasmid. This clone is epidemic and has been present in Cayenne since at least 2000 [11], the year that this first isolate was sampled.

* Corresponding author at: URMITE, UMR CNRS 7278, IRD 198, INSERM U1095, Faculté de Médecine, 27 Bd Jean Moulin, 13385 Marseille cedex 5, France. Tel.: +33 04 91 32 43 75; fax: +33 04 91 38 77 72.
E-mail addresses: felicettadamato@gmail.com (F. D'Amato), carole.eldin@gmail.com (C. Eldin), popigeorgiades@hotmail.com (K. Georgiades), soph.edouard@gmail.com (S. Edouard), jeremy.delerce@gmail.com (J. Delerce), labas-noemie@live.fr (N. Labas), didier.raoult@gmail.com (D. Raoult).
[1] These authors participated equally to the work and should be considered as first co-authors.

http://dx.doi.org/10.1016/j.cimid.2015.04.003
0147-9571/

We performed a comparative genomic analysis of one of these strains, Cb175, to investigate why this genotype causes more virulent acute forms than other strains that have been previously described.

2. Methods

2.1. Strain culture, basic genomic data and sequencing

The Cb175 strain was isolated in May 2012 from the cardiac valve of a 60-year-old patient living in French Guiana who underwent an operation for blood-culture negative endocarditis with cardiac dysfunction [11]. This strain and all other *C. burnetii* strains from patient samples were cultured at 35 °C on L929 cells using MEM (GIBCO, Invitrogen, Cergy-Pontoise, France) supplemented with 4% FBS (GIBCO) and 1% L-glutamine (GIBCO). Monolayers of cells and the supernatants from three 175 cm^2 flasks were harvested and incubated with 1% trypsin (GIBCO) for 1 h at 37 °C. The released bacteria were purified from L929 cell debris on a discontinuous Gastrografine (Schering, Lys-Lez-Lannoy, France) gradient (45%, 36% and 28%) and ultracentrifuged at 5000 Tars/min for 1 h 10 min. The genome of the culture-positive strain Cb175 was subjected to paired-end SOLiD sequencing (run accession ERR845240). DNA was extracted using a QIAamp® DNA Mini Kit protocol by performing two subsequent elutions in a total volume of 30 μl. The paired-end library was constructed from 1 μg purified genomic DNA (58 ng/μl) after quantification using the Quant-iT™ PicoGreen® dsDNA Assay Kit from Life Technologies. Sequencing was carried out to 50 × 35 bp using SOLiD™ V4 chemistry on one full slide that was associated with 95 other projects on an Applied Biosystems SOLiD4 machine. All 96 genomic DNA samples were barcoded with the module 1–96 barcodes provided by Life Technologies and were fragmented on a Covaris device. The concentration of the library was measured on the Qbit fluorometer as 20.3 nmol/l. Libraries were pooled in equimolar ratios and size-selected on the E-Gel iBase system at 240–270 bp. PCR was performed on the EZ beads automated Emulsifier, Amplifier and Enricher E80 using the full-scale template bead preparation kit according to the protocol provided by Life Technologies. A total of 708 million P2-positive beads were loaded onto the flow cell. The output paired read length was 85 bp (50 × 35 bp). Of the total of 39.8 Gb for the full slide, the *C. burnetii* Cb175 project yielded 7,529,649 barcoded paired reads, which add up to 640 Mbp. Among a total of 23,149,703 reads obtained for four *C. burnetii* strains sequenced in the same SOLiD slide, 21.7% were associated with Z3055, 32.5% with Cb175, 27% with Cb51 and 18.8% with the HenzerlingS strain.

2.2. Genotyping and finishing

Genotyping. Multi-spacer sequence typing (MST) was performed to define the genotype of the strain. This consisted of PCR and sequencing of the following 10 spacers of the *C. burnetii* genome: Cox2, Cox 5, Cox 6, Cox 18, Cox 20, Cox 22, Cox 37, Cox 51, Cox 56, Cox 57 and Cox 61 [3]. Subsequently, these results were confirmed by in silico genotyping. Using a web-based MST database (http://ifr48.timone.univ-mrs.fr/MSTCoxiella/mst), we aligned each Cox sequence with the Cb175 genome.

Finishing. The reads obtained were mapped against the reference Nine Mile I (NC_002971.3) genome using CLC Genomics Workbench 6.0 (copyright CLC bio) using the parameters length fraction = 0.5 and similarity fraction = 0.8, which allowed us to obtain a consensus of 1,992,640 bp with an average coverage of 154 and fraction of reference covered of 100%.

Then, gap finishing was performed by multiple rounds of PCR and sequencing using BigDye terminator chemistry on the ABI3730 sequencing machine (Applied Biosciences) to complete the consensus sequence. Standard PCR was conducted in a total volume of 20 μl containing 0.02 U Phusion High-Fidelity DNA Polymerase (Finnzyme, Thermo Scientific), 200 μM dNTPs, 1× reaction buffer, 0.5 μM of each amplification primer, and 0.5 μl template DNA. Finally, we performed a mismapping analysis in order discard regions incorrectly placed. During finishing step, a 6105 bp deletion region, compared to the reference genome of NM I, was detected.

2.3. Determination of missing region in other Coxiella burnetii strains from Guiana

A qPCR system was designed to target the 6105 bp deletion by choosing the sequences F-GTGACGTTTATGGTTACTCATG and R-CTCCGATGCGGATAAATCCTA on each side of the deleted region as the primers and 6-FAM-AATCCGCCGGAACAGTCGTCAAC-TAMRA as the probe targeting the 23 bp region flanking the deletion (Fig. 1). Using this strategy, a positive PCR product was only obtained in strains of *C. burnetii* that had this deleted region. Primer and probe specificity were verified in silico by a BLAST search of GenBank (National Center for Biotechnology Information, Bethesda, MD, USA) and by qPCR of a panel of 15 *C. burnetii* strains, including 5 strains of *C. burnetii* MST17 and 10 other genotypes of *C. burnetii* (Table 1). Then, a larger sample size was tested that included five DNA samples from 5 patients with Q fever from Cayenne [11], 1 DNA sample from the feces of a three-toed sloth, 2 DNA samples from ticks collected from a three-toed sloth from Cayenne [12] and 298 other DNA samples from patients diagnosed at our center with genotypes other than MST17.

2.4. Investigation of the genomic sequence corresponding to the missing region in the Guiana strain in other Coxiella burnetii genomes

The missing region in Cb175 was searched for in the following other *C. burnetii* strains sequenced by our laboratory: 21 strains sequenced using Illumina MiSeq technology, 13 strains sequenced using SOLiD and one strain sequenced using the 454 shotgun method. Reads obtained by sequencing were mapped against the missing region using stringent parameters (length fraction = 0.95 and similarity fraction = 0.95). An identical method was used for strain Cb109 [13], which has been sequenced in our laboratory, but does not have a completely annotated genome. Additionally, a BLASTN analysis [14] of the missing region was performed against the *C. burnetii* genomes available in GenBank, including the Dugway 5J108-111 [15], CbuG_Q212 [15,16], CbuK_Q154 [15], RSA 331, Nine Mile phase I (NMI) [15,16], Z3055 [17], Q321, MSU Goat Q177, Cb_C2 [18], Cb_B1 [18], EV-Cb_BK10 [18], Cb_O184 [18], EV-Cb_C13 [18] and Cb_B18 [18] strains.

2.5. Genetic variability

The Cb175 genomic sequence was compared with the 6 complete genomes available in GenBank: Dugway 5J108-111, CbuG_Q212, CbuK_Q154, RSA 331, Nine Mile, and Z3055 (for which we have submitted also plasmid sequence LN827801). The GenMark software [19] was used to predict and translate the coding regions (CDs). The CDs of each genome were concatenated and dereplicated by clustering using USEARCH [20] at an identity of 0.9 to reduce the size of dataset. TBLASTN [21] of the translated sequences was performed against each genome of the dataset and the query bit score for each genome was tabulated. The query bit score was divided by the maximum bit score for all genomes to calculate the blast score ratio [22,23], which could range from 1.0 (exact peptide

Please cite this article in press as: F. D'Amato, et al., Loss of TSS1 in hypervirulent *Coxiella burnetii* 175, the causative agent of Q fever in French Guiana, Comp Immunol Microbiol Infect Dis (2015), http://dx.doi.org/10.1016/j.cimid.2015.04.003

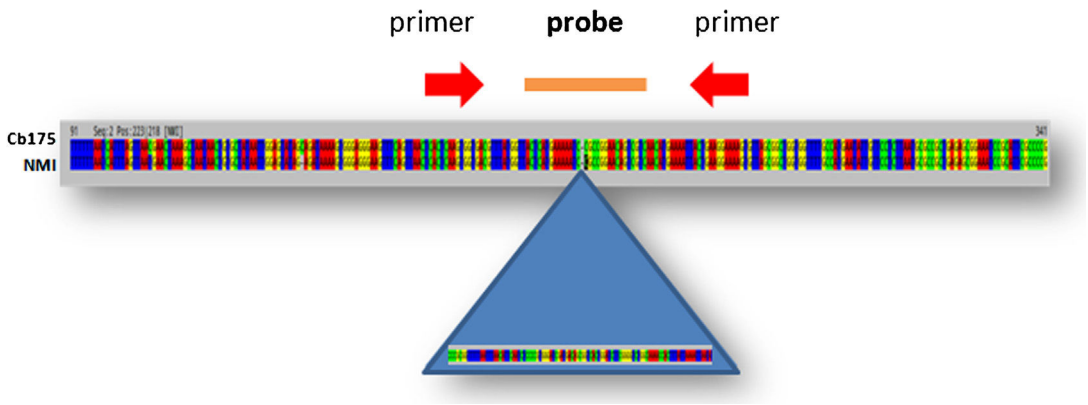


Fig. 1. The qPCR system used in this study. In orange, the probe targeting 23 bp of the Cb175 sequence containing the missing region. In red, primers flanking that region. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

match) to 0.0 (no significant alignment). Lost genes were found to have a blast score ratio ≤0.4 (having an identity ≤40% over 100% of the peptide length). A hierarchical clustering of strains based on genetic differences was obtained using the multi-experiment viewer (MeV) [24,25] by entering the value of the blast score ratio across groups. In this way, the Blast Score Ratio algorithm also allowed us to obtain an estimate of the genetic repertoire of the group of *C. burnetii* strains in our study, as previously described [17]. Therefore, genes with a blast score ratio >0.4 in some genomes, but not in all (i.e., genes that were found in some *C. burnetii* strains) have been classified as accessory, genes having a blast score ratio >0.4 in only one strain, but ≤0.4 in others were classified as unique, and genes having a blast score ratio >0.4 in all strains were classified as unique core genes. To obtain an estimate of the core genome/pangenome ratio, we considered: for the pangenome, the total number of genes before dereplication; for the core genome, the total of the genes of the core obtained by blast score ratio and genes from the filtered dataset of dereplication and not belonging to accessory or unique genes in multiple copies found in the same dataset.

The Cb175 and the reference NMI genomes were compared using the Probabilistic Variant Detection tool of the CLC Genomics Workbench 7.0 (Copyright CLC bio) with the parameters of length fraction = 0.7 and similarity fraction = 0.8 to identify Single- and Multiple-Nucleotide Variations (SNVs and MNVs), Insertions and Deletions. We confirmed or excluded larger InDels, when detected in the consensus sequence, by molecular biology during the finishing step.

Only mutations with coverage > 50 were considered. Protein sequences were aligned to verify whether point mutations were silent mutations or changed the amino acid sequence (non-synonymous mutations).

Genes of NMI and non-synonymous genes were classified by Clusters of Orthologous Groups (COG) using the WebMGA annotation tool [26]. Then, the proportion of different proteins for each COG category was calculated.

3. Results

3.1. General features

Cb175 is characterized by a 1,989,565 Mb chromosome (G + C content 42.6%) and a QpH1 plasmid of 37.388 bp (accession number HG825990 for both chromosome and plasmid). Based on PCR results, the Cb175 genome showed 100% identity and coverage with the sequences of Cox2.3, Cox5.8, Cox18.5, Cox20.7, Cox22.4, Cox37.1, Cox51.10, Cox56.4, Cox56.8, Cox57.6, and Cox61.2, suggesting that the genotype of Cb175, MST17, is similar to the other strains isolated in Guiana [7,11].

3.2. Blast score ratio and genetic variability

Genomic clustering using the blast score ratio algorithm demonstrated a high level of similarity between the strains, and also showed that Cb175 is in the cluster that includes strains NMI, RSA331 and Z3055, which harbor the plasmid QpH1 (Fig. 2). Based on the Blast score ratio algorithm, we estimated a core genome/pangenome ratio of 96%. Indeed, a total of 13,542 core genes (including 2270 unique core genes among 2568 total dereplicated genes), 498 accessory genes (including 210 unique genes),

Table 1
Panel of 15 *C. burnetii* strains used for RT-PCR.

Strain name	Nature of specimens	Clinical characteristic	Age/sex	Geographical area	Genotype	Ref.
Cb 179	Blood	Acute Q fever	44/M	Cayenne	MST17	[11]
Cb 181	Blood	Acute Q fever	55/F	Cayenne	MST17	[11]
Cb 182	Blood	Acute Q fever	47/M	Cayenne	MST17	[11]
Cb 77	Cardiac valve	Q fever endocarditis	40/M	Cayenne	MST17	[11]
Cb 176	Cardiac valve	Q fever Vascular infection	66/M	Cayenne	MST17	
Cb196	Cardiac valve	Q fever endocarditis	13/M	Saudi Arabia	MST51	[40]
Cb195	Cardiac valve	Q fever endocarditis	13/M	United states	MST54	
Cb165	Cardiac valve	Q fever endocarditis	75/M	France	MST1	
Cb173	Cardiac valve	Q fever endocarditis	66/M	France	MST8	
Cb163	Blood	Q fever endocarditis	12/F	Marseille, France	MST33	
Cb185	Placenta	Abortion	20/F	France	MST18	
Cb177	Cardiac valve	Q fever endocarditis	40/M	France	MST12	
Cb119	Cardiac valve	Q fever endocarditis	44/M	Senegal	MST19	
Dog Utad	Uterus			Canada	MST21	
Nine Mile	Tick			United states	MST16	

Please cite this article in press as: F. D'Amato, et al., Loss of TSS1 in hypervirulent *Coxiella burnetii* 175, the causative agent of Q fever in French Guiana, Comp Immunol Microbiol Infect Dis (2015), http://dx.doi.org/10.1016/j.cimid.2015.04.003

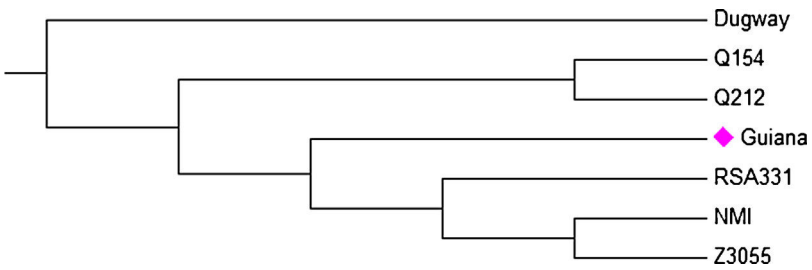


Fig. 2. Hierarchical clustering based on the blast score ratio across 7 whole analyzed genomes. The Guiana strain is in the same cluster of the other strains harboring plasmid QpH1, RSA331, Z3055 and NMI.

and 88 unique genes (not in multiple copies), were identified in these strains. Among the unique genes, 74 belonged to strain Dugway, 13 to strain Q212 and 1 unique gene to RSA331, as previously described [17]. Reads resulting unmapped against the Nine Mile I chromosome (NC_002971.3) have been assembled using CLC Genomics Workbench 7.0 with the parameters of minimum contig length = 85, length fraction = 0.5 and similarity fraction = 0.8. Then, the obtained contigs were aligned using BLASTN against the nr database and after only against plasmids' sequences database. Among 131 total contigs, 51 contigs showed a match with *C. burnetii* and all of these corresponded to a plasmid sequence, with more hits for the QpH1 plasmid.

The others remaining assembled contigs corresponded to *Mus musculus*, *Plasmodium yoelii yoelii*, *Francisella tularensis*, *Trypanosoma congolens*, *Staphylococcus epidermidis*, *Cricetulus griseus*, *Bacillus sp.*, *Staphylococcus epidermidis*, and *Influenza A virus*, representing the contaminants in our sample. No contig has been found to have a match with a bacterium close to *Coxiella* species. Therefore, no evidence suggested the presence of a unique gene specific to strain Cb175. However, we must consider that the small size of reads could represent a limit for assembly. This is, by the way, the reason for our choice to perform mapping to obtain chromosome and plasmid sequences rather than an assembly.

3.3. Point mutation analysis

A total of 1163 mutations between Cb175 and NMI were identified, among which 401 were located in intergenic regions and 762 were in coding regions. Among the 762 mutations in coding genes, 730 were SNPs (including 463 non-synonymous), 9 were insertion mutations and 23 were deletions (Supplementary Table 1). These mutated genes correspond to 397 non-synonymous genes. Mutated genes encoded proteins that are involved in the following processes: 137, metabolism; 26, replication; 4, transcription; 13, translation; 5, secretion system; 1, the two component system; 18, cellular processes, 3, defense mechanisms and 47 are membrane proteins, 9 are transporters, 4 are ankyrins, 1 is a chaperonin, 1 is a CRISPR and 128 are hypothetical proteins. Statistical analyses suggested a trend for an increase in the proportion of mutated genes that encode ankyrin repeat domain proteins compared with other proteins (4/9 vs. 393/1904, bilateral Mid-P test $p = 0.12$).

When proteins were classified by COG categories, the proportion of mutated proteins was significantly increased for category M (cell wall/membrane/envelope biogenesis) compared with the other categories (31/108 vs. 366/1805; bilateral chi^2 test $p = 0.04$). There was a trend toward an increase in the proportion of mutated proteins in category E (amino acid transport and metabolism, 23/86 vs. 374/1827; $p = 0.16$) and category H (coenzyme transport and metabolism, 21/77 vs. 376/1836; $p = 0.15$).

3.4. Missing region compared with the NMI' genome

PCR amplification of a gap to close with an expected size of 6869 bp yielded a band of 767 bp. Sequencing the 767 bp band followed by a BLASTN analysis [14] against the reference strain Nine Mile I showed that the sequence matched at 546 bp and 224 bp, flanking a missing region of 6105 bp in strain Cb175 (Fig. 3). The deleted region contains one pseudogene (CBU_2034) that corresponds to a hypothetical membrane-associated protein, one gene (CBU_2036) encoding a hypothetical protein, one pseudogene (CBU_2037) corresponding to a multidrug resistance ABC transporter ATP-binding and permease, one pseudogene (CBU_2038a) corresponding to a α-hemolysin translocation ATP-binding protein HlyB (*Escherichia coli*), and one gene (CBU_2040) encoding a HlyD family type I secretion membrane fusion protein. HlyA and HlyC, which are two important components of the *E. coli* Type 1 Secretion System (T1SS) hlyCABD operon, are not present in *C. burnetii.*

In all samples isolated from patients and ticks from French Guiana, the 6869 bp region was not amplified, but a 767 bp fragment similar to that of strain Cb175 was obtained. These results were corroborated with the qPCR analysis that was specifically designed to target the deleted region. Only DNA samples from patient and ticks from Cayenne tested positive in this qPCR assay, while those of all other genotypes tested negative, suggesting that the qPCR assay was specific for the MST 17 genotype.

3.5. No detection of the deletion in other sequenced strains and genomes available in GenBank

The deletion described earlier was not found in any other sequenced strains. The average coverage results of mapping and consensus size indicated that the 6105 bp region is present in all other C. *burnetii* genomes analyzed. Specifically, for strains sequenced using the SOLiD and 454 methods, we found an average coverage ranging from a minimum of 61 bp (for Cb121) to a maximum of 247 bp (for Cb13) with a consensus size of at least 5998 bp. For strains sequenced by Illumina, the minimum average value of coverage that we obtained was 54 bp (for Cb94) and the maximum value was 226 bp (as for Cb111) with a consensus size of at least 6089 bp. Moreover, the strains available in GenBank were not characterized by the loss of this region; each genome matched with a sequence identity value of 99–100%.

3.6. Alignment of the missing region against Legionella pneumophila genomes

Alignment of the 6105 bp sequence against *Legionella pneumophila* genomes revealed the presence of conserved regions in the lssB and HlyD genes. Specifically, in *Legionella pneumophila philadelphia1*, the following two genes were in the same region and situated next to each other: LPE509_01686

Please cite this article in press as: F. D'Amato, et al., Loss of TSS1 in hypervirulent *Coxiella burnetii* 175, the causative agent of Q fever in French Guiana, Comp Immunol Microbiol Infect Dis (2015), http://dx.doi.org/10.1016/j.cimid.2015.04.003

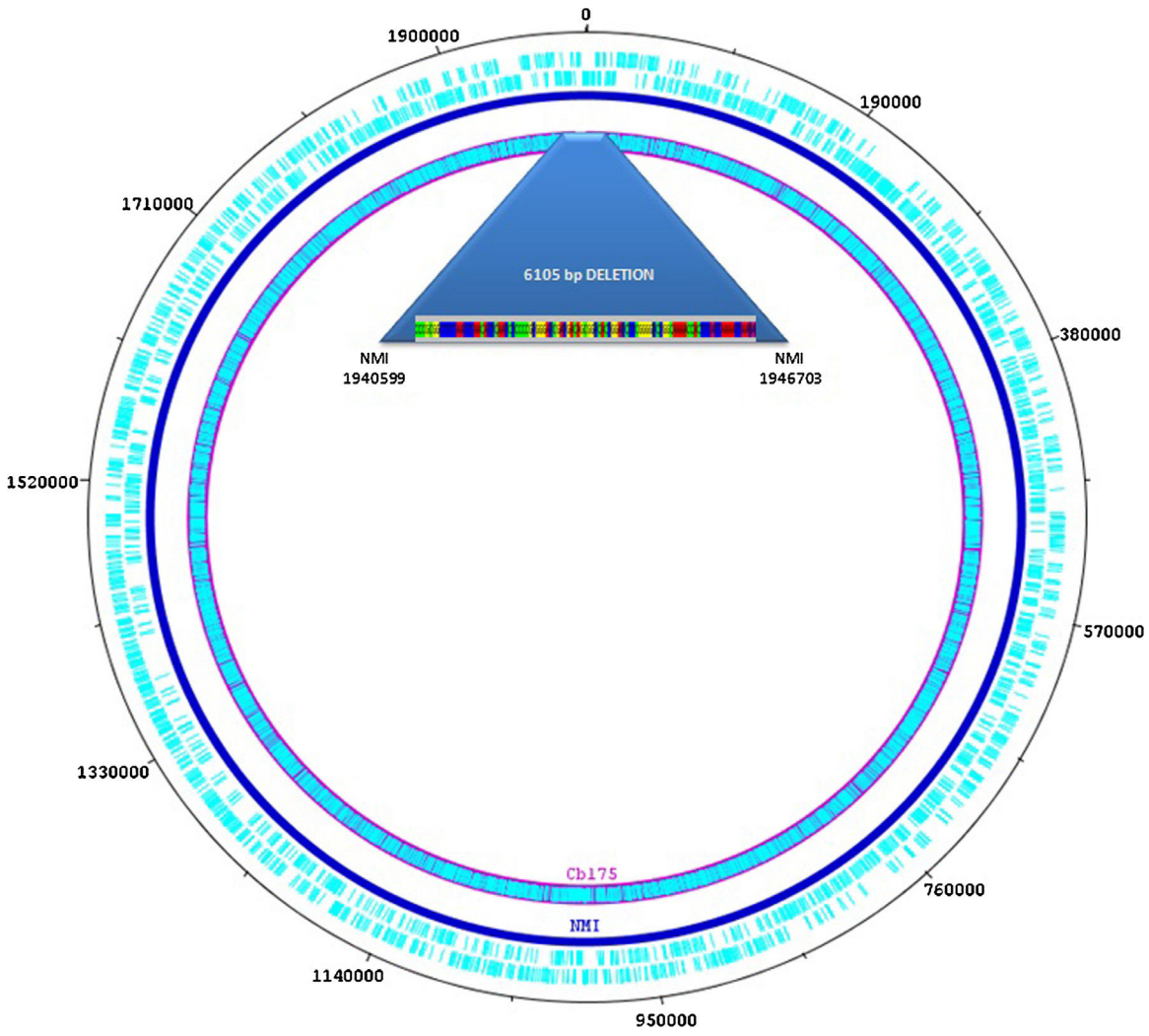


Fig. 3. Circular genomic comparison between Cb175 and NMI. In blue, NMI genome. In fuchsia, the Cb175'genome. On the Cb175 chromosome, the missing region is marked. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

HlyD family secretion protein (strain: Philadelphia 1, sub-species: pneumophila), sequence: NC_020521.1 (1791946.1793082, complement) and LPE509_01687 Legionella secretion system protein B (lssB) (strain: Philadelphia 1, sub-species: pneumophila), sequence: NC_020521.1 (1793086.1795071, complement). LssB (ABC-transporter-ATP binding) and lssD (HlyD family secretion protein) were also missing from the *Legionella longbeachae* compared with the *Legionella pneumophila* genomes [27,28].

4. Discussion

A comparative genomic analysis of strain Cb175, based on the Blast score ratio algorithm with 6 other available complete genomes of *C. burnetii* strains, showed low variability. Additionally, although the number of strains was relatively small and considering that the estimated ratio core genome/pangenome is high (96%), this study suggests that the pangenome of *C. burnetii* is closed [27].

C. burnetii 175 is an epidemic strain that causes a higher prevalence of acute Q fever than almost all other known genotypes of *C. burnetii* [7,10]. Genome analysis demonstrated a total of 397 mutated genes, including 137 that encode metabolic proteins and a trend for an increasing proportion of non-synonymous mutations in genes involved in ankyrin repeat domain proteins, amino acid metabolism, coenzyme transport and membrane or envelope biogenesis. [29]. A similar trend has been observed for strain Z3055 that is genotypically related to the epidemic strain from the outbreak in the Netherlands [17]. However, it is more likely that a big deletion has had phenotypic effects than these mutations.

Additionally, the size of the Cb175 genome (1,989,565 bp) was reduced compared with that of the Nine Mile I strain (1,995,275 bp). During the pregenomic era, the understanding of bacterial virulence was based on studies demonstrating that removal of certain genes from pathogenic species eliminated their capacity to infect hosts. Biased by an anthropocentric perspective, the term "virulence factors" was coined to describe such genes [30]. However, in the comparative genomics era, highly pathogenic bacteria, such as *Mycobacteria* and *Rickettsia*, have been shown to have reduced genome sizes [31–36]. Subsequently, it has been demonstrated that the evolution to hyper pathogenicity can be driven by genome reduction, resulting in the inactivation or deletion of non-virulence genes [37]. A recent study that compared the 12 most dangerous epidemic bacteria with their closest non-epidemic species found that epidemic species are characterized by reduced genomes and are accompanied by a significant reduction in ORF content and a gradual disappearance of genes [29].

By performing PCR and a BLASTN search against the reference *C. burnetii* strain Nine Mile I, we discovered that a 6105 bp region was missing in strain Cb175. To date, this deletion is only and specifically detected in *C. burnetii* strains from French Guiana (8/8 vs 0/298, Fisher's exact test, $p < 0.0000001$). This deleted region

Please cite this article in press as: F. D'Amato, et al., Loss of TSS1 in hypervirulent *Coxiella burnetii* 175, the causative agent of Q fever in French Guiana, Comp Immunol Microbiol Infect Dis (2015), http://dx.doi.org/10.1016/j.cimid.2015.04.003

contains proteins involved in the T1SS. Protein secretion systems play roles as communication ports with eukaryotic cells and are therefore considered to be an integral part of the bacterial virulence arsenal. However, it was recently shown that some of the most dangerous epidemic bacterial species contain significantly fewer secretion system proteins than their closest non-epidemic relatives [29].

The same genes have also been shown to be lost in *L. longbeachae* [27,28]. *Coxiella* is a genus that is closely related to *Legionella* [38]. Therefore, we compared the missing region of *C. burnetii* 175 with that of *L. pneumophila*, the primary human pathogenic bacterium of the genus [39]. In *L. pneumophila*, T1SS genes are located in the same region and T1SS is functional, playing a role in internalization into its host cell [39]. In *C. burnetii* Nine Mile, the HlyB protein was mutated and two pseudogenes remained: HlyA is absent and HlyD is still complete and possibly functional. However, the influence of the deletion that we found on the functionality in *C. burnetii* 175 cannot be determined in the absence of experimental studies on this subject. For *L. longbeachae*, both hypotheses of an alternative T1SS system (using a different secretion machinery) and of a non-functional T1SS can be found in the literature [28]. In our work concerning the highly virulent *C. burnetii* 175 strain, the only observable effect of this deletion is a genome reduction (in contrast to *L. longbeachae*), which is comparable to what can be found in most virulent and epidemic microorganisms.

5. Conclusion

C. burnetii 175, a representative strain of the agent of Q fever in French Guiana, causing 24% of community acquired pneumonia, exhibits genome reduction due to a deletion in T1SS, which is specific to guianan genotype 17 as showed by qPCR results. Indeed, missing region was only detected in 8 other *C. burnetii* strains isolated from this territory and is absent from the 298 other strains that we tested in qPCR. This genome reduction is consistent with previous findings that showed that virulent or epidemic strains do not exhibit more genes or more virulence factors, but instead can be characterized by progressive gene loss. The particularly high prevalence and clinical severity of acute Q fever in French Guiana may be linked to these genomic features.

Competing interests and funding

The authors declare that they have no competing interests. This work was supported by the Mediterranean Infection Foundation. The exact name of the sponsor in French is "Fondation de Coopération Scientifique Méditerrannée Infection" and the national French SIRET number is: 50198088200010.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.cimid.2015.04.003

References

[1] D. Raoult, T. Marrie, J. Mege, Natural history and pathophysiology of Q fever, Lancet Infect Dis 5 (2005) 219–226, http://dx.doi.org/10.1016/S1473-3099(05)70052-9, pii: S1473-3099(05)70052-9.
[2] D. Raoult, Chronic Q fever: expert opinion versus literature analysis and consensus, J Infect 65 (2012) 102–108, http://dx.doi.org/10.1016/j.jinf.2012.04.006, pii: S0163-4453(12)00106-5.
[3] O. Glazunova, V. Roux, O. Freylikman, Z. Sekeyova, G. Fournous, J. Tyczka, et al., *Coxiella burnetii* genotyping, Emerg Infect Dis 11 (2005) 1211–1217, http://dx.doi.org/10.3201/eid1108.041354
[4] D. Frankel, H. Richet, A. Renvoise, D. Raoult, Q fever in France, 1985–2009, Emerg Infect Dis 17 (2011) 350–356, http://dx.doi.org/10.3201/eid1703.100882
[5] M. Maurin, D. Raoult, Q fever, Clin Microbiol Rev 12 (1999) 518–553.
[6] C. Eldin, A. Mahamat, M. Demar, P. Abboud, F. Djossou, D. Raoult, Q fever in French Guiana, Am J Trop Med Hyg 91 (2014) 771–776, http://dx.doi.org/10.4269/ajtmh.14-0282, pii: ajtmh.14-0282.
[7] S. Edouard, A. Mahamat, M. Demar, P. Abboud, F. Djossou, D. Raoult, Comparison between emerging Q fever in French Guiana and endemic Q fever in Marseille, France, Am J Trop Med Hyg (2014), http://dx.doi.org/10.4269/ajtmh.13-0164, pii: ajtmh.13-0164.
[8] F. Pfaff, A. Francois, D. Hommel, I. Jeanne, J. Margery, G. Guillot, et al., Q fever in French Guiana: new trends, Emerg Infect Dis 4 (1998) 131–132, http://dx.doi.org/10.3201/eid0401.980124
[9] J. Gardon, J.M. Heraud, S. Laventure, A. Ladam, P. Capot, E. Fouquet, et al., Suburban transmission of Q fever in French Guiana: evidence of a wild reservoir, J Infect Dis 184 (2001) 278–284, http://dx.doi.org/10.1086/322034, pii: JID010098.
[10] L. Epelboin, C. Chesnais, C. Boulle, A.S. Drogoul, D. Raoult, F. Djossou, et al., Q fever pneumonia in French Guiana: prevalence, risk factors, and prognostic score, Clin Infect Dis 55 (2012) 67–74, http://dx.doi.org/10.1093/cid/cis288, pii: cis288.
[11] A. Mahamat, S. Edouard, M. Demar, P. Abboud, J.Y. Patrice, S.B. La, et al., Unique clone of *Coxiella burnetii* causing severe Q fever, French Guiana, Emerg Infect Dis 19 (2013) 1102–1104, http://dx.doi.org/10.3201/eid1907.130044
[12] B. Davoust, J.L. Marie, V. Pommier dS, J.M. Berenger, S. Edouard, D. Raoult, Three-toed sloth as putative reservoir of *Coxiella burnetii*, cayenne, French Guiana, Emerg Infect Dis 20 (2014) 1760–1761, http://dx.doi.org/10.3201/eid2010.140694
[13] L. Rouli, J.M. Rolain, F.A. El, C. Robert, D. Raoult, Genome sequence of *Coxiella burnetii* 109, a doxycycline-resistant clinical isolate, J Bacteriol 194 (2012) 6939, http://dx.doi.org/10.1128/JB. 01856-12, pii: 194/24/6939.
[14] S.F. Altschul, W. Gish, W. Miller, E.W. Myers, D.J. Lipman, Basic local alignment search tool, J Mol Biol 215 (1990) 403–410, http://dx.doi.org/10.1016/S0022-2836(05)80360-2, pii: S0022-2836(05)80360-2.
[15] P.A. Beare, N. Unsworth, M. Andoh, D.E. Voth, A. Omsland, S.D. Gilk, et al., Comparative genomics reveal extensive transposon-mediated genomic plasticity and diversity among potential effector proteins within the genus Coxiella, Infect Immun 77 (2009) 642–656, http://dx.doi.org/10.1128/IAI. 01141-08, pii: IAI.01141-08.
[16] R. Seshadri, I.T. Paulsen, J.A. Eisen, T.D. Read, K.E. Nelson, W.C. Nelson, et al., Complete genome sequence of the Q-fever pathogen *Coxiella burnetii*, Proc Natl Acad Sci U S A 100 (2003) 5455–5460, http://dx.doi.org/10.1073/pnas.0931379100, pii: 0931379100.
[17] F. D'Amato, L. Rouli, S. Edouard, J. Tyczka, M. Million, C. Robert, et al., The genome of *Coxiella burnetii* Z3055, a clone linked to the Netherlands Q fever outbreaks, provides evidence for the role of drift in the emergence of epidemic clones, Comp Immunol Microbiol Infect Dis (2014), http://dx.doi.org/10.1016/j.cimid.2014.08.003, pii: S0147-9571(14)00042-3.
[18] K. Sidi-Boumedine, R.J. Ellis, G. Adam, M. Prigent, O. Angen, A. Aspan, et al., Draft genome sequences of six ruminant *Coxiella Burnetii* isolates of European origin, Genome Announc (2014), http://dx.doi.org/10.1128/genomeA.00285-14, pii: 00285-14.
[19] J. Besemer, A. Lomsadze, M. Borodovsky, GeneMarkS: a self-training method for prediction of gene starts in microbial genomes: implications for finding sequence motifs in regulatory regions, Nucleic Acids Res 29 (2001) 2607–2618.
[20] R.C. Edgar, Search and clustering orders of magnitude faster than BLAST, Bioinformatics 26 (2010) 2460–2461, http://dx.doi.org/10.1093/bioinformatics/btq461, pii: btq461.
[21] E.M. Gertz, Y.K. Yu, R. Agarwala, A.A. Schaffer, S.F. Altschul, Composition-based statistics and translated nucleotide searches: improving the TBLASTN module of BLAST, BMC Biol 4 (2006) 41, http://dx.doi.org/10.1186/1741-7007-4-41, pii: 1741-7007-4-41.
[22] D.A. Rasko, G.S. Myers, J. Ravel, Visualization of comparative genomic analyses by BLAST score ratio, BMC Bioinformatics 6 (2005) 2, http://dx.doi.org/10.1186/1471-2105-6-2, pii: 1471-2105-6-2.
[23] J.W. Sahl, J.D. Gillece, J.M. Schupp, V.G. Waddell, E.M. Driebe, D.M. Engelthaler, et al., Evolution of a pathogen: a comparative genomics analysis identifies a genetic pathway to pathogenesis in Acinetobacter, PLOS ONE 8 (2013) e54287, http://dx.doi.org/10.1371/journal.pone.0054287, pii: PONE-D-12-31404.
[24] A.I. Saeed, V. Sharov, J. White, J. Li, W. Liang, N. Bhagabati, et al., TM4: a free, open-source system for microarray data management and analysis, Biotechniques 34 (2003) 374–378.
[25] A.I. Saeed, N.K. Bhagabati, J.C. Braisted, W. Liang, V. Sharov, E.A. Howe, et al., TM4 microarray software suite, Methods Enzymol 411 (2006) 134–193, http://dx.doi.org/10.1016/S0076-6879(06)11009-5, pii: S0076-6879(06)11009-5.
[26] S. Wu, Z. Zhu, L. Fu, B. Niu, W. Li, WebMGA: a customizable web server for fast metagenomic sequence analysis, BMC Genomics 12 (2011) 444, http://dx.doi.org/10.1186/1471-2164-12-444, pii: 1471-2164-12-444.
[27] C. Cazalet, L. Gomez-Valero, C. Rusniok, M. Lomma, D. Dervins-Ravault, H.J. Newton, et al., Analysis of the Legionella longbeachae genome and transcriptome uncovers unique strategies to cause Legionnaires' disease, PLoS Genet 6 (2010) e1000851, http://dx.doi.org/10.1371/journal.pgen.1000851
[28] N.A. Kozak, M. Buss, C.E. Lucas, M. Frace, D. Govil, T. Travis, et al., Virulence factors encoded by Legionella longbeachae identified on the basis of the genome sequence analysis of clinical isolate D-4968, J Bacteriol 192 (2010) 1030–1044, http://dx.doi.org/10.1128/JB. 01272-09, pii: JB.01272-09.
[29] K. Georgiades, D. Raoult, Genomes of the most dangerous epidemic bacteria have a virulence repertoire characterized by fewer genes but more

Please cite this article in press as: F. D'Amato, et al., Loss of TSS1 in hypervirulent *Coxiella burnetii* 175, the causative agent of Q fever in French Guiana, Comp Immunol Microbiol Infect Dis (2015), http://dx.doi.org/10.1016/j.cimid.2015.04.003

toxin–antitoxin modules, PLoS ONE 6 (2011) e17962, http://dx.doi.org/10.1371/journal.pone.0017962
[30] U. Dobrindt, B. Hochhut, U. Hentschel, J. Hacker, Genomic islands in pathogenic and environmental microorganisms, Nat Rev Microbiol 2 (2004) 414–424, http://dx.doi.org/10.1038/nrmicro884, pii: nrmicro884.
[31] J. Wixon, Featured organism: reductive evolution in bacteria: *Buchnera* sp., *Rickettsia prowazekii* and *Mycobacterium leprae*, Comp Funct Genomics 2 (2001) 44–48, http://dx.doi.org/10.1002/cfg.70
[32] S.T. Cole, K. Eiglmeier, J. Parkhill, K.D. James, N.R. Thomson, P.R. Wheeler, et al., Massive gene decay in the leprosy bacillus, Nature 409 (2001) 1007–1011, http://dx.doi.org/10.1038/35059006
[33] K.R. Sakharkar, P.K. Dhar, V.T. Chow, Genome reduction in prokaryotic obligatory intracellular parasites of humans: a comparative analysis, Int J Syst Evol Microbiol 54 (2004) 1937–1941, http://dx.doi.org/10.1099/ijs.0.63090-0, pii: 54/6/1937.
[34] J.O. Andersson, S.G. Andersson, Insights into the evolutionary process of genome degradation, Curr Opin Genet Dev 9 (1999) 664–671, pii: S0959-437X(99)00024-6.
[35] V. Merhej, D. Raoult, Rickettsial evolution in the light of comparative genomics, Biol Rev Camb Philos Soc 86 (2011) 379–405, http://dx.doi.org/10.1111/j.1469-185X.2010.00151.x, pii: BRV151.
[36] P.E. Fournier, K.K. El, Q. Leroy, C. Robert, B. Giumelli, P. Renesto, et al., Analysis of the *Rickettsia africae* genome reveals that virulence acquisition in *Rickettsia* species may be explained by genome reduction, BMC Genomics 10 (2009) 166, http://dx.doi.org/10.1186/1471-2164-10-166, pii: 1471-2164-10-166.
[37] V. Merhej, M. Royer-Carenzi, P. Pontarotti, D. Raoult, Massive comparative genomic analysis reveals convergent evolution of specialized bacteria, Biol Direct 4 (2009) 13, http://dx.doi.org/10.1186/1745-6150-4-13, pii: 1745-6150-4-13.
[38] W.G. Weisburg, M.E. Dobson, J.E. Samuel, G.A. Dasch, L.P. Mallavia, O. Baca, et al., Phylogenetic diversity of the Rickettsiae, J Bacteriol 171 (1989) 4202–4206.
[39] R. Takamatsu, H. Teruya, E. Takeshima, C. Ishikawa, K. Matsumoto, N. Mukaida, et al., Molecular characterization of *Legionella pneumophila*-induced interleukin-8 expression in T cells, BMC Microbiol 10 (2010) 1, http://dx.doi.org/10.1186/1471-2180-10-1, pii: 1471-2180-10-1.
[40] E. Angelakis, S. Johani, A. Ahsan, Z. Memish, D. Raoult, Q fever endocarditis and new *Coxiella burnetii* genotype, Saudi Arabia, Emerg Infect Dis 20 (2014) 726–728, http://dx.doi.org/10.3201/eid2004.131603.

Please cite this article in press as: F. D'Amato, et al., Loss of TSS1 in hypervirulent *Coxiella burnetii* 175, the causative agent of Q fever in French Guiana, Comp Immunol Microbiol Infect Dis (2015), http://dx.doi.org/10.1016/j.cimid.2015.04.003

Supplementary Table 1. Non-synonymous mutations. In "position" column, the position of mutation in NMI genome is indicated.

Position	type	NMI	Guiana	gene	aa change
6289	SNV	C	T	gyrB	Thr723Ile
9704	SNV	G	A	dacB	Asp49Asn
10185	SNV	C	A	dacB	Ser209Tyr
10441	SNV	A	C	dacB	Gln294His
15049	SNV	G	A	CBU_0015	Gly165Asp
18808	SNV	C	G	CBU_0021	His158Asp
19322	SNV	T	C	CBU_0021	Val329Ala
24651	SNV	G	A	CBU_0027	Ala25Val
32169	SNV	C	T	CBU_0035	Arg4His
32360	SNV	G	C	CBU_0036	Thr99Ser
37871	SNV	G	T	CBU_0040a	Ala38Glu
39339	SNV	C	G	CBU_0041	Glu294Asp
41579	SNV	C	T	CBU_0043	Ala58Val
46693	SNV	G	A	CBU_0049	Ser427Phe
47396	SNV	T	C	CBU_0049	Lys193Arg
49109	SNV	T	C	CBU_0051	Asn90Ser
50218	SNV	A	G	enhA.1	Ile245Val
50854	SNV	A	G	CBU_0054	Asn184Asp
53180	SNV	A	G	CBU_0056	Asn451Asp
53430	SNV	T	C	CBU_0056	Leu534Pro
53532	SNV	G	A	CBU_0056	Arg568His
66958	SNV	T	C	CBU_0073	Asp589Gly
75375	SNV	A	T	CBU_0083	Ile39Phe
81235	SNV	C	T	CBU_0088a	Ala126Val
94893	SNV	C	T	nadB	Asp413Asn
105854	SNV	A	G	CBU_0113	Val27Ala
107773	SNV	T	C	mraW	Ile187Thr
112286	SNV	A	C	murE	Asp58Ala
124392	SNV	T	G	murC	Asn416Lys
127888	SNV	C	A	ftsA	Gln292Lys
129425	SNV	T	A	ftsZ	Leu390Gln
132893	SNV	A	G	secA	Asp45Gly
138677	SNV	G	A	pilD	Ser253Leu
138732	SNV	A	C	pilD	Ser235Ala
139350	SNV	G	T	pilD	Leu29Ile
142153	Del	C	-	pilA	Gly107fs
143711	Del	T	-	CBU_0159	Asp363fs
147891	SNV	T	C	CB_0162	Ile313Met
155817	SNV	C	T	CBU_0173	Arg97Gln

156818	SNV	T	C	CBU_0175	Leu2Ser
157460	SNV	A	G	CBU_0175	Gln216Arg
162195	SNV	C	T	anmK	Glu41Lys
164047	SNV	A	G	tyrS	Lys116Arg
165134	SNV	C	T	CBU_0181a	Val21Ile
169011	SNV	A	G	CBU_23S Misc. feature: intervening sequence CbIVS-1	Glu81Gly
173015	SNV	C	T	CBU_0184	Cys248Tyr
186346	SNV	G	A	CBU_0198	Pro412Leu
187778	SNV	C	G	coaA	Gln20Glu
189699	SNV	C	T	CBU_0201	Glu115Lys
204255	SNV	G	A	CBU_0220a	Ala175Thr
204700	SNV	T	A	CBU_0220a	Leu323Gln
221268	SNV	A	G	rpsG	Lys190Arg
221924	SNV	A	G	fusA	Lys214Glu
226425	SNV	G	A	rplW	Glu8Lys
229649	SNV	G	T	rpsQ	Gly50Cys
238573	SNV	G	A	CBU_0265	Ala56Thr
247709	SNV	A	C	uvrA	Gln778His
249505	SNV	C	T	hemE	Ser319Phe
255364	SNV	C	A	CBU_0285	Ala85Glu
255414	SNV	A	C	CBU_0285	Lys102Gln
259563	SNV	G	A	CBU_0291b	Pro108Leu
259747	Del	A	-	CBU_0291b	Ser47fs
262037	SNV	T	C	CBU_0295	Thr403Ala
262895	SNV	C	T	CBU_0295	Ala117Thr
264973	SNV	G	T	murI	Gln274Lys
267453	SNV	T	G	CBU_0300	Leu269Val
277051	SNV	A	G	CBU_0311	Thr48Ala
277088	SNV	A	G	CBU_0311	Asn60Ser
279414	SNV	C	A	CBU_0313	Gly488*
280074	SNV	G	T	CBU_0313	Leu268Met
281307	SNV	C	T	CBU_0314	Gly160Asp
284753	SNV	A	C	enhA.2	Leu25Arg
288381	SNV	G	A	icmH	Gln110*
301343	SNV	T	C	thiDE	Val375Ala
302402	SNV	A	C	CBU_0335	*222Tyr
305808	SNV	A	G	pepN	Asn448Asp
308270	SNV	T	C	CBU_0340	Ser29Pro
320450	SNV	G	A	CBU_0354	Met9Ile
329984	SNV	C	T	CBU_0364	Val235Ile
333664	SNV	A	G	CBU_0369	Val9Ala
336657	SNV	G	A	CBU_0372	Ser19Leu

336965	SNV	G	A	CBU_0373	Thr140Met
336989	SNV	T	C	CBU_0373	His132Arg
343245	SNV	C	A	ispB	Ala260Ser
344151	SNV	G	T	tag	Ala22Ser
346500	SNV	C	T	CBU_0384b	Arg30*
350500	SNV	G	T	CBU_0388	Gln639His
351689	SNV	A	G	CBU_0388	Asn1036Asp
354637	SNV	T	G	mviN	Phe452Val
355044	SNV	C	T	ribF	Pro73Ser
361950	SNV	T	C	lspA	Leu100Ser
362558	SNV	C	G	CBU_0398	Ser133Trp
367127	SNV	C	T	CBU_0410	Val282Met
378315	SNV	A	G	CBU_0425	Leu135Pro
386573	SNV	G	A	CBU_0433	Gly94Ser
387244	SNV	A	C	CBU_0434a	Phe36Leu
388455	SNV	C	T	CBU_0436	Met99Ile
396464	SNV	A	G	ffh	Ile158Thr
400168	SNV	A	G	CBU_0455	Asn28Ser
400401	SNV	C	A	CBU_0455	Leu106Met
401187	SNV	C	A	CBU_0458	Pro23Asn
401744	SNV	G	A	rosB	Val82Ile
406851	SNV	G	A	pdhC	Glu172Lys
408501	SNV	A	G	lpdA	Ile281Val
418459	SNV	A	T	CBU_0476	Leu391Phe
422257	SNV	A	G	CBU_0482	Thr157Ala
423369	SNV	T	C	artM	Val37Ala
423371	SNV	T	C	artM	Phe38Leu
424407	SNV	G	A	CBU_0485	Gly155Glu
426722	SNV	C	T	CBU_0486	Arg354His
431521	SNV	C	T	CBU_0490	Asp407Asn
432426	SNV	C	T	CBU_0490	Arg105His
439202	SNV	T	G	CBU_0498	Leu215Val
444906	SNV	C	T	CBU_0504	Ser47Asn
462656	SNV	C	T	CBU_0519	Pro199Ser
463026	SNV	T	C	leuA	Ile192Val
464460	SNV	G	C	CBU_0521	Ala194Gly
469318	SNV	G	A	gyrA	Val722Ile
472250	SNV	T	G	cmk	Phe26Leu
475949	SNV	G	A	CBU_0530	Asp275Asn
479244	SNV	T	C	CBU_0534	Ile267Val
479708	SNV	C	T	CBU_0534	Gly112Asp
490861	SNV	T	C	CBU_0543	Val231Ala

493459	SNV	A	T	CBU_0544	Lys64*
493669	Ins	-	G	lemA	Gln20fs
507040	SNV	T	C	leuS	Ile532Val
509261	SNV	G	A	CBU_0560	Met163Ile
530177	SNV	C	T	CBU_0577	Arg193Trp
534955	SNV	A	G	CBU_0584	Asp17Gly
545537	SNV	T	A	CBU_0597	Leu89Phe
559519	SNV	C	T	ompH	Thr98Met
564539	SNV	C	G	CBU_0618	Met246Ile
582597	SNV	G	C	CBU_0634	Met407Ile
584107	SNV	C	A	CBU_0635	Gln274Lys
584924	SNV	T	C	CBU_0636	Val35Ala
590787	SNV	G	A	CBU_0641	Pro233Leu
594346	SNV	A	G	ribE	Asp25Gly
596000	SNV	T	G	ribA	Val355Gly
610018	SNV	C	A	CBU_0668	Arg41Leu
611040	SNV	C	T	rhlE	Arg160Trp
615744	SNV	T	A	CBU_0673	Thr27Ser
641957	SNV	C	A	CBU_0694	Ser18Ile
642408	SNV	A	G	CBU_0695	Val347Ala
643407	SNV	G	T	CBU_0695	Ala14Glu
644170	SNV	G	A	CBU_0696	Ala142Val
646795	SNV	G	T	CBU_0698	Asn61Lys
647282	SNV	T	G	CBU_0699	Leu35Val
648375	SNV	A	G	CBU_0699	Tyr399Cys
648843	SNV	A	G	CBU_0700	Phe564Leu
650854	SNV	G	A	cysQ-2	Ala165Val
651760	SNV	G	A	CBU_0702	Asp93Asn
655195	SNV	C	T	CBU_0706	Glu156Lys
674120	SNV	T	C	CBU_0728	Val263Ala
678315	SNV	C	T	CBU_0736	Gly208Asp
678741	SNV	G	A	CBU_0736	Pro66Leu
689329	SNV	T	C	CBU_0748	Thr69Ala
694715	SNV	T	C	CBU_0753	Met692Val
696424	SNV	T	G	CBU_0753	Asn122Thr
697519	SNV	C	T	CBU_0754	Asp122Asn
699955	SNV	G	A	CBU_0756	Pro28Ser
700274	SNV	C	T	rluD	Gly263Arg
703532	Ins	-	T	gacS	Pro6fs
712544	SNV	A	G	CBU_0768	Glu140Gly
721718	Ins	-	A	CBU_0777c	Val82fs
723893	SNV	G	T	CBU_0781	Arg338Leu
723895	SNV	T	C	CBU_0781	*339Arg

729912	SNV	A	C	CBU_0788	Gln59His
733633	SNV	A	G	CBU_0788	Lys1300Glu
734386	SNV	G	A	CBU_0788	Asp1551Asn
745172	SNV	C	T	folE	Arg30Lys
746421	SNV	A	G	CBU_0797	Ser260Pro
749450	SNV	T	C	CBU_0802	Lys189Glu
762003	SNV	T	C	prfC	Tyr242His
764530	SNV	A	G	CBU_0813	Lys5Glu
764749	SNV	G	T	CBU_0813	Glu78*
767721	SNV	T	C	aacA4	Lys42Glu
767772	SNV	A	G	aacA4	Tyr25His
768325	SNV	G	A	CBU_0818	Val90Ile
768601	SNV	A	G	CBU_0818	Ile182Val
770909	Del	A	-	CBU_0822	Leu356fs
784770	SNV	A	C	CBU_0833	Lys217Asn
785378	SNV	G	T	CBU_0833	Cys420Phe
787099	SNV	T	C	CBU_0835	Val56Ala
793060	SNV	A	G	asnB-2	Ser156Gly
793900	SNV	G	A	asnB-2	Val436Met
794911	SNV	T	C	CBU_0841	Ser133Pro
796642	SNV	T	C	wecB	Val255Ala
797110	SNV	C	T	CBU_0843	His42Tyr
805538	SNV	T	C	CBU_0850	Leu103Ser
813034	SNV	G	A	CBU_0856	Arg415Gln
814329	SNV	C	T	lpxK	Ala250Val
820121	SNV	T	C	CBU_0866	Ser106Pro
826642	SNV	C	T	aroC	Arg54Cys
842219	SNV	G	C	CBU_0890	Gln150Glu
847079	SNV	C	A	dedD	Asn105Lys
849361	SNV	G	A	purF	Gly450Asp
862515	SNV	T	C	CBU_0913	Arg2Gly
867986	SNV	G	A	CBU_0918	Asp373Asn
868163	SNV	G	A	CBU_0918	Val432Ile
873788	SNV	A	G	CBU_0923	Glu246Gly
876157	SNV	A	G	mmsB	Val229Ala
880017	SNV	T	C	CBU_0930	Asn164Ser
880248	SNV	C	T	CBU_0930	Gly87Glu
881159	SNV	T	C	glpD	Ile279Val
883599	SNV	C	T	gplK, CBU_0933	[Glu250Lys]
884108	SNV	G	A	CBU_0933	Ala80Val
886303	SNV	A	G	CBU_0936	Val39Ala
894521	SNV	G	A	CBU_0945	His37Tyr

895176	SNV	C	A	rhuM	Ala189Ser
897595	SNV	C	G	CBU_0948	Pro105Arg
897662	SNV	T	G	CBU_0948	Phe127Leu
902684	Del	GG	-	CBU_0952	Pro184fs
903402	Del	A	-	CBU_0953	Leu416fs
904528	SNV	T	C	CBU_0953	Ile41Val
904881	SNV	C	G	CBU_0954	Gly184Ala
905953	SNV	C	G	gacA.3	Val51Leu
913484	SNV	A	G	CBU_0959	Ser16Pro
922857	SNV	T	C	CBU_0972	Asn181Asp
922904	SNV	G	A	CBU_0972	Thr165Met
928151	SNV	A	G	CBU_0976	Lys204Glu
930429	Del	G	-	CBU_0978	Arg27fs
930870	SNV	A	G	CBU_0979	Phe94Ser
932026	SNV	C	T	CBU_0982	Leu69Phe
932500	SNV	G	A	CBU_0982	Asp227Asn
934400	SNV	A	C	CBU_0984	Lys217Gln
943054	SNV	A	G	ctaB	Gln279Arg
947720	SNV	C	G	lolC	Val68Leu
956471	SNV	G	C	CBU_1011	Arg119Gly
962755	SNV	A	G	CBU_1017	Lys174Glu
972233	SNV	A	G	CBU_1028	Trp232Arg
972595	SNV	C	T	CBU_1028	Gly111Asp
972861	SNV	A	C	CBU_1028	Phe22Leu
984290	SNV	C	A	cyoA	Leu8Phe
1000831	SNV	A	C	mutS	Lys428Gln
1004134	SNV	A	G	CBU_1058	Glu406Gly
1007863	SNV	C	T	CBU_1063	His65Tyr
1008216	Del	GT	-	CBU_1063	Lys182fs
1008219	Del	A	-	CBU_1063	Ser183fs
1009931	SNV	G	T	CBU_1065	Arg7Ser
1011950	SNV	T	C	ttcA	Asp202Gly
1014363	SNV	A	G	CBU_1067a	Met582Val
1016695	SNV	A	G	CBU_1071	Tyr138His
1017082	SNV	T	G	CBU_1071	Lys9Gln
1020423	SNV	A	G	CBU_1075	Phe79Ser
1020526	SNV	C	T	CBU_1075	Val45Ile
1022897	SNV	G	A	nagZ	Thr118Met
1024092	SNV	T	C	CBU_1079	Glu35Gly
1024280	SNV	C	T	CBU_1080	Gly159Glu
1026850	SNV	G	T	mutL	Leu389Met
1029390	SNV	T	C	CBU_1084	His246Arg
1031252	SNV	G	C	CBU_1085	His44Asp

1031348	SNV	T	C	CBU_1085	Ile12Val
1032928	SNV	A	G	CBU_1087	Val21Ala
1038484	SNV	C	T	CBU_1092	Arg4Trp
1044133	SNV	G	C	fumC	Glu151Gln
1048478	SNV	A	G	CBU_1100a	Asn8Asp
1050127	Del	AT	-	CBU_1103	Asp310fs
1050742	SNV	G	C	CBU_1103	Ile105Met
1050982	SNV	C	T	CBU_1103	Trp25*
1063573	SNV	T	C	CBU_1119	Asp122Gly
1066761	SNV	C	T	cbpA	Pro20Leu
1070913	SNV	G	T	nifS	His269Asn
1074035	SNV	T	C	CBU_1131	Ile154Met
1074220	SNV	G	C	CBU_1131	Pro93Ala
1075877	SNV	C	T	CBU_1134	Trp47*
1081001	SNV	G	A	enhA.4	Thr33Met
1084973	SNV	A	G	secD	Leu308Pro
1092507	SNV	G	A	mfd	Gly357Glu
1095793	SNV	A	G	CBU_1151	Tyr235His
1100457	SNV	A	G	trpBF	Asp11Gly
1101479	Del	C	-	trpBF	Pro352fs
1116436	SNV	T	C	CBU_1175	*438Trp
1116444	SNV	T	G	CBU_1175	Asn436His
1116445	SNV	G	C	CBU_1175	Asn435Lys
1116449	SNV	T	A	CBU_1175	Gln434Leu
1116503	SNV	G	A	CBU_1175	Ser416Phe
1116722	Del	A	-	CBU_1175	Leu343fs
1117220	SNV	C	G	CBU_1175	Gly177Ala
1119531	SNV	G	A	CBU_1177	Pro32Leu
1120054	SNV	G	A	CBU_1178	Glu15Lys
1123595	SNV	G	A	thiI	Thr53Ile
1127524	SNV	C	T	uvrC	Val220Ile
1132551	SNV	T	C	CBU_1189	Ile410Val
1143800	SNV	G	C	icd	Pro149Ala
1145079	SNV	T	C	queA	Lys81Glu
1146144	SNV	C	T	CBU_1202	Leu237Phe
1146375	SNV	C	T	CBU_1203	Arg9Cys
1146397	SNV	T	C	CBU_1203	Phe16Ser
1146748	SNV	C	T	CBU_1203	Ala133Val
1158971	Ins	-	C	CBU_1213	Ser635fs
1165404	SNV	G	T	CBU_1219	Tyr171*
1165679	SNV	T	G	CBU_1219	Lys80Gln
1167091	SNV	T	C	CBU_1221	Lys48Glu
1179615	SNV	G	A	CBU_1229	Val89Ile

1186016	Del	T	-	CBU_1231	Arg33fs
1186414	SNV	A	C	CBU_1233	Gln8His
1191785	SNV	G	A	gcp	Ala163Val
1191915	SNV	G	C	gcp	Leu120Val
1192607	SNV	A	G	mdh	Ser221Pro
1201315	SNV	A	T	CBU_1249	Val170Glu
1205529	SNV	A	G	CBU_1253b	Lys192Glu
1211257	SNV	T	C	CBU_1260	Val165Ala
1214660	SNV	C	T	lipB	Thr174Met
1216538	SNV	A	G	CBU_1267	His228Arg
1218388	SNV	G	A	CBU_1268	Ser261Leu
1226972	SNV	C	T	CBU_1276	Ala22Thr
1227650	SNV	G	A	eda	Gly185Arg
1237653	SNV	C	T	CBU_1284	Ala347Val
1237942	Del	A	-	CBU_1285	Lys8fs
1240257	SNV	G	C	CBU_1286	Arg303Thr
1243088	SNV	G	A	dnaJ	Ala230Val
1251517	SNV	G	A	recN	Arg350Gln
1255443	SNV	T	C	omlA	Ile73Thr
1257602	SNV	C	A	msrA	Thr271Lys
1260800	Ins	-	T	CBU_1310	Gly14fs
1265475	SNV	G	A	CBU_1316	Arg46Cys
1278420	SNV	C	A	CBU_1332a	[Pro16His]
1278505	SNV	A	G	CBU_1333	Ile15Val
1281633	SNV	T	C	CBU_1335	Ile264Thr
1285635	SNV	T	C	dnaE	Asp270Gly
1287235	SNV	T	C	ddl	Gln111Arg
1288512	SNV	G	C	CBU_1340	Arg40Gly
1291412	SNV	G	T	guaB	Asp95Glu
1295242	SNV	T	C	CBU_1346	Ser112Pro
1296606	SNV	C	T	CBU_1347	Val341Ile
1298772	SNV	C	G	CBU_1349	Phe230Leu
1299499	SNV	G	A	CBU_1349	Ala473Thr
1319890	SNV	T	C	CBU_1370	Lys196Glu
1320222	SNV	C	A	CBU_1370	Cys85Phe
1322024	SNV	T	C	CBU_1372	Glu17Gly
1323911	SNV	G	T	CBU_1374	Trp150Leu
1325003	SNV	C	T	relA	Glu373Lys
1326088	SNV	G	A	relA	Thr11Ile
1326494	SNV	G	A	CBU_1376	Val100Ile
1326869	SNV	A	G	CBU_1376	Ile225Val
1341613	SNV	G	T	CBU_1388d	Val7Phe
1347460	Del	C	-	CBU_1395	Arg140fs

1347699	SNV	G	C	CBU_1395	Phe60Leu
1357478	Del	G	-	sdhC	Thr19fs
1360945	SNV	G	A	gltA	Arg69Gln
1362783	SNV	C	T	CBU_1412	Gly108Glu
1362798	SNV	C	A	CBU_1412	Arg103Leu
1365843	SNV	C	T	thiL	Trp3*
1370964	SNV	G	C	radA	Ala194Gly
1372337	SNV	C	A	hemB	Met195Ile
1373138	SNV	G	A	CBU_1425	Thr85Met
1387939	SNV	C	T	nuoN	Ala186Thr
1398697	SNV	A	G	nuoE	Val13Ala
1401100	SNV	C	T	nuoB	Arg20His
1404461	SNV	A	C	CBU_1451	Ser274Arg
1410859	Del	T	-	CBU_1457	Asn390fs
1411726	SNV	G	A	CBU_1457	Ser101Leu
1414214	SNV	A	T	CBU_1460	Tyr246Asn
1418378	SNV	C	T	CBU_1463	Ala39Thr
1422385	SNV	C	T	CBU_1468	Ala816Thr
1432119	SNV	C	T	oxyR	Ala16Thr
1441871	SNV	T	C	gltX-2	Thr176Ala
1452355	SNV	G	A	recO	Pro223Leu
1467118	Del	C	-	CBU_1514	Thr172fs
1469990	SNV	T	C	secB	Thr91Ala
1470948	SNV	G	A	CBU_1521	Pro8Ser
1479456	SNV	C	A	CBU_1530	Asp556Tyr
1480187	Del	C	-	CBU_1530	Gly312fs
1480320	SNV	C	T	CBU_1530	Ala268Thr
1480487	SNV	C	T	CBU_1530	Gly212Glu
1487250	SNV	A	G	CBU_1537	Thr244Ala
1495493	SNV	T	C	lgt	His258Arg
1501809	SNV	G	A	nrdA	Ser728Asn
1516930	SNV	A	G	CBU_1569	Lys276Glu
1522001	SNV	A	G	tolQ	Lys72Glu
1522027	SNV	G	C	tolQ	Met80Ile
1523177	SNV	T	C	CBU_1576	Phe38Ser
1526559	SNV	T	C	CBU_1578	Ile259Val
1528937	SNV	A	C	CBU_1580	Ile302Leu
1530224	SNV	C	G	shaF	Val53Leu
1530785	SNV	T	C	shaD	Lys479Arg
1531746	SNV	C	T	shaD	Ala159Thr
1558085	SNV	A	C	CBU_1618	Asn17Thr
1558340	SNV	T	C	CBU_1618	Ile102Thr
1559313	SNV	C	T	CBU_1621	Arg42Trp

1566123	SNV	G	A	icmE	Ala639Val
1566280	SNV	C	T	icmE	Ala587Thr
1568414	SNV	C	G	icmK	Glu223Gln
1571010	SNV	A	G	icmN	Ile3Thr
1577648	SNV	G	A	CBU_1638	Gln285*
1586488	SNV	C	A	CBU_1647	Gly11Val
1586826	SNV	T	C	dotaA	Ile731Met
1588240	SNV	G	A	dotaA	Thr260Ile
1605863	SNV	A	G	CBU_1670	Tyr124His
1618590	SNV	A	G	CBU_1686	Ile290Val
1631387	Del	C	-	CBU_1699	Thr491fs
1634261	SNV	G	T	CBU_1699b	Ala46Glu
1635406	SNV	G	C	CBU_1701	Ser298Thr
1640880	SNV	G	A	CBU_1710	Val130Ile
1641474	SNV	G	A	CBU_1712	Ala61Thr
1649433	SNV	A	G	groES	Ser76Pro
1655720	SNV	G	A	dsbD	Val211Ile
1657298	SNV	T	C	CBU_1724	Glu646Gly
1659208	Del	T	-	CBU_1724	Lys9fs
1659735	Ins	-	C	CBU_1724a	Val27fs
1663771	SNV	G	T	hemF	Ala176Glu
1667516	SNV	A	G	CBU_1732	Ser311Gly
1672547	SNV	A	G	purM	Glu255Gly
1673928	SNV	A	G	hipB, CBU_1739	[Leu280Pro]
1676581	SNV	G	A	CBU_1742	Ala105Thr
1683390	SNV	C	A	CBU_1752	Leu114Ile
1683609	SNV	A	C	CBU_1752	Ile187Leu
1692582	SNV	C	A	CBU_1761	Trp167Cys
1695336	SNV	C	G	CBU_1763	Leu221Val
1696385	SNV	G	C	CBU_1764a	Lys82Asn
1698561	SNV	G	A	feoB	Leu244Phe
1704866	SNV	T	C	engB	Lys5Glu
1711232	SNV	T	G	pyk	Ser11Arg
1722316	Del	G	-	CBU_1790	Ser30fs
1724077	SNV	A	G	CBU_1790	Glu617Gly
1726727	SNV	T	G	polI	Lys759Gln
1734619	SNV	A	G	CBU_1800	Phe59Ile
1734621	SNV	A	T	CBU_1800	Phe59Ile
1736660	SNV	A	G	CBU_1805	Ile98Met
1740305	SNV	T	C	macA	Thr323Ala
1740361	SNV	T	C	macA	Gln304Arg
1740613	SNV	T	C	macA	Gln220Arg
1747085	SNV	C	T	efp	Arg10Cys

1752352	SNV	T	C	CBU_1821	Thr46Ala
1752650	SNV	C	T	sodC	Glu182Lys
1758844	SNV	T	C	lolB	Ser199Pro
1768546	SNV	T	G	CBU_1839	Gln350Pro
1769489	SNV	A	G	CBU_1839	Trp36Arg
1779941	SNV	A	G	CBU_1851	Lys593Arg
1784613	SNV	G	A	oppD	Pro367Ser
1784657	SNV	C	T	oppD	Arg352Lys
1785312	SNV	C	A	oppD	Glu134*
1786611	SNV	G	A	oppC	Leu206Phe
1791564	SNV	C	G	CBU_1863	[Gln342Glu]
1795570	SNV	G	T	parC	Asp329Glu
1797502	SNV	T	C	CBU_1869	Val124Ala
1804551	SNV	C	A	CBU_1877	Gln338His
1808422	SNV	A	G	hemL	Lys314Arg
1808763	SNV	G	T	hemL	Glu428*
1809415	SNV	A	G	CBU_1884	Thr25Ala
1809602	SNV	A	C	CBU_1884	Gln87Pro
1810236	SNV	T	C	ponA	Ser785Gly
1817354	SNV	C	G	CBU_1894a	Met9Ile
1822159	SNV	C	A	CBU_1901	Ser277Ile
1832538	SNV	A	G	glyS	Lys273Glu
1836665	SNV	C	T	CBU_1919	Thr30Met
1844270	SNV	A	G	parB	Thr120Ala
1846257	SNV	G	A	CBU_1928	Arg488His
1846624	SNV	A	T	CBU_1928	Gln610His
1846790	SNV	G	A	CBU_1928	Gly666Ser
1852966	SNV	A	G	CBU_1934	Glu700Gly
1853417	SNV	A	G	CBU_1934a	Thr14Ala
1854460	SNV	T	C	CBU_1936	Asn177Asp
1859852	SNV	A	G	atpA	Asn187Ser
1861280	SNV	G	A	atpG	Met114Ile
1865587	SNV	T	C	CBU_1947a	Ser30Pro
1865653	SNV	G	A	CBU_1947a	Gly52Arg
1865674	Ins	-	G	CBU_1947a	Arg59fs
1867649	SNV	T	C	CBU_1950	Met18Val
1870926	Ins	-	T	CBU_1954	Tyr85fs
1879809	SNV	C	T	hemK	Arg177Gln
1882534	SNV	G	A	hemA	Ala44Val
1883867	SNV	C	T	CBU_1967	Ala376Val
1887978	SNV	C	A	CBU_1977	Met342Ile
1902087	SNV	T	C	CBU_1995	Val13Ala
1902186	SNV	A	G	CBU_1995	His46Arg

1904146	Del	C	-	CBU_1998	Ile45fs
1904580	SNV	T	C	CBU_1998	Ile190Thr
1909480	SNV	G	A	CBU_2003a	Gly18Asp
1910455	SNV	C	G	CBU_2004	Asp83His
1911254	SNV	T	C	rstB	Asn245Asp
1913408	SNV	T	A	CBU_2007	Leu183Phe
1913413	SNV	A	C	CBU_2007	Phe182Val
1924675	SNV	T	C	CBU_2018	Leu21Pro
1925382	SNV	G	C	ubiB	Gly53Ala
1928907	SNV	G	T	CBU_2021	Thr379Lys
1929784	SNV	G	C	CBU_2021	His87Asp
1930023	SNV	T	C	CBU_2021	Tyr7Cys
1932269	SNV	C	T	CBU_2024	Leu144Phe
1933233	SNV	G	A	metC	Gly122Asp
1937273	SNV	A	G	metK	Asn133Asp
1939763	SNV	A	G	CBU_2032	p.Thr121Ala
1939887	SNV	T	C	CBU_2032	Val162Ala
1946697	SNV	C	G	CBU_2040	Ala299Gly
1946704	Ins	-	C	CBU_2040	Ile301fs
1947126	SNV	G	A	CBU_2040	Arg442Lys
1958227	SNV	T	C	CBU_2052	Asn186Ser
1962423	SNV	C	G	CBU_2056	Ala109Pro
1963882	SNV	G	A	CBU_2057	Ala11Val
1971520	SNV	G	C	CBU_2067	Pro159Ala
1974119	SNV	T	C	CBU_2070	Ile105Val
1976797	SNV	G	A	hemC	Ser57Asn
1983078	SNV	G	A	hemY	Met250Ile
1988532	SNV	A	G	CBU_2088	Asn144Ser
1992644	SNV	C	G	pckA	Met178Ile

CHAPITRE 3: Géotypage et séquençage des souches responsables de la fièvre Q au Canada et en Arabie Saoudite

Avant-propos

Les deux articles suivants, illustrent la contribution de la génomique à l'étude de l'épidémiologie de la fièvre Q et au concept de « géotyping » ou géotypage. Le géotypage peut être considéré un outil qui nous permet d'établir la distribution géographique des différents clones de *C .burnetii*. Cette méthode a permis d'identifier les génotypes globalement repartis parmi les cinq continents et les génotypes très spécifiques à une zone géographique donnée. Ceci est le cas de MST 17 en Guyane française et MST 33 en Hollande (discuté précédemment dans le texte). Également un seul génotype MST21 a été identifié au Canada responsable de tous les événements épidémiques décrits dans ce pays et un nouveau et unique génotype MST 51 a été détectée en Arabie Saoudite. Nous avons analysé deux génomes spécifiques, la souche DOG UTAD et la souche Cb196.

Le séquençage du génome de la souche DOG UTAD suivi par le « mapping » contre le génome de référence NMI, a permis d'obtenir un « draft » de la séquence que nous avons analysé, dans l'objectif d'en connaître les caractéristiques principales. La souche a été responsable d'une épidémie humaine familiale au Canada pendant la mise bas d'une chienne. Ce travail est basé sur l'isolement d'une souche de *C.burnetii* et sur le séquençage complet du génome et fournit des informations importantes sur l'évolution du génotype MST21, le seul détecté au Canada et causant des endocardites. Cette étude a également été utile pour avoir démontré le risque élevé de zoonose associé à l'accouchement d'animaux infectés, étant donné que les 3 individus exposés ont tous développé une pneumonie aiguë à Coxiella. En outre, la forte similitude avec la souche canadienne Q212 de même génotype, la taille du chromosome, l'absence de plasmide et seulement quelques mutations, a suggéré une évolution clonale récente. Ce travail a été publié dans le journal ***New Microbes and New Infections***.

Dans un autre travail nous avons séquencé et assemblé le génome de la souche Cb196 isolée en Arabie Saoudite d'une valve cardiaque d'un patient atteint d'une endocardite à *C.burnetii*.Le génotype de cette souche, MST51, n'a jamais été décrit ailleurs et il est le seul qui circule dans ce pays. Nous avons analysé ce « draft » génome, et nous avons trouvé des similitudes avec la souche Q154 ayant le génotype MST8, plus proche de MST51 au niveau phylogénétique que les génotypes des autres souches disponibles. Ce travail a été publié dans le journal ***Genome Announcements***.

Draft Genome Sequence of *Coxiella burnetii* Dog Utad, a Strain Isolated from a Dog-Related Outbreak of Q Fever

Felicetta D'AMATO[1], Matthieu MILLION[1], Sophie EDOUARD[1], Laetitia ROULI[1], Jeremy DELERCE[1], Nathalie DUCLOS[1], Catherine ROBERT[1], Tom MARRIE[2], Didier RAOULT[1]

1. Aix-Marseille Université, Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes, UM63, CNRS 7278, IRD 198, INSERM 1095, Marseille, France
2. Faculty of Medicine, Dalhousie University, 5849 University Avenue, Halifax, NS, Canada

*Corresponding author

Draft genome sequence of *Coxiella burnetii* Dog Utad, a strain isolated from a dog-related outbreak of Q fever

F. D'amato[1], M. Million[1], S. Edouard[1], J. Delerce[1], C. Robert[1], T. Marrie[2] and D. Raoult[1]

1) *Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes, UM63, CNRS 7278, IRD 198, INSERM 1095, Aix-Marseille Université, Marseille, France and* 2) *Faculty of Medicine, Dalhousie University, Halifax, Nova Scotia, Canada*

Abstract

Coxiella burnetii Dog Utad, with a 2 008 938 bp genome is a strain isolated from a parturient dog responsible for a human familial outbreak of acute Q fever in Nova Scotia, Canada. Its genotype, determined by multispacer typing, is 21; the only one found in Canada that includes Q212, which causes endocarditis. Only 107 single nucleotide polymorphisms and 16 INDELs differed from Q212, suggesting a recent clonal radiation.

Original Submission: 30 April 2014; **Accepted:** 19 May 2014
Article published online: 26 June 2014
New Microbe and New Infect 2014; **2:** 136–137

Corresponding author: D. Raoult, Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes, UM63, CNRS 7278, IRD 198, INSERM 1095, Aix-Marseille Université, Marseille, France
E-mail: Didier.raoult@gmail.com

Coxiella burnetii is a Gram-negative bacterium with a complex intracellular cycle belonging to the γ-proteobacteria [1]. To date, 13 genomes are available from the NCBI [2]. Our strain, Dog Utad, was isolated in Marseille from a dog responsible for a Q fever familial outbreak in Nova Scotia, Canada [3]. This female dog was a hound that had caught rabbits during its pregnancy and gave birth to four pups. All of them died within 24 h after birth. All three family members (mother, father and son) who were present during the delivery, developed pneumonia in the 2 weeks following parturition. Index case was the mother, who helped with the delivery and cleaned up afterward. Serological data confirmed acute Q fever in the mother and the son and *C. burnetii* was isolated by shell vial technique [4] from the dog uterus, which was removed 70 days after parturition, frozen and shipped to our laboratory [5].

Genotyping was performed on the strain using multi-spacer sequence typing (MST) [6], a technique based on the variability of ten intergenic sequences. The genotype MST 21 was identified and confirmed by BLAST of the genome [7] against Cox sequences (10). We compared the Dog Utad strain to the seven available strains. A comparison of the COG categories showed that Dog Utad follows the same trend as the other available *C. burnetii* genomes, but with more similarities with CbuG_Q212, a genome previously deposited in GenBank corresponding to a strain from a Canadian man presenting with Q fever endocarditis and having the same genotype (MST21).

Coxiella burnetii Dog Utad reads best covered (98%) the *Coxiella burnetii* CbuG_Q212 genome with a maximum coverage of 1482 and an average coverage of 173. Moreover, the final Dog Utad sequence was the same size as CbuG_Q212 (2 Mb) and there were only 80 gaps (for a total of 52 047 bp). The genome is characterized by a consensus sequence of 2 008 938 bp (G+C content 44%). It encodes 1896 proteins and carries 44 tRNA and one ribosomal operon.

To the best of our knowledge after analysing 335 strains with available MST from around the world, MST21 was the only genotype identified in Canada. Two other strains from humans presenting with Q fever endocarditis, two from cats and one other from a dog from Canada were also identified as MST21. The uniformity of geographical and genotypic criteria allows us to define the Canadian geotype, which corresponds to the genotype MST21, infects cats, dogs and humans and is responsible for acute Q fever and endocarditis in the Canadian population. It was also found in two French patients [6] and in Alberta, Canada.

We found only 123 mutations (70 of them in putative open reading frames (ORF)) when compared with CbuG_Q212: 107 single nucleotide polymorphisms (SNPs; 67 in ORF), eight insertions (one ORF) and eight deletions (two in ORF). Forty-seven of the 70 are non-synonymous mutations, corresponding to 44 mutated genes. These genes encode for 17 proteins involved in metabolism, five transporters, three membrane proteins, three proteins of signal transduction, three translation proteins, two transcription proteins, two chaperone proteins, three hypothetical proteins, one type IV secretion system protein, one DNA replication protein, one protein involved in cellular processes, one organic solvent tolerance protein, one isomerase and one stress protein (Table 1). This very low number of SNPs suggests a very short genetic distance between these two genomes, suggesting a recent clonal radiation of *C. burnetii* MST21 in Canada.

TABLE 1. Non-synonymous point mutations in *Coxiella burnetii* Dog Utad compared with *Coxiella burnetii* Q212

Position[a]	Type	Gene ID	Annotation	Nucleotide change (Q212→Dog Utad)	Amino acid change (Q212→Dog Utad)
26219	SNP	CbuG_0030	Hypothetical protein	C→A	Ala→Glu
57338	SNP	CbuG_0066	Thioredoxin peroxidase	A→C	Ser→Ala
77850	SNP	CbuG_0091	Transporter-sodium dependent	A→C	Thr→Pro
100854	SNP	CbuG_0107	Biotin carboxylase	G→C	Gly→Ala
329943	SNP	CbuG_0357	Ribosomal protein α-L glutamate ligase	G→A	Gly→Glu
366913	SNP	CbuG_0400	IcmJ	G→A	Glu→Lys
366926	SNP	CbuG_0400	IcmJ	A→G	Glu→Gly
393251	SNP	CbuG_0426	Sodium/proton antiporter protein	T→C	DEL20aa
503017	SNP	CbuG_0540	Exported membrane spanning protein	T→A	Phe→Tyr
504897	SNP	CbuG_0540	Exported membrane spanning protein	A→G	Thr→Ala
543181	SNP	CbuG_0577	Bacterial protein translation initiation factor 2	G→A	Val→Met
610658	SNP	CbuG_0645	Outer membrane lipoprotein	C→T	Met→Leu
719445	SNP	CbuG_0753	Na^+/H^+ antiporter	T→C	Gln→Arg
736723	SNP	CbuG_0769	Malate dehydrogenase	G→A	Arg→Gln
763310	SNP	CbuG_0791	Phosphoglycolate phosphatase	T→C	Phe→Ser
803836	SNP	CbuG_0828	Glycine-rich RNA-binding protein	T→C	Met→Leu
836313	SNP	CbuG_0860	Transcription-repair coupling factor	G→T	Ser→Tyr
894420	SNP	CbuG_0919	Two component system histidine kinase	T→G	Ile→Ser
914733	SNP	CbuG_0941	ABC transporter	T→G	Leu→Phe
928574	SNP	CbuG_0953	Aspartokinase	C→T	Pro→Leu
974371	SNP	CbuG_1002	Carboxylesterase	A→T	Glu→Asp
1006992	SNP	CbuG_1039	Cytochrome D ubiquinol oxidase subunit I	T→C	Ser→Gly
1050734	SNP	CbuG_1088	Endonuclease/exonuclease/phosphatase family protein	C→G	Ala→Pro
1052481	SNP	CbuG_1090	UDP-*N*-acetyl glucosamine 6 dehydrogenase	C→T	Ala→Val
1065901	SNP	CbuG_1106	Colicin V production protein	A→C	Leu→Trp
1076084	SNP	CbuG_1117	Aminoglycoside *N*-6′ acetyltransferase	G→A	Ala→Val
1164183	SNP	CbuG_1199	Ribosomal-protein-S18-alanine acetyltransferase	T→C	Asn→Asp
1190949	SNP	CbuG_1221	Response regulator	A→C	Val→Gly
1292064	SNP	CbuG_1326	NAD-dependent epimerase/dehydratase family	G→T	Thr→Lys
1314867	SNP	CbuG_1357	GTP cyclohydrolase II	T→G	Glu→Ala
1549656	SNP	CbuG_1595	ATP-dependent DNA helicase Rec G	G→T	Ser→Arg
1615472	SNP	CbuG_1668	Membrane alanine aminopeptidase	A→G	Leu→Pro
1662811	SNP	CbuG_1713	Hypothetical protein	T→C	Phe→Xaa
1790891	SNP	CbuG_1860	Type 4 pili biogenesis protein (nucleotide-binding protein)	G→A	Glu→Lys
1802885	SNP	CbuG_1874	UDP-3-*O*-[3-hydroxymyristoyl] *N*-acetylglucosamine deacetylase	C→T	Gly→Asp
1807487	SNP	CbuG_1879	UDP-*N*-acetylenolpyruvoylglucosamine reductase	T→G	Asn→His
1899207	SNP	CbuG_1975	Multidrug resistance protein D	G→T	Gly→Cys
1902387	SNP	CbuG_1983	Carboxylesterase	T→G	Ser→Ala
1904423	SNP	CbuG_1985	Organic solvent tolerance protein	A→G	Tyr→Cys
1906869	SNP	CbuG_1985	Organic solvent tolerance protein	A→C	Glu→Asp
1907022	SNP	CbuG_1987	Peptidyl-prolyl *cis-trans* isomerase	G→A	Val→Ile
1910350	SNP	CbuG_1990	Universal stress protein A	G→A	Ala→Thr
1933346	SNP	CbuG_2020	ATP-dependent endopeptidase hsl proteolytic subunit	C→G	Phe→Leu
1934809	SNP	CbuG_2021	ATP-dependent endopeptidase hsl ATP-binding subunit	T→C	Met→Thr
1185060	Insertion	CbuG_1215	CoA-transferase family III protein	–TGG	INS His
479626	Deletion	CbuG_0513	NADH quinone oxidoreductase subunit L	G–	DEL5aa
491265	Deletion	CbuG_0526	Hypothetical protein	T–	DEL1aa

Single nucleotide polymorphisms and INDELs after comparison between *Coxiella burnetii* Dog_Utad and *Coxiella burnetii* Cbu_Q212.
[a]Position in *C. burnetii* Cbu_Q212.

Nucleotide Sequence Accession Numbers

Strain Dog Utad has been deposited in GenBank under the project accession number PRJEB4294. The version described in this article is the first version, PRJEB4294.

Acknowledgements

We thank Nathalie Duclos for her effective collaboration on the *Coxiella burnetii* pangenome project and Laetitia Rouli for her help in bioinformatics analysis.

Conflict of Interest

None declared.

References

1. Maurin M, Raoult D. Q fever. *Clin Microbiol Rev* 1999; 12: 518–553.
2. Beare PA, Unsworth N, Andoh M *et al.* Comparative genomics reveal extensive transposon-mediated genomic plasticity and diversity among potential effector proteins within the genus *Coxiella*. *Infect Immun* 2009; 77: 642–656.
3. Buhariwalla F, Cann B, Marrie TJ. A dog-related outbreak of Q fever. *Clin Infect Dis* 1996; 23: 753–755.
4. Raoult D, Torres H, Drancourt M. Shell-vial assay: evaluation of a new technique for determining antibiotic susceptibility, tested in 13 isolates of *Coxiella burnetii*. *Antimicrob Agents Chemother* 1991; 35: 2070–2077.
5. Buhariwalla F, Cann B, Marrie TJ. A dog-related outbreak of Q fever. *Clin Infect Dis* 1996; 23: 753–755.
6. Glazunova O, Roux V, Freylikman O *et al. Coxiella burnetii* genotyping. *Emerg Infect Dis* 2005; 11: 1211–1217.
7. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol* 1990; 215: 403–410.

Draft Genome Sequence of *Coxiella burnetii* Cb196 from Saudi Arabia, a strain with a novel genotype

Felicetta D'AMATO[1], Noémie LABAS, Azhar EI[2,3], Didier RAOULT[1,2*]

1. Aix-Marseille Université, Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes, UM63, CNRS 7278, IRD 198, INSERM 1095, Marseille, France
2. Special Infectious Agents Unit, King Fahd Medical Research Center, King Abdulaziz University, Jeddah, Saudi Arabia ; Department of Medical Laboratory Technology, Faculty of Applied Medical Sciences, King Abdulaziz University, Jeddah, Saudi Arabia.
3. Department of Medical Laboratory Technology, Faculty of Applied Medical Sciences, King Abdulaziz University, Jeddah, Saudi Arabia

genomeA
Journals.ASM.org

Draft Genome Sequence of *Coxiella burnetii* Strain Cb196, an Agent of Endocarditis in Saudi Arabia

Felicetta D'Amato,[a] Catherine Robert,[a] E. I. Azhar,[b,c] Pierre-Edouard Fournier,[a] Didier Raoult[a,b,c]

Aix-Marseille Université, Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes, UM63, CNRS 7278, IRD 198, INSERM 1095, Marseille, France[a]; Special Infectious Agents Unit, King Fahd Medical Research Center, King Abdulaziz University, Jeddah, Saudi Arabia[b]; Department of Medical Laboratory Technology, Faculty of Applied Medical Sciences, King Abdulaziz University, Jeddah, Saudi Arabia[c]

***Coxiella burnetii* Cb196, with a 2,006,415-bp genome, is a strain isolated from a 45-year-old man in Saudi Arabia with endocarditis. It belongs to the genotype MST51, which was detected for the first time only in this country. Cb196 shows more similarity to *C. burnetii* CbuK_Q154, belonging to genotype 8, which was phylogenetically close to MST51.**

Received 2 October 2014 **Accepted** 16 October 2014 **Published** 26 November 2014

Citation D'Amato F, Robert C, Azhar EI, Fournier P-E, Raoult D. 2014. Draft genome sequence of *Coxiella burnetii* strain Cb196, an agent of endocarditis in Saudi Arabia. Genome Announc. 2(6):e01180-14. doi:10.1128/genomeA.01180-14.

Address correspondence to Didier Raoult, didier.raoult@gmail.com.

Coxiella burnetii is a zoonotic and strictly intracellular Gram-negative bacterium belonging to the *Gammaproteobacteria*, and it is the agent of Q fever (1).

We isolated *C. burnetii* strain Cb196 from the cardiac valve, received in September 2012, of a 45-year-old man in Saudi Arabia with aortic endocarditis (2). Q fever endocarditis has been described in this country, both in children and adults (2–5).

Strain Cb196 was cultured in human embryonic lung cells using the shell vial method (6). Multispacer genotyping (MST), based on the variability of 10 intergenic sequences (7, 8), was performed and enabled the identification of a unique genotype, MST51 (2).

Sequencing was performed using the MiSeq technology with the paired-end and bar code strategies, according to the manufacturer's instructions (Illumina, San Diego, CA, USA). Briefly, the library was mixed with 13 other genomic projects constructed according to the Nextera XT library kit (Illumina). The libraries were normalized and pooled into one single-strand library that was loaded onto the reagent cartridge and then onto the instrument along with the flow cell. Automated cluster generation and paired-end sequencing with dual-index reads were performed in a single 39-h run in 2 × 250-bp. A total output of 8.9 Gb was obtained with 1,023,000/mm^2 density and 89.6% (19,294,000 clusters) of the clusters passing quality control filters. Within this pooled run, the index representation of Cb196 was determined to 1.31%. The 226,405 paired-end reads were filtered according to the read qualities. We followed the A5-miseq pipeline (9) for read trimming, correction, contig assembly, and scaffolding. We performed a further step of misassembling correction using the r2cat tool (10) and then reordered the scaffolds with Mauve (11) using Cbuk_Q154 (12) as the reference genome, having genotype 8 phylogenetically close to MST51.

Sequencing resulted in 3 scaffolds of 507,295 bp, 1,474,541 bp, and 24,579 bp (with 99, 100, and 99% coverage of the *C. burnetii* CbuK_Q154 genome, respectively). The genome size is 2,006,415 bp (G+C%, 42.9%), which is very similar to that of strain CbuK_Q154. The genome contains genes encoding 1,898 proteins, 44 tRNAs, and one ribosomal operon.

Compared to another 7 complete *C. burnetii* genomes, strain Cb196 had a similar gene function distribution. Strain Cb196, as CbuK_Q154, harbors the QpRS plasmid (39,312 bp).

In comparison with strain CbuK_Q154, strain Cb196 had 2,356 mutations, including 2,018 single-nucleotide polymorphisms (SNPs) (916 nonsynonymous among 1,358 SNPs in putative open reading frames [ORFs]), 171 insertions (57 in ORFs), and 167 deletions (63 in ORFs). A total of 674 genes were mutated. We classified the products of the nonsynonymous genes in COG categories. Forty-one proteins were classified in category J, 6 in K, 37 in L, 5 in D, 14 in V, 12 in T, 46 in M, 8 in U, 21 in O, 24 in C, 27 in G, 33 in E, 9 in F, 27 in H, 14 in I, 16 in P, 8 in Q, 58 in R, 42 in S, 34 belonged to multiple classes, and 192 were not annotated.

Statistical tests showed that for categories V and M, the proportions of mutated proteins were higher than for other protein categories (14/20 versus 660/1,878; bilateral chi-square test, $P = 0.001$; and 46/104 versus 628/1,794; $P = 0.056$, respectively).

Nucleotide sequence accession numbers. Strain Cb196_Saudi_Arabia has been deposited in GenBank under the project accession no. PRJEB7092, and the accession numbers are CCXO01000001 to CCXO01000003. The version described in this article is the first version.

ACKNOWLEDGMENTS

This work was supported by the IHU Mediterranée Infection and the French National referral center for Q fever.

REFERENCES

1. **Maurin M, Raoult D.** 1999. Q fever. Clin. Microbiol. Rev. **12:**518–553.
2. **Angelakis E, Johani S, Ahsan A, Memish Z, Raoult D.** 2014. Q fever endocarditis and new *Coxiella burnetii* genotype, Saudi Arabia. Emerg. Infect. Dis. **20:**726–728. http://dx.doi.org/10.3201/eid2004.131603.
3. **Ross PJ, JacobsoJn J, Muir JR.** 1983. Q fever endocarditis of porcine xenograft valves. Am. Heart J. **105:**151–153. http://dx.doi.org/10.1016/0002-8703(83)90293-4.

4. **Al-Hajjar S, Hussain Qadri SM, Al-Sabban E, Jäger C.** 1997. *Coxiella burnetii* endocarditis in a child. Pediatr. Infect. Dis. J. **16:**911–913. http://dx.doi.org/10.1097/00006454-199709000-00020.
5. **Saginur R, Silver SS, Bonin R, Carlier M, Orizaga M.** 1985. Q-fever endocarditis. CMAJ **133:**1228–1230.
6. **Angelakis E, Richet H, Rolain JM, La Scola B, Raoult D.** 2012. Comparison of real-time quantitative PCR and culture for the diagnosis of emerging rickettsioses. PLoS Negl. Trop Dis. **6:**e1540. http://dx.doi.org/10.1371/journal.pntd.0001540.
7. **Glazunova O, Roux V, Freylikman O, Sekeyova Z, Fournous G, Tyczka J, Tokarevich N, Kovacava E, Marrie TJ, Raoult D.** 2005. *Coxiella burnetii* genotyping. Emerg. Infect. Dis. **11:**1211–1217. http://wwwnc.cdc.gov/eid/article/11/8/04-1354_article.
8. **Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ.** 1990. Basic local alignment search tool. J. Mol. Biol. **215:**403–410. http://dx.doi.org/10.1016/S0022-2836(05)80360-2.
9. **Tritt A, Eisen JA, Facciotti MT, Darling AE.** 2012. An integrated pipeline for *de novo* assembly of microbial genomes. PLoS One **7:**e42304. http://dx.doi.org/10.1371/journal.pone.0042304.
10. **Husemann P, Stoye J.** 2010. r2cat: synteny plots and comparative assembly. Bioinformatics **26:**570–571. http://dx.doi.org/10.1093/bioinformatics/btp690.
11. **Darling AE, Mau B, Perna NT.** 2010. progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. PLoS One **5:**e11147. http://dx.doi.org/10.1371/journal.pone.0011147.
12. **Beare PA, Unsworth N, Andoh M, Voth DE, Omsland A, Gilk SD, Williams KP, Sobral BW, Kupko JJ III, Porcella SF, Samuel JE, Heinzen RA.** 2009. Comparative genomics reveal extensive transposon-mediated genomic plasticity and diversity among potential effector proteins within the genus *Coxiella*. Infect. Immun. **77:**642–656. http://dx.doi.org/10.1128/IAI.01141-08.

CHAPITRE 4: Fièvre Q et grossesse

Avant-propos

Au cours de ces trois années, j'ai également eu l'occasion de participer à deux travaux qui ont porté sur l'étude de certains aspects importants de la fièvre Q pendant la grossesse.

Dans le premier travail publié dans le journal ***European Journal of Clinical Microbiology & Infectious Diseases***, je me suis occupée du génotypage des souches de *C. burnetii* isolées à partir des placentas et d'autres souches isolées des patients au sein de notre laboratoire. Nous avons comparé le type de plasmide entre les souches associées à l'avortement et d'autres souches cliniques, pour déterminer s'il y a un lien entre le génotype et l'avortement chez les humains.

Les résultats de ces analyses ont montré que les souches de *C. burnetii* isolées des placentas de femmes infectées et qui ont subi un avortement, hébergeaient principalement le plasmide QpDV. Cela suggère que la fièvre Q conduit le plus souvent à un avortement si la souche de *C. burnetii* responsable de l'infection héberge ce plasmide (QpDV). Cette spécificité peut être liée au géotyping : les complications obstétricales semblent être corrélées à la zone géographique et comme nous avons discuté dans les chapitres précédents, il existe des génotypes ayant une distribution géographique limitée. Nous avons montré une fois de plus l'importance du géotyping parce que les informations concernant la distribution géographique des souches, si disponibles, peuvent aider les médecins à élaborer des stratégies de gestion pour les patients.

Dans le deuxième travail, je me suis occupée du « mapping » de la séquence du génome de la souche Cb185 isolée du placenta d'une patiente de 20 ans avec un plasmide de type QpH1. Cette jeune femme avait subi un avortement spontané suite à la fièvre Q, juste avant le début du traitement antibiotique. *C.burnetii* a été identifié par PCR dans le placenta et le fœtus. L'analyse du fœtus a montré

un grand omphalocèle contenant le foie et l'intestin grêle. Ce travail a démontré qu'il existe un lien de causalité entre l'infection maternelle par *C. burnetii* d'un coté et la mort et les malformations fœtales d'un autre coté. L'étude a également démontré que le pronostic est plus mauvais lorsque l'infection a lieu dans le premier trimestre de la grossesse, comme dans ce cas. Ce travail a fait l'objet d'un article publié dans le journal ***Clinical Infectious Disease***.

Q fever and pregnancy: disease, prevention, and strain specificity.

Angelakis E1, Million M, D'Amato F, Rouli L, Richet H, Stein A, Rolain JM, Raoult D.

[1]Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes, Faculté de Médecine, CNRS UMR 7278, IRD 198, Aix-Marseille Université, 27 Bd. Jean Moulin, 13005 Marseille, France.

Eur J Clin Microbiol Infect Dis (2013) 32:361–368
DOI 10.1007/s10096-012-1750-3

ARTICLE

Q fever and pregnancy: disease, prevention, and strain specificity

E. Angelakis · M. Million · F. D'Amato · L. Rouli · H. Richet · A. Stein · J.-M. Rolain · D. Raoult

Received: 12 July 2012 / Accepted: 5 September 2012 / Published online: 28 September 2012

Abstract The link between fetal morbidity and Q fever and the necessity of long-term antibiotics for *Coxiella burnetii* infection during pregnancy have been recently questioned in the Netherlands, where the clone responsible for the Q fever outbreak harbors the QpH1 plasmid. In this context, we assessed pregnancy outcomes according to antibiotic administration in a new series and compared the plasmid type between isolates associated with abortion and other clinical isolates to determine if there is a link between genotype and abortion in humans. All French patients who received a diagnosis of Q fever during pregnancy at the French National Referral Centre for Q Fever from 2006 through July 2011 were included. On the other hand, the plasmid types of 160 clinical isolates, including seven isolates from patients who experienced an abortion, were compared. The differences between the QpDV and QpH1 plasmid sequences were analyzed. Acute Q fever was a cause of fetal morbidity, and the absence of long-term cotrimoxazole therapy was associated with fetal death ($p<0.0001$). Genotypic analysis showed that the QpDV plasmid was more frequent in isolates associated with abortion ($p=0.03$). A comparison of the plasmid sequences revealed that four QpDV proteins had no direct counterparts in QpH1, with two whose functions were not present in QpH1. The different obstetrical morbidity of *C. burnetii* relative to different geographical areas could be related to strain specificity, possibly based on differences in plasmid sequences, or to a failure of public health authorities to detect early miscarriages.

Electronic supplementary material The online version of this article (doi:10.1007/s10096-012-1750-3) contains supplementary material, which is available to authorized users.

E. Angelakis and M. Million contributed equally to this work.

E. Angelakis · M. Million · F. D'Amato · L. Rouli · H. Richet · A. Stein · J.-M. Rolain · D. Raoult (✉)
Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes, Faculté de Médecine, CNRS UMR 7278, IRD 198, Aix-Marseille Université,
27 Bd. Jean Moulin,
13005 Marseille, France
e-mail: didier.raoult@gmail.com

Introduction

Q fever during pregnancy has been linked to poor obstetrical outcomes in southern France, Canada, Scotland, and Spain [1–4], primarily when the disease is acquired during the first trimester [5, 6]. The causative bacterium, *Coxiella burnetii*, has been cultivated from birth products and fetal viscera [7]. Long-term treatment with cotrimoxazole was shown to improve the prognosis in a study of an obstetric cohort [6]; however, this approach was questioned during the large 2007–2010 outbreak in the Netherlands.

The human outbreak in the Netherlands was caused by a single *C. burnetii* clone (CbNL01) made up of microvariants of a hypervirulent strain that was introduced into the Dutch animal population [8, 9]. In the Netherlands, some authors report that there is no evidence of adverse effects on pregnancy outcomes among asymptomatic Dutch pregnant women whose blood contains anti-phase II IgM and IgG antibodies during early pregnancy [10]. However, in the same country, one symptomatic chronic infection during pregnancy was diagnosed and successfully treated [11]. Additionally, 7 of 536 pregnant women identified in a screening trial had serological profiles suggestive of acute *C. burnetii* infection and were treated with antibiotics [12]. Histopathological analyses of three of five placentas from these patients revealed abnormalities, including maternal vascular underperfusion, low-grade chronic villitis, and placental hypoplasia, and the clinical outcome was uncomplicated at term in only one case [12]. Therefore, it is not clear if Q fever is associated with less frequent or less severe obstetrical outcomes in the Netherlands, even

though this hypothesis has been supported by some authors [10]. Moreover, most of the complications of Q fever, such as early spontaneous abortion, have not been analyzed in the Netherlands.

Because the clinical presentation of acute Q fever depends both on host factors and on the infecting strain [13–17], we wished to test the putative role of the strain in the obstetrical context. The QpDV plasmid has been previously observed in the placentas of pregnant women from France who spontaneously aborted [14] and in the placenta of a woman in Spain with a good outcome [18]. In the literature, we found no QpH1, QpDG, QpRS, or plasmidless strains isolated from human abortion products [16, 19, 20]. In the Netherlands, the clone responsible for the outbreak harbors the QpH1 plasmid [13].

In the present study, we first examined patients from France who received a diagnosis of Q fever during pregnancy from 2006 to 2011 in our center to test whether long-term antibiotic administration during pregnancy could result in better outcomes. As a new hypothesis, we test if *C. burnetii* strains isolated from aborted placentas in our laboratory and in the literature were associated with a specific plasmid type (QpDV). Finally, we compared the QpDV and QpH1 plasmid sequences to identify a genetic basis for the possible difference in obstetrical morbidity between the Dutch epidemic strain (QpH1) and strains known to be associated with fetal loss (QpDV).

Methods

Patients

All French patients who received a diagnosis of Q fever during pregnancy at the French National Reference Center for Rickettsioses (Marseille, France) from 2006 through July 2011 were included in the clinical study. None of these cases have been previously reported. The clinical and epidemiological data for these patients were obtained as previously described [6]. Other infectious diseases were excluded based on negative blood cultures and the lack of serological evidence of active infection with *Toxoplasma gondii*, rubella virus, cytomegalovirus, human immunodeficiency virus, hepatitis B, influenza viruses, parvovirus B19, adenovirus, *Chlamydia* species, and *Mycoplasma pneumonia*. Patients with missing or incomplete treatment or obstetric data were excluded from the study. The outcome of the pregnancy was categorized as abortion, intrauterine fetal death (IUFD), intrauterine growth retardation (IUGR), premature delivery (<37 weeks), or normal. A spontaneous abortion was defined as the spontaneous expulsion of the embryo or the fetus prior to the 20th week of gestation. IUFD was defined as the intrauterine death of the fetus regardless of the duration of gestation. Oligoamnios was defined as the ultrasonographic measure of an amniotic index of ≤5 cm [21]. IUGR was defined as a fetal birth weight below the 10th percentile for the gestational age based on the reference growth curves [22]. Placental histopathology and the results of autopsy from aborted fetuses were not systematically collected and were reported only when available. Long-term cotrimoxazole treatment was defined as the oral administration of trimethoprim–sulfamethoxazole during at least 5 weeks of pregnancy [6].

Diagnosis of *Coxiella burnetii* infection and multi-spacer sequence typing

Sera were tested by immunofluorescence for *C. burnetii* phase I and II antigens, as previously described [6]. Serum with a phase II IgG titer ≥200 and a phase II IgM titer ≥50 or a positive polymerase chain reaction (PCR) assay for a blood sample was predictive of acute Q fever [6]. If the phase I IgG titer was ≥800, persistent uterine infection was suspected. We determined the multi-spacer sequence typing (MST) genotypes of *C. burnetii* in placentas and of other human *C. burnetii* isolates from our laboratory, as previously described [14]. To increase the sample size, we included strains of *C. burnetii* from the literature isolated from human samples for which the origin of the strain and the plasmid type were specified.

Statistical analysis

In a first clinical analysis, we tested whether long-term treatment with cotrimoxazole was associated with a better prognosis in a cohort of pregnant women infected during pregnancy and followed in our center. In a second genotypic analysis, of all 160 strains analyzed, we compared the proportion of strains isolated from aborted placenta between strains carrying the QpDV plasmid and strains not carrying this plasmid. A bilateral Barnard exact test [23] was used to test associations in 2×2 tables. This test was preferred to the classical exact Fisher test because it is more powerful [24, 25]. A p-value <0.05 was considered to be significant. The results were summarized using a correlation circle of primary component analysis obtained using XLSTAT 2012 (Addinsoft, Paris, France).

Comparison of the QpDV and QpH1 plasmid sequences

As a first step to identify the genomic basis explaining a putative strain specificity of obstetrical outcome, we compare the QpDV and QpH1 plasmid sequences. Two QpH1 plasmids were available from the National Center for Biotechnology Information (NCBI) (NC_002118.1 *Coxiella burnetii* Nine Mile phase I and NC_010115.1 *Coxiella burnetii* RSA 331) and one sequence was available for QpDV (NC_002131.1 *Coxiella burnetii* R1140). We started

Table 1 Characteristics of pregnant patients with Q fever

Patient	Age (years)	Year of diagnosis	Clinical manifestation	Time of infection	Time of diagnosis	Treatment start	Serological profile	Treatment	Duration of treatment	Pregnancy outcome	Test for placentitis	Postpartum treatment
1	27	2006	Fatigue and myalgias	2nd Tr	2nd Tr	2nd Tr	A, C	Cotrimoxazole	3 months	IUGR, OL, bilateral renal agenesis	NP	Yes
2	27	2006	Asymptomatic, previous abortions	1st Tr	1st Tr	1st Tr	A	Cotrimoxazole	6 months	Normal	NP	Yes
3	27	2006	Asymptomatic, previous abortions	1st Tr	1st Tr	1st Tr	A, C	Cotrimoxazole	6 months	Normal	Negative	Yes
4	28	2006	Asymptomatic	2nd Tr	2nd Tr	–	C	No		IUFD, OL	NP	Yes
5	28	2006	Asymptomatic	Unknown/frequent contact with goats	1st Tr	–	A, C	No		Abortion	NP	Yes
6	26	2006	Fever and hepatitis	2nd Tr	2nd Tr	2nd Tr	A	Erythromycin	1 week	IUFD, hydrops fetalis	Negative	Yes
7	35	2006	Fever	2nd Tr	2nd Tr	2nd Tr	C	Cotrimoxazole	3 months	LBW	Negative	LC
8	28	2006	Asymptomatic	Unknown/frequent contact with goats	1st Tr	–	A, C	No		Abortion	Negative	Yes
9	31	2006	Fever	Unknown	2nd Tr	2nd Tr	A	Cotrimoxazole	2 weeks	Abortion	Negative	Yes
10	35	2006	Asymptomatic	Unknown/frequent contact with goats	1st Tr	1st Tr	A	Cotrimoxazole	4 weeks	Abortion	Negative	Yes
11	30	2007	Fever	2nd Tr	2nd Tr	2nd Tr	A	Cotrimoxazole	4 months	Premature delivery (34 weeks)	NP	No
12	36	2007	Asymptomatic	Unknown	1st Tr	–	A, C	No		Abortion	NP	LC
13	25	2007	Asymptomatic	Unknown	2nd Tr	–	C	No		IUFD	NP	Yes
14	30	2007	Fever	1st Tr	1st Tr	1st Tr	A, C	Cotrimoxazole	6 months	Normal	Negative	Yes
15	24	2007	Asymptomatic, previous abortions	1st Tr	2nd Tr	2nd Tr	C	Cotrimoxazole	1 month	IUFD	Negative	Yes
16	30	2007	Asymptomatic, previous abortions	1st Tr	1st Tr	1st Tr	A, C	Cotrimoxazole	6 months	Premature delivery (31 weeks)	Negative	Yes
17	31	2007	Asymptomatic	Unknown/frequent contact with goats	3rd Tr	–	C	No		Premature delivery (33 weeks)	Negative	Yes
18	36	2008	Fever	2nd Tr	2nd Tr	2nd Tr	A, C	Cotrimoxazole	3 months	Normal	Negative	No
19	28	2008	Fever	2nd Tr	2nd Tr	2nd Tr	A	Cotrimoxazole	3 months	Normal	NP	No
20	28	2008	Asymptomatic	1st Tr	2nd Tr	–	A	No		IUFD	NP	LC
21	36	2008	Hepatitis	2nd Tr	2nd Tr	2nd Tr	A, C	Cotrimoxazole	6 months	Normal	Negative	Yes
22	27	2008	Fever	2nd Tr	2nd Tr	2nd Tr	A	Cotrimoxazole	3 months	Normal	NP	Yes
23	34	2009	Asymptomatic	2nd Tr	2nd Tr	2nd Tr	A	Cotrimoxazole	2 months	Premature delivery (32 weeks)	Negative	No
24	34	2010	Asymptomatic	Unknown	2nd Tr	–	C	No		IUFD	NP	Yes
25	42	2010	Asymptomatic	Unknown	1st Tr	–	A	No		Abortion	Negative	Yes
26	29	2010	Fever, hepatitis, fatigue and myalgias	1st Tr	1st Tr	1st Tr	O, A, C	Cotrimoxazole	6 months	Normal	NP	No
27	40	2010	Asymptomatic	Unknown/frequent contact with goats	2nd Tr	–	C	No		IUFD	NP	Yes

Table 1 (continued)

Patient	Age (years)	Year of diagnosis	Clinical manifestation	Time of infection	Time of diagnosis	Treatment start	Serological profile	Treatment	Duration of treatment	Pregnancy outcome	Test for placentitis	Postpartum treatment
28	42	2011	Fever, lymphadenopathy	1st Tr	1st Tr	1st Tr	A	Amoxicillin+ azithromycin	1 week	Abortion	NP	Yes
29	30	2011	Fever, hepatitis	1st Tr	1st Tr	1st Tr	A, C	Cotrimoxazole	5 months	IUFD	NP	No
30	26	2011	Asymptomatic	1st Tr	1st Tr	–	A	No		Abortion	NP	Yes

A, "acute" profile (IgM anti-phase I titer ≥1:50; IgG anti-phase II titer ≥1:200); *C*, "chronic" profile (IgG anti-phase I and II titers ≥1:1,600; IgA anti-phase I and II titers ≥1:100); *O*, "seroconversion"; *IUGR* intrauterine growth retardation; *IUFD* intrauterine fetal death; *OL* oligoamnios; *LBW* low birth weight; *Tr* trimester; *NP* not performed; *LC* lost contact

with the annotation of the proteins from plasmid QpDV, which were all hypothetical, and we used an 80 % threshold for "query coverage". In the following analysis, we only compared sequences to QpDV and the Nine Mile QpH1.

Literature analysis

The PubMed, MEDLINE, Google Scholar, and Google databases were searched for articles unrestricted by date or language. The following search terms were used: *C. burnetii*, genotyping, plasmid, QpH1, QpRS, QpDV, and associated author names. The exclusion criteria included the study of animals and ticks and a lack of information on the type of plasmid or the origin of the isolate.

Results

Q fever during pregnancy, antibiotic use, and outcome

A total of 46 new cases of Q fever during pregnancy were diagnosed and followed in our center during the study, but the obstetric and/or treatment data for 11 patients could not be collected and these patients were excluded from the study. In addition, five women were infected with Q fever prior to becoming pregnant and were also excluded. In total, 30 pregnant women were included in the study (Table 1). Seventeen (56 %) patients were asymptomatic when diagnosed Q fever; in 11 of them (64 %), the diagnosis was made retrospectively after abortion or after a premature delivery, in five patients (29 %), Q fever was suspected because of a history of previous abortions, and one patient had frequent contact with goats. Thirteen (43 %) patients presented clinical manifestations that led to the diagnosis of Q fever. The most common manifestation was fever (84 %), followed by hepatitis (30 %), fatigue-myalgias (15 %), and lymphadenopathy (7 %). A placental analysis was also performed for 14 patients. All samples were negative for *C. burnetii* (Table 1). Obstetric complications were observed in all 16 patients who did not receive long-term cotrimoxazole therapy and only one delivered a live newborn. In contrast, of the 14 patients who received long-term cotrimoxazole therapy, 8 (57 %) patients delivered normally, 3 (21 %) patients delivered prematurely, and one patient delivered a low-birth-weight newborn; the fetus died in only 2 (14 %) cases. A significantly greater number of pregnant women delivered a living newborn when treated with cotrimoxazole [odds ratio (OR) =39; 95 % confidence interval (CI) 3.8–400; p=0.0002]. There was no evidence of a link between a live newborn and the occurrence of infection during the first trimester and cotrimoxazole treatment (p=0.11), but the occurrence of infection during the second trimester and cotrimoxazole treatment was significantly

Table 2 The types of plasmids harbored by the 160 *Coxiella burnetii* isolates for which the information were available

	QpDV	QpH1	QpRS	Total[a]
Number of placentas from France	3	2	1	6
Number of placentas from elsewhere	1	0	0	1
Number of isolates from France	33	38	41	112
Number of isolates from elsewhere	2	33	7	42
Total	39	73	49	

[a] One isolate (CB13) harbored both QpDV and QpH1 sequences and have been previously reported in [14]

associated with a live newborn (OR undefined because none of the untreated women delivered a live newborn; $p=0.009$). Long-term cotrimoxazole therapy was also highly associated with a live newborn (OR=195; 95 % CI [11–3,437]; $p=0.000003$).

Coxiella burnetii genotyping

A total of 22 human *C. burnetii* isolates from France and 13 isolates from other countries were genotyped (see Supplementary information). One isolate was obtained from a human placenta and was identified as genotype 22. Nine human isolates from six different cities in the Netherlands were obtained from sputum, heart valve, and bronchoalveolar lavage samples; eight of these isolates were identified as genotype MST33, and one novel genotype was found in a valve sample from a patient from Standdaarbuiten. All nine isolates harbored the QpH1 plasmid. To our knowledge, no strain was isolated from a placenta during the Dutch outbreak [12].

Overall, the *C. burnetii* plasmid type has been determined in 160 human isolates for which origin of the isolate was known: 35 isolates in this study and 125 isolates in previous reports [13, 14, 26, 27] (see Supplementary information). A total of 117 *C. burnetii* strains were obtained from France, and 43 strains were obtained from other countries. The clinical manifestation, including acute Q fever, endocarditis, vascular infection, or abortion, was identified for 144 strains (see Supplementary information). Six of the isolates from France (three QpDV, two QpH1, and one QpRS) and one from Spain (QpDV) were obtained from the placentas of women who aborted (Table 2). These isolates corresponded to genotypes MST1 (three strains, QpDV), MST8 (one strain,

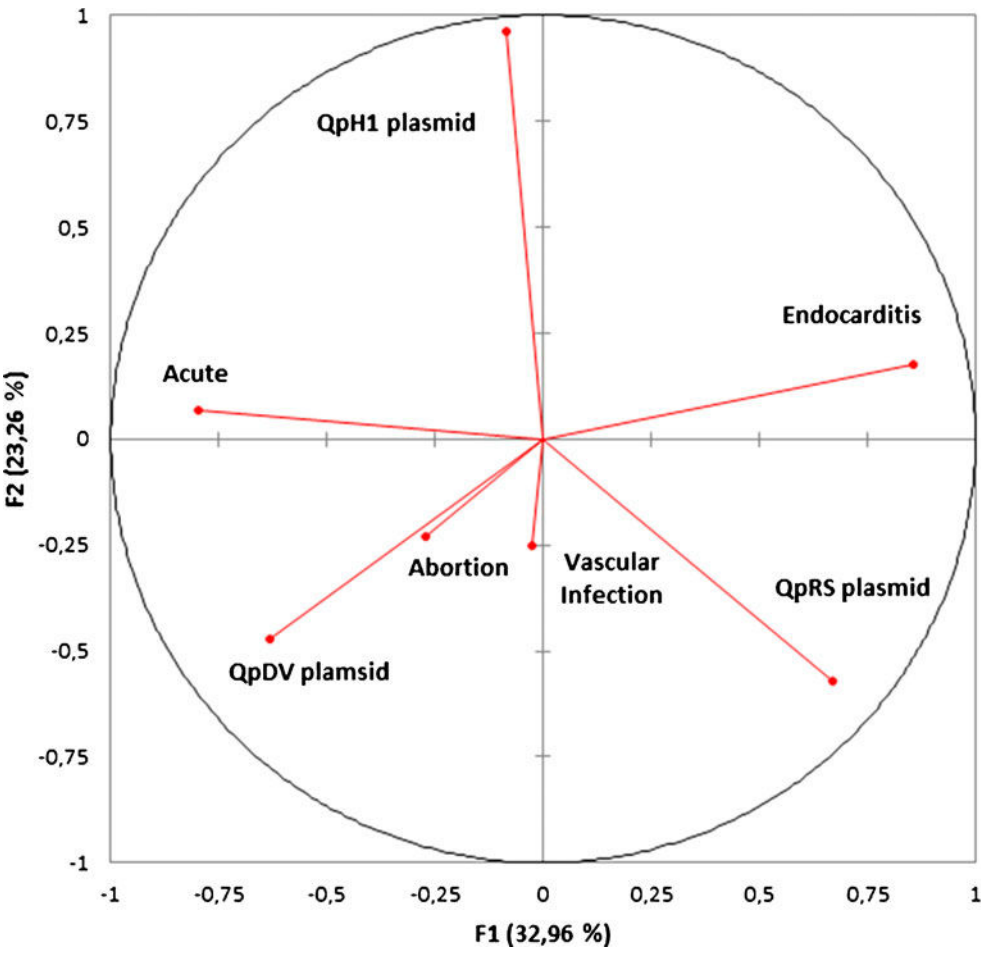


Fig. 1 Principal component analysis (PCA) including plasmid types and clinical isolates. These analyses were performed using 119 *Coxiella burnetii* isolates from patients with acute Q fever, endocarditis, vascular infections, or abortion for whom the plasmid type was available. All isolates harbored the QpDV, QpH1, or QpRS plasmid. Correlation circle of PCA using XLSTAT 2012 (Addinsoft, Paris, France) showed that the QpDV plasmid type was associated with abortion (Pearson coefficient, 0.18; $p<0.05$)

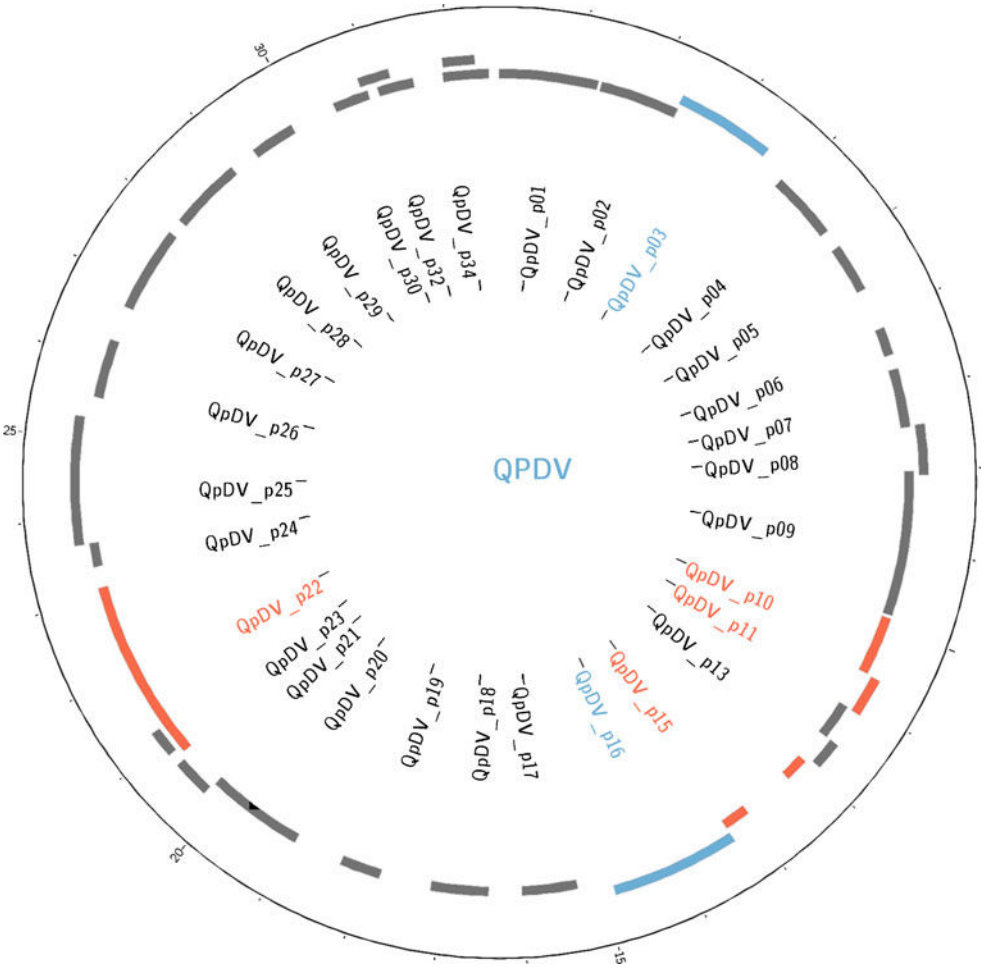


Fig. 2 A comparison of the QpDV and QpH1 plasmid sequences. We constructed this figure using Circos software [30]. Of the 34 proteins encoded by QpDV, 27 had a homolog in QpH1. These proteins are represented in *dark gray* on the circular map. The four proteins in *red* have no counterparts in QpH1, but they are either hypothetical or have functions that are present in QpH1. The two proteins in *blue* have no counterparts in QpH1, and their functions are not present in this plasmid

QpRS), and MST20 and MST22 (one strain each, QpH1). The proportion of strains isolated from placenta after abortion was higher for strains carrying the QpDV plasmid compared to strains not carrying this plasmid [4/39 vs. 3/121, respectively; bilateral exact Barnard test (p=0.033), Fig. 1].

Comparison of the QpDV and QpH1 plasmid sequences

Of the 34 proteins identified in QpDV, 27 had a homolog in QpH1. Among the proteins for which we were able to assign a putative function, only four had no counterparts in QpH1 (Fig. 2). These four proteins encode the following functions: "plasmid partition protein A", "E surface protein", "DNA helicase I putative protein traI" and "Rhs protein family". However, we noted that the last two functions were present in QpH1. "Plasmid partition protein A" and "E surface protein" were the only two divergent functions identified for the two plasmids; however, other hypothetical proteins were also present, and, therefore, these functions could still be found in QpH1. QpH1 and QpDV were very similar with respect to the GC% (39.3 % vs. 39.5 %, respectively), size (0.04 Mb vs. 0.03 Mb, respectively), and clusters of orthologous groups (COGs) distributions of their proteins. In conclusion, four QpDV proteins had no counterparts in QpH1, but these proteins were either hypothetical or had a function present in QpH1. Two proteins had no counterparts in QpH1, and their functions were not present in this plasmid and two of the four proteins had functions not present in this plasmid.

Discussion

Our study confirmed that Q fever during pregnancy is a significant cause of morbidity in France and can result in spontaneous abortion in pregnant women. To date, teratogenicity has never been associated with *C. burnetii* infection [6], and the bilateral renal agenesis presented in case 1 can be sporadic, with no evidence of a direct relation with *C. burnetii* infection. Moreover, we confirmed that long-term cotrimoxazole therapy (≥5 weeks) prevents complications and is significantly associated with a living newborn [6].

As already reported in the literature [14], we found that the clinical manifestations of Q fever depended, at least in part, on the *C. burnetii* genotype, with strains carrying the QpDV plasmid being more frequently associated with acute Q fever ([14] and see Supplementary information). We, therefore,

tested the hypothesis that obstetric complications could be associated with a particular genotype and, specifically, QpDV, which has been associated with obstetrical complications in the literature [14]. Using a two-tailed Barnard test, our analysis revealed that *C. burnetii* strains isolated from the placentas of infected women who experienced abortion primarily harbored the QpDV plasmid ($p=0.033$), in favor of such a hypothesis.

In addition, *C. burnetii* isolates associated with the QpH1 plasmid have been shown to have fewer deleted genes than isolates harboring the QpRS and QpDV plasmids [13, 14]. Finally, QpDV plasmid harbors sequences coding for four proteins not found in QpH1, which could explain the differences in clinical expression and obstetrical outcomes. However, as the plasmid type is associated with the genetic chromosome content [15], only the ongoing pangenome analysis of *C. burnetii* will determine the comprehensive genomic basis for the difference in virulence between strains in pregnancy.

One genotype of *C. burnetii* was found predominantly in dairy goats and sheep in the southern part of the Netherlands, and this genotype has also been isolated from a human [28]. We found the same MST33 genotype in several human samples obtained from six different cities in the Netherlands. A novel genotype isolated from a valve sample from a patient in Standdaarbuiten appears to be unrelated to MST33 and could represent an incidental case of Q fever unrelated to the outbreak. This novel genotype might be related to the sporadic cases of Q fever that occurred in the Netherlands prior to the current outbreak. The underreported obstetrical morbidity in the Netherlands relative to other geographical areas could be related to strain specificity, which may be based on differences in plasmid sequences or to a failure of the public authorities to detect this type of morbidity. Moreover, while the present study relies on the retrospective analysis of a cohort of patients diagnosed with Q fever, the Dutch study [10] was based on a large screening program. This is why a selection bias cannot be ruled out.

In this work, we confirm *C. burnetii* as a cause of severe obstetrical outcome and highlight the benefit of long-term cotrimoxazole therapy for patients diagnosed with Q fever while pregnant. We found that *C. burnetii* strains isolated from the placentas of infected women who experienced abortion primarily harbored the QpDV plasmid, suggesting that Q fever more often leads to an abortion if the infecting strain of *C. burnetii* harbors a QpDV plasmid. This strain specificity could explain why Q fever has been linked with obstetrical complications depending on geographical area, since many genotypes appear to have limited geographic distributions [14, 29]. If available, information on locally circulating strains may assist physicians in developing management strategies for patients. We will continue to propose our current treatment strategy of long-term cotrimoxazole therapy for the full duration of pregnancy in pregnant women diagnosed with acute Q fever in France, Spain, and Eastern Europe.

Acknowledgments We thank Patrizia Carrieri for her precious help with the statistical analyses.

Funding This work was supported by the French National Referral Centre for Q Fever.

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Jover-Díaz F, Robert-Gates J, Andreu-Gimenez L, Merino-Sanchez J (2001) Q fever during pregnancy: an emerging cause of prematurity and abortion. Infect Dis Obstet Gynecol 9(1):47–49
2. Stein A, Raoult D (1998) Q fever during pregnancy: a public health problem in southern France. Clin Infect Dis 27(3):592–596
3. Langley JM, Marrie TJ, Leblanc JC, Almudevar A, Resch L, Raoult D (2003) *Coxiella burnetii* seropositivity in parturient women is associated with adverse pregnancy outcomes. Am J Obstet Gynecol 189(1):228–232
4. Ellis ME, Smith CC, Moffat MA (1983) Chronic or fatal Q-fever infection: a review of 16 patients seen in North-East Scotland (1967–80). Q J Med 52(205):54–66
5. Carcopino X, Raoult D, Bretelle F, Boubli L, Stein A (2009) Q Fever during pregnancy: a cause of poor fetal and maternal outcome. Ann N Y Acad Sci 1166:79–89
6. Carcopino X, Raoult D, Bretelle F, Boubli L, Stein A (2007) Managing Q fever during pregnancy: the benefits of long-term cotrimoxazole therapy. Clin Infect Dis 45(5):548–555
7. Raoult D, Stein D (1994) Q fever during pregnancy—a risk for women, fetuses, and obstetricians (letter). N Engl J Med 330:371
8. Roest HI, Ruuls RC, Tilburg JJ, Nabuurs-Franssen MH, Klaassen CH, Vellema P et al (2011) Molecular epidemiology of *Coxiella burnetii* from ruminants in Q fever outbreak, the Netherlands. Emerg Infect Dis 17(4):668–675
9. Klaassen CH, Nabuurs-Franssen MH, Tilburg JJ, Hamans MA, Horrevorts AM (2009) Multigenotype Q fever outbreak, the Netherlands. Emerg Infect Dis 15(4):613–614
10. van der Hoek W, Meekelenkamp JC, Leenders AC, Wijers N, Notermans DW, Hukkelhoven CW (2011) Antibodies against *Coxiella burnetii* and pregnancy outcome during the 2007–2008 Q fever outbreaks in The Netherlands. BMC Infect Dis 11:44
11. Munster JM, Hamilton CJ, Leenders AC, Lestrade PJ (2011) Chronic Q fever during pregnancy. Ned Tijdschr Geneeskd 155: A2781
12. Munster JM, Leenders AC, Hamilton CJ, Hak E, Aarnoudse JG, Timmer A (2012) Placental histopathology after *Coxiella burnetii* infection during pregnancy. Placenta 33(2):128–131
13. Leroy Q, Armougom F, Barbry P, Raoult D (2011) Genomotyping of *Coxiella burnetii* using microarrays reveals a conserved genomotype for hard tick isolates. PLoS One 6(10):e25781
14. Glazunova O, Roux V, Freylikman O, Sekeyova Z, Fournous G, Tyczka J et al (2005) *Coxiella burnetii* genotyping. Emerg Infect Dis 11(8):1211–1217
15. Beare PA, Samuel JE, Howe D, Virtaneva K, Porcella SF, Heinzen RA (2006) Genetic diversity of the Q fever agent, *Coxiella burnetii*, assessed by microarray-based whole-genome comparisons. J Bacteriol 188(7):2309–2324
16. Samuel JE, Frazier ME, Mallavia LP (1985) Correlation of plasmid type and disease caused by *Coxiella burnetii*. Infect Immun 49 (3):775–779

17. Angelakis E, Raoult D (2010) Q Fever. Vet Microbiol 140(3–4):297–309
18. Téllez A, Sanz Moreno J, Valkova D, Domingo C, Anda P, de Ory F et al (1998) Q fever in pregnancy: case report after a 2-year follow-up. J Infect 37(1):79–81
19. Hendrix LR, Samuel JE, Mallavia LP (1991) Differentiation of *Coxiella burnetii* isolates by analysis of restriction-endonuclease-digested DNA separated by SDS-PAGE. J Gen Microbiol 137 (2):269–276
20. Valková D, Kazár J (1995) A new plasmid (QpDV) common to *Coxiella burnetii* isolates associated with acute and chronic Q fever. FEMS Microbiol Lett 125(2–3):275–280
21. Phelan JP, Ahn MO, Smith CV, Rutherford SE, Anderson E (1987) Amniotic fluid index measurements during pregnancy. J Reprod Med 32(8):601–604
22. Mamelle N, Munoz F, Grandjean H (1996) Fetal growth from the AUDIPOG study. I. Establishment of reference curves. J Gynecol Obstet Biol Reprod (Paris) 25(1):61–70
23. Barnard GA (1945) A new test for 2×2 tables. Nature 156:177
24. Mehta CR, Senchaudhuri P (2003) Conditional versus unconditional exact tests for comparing two binomials. Cytel Software corporation, 675 Massachusetts Avenue, Cambridge, MA 02139, USA. Available online at: http://www.cytel.com/papers/twobinomials.pdf. 4 September 2003
25. Lydersen S, Fagerland MW, Laake P (2009) Recommended tests for association in 2×2 tables. Stat Med 28(7):1159–1175
26. Arricau-Bouvery N, Hauck Y, Bejaoui A, Frangoulidis D, Bodier CC, Souriau A et al (2006) Molecular characterization of *Coxiella burnetii* isolates by infrequent restriction site-PCR and MLVA typing. BMC Microbiol 6:38
27. Hotta A, Kawamura M, To H, Andoh M, Yamaguchi T, Fukushi H et al (2003) Use of monoclonal antibodies to lipopolysaccharide for antigenic analysis of *Coxiella burnetii*. J Clin Microbiol 41 (4):1747–1749
28. Hogerwerf L, van den Brom R, Roest HI, Bouma A, Vellema P, Pieterse M et al (2011) Reduction of *Coxiella burnetii* prevalence by vaccination of goats and sheep, The Netherlands. Emerg Infect Dis 17(3):379–386
29. Hornstra HM, Priestley RA, Georgia SM, Kachur S, Birdsell DN, Hilsabeck R et al (2011) Rapid typing of *Coxiella burnetii*. PLoS One 6(11):e26201
30. Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D et al (2009) Circos: an information aesthetic for comparative genomics. Genome Res 19(9):1639–1645

Isolate	Origin	IbaSputum	Valve vegetation	BloodSerum	Thrombus/Aneurism	Placenta	Bone	Node	Spleen bscess	Disease, symptoms, clinical status	Abortion	Acute	Endocarditis	Vascular	Bone	Spleen	Node	Unsp chronic	Unknown	Geographic source	France	Swiss	Spain	Italy	Netherland	Asia	urss	East Europe	Africa	America	MST	MST1	MST 2	MST 3	MST 4	MST 5	MST 6	MST 7	MST 8	MST 9	MST 10	MST 11	MST 12	MST 13	MST 14	MST 15	MST 16	MST 17	MST 18	MST 19	MST 20	MST 21	MST 22	MST 23	MST 26	MST 29	MST33	Plasmid type	Plasmid QpH1	Plasmid QpDV	Plasmid QpRS	Reference
CB 36	Human placenta	0	0	0	0	1	0	0	0	Abortion, placenta	1	0	0	0	0	0	0	0	0	Martigues, France, 1992	1	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Qp DV	0	1	0	Glazunova et al.
CB 87	Human placenta	0	0	0	0	1	0	0	0	Abortion, placenta	1	0	0	0	0	0	0	0	0	Martigues, France, 1999	1	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Qp DV	0	1	0	Glazunova et al.
CB 89	Human placenta	0	0	0	0	1	0	0	0	Abortion, placenta	1	0	0	0	0	0	0	0	0	Martigues, France, 2000	1	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Qp DV	0	1	0	Glazunova et al.
CB 108	Human blood	0	0	1	0	0	0	0	0	Acute, blood	0	1	0	0	0	0	0	0	0	Marseille, France, 2001	1	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Qp DV	0	1	0	Glazunova et al.
CB 1	Human heart valve	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Istres, France, 1989	1	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Qp DV	0	1	0	Glazunova et al.
CB 3	Human heart valve	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Marseille, France	1	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Qp DV	0	1	0	Glazunova et al.
CB 38	Human blood	0	0	1	0	0	0	0	0	Acute, blood	0	1	0	0	0	0	0	0	0	Marseille, France,	1	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Qp DV	0	1	0	Glazunova

																				1992																																										et al.
CB 41	Human heart valve	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Marseille, France, 1993	1	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Qp DV	0	1	0	Glazunova et al.
CB 60	Human blood	0	0	1	0	0	0	0	0	Acute, blood	0	1	0	0	0	0	0	0	0	Marseille, France, 1996	1	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Qp DV	0	1	0	Glazunova et al.
CB 63	Human heart valve	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Marseille, France, 1994	1	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Qp DV	0	1	0	Glazunova et al.
CB 75	Human blood	0	0	1	0	0	0	0	0	Chronic, blood	0	0	0	0	0	0	0	1	0	Marseille, France, 1998	1	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Qp DV	0	1	0	Glazunova et al.
CB 82	Human blood	0	0	1	0	0	0	0	0	Acute, blood	0	1	0	0	0	0	0	0	0	Marseille, France, 1999	1	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Qp DV	0	1	0	Glazunova et al.
CB 86	Human blood	0	0	1	0	0	0	0	0	Chronic, blood	0	0	0	0	0	0	0	1	0	Marseille, France, 1999	1	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Qp DV	0	1	0	Glazunova et al.
CB 94	Human blood	0	0	1	0	0	0	0	0	Acute, blood	0	1	0	0	0	0	0	0	0	Aix en Provence, France, 2000	1	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Qp DV	0	1	0	Glazunova et al.
CB 97	Human	0	0	1	0	0	0	0	0	Acute, blood	0	1	0	0	0	0	0	0	0	Marseille,	1	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Qp DV	0	1	0	Glazuno

	blood																			France, 2000																																										va et al.
CB110	Human blood	0	0	1	0	0	0	0	0	Chronic, blood	0	0	0	0	0	0	0	1	0	Marseille, France, 2002	1	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Qp DV	0	1	0	Glazunova et al.
CB28	Human blood	0	0	1	0	0	0	0	0	Acute, blood	0	1	0	0	0	0	0	0	0	Salon de Provence, France, 1992	1	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Qp DV	0	1	0	Glazunova et al.
CB26	Human blood	0	0	1	0	0	0	0	0	Acute, blood	0	1	0	0	0	0	0	0	0	Marseille, France, 1992	1	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Qp DV	0	1	0	Glazunova et al.
CB64	Human blood	0	0	1	0	0	0	0	0	Acute, blood	0	1	0	0	0	0	0	0	0	Martigues, France, 1996	1	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Qp DV	0	1	0	Glazunova et al.
CB39	Human blood	0	0	1	0	0	0	0	0	Acute, blood	0	1	0	0	0	0	0	0	0	Marseille, France, 1992	1	0	0	0	0	0	0	0	0	0	2	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Qp DV	0	1	0	Glazunova et al.
CB118	Human heart valve	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Marseille, France, 2004	1	0	0	0	0	0	0	0	0	0	3	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Qp DV	0	1	0	Glazunova et al.
CB62	Human blood	0	0	1	0	0	0	0	0	Acute, blood	0	1	0	0	0	0	0	0	0	Martigues, France, 1996	1	0	0	0	0	0	0	0	0	0	4	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Qp DV	0	1	0	Glazunova et al.

CB 12	Human blood	0	0	1	0	0	0	0	0	Acute, blood	0	1	0	0	0	0	0	0	0	Aix en Provence, France	1	0	0	0	0	0	0	0	0	0	4	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Qp DV	0	1	0	Glazunova et al.
CB 57	Human blood	0	0	1	0	0	0	0	0	Acute, blood	0	1	0	0	0	0	0	0	0	Martigues, France, 1996	1	0	0	0	0	0	0	0	0	0	4	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Qp DV	0	1	0	Glazunova et al.
CB 54	Human blood	0	0	1	0	0	0	0	0	Acute, blood	0	1	0	0	0	0	0	0	0	Aix en Provence, France, 1996	1	0	0	0	0	0	0	0	0	0	4	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Qp DV	0	1	0	Glazunova et al.
MANc	Human blood	0	0	1	0	0	0	0	0	Vascular infection (Aortic aneurysm), blood	0	0	0	1	0	0	0	0	0	France	1	0	0	0	0	0	0	0	0	0																												Qp DV	0	1	0	Hotta et al.
MEc	Human heart valve	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	France	1	0	0	0	0	0	0	0	0	0																												Qp DV	0	1	0	Hotta et al.
Cb 123	Human blood	0	0	1	0	0	0	0	0	Chronic, blood	0	0	0	0	0	0	0	1	0	Marseille, France, 2004	1	0	0	0	0	0	0	0	0	0																												Qp DV	0	1	0	This study
Cb 128	Human heart valve	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Marseille, France, 2005	1	0	0	0	0	0	0	0	0	0																												Qp DV	0	1	0	This study
Cb 138	Human heart valve	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Lyon, France, 2007	1	0	0	0	0	0	0	0	0	0																												Qp DV	0	1	0	This study
Cb 139	Human heart valve	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Tours, France, 2007	1	0	0	0	0	0	0	0	0	0																												Qp DV	0	1	0	This study
Cb 141	Human heart valve	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Marseille, France 2008	1	0	0	0	0	0	0	0	0	0																												Qp DV	0	1	0	This study

Cb142	Human heart valve	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Angers, France, 2008	1	0	0	0	0	0	0	0	0	0																												QpDV	0	1	0	This study
Cb133	Human blood	0	0	1	0	0	0	0	0	Chronic, blood	0	0	0	0	0	0	0	1	0	Marseille, France	1	0	0	0	0	0	0	0	0	0																												QpDV	0	1	0	This study
Cb145	Thrombus	0	0	0	1	0	0	0	0	Chronic, blood	0	0	0	0	0	0	0	1	0	Grenoble, France, 2008	1	0	0	0	0	0	0	0	0	0																												QpDV	0	1	0	This study
CB51	Human placenta	0	0	0	0	1	0	0	0	Abortion, placenta	1	0	0	0	0	0	0	0	0	Madrid, Spain, 1996	0	0	1	0	0	0	0	0	0	0	4	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	QpDV	0	1	0	Glazunova et al.
RT-Schperling	Human blood	0	0	1	0	0	0	0	0	Fever, blood									1	Kyrgyzstan, 1955	0	0	0	0	0	0	1	0	0	0	2	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	QpDV	0	1	0	Glazunova et al.
R1140	Human Blood	0	0	1	0	0	0	0	0	Pneumonia, blood	0	1	0	0	0	0	0	0	0	Russia	0	0	0	0	0	0	1	0	0	0																												QpDV	0	1	0	Arricau-Bouvery et al.
CB40	Human heart valve	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Paris, France, 1993	1	0	0	0	0	0	0	0	0	0	11	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	QpH1	1	0	0	Glazunova et al.
CB46	Valvular prosthesis	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Paris, France, 1993	1	0	0	0	0	0	0	0	0	0	11	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	QpH1	1	0	0	Glazunova et al.
CB5	Human blood	0	0	1	0	0	0	0	0	Chronic, blood	0	0	0	0	0	0	0	1	0	Paris, France, 1990	1	0	0	0	0	0	0	0	0	0	12	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	QpH1	1	0	0	Glazunova et

																																																													al.
CB 6	Human blood	0	0	1	0	0	0	0	0	Chronic, blood	0	0	0	0	0	0	0	1	0	Paris, France	1	0	0	0	0	0	0	0	0	0	12	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Qp H1	1	0	0	Glazunova et al.
CB 42	Valvular prosthesis	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Toulouse, France, 1993	1	0	0	0	0	0	0	0	0	0	12	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Qp H1	1	0	0	Glazunova et al.
CB 52	Valvular prosthesis	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Paris, France, 1995	1	0	0	0	0	0	0	0	0	0	12	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Qp H1	1	0	0	Glazunova et al.
CB 56	Human heart valve	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Paris, France, 1996	1	0	0	0	0	0	0	0	0	0	12	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Qp H1	1	0	0	Glazunova et al.
CB 58	Spleen abscess	0	0	0	0	0	0	0	1	Chronic, spleen abscess	0	0	0	0	0	1	0	0	0	Lyon, France, 1996	1	0	0	0	0	0	0	0	0	0	12	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Qp H1	1	0	0	Glazunova et al.
CB 76	Human heart valve	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Paris, France, 1998	1	0	0	0	0	0	0	0	0	0	12	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Qp H1	1	0	0	Glazunova et al.
CB 105	Human heart valve	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Montpellier, France, 2001	1	0	0	0	0	0	0	0	0	0	12	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Qp H1	1	0	0	Glazunova et al.
CB 33	Human heart	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Clermont-Ferrand,	1	0	0	0	0	0	0	0	0	0	12	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Qp H1	1	0	0	Glazunova

	valve																			France, 1992																																										et al.
CB 2	Human blood	0	0	1	0	0	0	0	0	Immunodepression, blood									1	Toulouse, France,	1	0	0	0	0	0	0	0	0	0	13	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	Qp H1	1	0	0	Glazunova et al.
CB 85	Human blood	0	0	1	0	0	0	0	0	Chronic, blood	0	0	0	0	0	0	0	1	0	Tours, France, 1999	1	0	0	0	0	0	0	0	0	0	14	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	Qp H1	1	0	0	Glazunova et al.
CB 74	Valvular prosthesis	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Toulouse, France, 1998	1	0	0	0	0	0	0	0	0	0	14	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	Qp H1	1	0	0	Glazunova et al.
CB 59	Aneurysm	0	0	0	1	0	0	0	0	Vascular infection, aneurysm	0	0	0	1	0	0	0	0	0	Saint-Etienne, France, 1996	1	0	0	0	0	0	0	0	0	0	14	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	Qp H1	1	0	0	Glazunova et al.
CB 80	Node	0	0	0	0	0	0	1	0	Chronic, node	0	0	0	0	0	0	1	0	0	Niort, France, 1999	1	0	0	0	0	0	0	0	0	0	14	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	Qp H1	1	0	0	Glazunova et al.
CB 102	Valvular prosthesis	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Poitiers, France, 2001	1	0	0	0	0	0	0	0	0	0	15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	Qp H1	1	0	0	Glazunova et al.
CB 11	Human blood	0	0	1	0	0	0	0	0	Acute, blood	0	1	0	0	0	0	0	0	0	Marseille, France	1	0	0	0	0	0	0	0	0	0	16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	Qp H1	1	0	0	Glazunova et al.
CB 23	Human	0	0	1	0	0	0	0	0	Chronic, blood	0	0	0	0	0	0	0	1	0	Clermont-	1	0	0	0	0	0	0	0	0	0	16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	Qp H1	1	0	0	Glazuno

	blood																			Ferrand, France, 1988																																										va et al.
CB 25	Human blood	0	0	1	0	0	0	0	0	Acute, blood	0	1	0	0	0	0	0	0	0	Paris, France, 1991	1	0	0	0	0	0	0	0	0	0	16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	Qp H1	1	0	0	Glazunova et al.
CB 77	Human heart valve	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Paris, France, 1998	1	0	0	0	0	0	0	0	0	0	17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	Qp H1	1	0	0	Glazunova et al.
CB 100	Human blood	0	0	1	0	0	0	0	0	Chronic, blood	0	0	0	0	0	0	0	1	0	Strasbourg, France	1	0	0	0	0	0	0	0	0	0	18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	Qp H1	1	0	0	Glazunova et al.
CB 48	Human placenta	0	0	0	0	1	0	0	0	Abortion, placenta	1	0	0	0	0	0	0	0	0	Grenoble, France, 1994	1	0	0	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	Qp H1	1	0	0	Glazunova et al.
CB 50	Valvular prosthesis	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Paris, France, 1994	1	0	0	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	Qp H1	1	0	0	Glazunova et al.
CB 66	Aneurysm	0	0	0	1	0	0	0	0	Vascular infection, aneurysm	0	0	0	1	0	0	0	0	0	Marseille, France, 1996	1	0	0	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	Qp H1	1	0	0	Glazunova et al.
CB 72	Valvular prosthesis	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Paris, France, 1996	1	0	0	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	Qp H1	1	0	0	Glazunova et al.

CB 78	Valvular prosthesis	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Marseille, France, 1998	1	0	0	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	Qp H1	1	0	0	Glazunova et al.
CB 88	Human heart valve	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Lyon, France, 1999	1	0	0	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	Qp H1	1	0	0	Glazunova et al.
CB 90	Human heart valve	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Lyon, France, 2000	1	0	0	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	Qp H1	1	0	0	Glazunova et al.
CB 4	Human blood	0	0	1	0	0	0	0	0	Chronic, blood	0	0	0	0	0	0	0	1	0	Montpellier, France, 1988	1	0	0	0	0	0	0	0	0	0	21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	Qp H1	1	0	0	Glazunova et al.
CB 7	Human heart valve	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Marseille, France	1	0	0	0	0	0	0	0	0	0	21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	Qp H1	1	0	0	Glazunova et al.
CB 69	Vegetation	0	1	0	0	0	0	0	0	Endocarditis, vegetation	0	0	1	0	0	0	0	0	0	Toulouse, France	1	0	0	0	0	0	0	0	0	0	13	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	Qp H1	1	0	0	Leroy et al.
Fr0 15	Human heart valve	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Toulouse, France	1	0	0	0	0	0	0	0	0	0	33	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	Qp H1	1	0	0	This study
Fr0 51	Human heart valve	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Paris, France	1	0	0	0	0	0	0	0	0	0	33	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	Qp H1	1	0	0	This study
Cb 120	Human blood	0	0	1	0	0	0	0	0	Chronic, blood	0	0	0	0	0	0	0	1	0	Villeneuve ,France, 2004	1	0	0	0	0	0	0	0	0	0																												Qp H1	1	0	0	This study

Cb130	Human blood	0	0	1	0	0	0	0	0	Non available									1	Tours, France	1	0	0	0	0	0	0	0	0	0																												QpH1	1	0	0	This study
Cb132	Human heart valve	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Toulouse, France	1	0	0	0	0	0	0	0	0	0																												QpH1	1	0	0	This study
Cb143	Human heart valve	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Angers, France, 2008	1	0	0	0	0	0	0	0	0	0																												QpH1	1	0	0	This study
CB122	Human placenta	0	0	0	0	1	0	0	0	Abortion, placenta	1	0	0	0	0	0	0	0	0	Limoge, France	1	0	0	0	0	0	0	0	0	0	22	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	QpH1	1	0	0	This study
CB13	Human blood	0	0	1	0	0	0	0	0	Chronic, blood	0	0	0	0	0	0	0	1	0	Paris, France	1	0	0	0	0	0	0	0	0	0	11	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	QpH1/QpDV	1	1	0	Glazunova et al.
Dyer	Human blood	0	0	1	0	0	0	0	0	Acute, blood	0	1	0	0	0	0	0	0	0	USA, 1938	0	0	0	0	0	0	0	0	0	1	16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	QpH1	1	0	0	Glazunova et al.
CB112	Human heart valve	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Zurich, Switzerland, 2003	0	1	0	0	0	0	0	0	0	0	12	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	QpH1	1	0	0	Glazunova et al.
CB109	Human heart valve	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Berlin, Germany, 2002	0	0	0	0	0	0	0	1	0	0	12	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	QpH1	1	0	0	Glazunova et al.
Brasov	Human	0	0	0	0	0	0	0	0	Acute, unknown	0	1	0	0	0	0	0	0	0	Romania	0	0	0	0	0	0	0	1	0	0	16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	QpH1	1	0	0	Glazunova et al.

Bangui	Human blood	0	0	1	0	0	0	0	0	Acute, blood	0	1	0	0	0	0	0	0	0	Central Africa	0	0	0	0	0	0	0	0	1	0	16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	Qp H1	1	0	0	Glazunova et al.
Dyer	Human blood	0	0	1	0	0	0	0	0	Acute, blood	0	1	0	0	0	0	0	0	0	USA, 1938	0	0	0	0	0	0	0	0	0	1	16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	Qp H1	1	0	0	Glazunova et al.
Henzerling	Human blood	0	0	1	0	0	0	0	0	Acute, blood	0	1	0	0	0	0	0	0	0	Italy/Slovakia, 1945	0	0	0	1	0	0	0	0	0	0	18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	Qp H1	1	0	0	Glazunova et al.
Balaceanu	Human	0	0	0	0	0	0	0	0	Acute, unknown	0	1	0	0	0	0	0	0	0	Romania	0	0	0	0	0	0	0	1	0	0	18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	Qp H1	1	0	0	Glazunova et al.
Geier	Human	0	0	0	0	0	0	0	0	Acute, unknown	0	1	0	0	0	0	0	0	0	Romania	0	0	0	0	0	0	0	1	0	0	18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	Qp H1	1	0	0	Glazunova et al.
Heizberg	Human	0	0	0	0	0	0	0	0	Acute, unknown	0	1	0	0	0	0	0	0	0	Greece	0	0	0	0	0	0	0	1	0	0	18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	Qp H1	1	0	0	Glazunova et al.
Cs-Florian	Human blood	0	0	1	0	0	0	0	0	Non available									1	Slovakia, 1956	0	0	0	0	0	0	0	1	0	0	18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	Qp H1	1	0	0	Glazunova et al.
CB119	Human heart valve	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Senegal, 2004	0	0	0	0	0	0	0	0	1	0	19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	Qp H1	1	0	0	Glazunova et

																																																														al.
Q229	Human heart valve	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Nova Scotia, Canada, 1982	0	0	0	0	0	0	0	0	0	1	21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	Qp H1	1	0	0	Glazunova et al.
Oufa-1	Human blood	0	0	1	0	0	0	0	0	Non available									1	Oufa, Russia	0	0	0	0	0	0	1	0	0	0	23	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	Qp H1	1	0	0	Glazunova et al.
Vologda-2	Human blood	0	0	1	0	0	0	0	0	Non available									1	Vologda, Russia, 1987	0	0	0	0	0	0	1	0	0	0	23	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	Qp H1	1	0	0	Glazunova et al.
Banguia	Human serum	0	0	1	0	0	0	0	0	Acute, blood	0	1	0	0	0	0	0	0	0	Central America	0	0	0	0	0	0	0	0	0	1																												Qp H1	1	0	0	Hotta et al.
Henzerlinga	Human serum	0	0	1	0	0	0	0	0	Acute, blood	0	1	0	0	0	0	0	0	0	Italy	0	0	0	1	0	0	0	0	0	0																												Qp H1	1	0	0	Hotta et al.
307e	Human serum	0	0	1	0	0	0	0	0	Acute, blood	0	1	0	0	0	0	0	0	0	Japan	0	0	0	0	0	1	0	0	0	0																												Qp H1	1	0	0	Hotta et al.
HzR	Blood	0	0	1	0	0	0	0	0	Non available									1	Italy	0	0	0	1	0	0	0	0	0	0	29	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	Qp H1	1	0	0	Leroy et al.
HzS	Blood	0	0	1	0	0	0	0	0	Non available									1	Italy	0	0	0	1	0	0	0	0	0	0	18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	Qp H1	1	0	0	Leroy et al.
E070891	Bronchoalveolar lavage	1	0	0	0	0	0	0	0	Non available									1	Tilburg, Netherlands	0	0	0	0	1	0	0	0	0	0	33	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	Qp H1	1	0	0	This study
B62614	Sputum	1	0	0	0	0	0	0	0	Non available									1	Houten, Netherlands	0	0	0	0	1	0	0	0	0	0	33	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	Qp H1	1	0	0	This stud

9																																																														y
B627833	Bronchoalveolar lavage	1	0	0	0	0	0	0	0	Non available									1	A'foot, Netherlands	0	0	0	0	1	0	0	0	0	0	33	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	Qp H1	1	0	0	This study
B6248808	Sputum	1	0	0	0	0	0	0	0	Non available									1	Houten, Netherlands	0	0	0	0	1	0	0	0	0	0	33	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	Qp H1	1	0	0	This study
B6338881	Sputum	1	0	0	0	0	0	0	0	Non available									1	Houten, Netherlands	0	0	0	0	1	0	0	0	0	0	33	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	Qp H1	1	0	0	This study
90608	Sputum	1	0	0	0	0	0	0	0	Non available									1	Ravenstein, Netherlands	0	0	0	0	1	0	0	0	0	0	33	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	Qp H1	1	0	0	This study
90427	Sputum	1	0	0	0	0	0	0	0	Non available									1	Wijchen, Netherlands	0	0	0	0	1	0	0	0	0	0	33	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	Qp H1	1	0	0	This study
V09028610	Human heart valve	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Druten, Netherlands	0	0	0	0	1	0	0	0	0	0	33	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	Qp H1	1	0	0	This study
09M073375	Human heart valve	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Standdaarbuiten, Netherlands	0	0	0	0	1	0	0	0	0	0	New																											Qp H1	1	0	0	This study
Cb115	Human heart valve	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Alberta, Canada	0	0	0	0	0	0	0	0	0	1	21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	Qp H1	1	0	0	This study
Cb121	Human heart valve	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Basel, Switzerland, 2004	0	1	0	0	0	0	0	0	0	0																												Qp H1	1	0	0	This study
Cb137	Human heart valve	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Minnesota USA, 2007	0	0	0	0	0	0	0	0	0	1																												Qp H1	1	0	0	This study
Cb144	Human heart valve	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Senegal, 2008	0	0	0	0	0	0	0	0	1	0																												Qp H1	1	0	0	This study

CB 93	Human placenta	0	0	0	0	1	0	0	0	Abortion, placenta	1	0	0	0	0	0	0	0	0	Dreux, France, 2000	1	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Qp RS	0	0	1	Glazunova et al.
CB 20	Human blood	0	0	1	0	0	0	0	0	Acute, blood	0	1	0	0	0	0	0	0	0	Salon de Provence, France, 1991	1	0	0	0	0	0	0	0	0	0	4	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Qp RS	0	0	1	Glazunova et al.
CB 111	Human heart valve	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Marseille, France, 2003	1	0	0	0	0	0	0	0	0	0	5	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Qp RS	0	0	1	Glazunova et al.
CB 35	Human heart valve	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Paris, France, 1992	1	0	0	0	0	0	0	0	0	0	5	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Qp RS	0	0	1	Glazunova et al.
CB 45	Human heart valve	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Paris, France, 1993	1	0	0	0	0	0	0	0	0	0	6	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Qp RS	0	0	1	Glazunova et al.
CB 43	Human heart valve	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Paris, France, 1993	1	0	0	0	0	0	0	0	0	0	7	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Qp RS	0	0	1	Glazunova et al.
CB 10	Human heart valve	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Grenoble, France	1	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Qp RS	0	0	1	Glazunova et al.
CB 9	Human blood	0	0	1	0	0	0	0	0	Chronic, blood	0	0	0	0	0	0	0	1	0	Lyon, France	1	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Qp RS	0	0	1	Glazunova et

																																																												al.
CB 31	Human heart valve	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Marseille, France, 1992	1	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	QpRS	0	0	1	Glazunova et al.
CB 34	Human blood	0	0	1	0	0	0	0	0	Chronic, blood	0	0	0	0	0	0	0	1	0	Marseille, France, 1992	1	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	QpRS	0	0	1	Glazunova et al.
CB 44	Human heart valve	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Créteil, France, 1993	1	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	QpRS	0	0	1	Glazunova et al.
CB 53	Aneurysm	0	0	0	1	0	0	0	0	Vascular infection, aneurysm	0	0	0	1	0	0	0	0	0	Marseille, France, 1995	1	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	QpRS	0	0	1	Glazunova et al.
CB 61	Valvular prosthesis	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Marseille, France, 1996	1	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	QpRS	0	0	1	Glazunova et al.
CB 70	Human heart valve	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Grenoble, France, 1997	1	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	QpRS	0	0	1	Glazunova et al.
CB 71	Valvular prosthesis	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Saint-Laurent du Var, France, 1997	1	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	QpRS	0	0	1	Glazunova et al.
CB 73	Valvular prosth	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Marseille, France,	1	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	QpRS	0	0	1	Glazunova

	esis																			1998																																										et al.
CB 91	Valvular prosthesis	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Marseille, France, 2000	1	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Qp RS	0	0	1	Glazunova et al.
CB 95	Human blood	0	0	1	0	0	0	0	0	Chronic, blood	0	0	0	0	0	0	0	1	0	Marseille, France, 2000	1	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Qp RS	0	0	1	Glazunova et al.
CB 116	Aneurysm	0	0	0	1	0	0	0	0	Vascular infection, aneurysm	0	0	0	1	0	0	0	0	0	Marseille, France, 2003	1	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Qp RS	0	0	1	Glazunova et al.
CB 107	Human heart valve	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Tours, France, 2001	1	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Qp RS	0	0	1	Glazunova et al.
CB 96	Human heart valve	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Marseille, France, 2000	1	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Qp RS	0	0	1	Glazunova et al.
CB 114	Human heart valve	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Marseille, France, 2003	1	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Qp RS	0	0	1	Glazunova et al.
CB 15	Human heart valve	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Lyon, France, 1991	1	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Qp RS	0	0	1	Glazunova et al.
CB 99	Human	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Marseille,	1	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Qp RS	0	0	1	Glazuno

	heart valve																			France, 2000																																										va et al.
CB106	Human heart valve	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Toulouse, France, 2001	1	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Qp RS	0	0	1	Glazunova et al.
CB79	Human heart valve	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Paris, France, 1999	1	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Qp RS	0	0	1	Glazunova et al.
CB98	Human heart valve	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Marseille, France, 2000	1	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Qp RS	0	0	1	Glazunova et al.
CB8	Human heart valve	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Marseille, France, 1990	1	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Qp RS	0	0	1	Glazunova et al.
CB117	Human heart valve	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Marseille, France, 2004	1	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Qp RS	0	0	1	Glazunova et al.
CB32	Human heart valve	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Lyon, France, 1992	1	0	0	0	0	0	0	0	0	0	9	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Qp RS	0	0	1	Glazunova et al.
CB92	Human heart valve	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Marseille, France, 2000	1	0	0	0	0	0	0	0	0	0	9	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Qp RS	0	0	1	Glazunova et al.

CB49	Human heart valve	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Marseille, France, 1994	1	0	0	0	0	0	0	0	0	0	9	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	QpRS	0	0	1	Glazunova et al.
CB65	Human heart valve	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Marseille, France, 1996	1	0	0	0	0	0	0	0	0	0	10	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	QpRS	0	0	1	Glazunova et al.
f2	Human blood	0	0	1	0	0	0	0	0	Hepatitis, blood									1	France	1	0	0	0	0	0	0	0	0	0																												QpRS	0	0	1	Unknown
f4	Human blood	0	0	1	0	0	0	0	0	Endocarditis, blood	0	0	1	0	0	0	0	0	0	France	1	0	0	0	0	0	0	0	0	0																												QpRS	0	0	1	Unknown
Cb16	Aneurysm	0	0	0	1	0	0	0	0	Vascular infection, aneurysm	0	0	0	1	0	0	0	0	0	France, 1993	1	0	0	0	0	0	0	0	0	0																												QpRS	0	0	1	This study
Cb84	Human blood	0	0	1	0	0	0	0	0	Chronic, blood	0	0	0	0	0	0	0	1	0	France	1	0	0	0	0	0	0	0	0	0																												QpRS	0	0	1	This study
Cb127	Aneurysm	0	0	0	1	0	0	0	0	Vascular infection, aneurysm	0	0	0	1	0	0	0	0	0	Poitiers, France, 2005	1	0	0	0	0	0	0	0	0	0																												QpRS	0	0	1	This study
Cb131	Human bone	0	0	0	0	0	1	0	0	Osteitis, bone	0	0	0	0	1	0	0	0	0	Toulouse, France	1	0	0	0	0	0	0	0	0	0	9	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	QpRS	0	0	1	This study
Cb134	Human heart valve	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Marseille, France	1	0	0	0	0	0	0	0	0	0																												QpRS	0	0	1	This study
Cb135	Human heart valve	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Paris, France	1	0	0	0	0	0	0	0	0	0																												QpRS	0	0	1	This study
Cb151	Human heart	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Marseille, France,	1	0	0	0	0	0	0	0	0	0																												QpRS	0	0	1	This stud

	valve																			2009																																										y
Leningrad-2	Human blood	0	0	1	0	0	0	0	0	Non available									1	Leningrad, Russia, 1955	0	0	0	0	0	0	1	0	0	0	7	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Qp RS	0	0	1	Glazunova et al.
Leningrad-4	Human blood	0	0	1	0	0	0	0	0	Non available									1	Leningrad, Russia, 1957	0	0	0	0	0	0	1	0	0	0	7	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Qp RS	0	0	1	Glazunova et al.
CB 81	Human heart valve	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Madrid, Spain, 1999	0	0	1	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Qp RS	0	0	1	Glazunova et al.
CB 47	Human heart valve	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	Barcelona, Spain, 1994	0	0	1	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Qp RS	0	0	1	Glazunova et al.
Termez	Human blood	0	0	1	0	0	0	0	0	Non available									1	Uzbekistan, 1952	0	0	0	0	0	0	1	0	0	0	26	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	Qp RS	0	0	1	Glazunova et al.
KA Vd	Human heart valve	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	USA	0	0	0	0	0	0	0	0	0	1																												Qp RS	0	0	1	Hotta et al.
PA Vd	Human heart valve	0	1	0	0	0	0	0	0	Endocarditis, valve	0	0	1	0	0	0	0	0	0	USA	0	0	0	0	0	0	0	0	0	1																												Qp RS	0	0	1	Hotta et al.

Reevaluation of the risk of fetal death and malformation after Q Fever.

Million M[1], Roblot F[2], Carles D[3], D'Amato F[1], Protopopescu C[4], Carrieri MP[4], Raoult D[1]

[1]Aix-Marseille Université, Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes, UM63, CNRS 7278, IRD 198, INSERM 1095, Marseille.

[2]Service de Maladies Infectieuses, Hôpital La Milétrie, INSERM 1070, Poitiers.

[3]Service d'Anatomie Pathologique, Université de Bordeaux.

[4]INSERM, UMR912 (SESSTIM) Aix Marseille Université, UMR_S912, IRD ORS PACA, Observatoire Régional de la Santé Provence-Alpes-Côte-d'Azur, Marseille, France.

Reevaluation of the Risk of Fetal Death and Malformation After Q Fever

Matthieu Million,[1] France Roblot,[2] Dominiques Carles,[3] Felicetta D'Amato,[1] Camelia Protopopescu,[4,5,6] Maria Patrizia Carrieri,[4,5,6] and Didier Raoult[1]

[1]Aix-Marseille Université, Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes, UM63, CNRS 7278, IRD 198, INSERM 1095, Marseille; [2]Service de Maladies Infectieuses, Hôpital La Milétrie, INSERM 1070, Poitiers; [3]Service d'Anatomie Pathologique, Université de Bordeaux; and [4]INSERM, UMR912 (SESSTIM), [5]Aix Marseille Université, UMR_S912, IRD, and [6]ORS PACA, Observatoire Régional de la Santé Provence-Alpes-Côte-d'Azur, Marseille, France

A meta-analysis of 136 Q fever pregnancies, including 4 new cases and 7 population-based serological studies, revealed significant increases in fetal death and malformation after Q fever during pregnancy. This poor obstetric outcome is prevented by antibiotic treatment.

Keywords. Q fever; *Coxiella burnetii*; pregnancy; fetal death; fetal malformation.

Q fever is a zoonosis caused by the bacterium *Coxiella burnetii* and is associated with epizootic abortion in ungulates. In humans, the role of Q fever during pregnancy has been recently questioned because of the discrepancy between the high risk of obstetric complications among women infected with Q fever in published case series [1, 2] and the absence of increased risk of adverse pregnancy outcomes in population-based serological studies [3–6]. To elucidate the role of Q fever in fetal death and malformation, we cultured the bacterium from a new malformed aborted fetus. We performed a meta-analysis to study the links between Q fever and fetal death and malformation and analyzed study design bias by comparing clinical case series and population-based serological studies. We also examined research center bias using subgroup analyses differentiating studies from our and other centers.

Received 8 November 2013; accepted 10 April 2014; electronically published 18 April 2014.
Correspondence: Didier Raoult, MD, PhD, Unité des Rickettsies, CNRS UMR 6020, Université de la Méditerranée, 27 Boulevard Jean Moulin, Marseille 13385, France (didier.raoult@gmail.com).

Clinical Infectious Diseases 2014;59(2):256–60

DOI: 10.1093/cid/ciu259

METHODS

A literature search, unrestricted by date or language, was performed using PubMed, Google, and Google Scholar with the following keywords: human, *Coxiella burnetii*, Q fever, pregnancy, pregnant, fetal death, abortion, and fetal malformation.

Clinical Case Series of Q Fever During Pregnancy

A clinical case was defined as an infection diagnosed by a medical doctor. Women were considered to be "treated" if they received at least 1 of the following antibiotics: cotrimoxazole, tetracyclines, macrolides, or fluoroquinolones. Fetal death was defined as the sum of spontaneous abortions, stillbirths, intrauterine fetal deaths, and deaths in the hours following delivery. Fetal malformations were defined according to the International Clearinghouse for Birth Defect Surveillance and Research [7]. Fetal death rates were compared with the rate in the general population estimated by the largest existing population-based study [8]. Fetal malformations were compared with 2 French registries [7]. The cumulative incidence ratios were estimated by age-controlled Mantel-Haenszel stratification using Stata statistical software, version 12 (StataCorp, College Station, Texas). The cumulative incidence ratio of fetal malformation was computed without age stratification because of the low number of cases. Meta-analysis was performed with the Mantel-Haenszel method and a random-effects model using Review Manager version 5.2 (Cochrane Collaboration, Copenhagen, Denmark). Heterogeneity was assessed using I^2 analysis. Publication bias was estimated by funnel plot, Egger regression intercept, and Duval and Tweedie trim and fill analyses.

Population-Based Serological Studies

Population-based serological studies were defined as studies in which an investigator screened sera from pregnant women for *C. burnetii*–positive serology (using investigator-defined criteria) and evaluated the results for an association with fetal death. A meta-analysis was performed as described above. Study design bias was tested by comparing the summary odds ratios of the association between Q fever and fetal death in both clinical case series and population-based serological studies. The methods are described in further detail in the Supplementary Data.

RESULTS

In April 2012, a 20-year-old woman presented with fever, cough, and flu-like symptoms. A 6-week pregnancy was

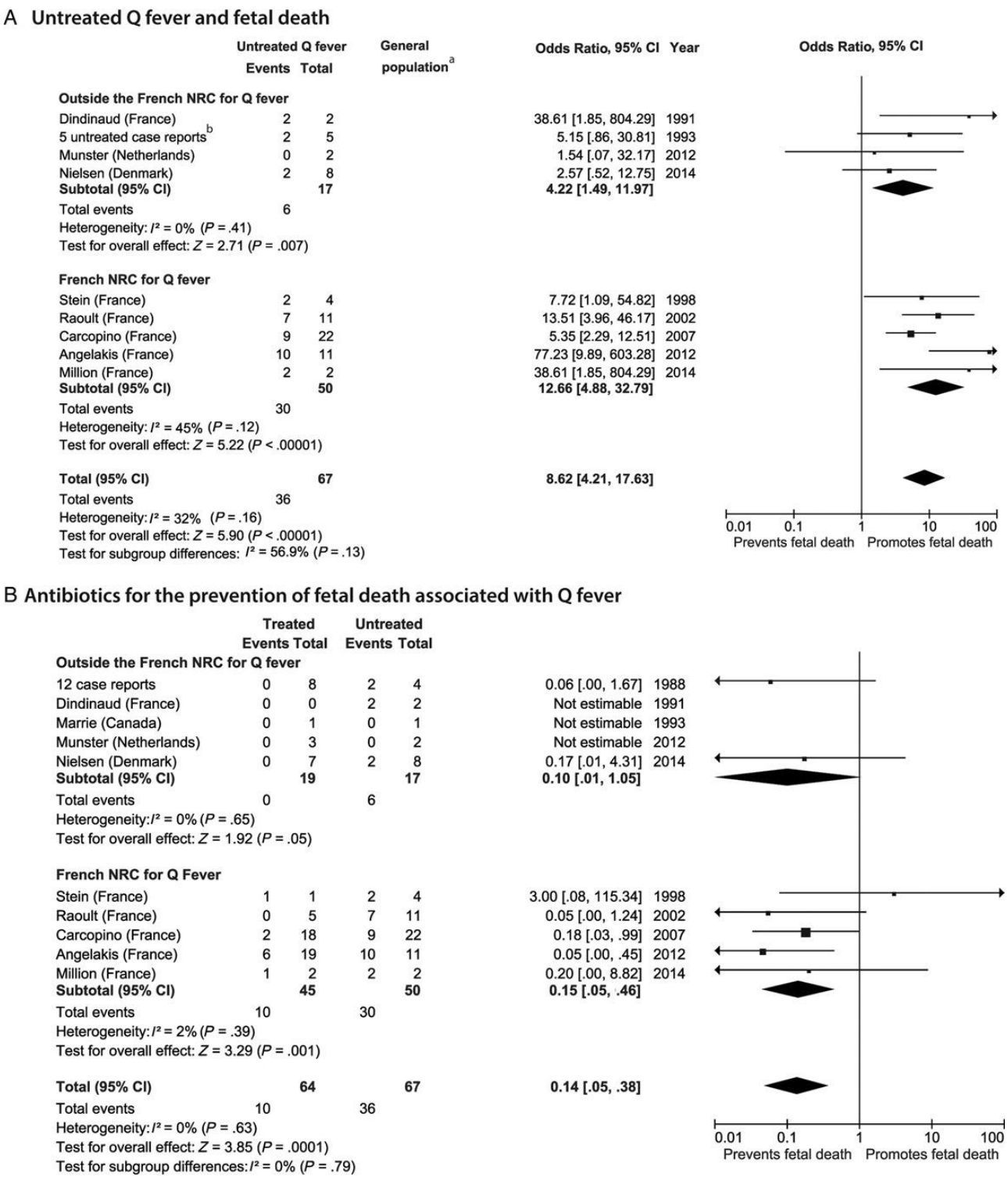


Figure 1. Q fever and pregnancy. *A*, Significantly increased fetal death rate was found in untreated Q fever inside (62%; odds ratio [OR], 12.7; $P < .00001$) or outside (35%; OR, 4.22; $P = .007$) the French national referral center (NRC) for Q fever vs the general population (11%) [8]. The difference was not significant between our center and the other centers ($P = .16$). [a]Data from the largest population-based study published to date have been considered as the reference (104 840 fetal deaths for 914 485 pregnancies [11%]) [8]. [b]Including 1 untreated case from the 2 cases reported by Marrie et al [9]. *B*, Antibiotic treatment significantly reduces the death rate when administered to pregnant women with Q fever inside (OR, 0.15) or outside (OR, 0.10) the French national referral center for Q fever.

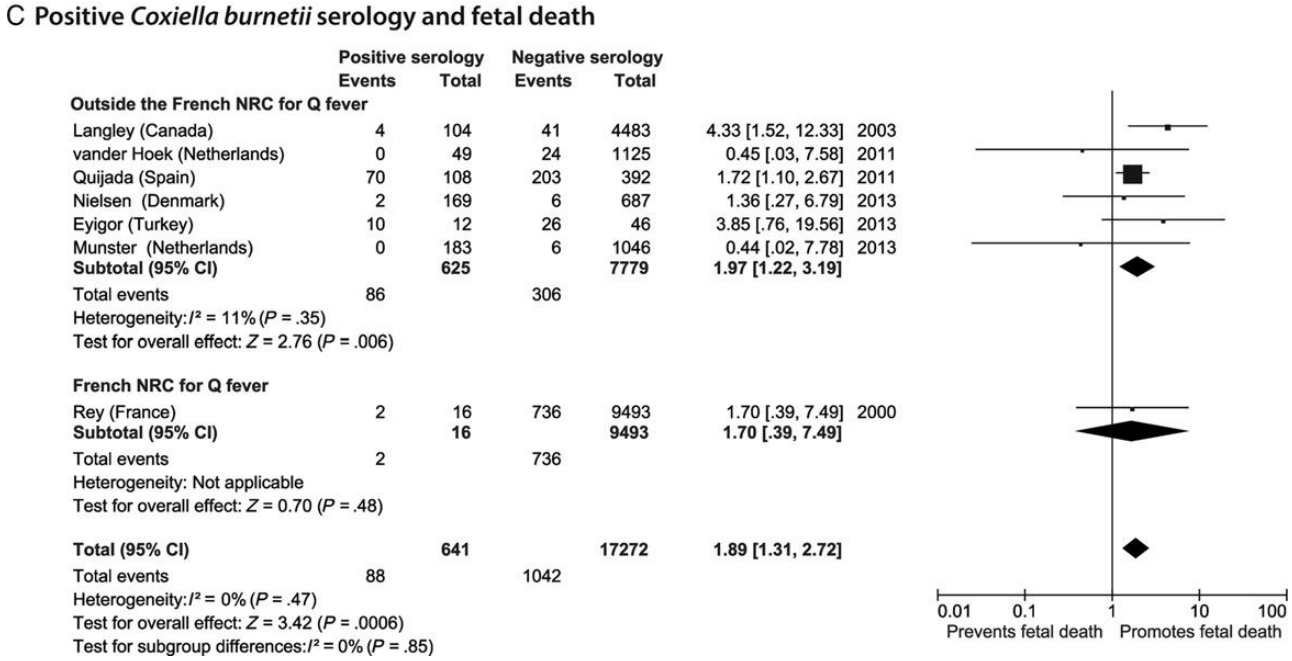


Figure 1 continued. *C*, A significant association between positive *Coxiella burnetii* serology and fetal loss was found inside (OR, 1.7; $P = .48$) and outside (OR, 2.0; $P = .006$) the French NRC for Q fever. Summary size effect of the link between Q fever and fetal death was significantly different between clinical case series (*A*) and population-based serological studies (*C*) (OR, 8.6; 95% confidence interval [CI], 4.2–17.6 vs OR, 1.9; 95% CI, 1.3–2.7; $P < .05$), evidencing a study design bias.

diagnosed serendipitously. Two months later, an anterior wall-closing defect was observed in the fetus, whereas the mother was completely asymptomatic. Amniocentesis identified a normal karyotype. At 18 weeks of pregnancy, serology was positive for Q fever, and cotrimoxazole was prescribed. However, the patient experienced spontaneous abortion just before the initiation of treatment. Analysis of the fetus revealed a large omphalocele containing the liver and small bowel, along with adrenal hypoplasia. Placental examination revealed ischemic lesions with acute intervillitis and microabscesses. The placenta and fetus were positive for *C. burnetii*–specific polymerase chain reaction; the bacterium was isolated from the placenta (minimal inhibitory concentration of 0.25 g/L for doxycycline), and the strain was sequenced (NCBI BioProject: PRJEB4272).

Analysis of all published cases and 4 new cases diagnosed in our center identified 136 relevant cases of pregnant women with Q fever in 10 case series and 13 case reports. Controlling for age, the fetal death rate in the untreated Q fever population was higher than that in the general population (cumulative incidence ratio of 3.9; 95% confidence interval [CI], 3.2–4.8; $P < .001$), but no significant difference was observed between the treated Q fever population and the general population (cumulative incidence ratio = 1.2; 95% CI, .7–2.1; $P = .48$). Among the 136 cases, 7 (5%) malformed fetuses were observed including hypospadias (2 cases, 1 associated with hydrocele), Potter syndrome with bilateral agenesis (1), congenital hydronephrosis (1), syndactyly (1), and the case described in the present study. Nielsen et al [2] observed a severely malformed fetus, but details were not provided. The cumulative incidence ratio of fetal malformations for Q fever patients compared with the general population was estimated as 16.1 (95% CI, 6.1–42.7; $P = .0001$).

Meta-analysis also revealed a significant increase in the fetal death rate in untreated Q fever pregnancies compared with pregnancies in the general population (53% and 11%, respectively; odds ratio [OR], 8.6; 95% CI, 4.2–17.6; $P < .00001$; $I^2 = 32\%$; Figure 1*A*). After the exclusion of 1 outlier [1], this increase remained significant, and publication bias was unlikely (Egger test, $P = .35$). Based on Duval and Tweedie trim and fill analysis, 1 study was added, resulting in a similar increase in death rate (45%; $P < .05$). This increased fetal death rate was also observed in Q fever pregnancies diagnosed outside of our center compared with the general population (35% and 11%, respectively; OR, 4.2; 95% CI, 1.5–12.0; $P = .007$; $I^2 = 0\%$).

The preventive effects of antibiotics on fetal death in Q fever during pregnancy were consistent in case series published by our center (OR, 0.15; 95% CI, .05–.46; $P = .001$; $I^2 = 2\%$) and by other centers (OR, 0.10; 95% CI, .01–1.05; $P = .05$; $I^2 = 0\%$, Figure 1*B*). Publication bias was unlikely based on analyses including funnel plot, Egger test ($P = .66$), and Duval and Tweedie trim and fill procedure.

The association between *C. burnetii*–positive serology and fetal death was analyzed in 7 population-based serological studies and 17 913 pregnancies after the exclusion of 1 outlier [10] (I^2 decreased from 47% to 0%). Significant associations were observed between *C. burnetii*–positive serology and fetal death both with (OR, 1.9; 95% CI, 1.3–2.7; $P = .0006$) and without the only study from our center (OR, 2.0; 95% CI, 1.2–3.2; $P = .006$; Figure 1*C*). Publication bias was unlikely based on Egger test ($P = .9$) and Duval and Tweedie trim and fill procedure. Overall, the risk of fetal death was significantly higher ($P < .05$) in Q fever case series (OR, 8.6; 95% CI, 4.2–17.6) than in seropositive pregnant women screened in population-based studies (OR, 1.9; 95% CI, 1.3–2.7). The results are described in further detail in the Supplementary Data.

DISCUSSION

Coxiella burnetii infection during pregnancy is a cause of poor obstetric prognosis based on the Bradford Hill criteria, including a 4- and 16-fold increased risk for fetal death and birth defects, respectively, and because fetal death has been observed in at least 7 different countries for >30 years. A temporal effect was observed because the prognosis is worse when the infection occurred in the first trimester, as in our case. The specificity of the disease and the underlying mechanism are supported by the typical placentitis aspects, including chronic intervillitis and microthrombus. Antiphospholipid antibodies, which are a key pathogenic factor in Q fever endocarditis [11] and are associated with early abortion [12], may also play a critical role in first-trimester fetal deaths. As seen in our patient, fetal infection might play a role in malformation and have an effect later in pregnancy. The biological gradient and reversibility of the disease are supported by the effectiveness of antibiotics. The fetal death and malformation data in animals meet the experimental criteria [13]. Finally, the consistency between current epidemiologic results and analysis of the fetus in this study is the strongest argument for a causal link between *C. burnetii* maternal infection and fetal death and malformation.

The absence of evidence of increased risk of abortion among women from the Netherlands [3, 4] or Denmark [5, 6] with *C. burnetii*–positive serology may be due to strain specificity, a bias in reporting or, most likely, study design (as shown in this study). Notably, the same group studying the same population observed no increased risk of adverse pregnancy outcomes in population-based serological studies [5, 6] but observed poor obstetric outcomes in Q fever case series [2, 14]. Although there is a risk of overrepresentation of pregnant women experiencing fetal death or malformations, 70% of the cases are from France, where Q fever screening in pregnant women (irrespective of problems during pregnancy) is more frequently performed on women living in at-risk areas. Therefore, the overrepresentation remains low and is less likely to affect estimates of the association of Q fever and pregnancy outcomes.

Therefore, it might be reasonable to screen all pregnant women who consult a doctor in endemic areas and to treat all pregnant women who have *C. burnetii*–positive serologies indicating either primary infection (immunoglobulin G [IgG] titers ≥200 and immunoglobulin M [IgM] titers ≥50 in phase II) or persistent infection (phase I IgG titers ≥800).

Supplementary Data

Supplementary materials are available at *Clinical Infectious Diseases* online (http://cid.oxfordjournals.org). Supplementary materials consist of data provided by the author that are published to benefit the reader. The posted materials are not copyedited. The contents of all supplementary data are the sole responsibility of the authors. Questions or messages regarding errors should be addressed to the author.

Notes

Acknowledgments. We thank Dr Goua for kindly providing Supplementary Figure 1; Dr Carcopino (obstetrician) for relevant comments on the manuscript; Dr Zandotti, Dr Baccard, and the French National Referral Laboratory for Epstein Barr-virus for helping us to identify a false-positive serology for this virus; and Mical Paul for expert assistance in the meta-analysis. We also thank the mother of the fetus for providing written consent for publication of this case and all related data, and Nielsen et al and the editor of *Emerging Infectious Diseases* for consenting to the use of data from an unpublished study currently in press [2].

Financial support. This work was supported by the French National Referral Center for Q fever.

Potential conflicts of interest. All authors: No reported conflicts.

All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

References

1. Angelakis E, Million M, D'Amato F, et al. Q fever and pregnancy: disease, prevention, and strain specificity. Eur J Clin Microbiol Infect Dis **2013**; 32:361–8.
2. Nielsen S, Molbak K, Henriksen T, Krogfelt K, Larsen C, Villumsen S. Q fever during pregnancy and maternofoetal consequences: a case series from Denmark. Emerg Infect Dis **2014**. In press.
3. Munster JM, Leenders AC, Hamilton CJ, et al. Routine screening for *Coxiella burnetii* infection during pregnancy: a clustered randomised controlled trial during an outbreak, the Netherlands, 2010. Euro Surveill **2013**; 18.
4. van der Hoek W, Meekelenkamp JC, Leenders AC, Wijers N, Notermans DW, Hukkelhoven CW. Antibodies against *Coxiella burnetii* and pregnancy outcome during the 2007–2008 Q fever outbreaks in the Netherlands. BMC Infect Dis **2011**; 11:44.
5. Nielsen SY, Andersen AM, Molbak K, et al. No excess risk of adverse pregnancy outcomes among women with serological markers of previous infection with *Coxiella burnetii*: evidence from the Danish National Birth Cohort. BMC Infect Dis **2013**; 13:87.
6. Nielsen SY, Hjollund NH, Andersen AM, et al. Presence of antibodies against *Coxiella burnetii* and risk of spontaneous abortion: a nested case-control study. PLoS One **2012**; 7:e31909.
7. International Clearinghouse for Birth Defects Surveillance and Research. Annual report 2010. Roma, Italy: The Centre of the International Clearinghouse for Birth Defects Surveillance and Research, **2010**. Available at: http://www.icbdsr.org/filebank/documents/ar2005/Report2010.pdf. Accessed 1 October 2013.

8. Nybo Andersen AM, Wohlfahrt J, Christens P, Olsen J, Melbye M. Maternal age and fetal loss: population based register linkage study. BMJ **2000**; 320:1708–12.
9. Marrie TJ. Q fever in pregnancy: report of two cases. Infect Dis Clin Pract. **1993**; 2:207–9.
10. Baud D, Peter O, Langel C, Regan L, Greub G. Seroprevalence of *Coxiella burnetii* and *Brucella abortus* among pregnant women. Clin Microbiol Infect **2009**; 15:499–501.
11. Million M, Raoult D. The pathogenesis of the antiphospholipid syndrome. N Engl J Med **2013**; 368:2335.
12. Giannakopoulos B, Krilis SA. The pathogenesis of the antiphospholipid syndrome. N Engl J Med **2013**; 368:1033–44.
13. Giroud A, Giroud P, Martinet M, Deluchat C. Inapparent maternal infection by *Coxiella burnetii* and fetal repercussions. Teratology **1968**; 1:257–62.
14. Nielsen SY, Molbak K, Hjollund NH. Q-fever caused spontaneous abortion. Ugeskr Laeger **2011**; 173:2053–4.

REEVALUATION OF THE RISK OF FETAL DEATH AND MALFORMATION AFTER Q FEVER

SUPPLEMENTARY DATA

METHODS

Diagnosis of Q fever in the present case

Serology, PCR, culture and sequencing were performed as previously reported [1].

Sequencing of the *Coxiella burnetii* strain was performed as previously reported [2].

Fetal death and malformation rate in untreated Q fever pregnancies compared to the general population and effectiveness of antibiotics for the prevention of fetal death associated with Q fever during pregnancy

Screening for clinical cases of Q fever during pregnancy

A clinical case of Q fever during pregnancy was defined as a case of infection diagnosed by a medical doctor based on clinical and microbiological data. To limit the publication bias for our cases, we add all the 4 cases diagnosed at the French National Referral center since our last publication (2011), and all cases before 2011 were systematically published in one of our 4 cases series [3-6]. In addition to our cases, we screened all case series and case reports of Q fever during pregnancy in the literature, regardless of the date and publication language, using PubMed, Google and Google scholar. The following keywords were used: human, *Coxiella burnetii*, Q fever, pregnancy, pregnant, fetal death, abortion and fetal malformation. Inclusion criteria were the articles for which a fetal outcome and the mother's age were mentioned, regardless of treatment or trimester of infection. Exclusion criteria were cases considered to be possible duplicates due to mention of an identical center and year of pregnancy, age of the mother and outcome and cases for whom another pathogen could be suspected as responsible

of fetal death or fetal malformation. In case of duplicates, the case was included in the oldest case series and excluded from the latest. Other exclusion criteria include cases for whom another cause of fetal death was found (coinfection with a pathogen known to be responsible for fetal death and/or malformation) and pregnancy with more than one fetus (twin, triplets). Collected data included country, cases diagnosed in our center or somewhere else, year of diagnosis, age of the mother, occurrence of fetal death, fetal malformation, and treatment. Several pregnancies could be included for the same woman.

Treatment

Women were considered as "treated" if they received at least one of the following antibiotics: cotrimoxazole, tetracyclines, macrolides and fluoroquinolones. Women for whom fetal death occurred before 7 days of treatment were considered as "without treatment". Cases for which this information was not available were excluded from the analysis on treatment efficacy.

Fetal death

Fetal death was defined as the sum of spontaneous abortion, stillbirth, intra-uterine fetal death and death in the hours following the delivery.

Fetal malformation

Fetal malformations were defined according to the international clearinghouse for birth defect surveillance and research [7]. Oligoamnios and hydrops fetalis were not considered fetal malformations.

General population

Fetal death was compared to the largest population based register study published to date, which analyzed the association between maternal age and fetal death and included 634,272 Danish women and 1,221,546 related pregnancy outcomes [8]. Ectopic pregnancies and induced abortions were excluded.

Fetal malformations were compared to two French registries [7]: the Paris monitoring system and the REMERA monitoring system, which corresponds to the central-east France Register of Congenital Malformation. Stillbirths of 22 gestation weeks or more were included. These 2 registries represent 94,000 annual births, which is approximately 13% of all births in France.

Association between a positive Q fever serology and fetal death

Screening of population-based serological studies

Population-based serological studies were defined as studies in which an investigator tested several sera for *Coxiella burnetii* antibodies and tested an association with fetal death. Instead of clinical cases, no medical diagnosis was performed to affirm infection. This difference is extremely important in Q fever, where the distinction between positive serology and infection is relatively difficult in some cases. We screened all studies when searching for an association between a positive Q fever serology during pregnancy and fetal death in the literature, regardless of the date and publication language, using PubMed, Google and Google Scholar. The following keywords were used: human, *Coxiella burnetii*, Q fever, serology, pregnancy, pregnant, fetal death, abortion and fetal malformation. Inclusion criteria were the articles for which a fetal outcome was mentioned regardless of treatment or trimester of infection. Exclusion criteria included studies corresponding to possible duplicates due to mention of an identical center. In cases of possible duplicates, the latest population-based serological study was included, considering that data from the oldest study were part of the more recent study.

Collected data included country, studies from our center or somewhere else and number of fetal death in each group with positive or negative Q fever serology.

Positive serology

Because techniques and thresholds differ between each study, a serology was considered positive according to the authors' criteria. However, the lack of standardization of techniques (ELISA, immunofluorescence, complement fixation test), thresholds and isotypes (IgG, IgM, IgA) needed to consider an active infection with *Coxiella burnetii* and the limited predictive value of serology alone requires interpreting the results of these population-based serological studies very carefully.

Fetal death

Fetal death was defined as the sum of spontaneous abortion, stillbirth, intra-uterine fetal death and death in the hours following the delivery.

Statistical analysis

Fetal death and malformation rate in untreated Q fever pregnancies compared to the general population and effectiveness of antibiotics for the prevention of fetal death associated with Q fever during pregnancy

Cumulative incidence ratios estimates and 95% confidence intervals measuring the associations between fetal death and the different exposures for the total, untreated and treated Q fever population compared to the general population, as defined above, were obtained via Mantel-Haenszel stratification [9] while controlling for the potential confounder effect of age. If the Mantel-Haenszel test of homogeneity across strata was not rejected in each of the three populations, combined (pooled) cumulative incidence ratios were computed using Mantel-

Haenszel weights. The Cochran-Mantel-Haenszel Chi-2 test [9] was used to assess the significance of the combined cumulative incidence ratios. In a sensitivity analysis, cumulative incidence ratios were alternatively estimated using robust Poisson regression [22], adjusted for the potential confounder effect of age. The cumulative incidence ratio of fetal malformation and the corresponding 2-sided Fisher's exact test p-value were computed for total Q fever patients vs. the general population without age-stratification, as the number of cases was very small. Analyses were performed using Stata statistical software, version SE 12.1 [10].

To better assess the center and publication biases, we performed a meta-analysis of all published case series, discriminating between case series published by our center and outside our center. Thirteen case reports were reported outside our center and were pooled as one case series. We first used meta-analysis to assess the fetal death rate in Q fever case series and to compare the fetal death rate between untreated cases from the French NRC, from other centers and from the general population. To give an estimate of the percentage of fetal death, we performed an event rate meta-analysis calculating the pooled event rate in each of these three groups and then calculated an odds ratio, taking the general population as the control group. We also used meta-analysis to assess the effectiveness of antibiotics in the prevention of fetal death in Q fever case series from the French NRC and from other centers. We evaluated whether this effectiveness was different between these two groups.

Association between positive Q fever serology and fetal death in population-based studies

Finally, another meta-analysis was performed to test if positive Q fever serology was associated with fetal death in population-based serological studies. Once again, we used a subgroup analysis to test whether the results were consistent inside and outside the French national referral center for Q fever.

Assessment of the study design bias

A study design bias was tested comparing the summary odds ratio of the association between Q fever and fetal death as assessed by clinical case series or by population-based serological studies. We used a continuous variable as a type of outcome and an inverse variance method. The effect measure was the mean difference. Sample size was the number of cases for case series and the number of serologies for population-based serological studies.

Meta-analyses

For the meta-analyses, we used a Mantel-Haenszel method with a random effects model. Odds ratio was defined as the effect measure because the software could not calculate cumulative incidence ratios, but odds ratios were, in our specific case (low incidence of the event), good estimates of cumulative incidence ratios. Heterogeneity was assessed by I^2. Publication bias was estimated by examination of the funnel plot, Egger's regression intercept and Duval and Tweedie's trim and fill procedure. Meta-analyses were performed using comprehensive meta-analysis software version 2 and Revman version 5.1 [11].

RESULTS

Case report

Serological history of the present malformed case is detailed in supplementary Table 1. Macroscopic examination, echographic results and genome of the *Coxiella burnetii* strain isolated from the placenta are detailed in Supplementary Figure 1.

Fetal death and malformation rate in untreated Q fever pregnancies compared to the general population and effectiveness of antibiotics for the prevention of fetal death associated with Q fever during pregnancy

We examined 35 studies that included clinical cases of Q fever during pregnancy (Supplementary Table 2 & Supplementary Figure 2). Ten were excluded, including 8 for which clinical individual data (age of the mother or fetal outcome) were not available and 2 possible duplicates. Twenty-five articles corresponding to 154 pregnancies were included in our review (Supplementary Table 3). Fourteen pregnancies were considered possible duplicates because they were published by the same center and with the same year of pregnancy, maternal age and fetal outcome. Fetal death was not attributable to Q fever with certainty in 4 cases due to *Chlamydia* co-infection [12], rubella co-infection [13] and the presence of twins [14] and triplets [15]. Finally, 136 pregnancies were included from 23 articles (10 case series and 13 case reports, Supplementary Table 4 & 6). The number of cases was higher in case series from the French NRC (median 16 [range 4-40]) that case series from other centers (4 [2-15], $P = .09$). Ninety-five cases (70%) were reported by the French NRC, and 41 (30%) by other centers. No treatment was administered in 67 cases (pregnancies), and 63 had been treated before delivery or fetal death (50 cotrimoxazole, 5 erythromycin alone, 3 azithromycin, 2 tetracycline, 1 ciprofloxacin, and two cases with association of erythromycin and rifampicin [16, 17]). It was not possible to determine whether an effective treatment for *Coxiella burnetii* was administered and if it was administered between the delivery and fetal death in 5 cases from 2 articles [13, 18]. The fetal death rate in the general population was estimated to be 11.5% (104840/914485) and increased with age (Supplementary Table 5 & Supplementary Figure 3). After controlling for age (Supplementary Table 7), Q fever during pregnancy was associated with a significant increase in fetal death (Supplementary Table 8). Using event rate meta-analysis, the pooled fetal death rate was 0.35 (95%CI 0.16-0.61)

outside our center, 0.62 (0.39-0.80) inside our center and 0.115 (0.114-0.115) for the general population. The difference between the case series from our center and other centers was not significant (Mixed effect analysis, Q-value 2.3, *P* = .13). When comparing case series from our center and other centers to the general population, both were associated with a significant fetal death rate increase (Outside our center, Q-value 7.3, P = .007 and inside our center, 28.4, P < .001). To improve the clarity of the MS, we finally chose to perform a comparative meta-analysis calculating odds ratio for each Q fever case series compared to data from 914485 pregnancies the general population [8] taken as a reference for each included studies (Figure 1, up). A funnel plot of studies included in the meta-analysis assessing the fetal death rate in untreated Q fever pregnancies compared to the general population is shown in Supplementary Figure 4. A funnel plot of studies included in the meta-analysis assessing the effectiveness of antibiotics for the prevention of fetal death associated with Q fever during pregnancy is shown in Supplementary Figure 5. A funnel plot of population-based serological studies included in the meta-analysis assessing the fetal death rate in women with a positive serology is shown in Supplementary Figure 6.

Malformation rate in untreated Q fever pregnancies compared to the general population

The comparison of the malformation rate in untreated Q fever pregnancies found an increased cumulative incidence ratio compared to the pooled data of two French registries (Supplementary Table 9).

Association between a positive Q fever serology and fetal death

We identified 11 studies analyzing an association between positive serology for Q fever during pregnancy and fetal outcome (Supplementary Table 2 & Supplementary Figure 2). Three studies were excluded because fetal death was not assessed in 2 studies [19, 20], and

one study was a possible duplicate because the same author published a second population-based serological study on the same population [21, 22] one year later (the first study was excluded). Eight studies were included in the meta-analysis (Supplementary Table 10), but one study was excluded [23] because it was an outlier when examined by funnel plot (Supplementary Figure 6) and was associated with a substantial (I^2=47%) heterogeneity. After exclusion of this study, heterogeneity was cancelled (I^2 = 0%, P = .47). In consequence, 7 population-based serological studies were ultimately included in the meta-analysis corresponding to 17,913 serologies (641 positive corresponding to 88 fetal deaths, and 17,272 negative corresponding to 1042 fetal deaths).

Supplementary Table 1. Serological and clinical evolution of the patient detailed in our MS

		Phase I			Phase II			PCR on the
Date		**IgG**	**IgM**	**IgA**	**IgG**	**IgM**	**IgA**	**Serum**
03/05/2012	Fever and pregnancy diagnosis	0	25	0	0	50	0	Negative
20/06/2012	Amniocentesis, no symptoms	800	200	50	1600	400	100	Negative
20/07/2012	Expert consultation and Q fever diagnosis	12800	100	25	25600	200	50	Negative
22/08/2012	Follow-up on month after delivery	3200	0	0	6400	0	0	NP

PCR: polymerase chain reaction. NP: not performed. Serology of May and June, 2012 were conducted afterwards. The titers of IgG, IgM and IgA against phase I and phase II *C. burnetii* antigens were quantified using an indirect immunofluorescence assay on the serum of the mother [40]. The sera were incubated with an RF absorbent (Siemens, Marburg, Germany) before IgM and IgA titrations to prevent the presence of rheumatoid factors.

Supplementary Table 2. Screened studies, included studies and reasons for exclusion of excluded studies

See enclosed Excel file. This table describes all studies obtained from the literature search (see methods in the supplementary data) and the present work. Studies included, excluded and reasons for exclusion are mentioned.

Supplementary Table 3. Screened patients from included studies, included patients and reasons for exclusion of excluded patients from included studies

See enclosed Excel file. This table describes all patients obtained from included studies with clinical data (at least age of the mother), if they are included in our "published Q fever pregnant population", and if excluded, the reason for exclusion.

Supplementary Table 4. Patients included in the "published Q fever pregnant population" to be compared to the general population

See enclosed Excel file. This table details all patients included in our "published Q fever population" with clinical data (at least age of the mother). Patients for whom information on treatment was available were analyzed in a second analysis comparing the pregnancy outcome of treated and untreated patients.

Supplementary Table 5. Fetal death rate in the general population as estimated by the largest published population-based study to date: Nybo Andersen, BMJ, 2000

See enclosed Excel file. This table details the fetal loss ratio according to age in the general population. These data were calculated from the general population data as assessed by the largest published population-based study to date on fetal loss (Nybo Andersen, BMJ, 2000 [8]).

Table S2

Reference	Clinical cases studies screened	Clinical studies included	Clinical studies excluded because clinical individual data not available (age of the mother, fetal outcome)	Clinical study excluded because duplicate	Seroprevalence studies screened	Seroprevalence studies included	Seroprevalence studies excluded because fetal death not assessed	Seroprevalence studies excluded because possible duplicate	Sum	Remarks
Angelakis, Eur J Clin Microbiol Infect Dis, 2012	1	1							1	
Babudieri, Adv Vet Sci, 1959	1		1						1	Age of each pregnant woman not mentioned
Baud, Clin Microbiol Infect, 2008					1	1			2	
Bental, Clin Infect Dis, 1995	1	1							1	
Bertaud, Bull Fed Soc Gynecol Obstet Lang Fr, 1953	1		1						1	One case, title only, outcome unknown
Boden, BMC Infect Dis, 2012	1		1						1	Age of each pregnant woman not mentioned
Carcopino, Clin Infect Dis, 2007	1	1							1	
Cerar, Wien Klin Wochenschr, 2009	1	1							1	
Denman, Int Med J, 2009	1	1							1	
Dindinaud, J Gynecol Obstet Biol Reprod, 1991	1	1							1	
Ellis, QJ Med, 1983	1	1							1	
Eyigor, Mikrobiyol Bul, 2013					1	1			2	
Fiset, Am J Epidemiol, 1974					1		1		2	Seroprevalence studies on newborns
Friedland, Lancet, 1994	1	1							1	
Gaburro, Mal Infect Parasit, 1956	1	1							1	
Hellmeyer, Z	1	1							1	

Geburtshilfe Neonatol, 2002										
Jover-Diaz, Infect Dis Obstet Gynecol, 2001	1	1							1	
Kaplan, Acta obstetrica et gynecologica scandinavica, 1995	1	1							1	
Langley, Am J Obstet Gynecol, 2003					1	1			2	
Ludlam, J Infect, 1997	1	1							1	
Marrie, Infect Dis Clin Pract, 1993	1	1							1	
McGivern, Br J Obstet Gynaecol, 1988	1	1							1	
Michev, Akush Gynekol, 1981	1		1						1	Spontaneous abortion, only title available, no age of the mother and treatement unavailable
Million, present work	1	1							1	
Minchev, Akush Ginekol, 1983	1		1						1	
Munster, Ned Tijdschr Geneeskd, 2011	1	1							1	
Munster, Placenta, 2012	1	1							1	
Munster, Eurosurveillance, 2013					1	1			2	
Nielsen, BMC Infect Dis, 2013					1	1			2	
Nielsen, PlosOne, 2012					1			1	2	Possible duplicated data with Nielsen, BMC Infectious Diseases, 2013
Nielsen, Ugeskr Laeger, 2011	1			1				1	2	Possible duplicate of case 2 from Nielsen, Emerging Infect Dis, In press
Nielsen, Emerg Infect Dis, 2014 in press	1	1							1	
Numazaki, Microbes and Infection, 2000					1		1		2	

Quijada, Clin Microbiol Infect, 2011					1	1			2	
Raoult, Arch Intern Med, 2002	1	1							1	
Raoult, N Eng J Med, 1994	1			1				1	2	Case detailed in Stein, Clin Infect Dis, 1997 (case 1) and Raoult, Arch Intern Med, 2002 (Case 2). For our analysis, this case was considered part of the series from Stein, Clin Infect Dis, 1998.
Rey, Eur J Obstet Gynecol Reprod Biol, 2000					1	1			2	
Riechman, Am J Med, 1988	1	1							1	
Shinar, Obstet Gynecol, 2012	1	1							1	
Stein, Clin Infect Dis, 1998	1	1							1	
Syrucek, J Hyg Epidemiol Microbiol Immunol, 1958	1	1							1	
Tellez, JInfect, 1998	1	1							1	
Tissot-Dupont, Clin Infect Dis, 2007	1		1						1	Age of each pregnant woman not mentioned
Vaidya, J Clin Microbiol, 2008	1		1						1	Analysis of samples collected from women with spontaneous abortion
van der hoek, BMC Infect Dis, 2011					1	1			2	
Wagstaff, Public Health Report, 1959	1		1						1	Age of each pregnant woman not mentioned
Total	35	25	8	2	11	8	2	3	59	

Table S3

Number of screened case in our review	Article included	Reference	Included YES/NO	Not included because possible duplicate	Not included because fetal loss possibly related to other cause	Country	French NRC	Year Dg	Age	Fetal Loss due to Coxiella burnetii	Treatment	Case series	Possible duplicate of	Cotrimoxazole	Tetracycline	Erythromycin	Quinolones	Rifampin	Azithromycin	Clarithromycin	Sum treatment
1	1	Angelakis, EJCMID, 2012 - case 1	1	0	0	France	1	2006	27	0	1	1		1							1
2	1	Angelakis, EJCMID, 2012 - case 2	1	0	0	France	1	2006	27	0	1	1		1							1
3	1	Angelakis, EJCMID, 2012 - case 3	1	0	0	France	1	2006	27	0	1	1		1							1
4	1	Angelakis, EJCMID, 2012 - case 4	1	0	0	France	1	2006	28	0	0	1									0
5	1	Angelakis, EJCMID, 2012 - case 5	1	0	0	France	1	2006	28	1	0	1									0
6	1	Angelakis, EJCMID, 2012 - case 6	1	0	0	France	1	2006	26	1	1	1				1					1
7	1	Angelakis, EJCMID, 2012 - case 7	1	0	0	France	1	2006	35	0	1	1		1							1
8	1	Angelakis, EJCMID, 2012 - case 8	1	0	0	France	1	2006	28	1	0	1									0
9	1	Angelakis, EJCMID, 2012 - case 9	1	0	0	France	1	2006	31	1	1	1		1							1
10	1	Angelakis, EJCMID, 2012 - case 10	1	0	0	France	1	2006	35	1	1	1		1							1
11	1	Angelakis, EJCMID, 2012 - case 11	1	0	0	France	1	2007	30	0	1	1		1							1
12	1	Angelakis, EJCMID, 2012 - case 12	1	0	0	France	1	2007	36	1	0	1									0
13	1	Angelakis, EJCMID, 2012 - case 13	1	0	0	France	1	2007	25	1	0	1									0
14	1	Angelakis, EJCMID, 2012 - case 14	1	0	0	France	1	2007	30	0	1	1		1							1
15	1	Angelakis, EJCMID, 2012 - case 15	1	0	0	France	1	2007	24	1	1	1		1							1
16	1	Angelakis, EJCMID, 2012 - case 16	1	0	0	France	1	2007	30	0	1	1		1							1
17	1	Angelakis, EJCMID, 2012 - case 17	1	0	0	France	1	2007	31	0	0	1									0

18	1	Angelakis, EJCMID, 2012 - case 18	1	0	0	France	1	2008	36	0	1	1		1						1
19	1	Angelakis, EJCMID, 2012 - case 19	1	0	0	France	1	2008	28	0	1	1		1						1
20	1	Angelakis, EJCMID, 2012 - case 20	1	0	0	France	1	2008	28	1	0	1								0
21	1	Angelakis, EJCMID, 2012 - case 21	1	0	0	France	1	2008	36	0	1	1		1						1
22	1	Angelakis, EJCMID, 2012 - case 22	1	0	0	France	1	2008	27	0	1	1		1						1
23	1	Angelakis, EJCMID, 2012 - case 23	1	0	0	France	1	2009	34	0	1	1		1						1
24	1	Angelakis, EJCMID, 2012 - case 24	1	0	0	France	1	2010	34	1	0	1								0
25	1	Angelakis, EJCMID, 2012 - case 25	1	0	0	France	1	2010	42	1	0	1								0
26	1	Angelakis, EJCMID, 2012 - case 26	1	0	0	France	1	2010	29	0	1	1		1						1
27	1	Angelakis, EJCMID, 2012 - case 27	1	0	0	France	1	2010	40	1	0	1								0
28	1	Angelakis, EJCMID, 2012 - case 28	1	0	0	France	1	2011	42	1	1	1						1		1
29	1	Angelakis, EJCMID, 2012 - case 29	1	0	0	France	1	2011	30	1	1	1		1						1
30	1	Angelakis, EJCMID, 2012 - case 30	1	0	0	France	1	2011	26	1	0	1								0
31	2	Bental, Clin Infect Dis, 1995	1	0	0	Israel	0	1995	28	0	1	0				1		1		2
32	3	Carcopino, Clin Infect Dis, 2007 - case 1	0	1	0	France	1	1991	30	0	0	1	Raoult, Arch Intern Med, 2002 - case 1							0
33	3	Carcopino, Clin Infect Dis, 2007 - case 2	0	1	0	France	1	1992	26	1	1	1	Stein, Clin Infect Dis, 1998 - case 1	1						1

34	3	Carcopino, Clin Infect Dis, 2007 - case 3	0	1	0	France	1	1993	18	0	0	1	Stein, Clin Infect Dis, 1998 - case 3								0
35	3	Carcopino, Clin Infect Dis, 2007 - case 4	1	0	0	France	1	1994	22	0	0	1									0
36	3	Carcopino, Clin Infect Dis, 2007 - case 5	0	1	0	France	1	1994	39	1	0	1	Raoult, Arch Intern Med, 2002 - case 5								0
37	3	Carcopino, Clin Infect Dis, 2007 - case 6	0	1	0	France	1	1994	31	1	0	1	Raoult, Arch Intern Med, 2002 - case 4								0
38	3	Carcopino, Clin Infect Dis, 2007 - case 7	1	0	0	France	1	1995	37	1	0	1									0
39	3	Carcopino, Clin Infect Dis, 2007 - case 8	1	0	0	France	1	1995	42	1	0	1									0
40	3	Carcopino, Clin Infect Dis, 2007 - case 9	0	1	0	France	1	1995	34	1	0	1	Raoult, Arch Intern Med, 2002 - case 7								0
41	3	Carcopino, Clin Infect Dis, 2007 - case 10	1	0	0	France	1	1996	25	0	1	1		1							1
42	3	Carcopino, Clin Infect Dis, 2007 -	1	0	0	France	1	1996	25	1	0	1									0

		case 11																		
43	3	Carcopino, Clin Infect Dis, 2007 - case 12	0	1	0	France	1	1997	26	1	0	1	Raoult, Arch Intern Med, 2002 - case 10								0
44	3	Carcopino, Clin Infect Dis, 2007 - case 13	1	0	0	France	1	1997	34	0	0	1									0
45	3	Carcopino, Clin Infect Dis, 2007 - case 14	0	1	0	France	1	1997	25	1	0	1	Raoult, Arch Intern Med, 2002 - case 9								0
46	3	Carcopino, Clin Infect Dis, 2007 - case 15	1	0	0	France	1	1998	33	1	1	1		1							1
47	3	Carcopino, Clin Infect Dis, 2007 - case 16	0	1	0	France	1	1999	36	0	1	1	Raoult, Arch Intern Med, 2002 - case 13	1							1
48	3	Carcopino, Clin Infect Dis, 2007 - case 17	0	1	0	France	1	1999	34	0	0	1	Raoult, Arch Intern Med, 2002 - case 14								0
49	3	Carcopino, Clin Infect Dis, 2007 - case 18	0	1	0	France	1	1999	35	1	0	1	Raoult, Arch Intern								0

													Med, 2002 - case 16							
50	3	Carcopino, Clin Infect Dis, 2007 - case 19	0	1	0	France	1	1999	23	0	1	1	Raoult, Arch Intern Med, 2002 - case 15	1						1
51	3	Carcopino, Clin Infect Dis, 2007 - case 20	1	0	0	France	1	1999	24	0	0	1								0
52	3	Carcopino, Clin Infect Dis, 2007 - case 21	1	0	0	France	1	1999	23	0	0	1								0
53	3	Carcopino, Clin Infect Dis, 2007 - case 22	1	0	0	France	1	1999	30	0	1	1		1						1
54	3	Carcopino, Clin Infect Dis, 2007 - case 23	0	1	0	France	1	2000	42	0	1	1	Raoult, Arch Intern Med, 2002 - case 17	1						1
55	3	Carcopino, Clin Infect Dis, 2007 - case 24	1	0	0	France	1	2000	25	0	0	1								0
56	3	Carcopino, Clin Infect Dis, 2007 - case 25	1	0	0	France	1	2001	28	0	0	1								0
57	3	Carcopino, Clin Infect Dis, 2007 - case 26	1	0	0	France	1	2001	25	0	0	1								0
58	3	Carcopino, Clin Infect Dis, 2007 - case 27	1	0	0	France	1	2001	31	1	0	1								0
59	3	Carcopino, Clin Infect Dis, 2007 - case 28	1	0	0	France	1	2001	29	0	1	1		1						1
60	3	Carcopino, Clin Infect Dis, 2007 -	1	0	0	France	1	2002	21	0	0	1								0

		case 29																			
61	3	Carcopino, Clin Infect Dis, 2007 - case 30	1	0	0	France	1	2002	24	1	0	1									0
62	3	Carcopino, Clin Infect Dis, 2007 - case 31	1	0	0	France	1	2002	37	0	0	1									0
63	3	Carcopino, Clin Infect Dis, 2007 - case 32	1	0	0	France	1	2002	29	0	1	1		1							1
64	3	Carcopino, Clin Infect Dis, 2007 - case 33	1	0	0	France	1	2002	22	0	1	1		1							1
65	3	Carcopino, Clin Infect Dis, 2007 - case 34	1	0	0	France	1	2002	27	0	1	1		1							1
66	3	Carcopino, Clin Infect Dis, 2007 - case 35	1	0	0	France	1	2002	34	0	1	1		1							1
67	3	Carcopino, Clin Infect Dis, 2007 - case 36	1	0	0	France	1	2002	24	0	1	1		1							1
68	3	Carcopino, Clin Infect Dis, 2007 - case 37	1	0	0	France	1	2003	28	0	1	1		1							1
69	3	Carcopino, Clin Infect Dis, 2007 - case 38	1	0	0	France	1	2003	34	0	1	1		1							1
70	3	Carcopino, Clin Infect Dis, 2007 - case 39	1	0	0	France	1	2003	23	1	0	1									0
71	3	Carcopino, Clin Infect Dis, 2007 - case 40	1	0	0	France	1	2003	26	0	1	1		1							1
72	3	Carcopino, Clin Infect Dis, 2007 - case 41	1	0	0	France	1	2003	34	0	1	1		1							1
73	3	Carcopino, Clin Infect Dis, 2007 - case 42	1	0	0	France	1	2003	32	0	1	1		1							1
74	3	Carcopino, Clin Infect Dis, 2007 - case 43	1	0	0	France	1	2003	29	0	1	1		1							1
75	3	Carcopino, Clin Infect Dis, 2007 - case 44	1	0	0	France	1	2004	28	0	0	1									0
76	3	Carcopino, Clin	1	0	0	Fran	1	2004	32	0	0	1									0

		Infect Dis, 2007 - case 45				ce														
77	3	Carcopino, Clin Infect Dis, 2007 - case 46	1	0	0	France	1	2004	36	0	0	1								0
78	3	Carcopino, Clin Infect Dis, 2007 - case 47	1	0	0	France	1	2004	33	1	1	1		1						1
79	3	Carcopino, Clin Infect Dis, 2007 - case 48	1	0	0	France	1	2005	27	1	0	1								0
80	3	Carcopino, Clin Infect Dis, 2007 - case 49	1	0	0	France	1	2005	27	1	0	1								0
81	3	Carcopino, Clin Infect Dis, 2007 - case 50	1	0	0	France	1	2005	26	0	1	1		1						1
82	3	Carcopino, Clin Infect Dis, 2007 - case 51	1	0	0	France	1	2005	30	0	1	1		1						1
83	3	Carcopino, Clin Infect Dis, 2007 - case 52	1	0	0	France	1	2005	35	1	0	1								0
84	3	Carcopino, Clin Infect Dis, 2007 - case 53	1	0	0	France	1	2005	26	0	0	1								0
85	4	Cerar, Wien Klin Wochenschr, 2009	1	0	0	Slovenia	0	2008	29	0	1	0							1	1
86	5	Denman, Int Med J, 2009	1	0	0	Australia	0	2009	22	0	1	0		1						1
87	6	Dindinaud, J Gynecol Obstet Biol Reprod, 1991 - case 1	1	0	0	France	0	1991	22	1	1	1			1					1
88	6	Dindinaud, J Gynecol Obstet Biol Reprod, 1991 - case 2	1	0	0	France	0	1991	34	1	1	1			1					1
89	7	Ellis, Q J Med, 1983	1	0	0	UK	0	1983	19	1	1	0			1					1
90	8	Friedland, Lancet, 1994	1	0	0	UK	0	1994	26	1	0	0								0
91	9	Gaburro, Mal Infect Parasit, 1956	0	0	1	Italy	0	1956	37	ND (triplets)	NA	0								0
92	10	Hellmeyer, Z Geburtshilfe Neonatol, 2002	1	0	0	Germany	0	2002	27	0	1	0						1		1	2

93	11	Jover-Diaz, Infect Dis Obstet Gynecol, 2001	1	0	0	Spain	0	2001	18	0	1	0								1	1
94	12	Kaplan, Acta Obstet Gynecol Scand, 1995	1	0	0	Israel	0	1995	24	1	0	0									0
95	13	Ludlam, J Infect, 1997	1	0	0	UK	0	1997	26	0	1	0					1				1
96	14	Marrie, Infect Dis Clin Pract, 1993 - case 1	1	0	0	Canada	0	1993	38	0	0	1									0
97	14	Marrie, Infect Dis Clin Pract, 1993 - case 2	1	0	0	Canada	0	1993	27	0	1	1			1						1
98	15	McGivern, Br J Obstet Gynaecol, 1988	0	0	1	UK	0	1988	30	ND (Chlamydia)	NA	0									0
99	16	Million, present work - case 1	1	0	0	France	1	2011	29	0	1	1		1							1
100	16	Million, present work - case 2	1	0	0	France	1	2011	42	1	0	1									0
101	16	Million, present work - case 3 - case detailed in this MS	1	0	0	France	1	2012	20	1	0	1									0
102	16	Million, present work - case 4	1	0	0	France	1	2013	41	1	0	1									
103	17	Munster, Ned Tijdschr Geneeskd, 2011	1	0	0	Holland	0	2011	42	0	1	0				1					1
104	18	Munster, Placenta, 2012 - case 1	1	0	0	Holland	0	2009	31	0	1	1				1					1
105	18	Munster, Placenta, 2012 - case 2	1	0	0	Holland	0	2009	32	0	1	1				1					1
106	18	Munster, Placenta, 2012 - case 3	1	0	0	Holland	0	2009	31	0	1	1				1					1
107	18	Munster, Placenta, 2012 - case 4	1	0	0	Holland	0	2009	34	0	0	1									0
108	18	Munster, Placenta, 2012 - case 5	1	0	0	Holland	0	2009	33	0	0	1									0
109	19	Nielsen, Emerging Infectious Diseases, 2014 - patient 1 pregnancy 1	1	0	0	Denmark	0	2007-2011	33	0	1	1		1							1
110	19	Nielsen, Emerging Infectious Diseases, 2014 - patient 2 pregnancy 1	1	0	0	Denmark	0	2007-2011	40	1	0	1									0

111	19	Nielsen, Emerging Infectious Diseases, 2014 - patient 3 pregnancy 1	1	0	0	Denmark	0	2007-2011	30	0	1	1									0
112	19	Nielsen, Emerging Infectious Diseases, 2014 - patient 4 pregnancy 1	0	0	1	Denmark	0	2007-2011	34	ND (single fetal demise)	0	1									0
113	19	Nielsen, Emerging Infectious Diseases, 2014 - patient 5 pregnancy 1	1	0	0	Denmark	0	2007-2011	32	0	0	1									0
114	19	Nielsen, Emerging Infectious Diseases, 2014 - patient 5 pregnancy 2	1	0	0	Denmark	0	2007-2011	33	0	0	1									0
115	19	Nielsen, Emerging Infectious Diseases, 2014 - patient 6	1	0	0	Denmark	0	2007-2011	26	0	0	1									0
116	19	Nielsen, Emerging Infectious Diseases, 2014 - patient 7	1	0	0	Denmark	0	2007-2011	32	0	1	1									0
117	19	Nielsen, Emerging Infectious Diseases, 2014 - patient 8 pregnancy 1	1	0	0	Denmark	0	2007-2011	24	1	0	1									0
118	19	Nielsen, Emerging Infectious Diseases, 2014 - patient 8 pregnancy 2	1	0	0	Denmark	0	2007-2011	26	0	1	1									0
119	19	Nielsen, Emerging Infectious Diseases, 2014 - patient 9 pregnancy 1	1	0	0	Denmark	0	2007-2011	30	0	1	1									0
120	19	Nielsen, Emerging Infectious Diseases, 2014 - patient 9 pregnancy 2	1	0	0	Denmark	0	2007-2011	33	0	0	1									0
121	19	Nielsen, Emerging Infectious Diseases, 2014 - patient 10 pregnancy 1	1	0	0	Denmark	0	2007-2011	30	0	1	1									
122	19	Nielsen, Emerging Infectious Diseases, 2014 - patient 10 pregnancy 2	1	0	0	Denmark	0	2007-2011	33	0	0	1									
123	19	Nielsen, Emerging Infectious Diseases,	1	0	0	Denmar	0	2007-2011	30	0	0	1									

		2014 - patient 11				k															
124	19	Nielsen, Emerging Infectious Diseases, 2014 - patient 12	1	0	0	Denmark	0	2007-2011	31	0	1	1									
125	20	Raoult, Arch Intern Med, 2002 - case 1	1	0	0	France	1	1991	30	0	0	1									0
126	20	Raoult, Arch Intern Med, 2002 - case 2	0	1	0	France	1	1992	26	1	1	1	Stein, Clin Infect Dis, 1998 - case 1								0
127	20	Raoult, Arch Intern Med, 2002 - case 3	1	0	0	France	1	1993	32	0	0	1									0
128	20	Raoult, Arch Intern Med, 2002 - case 4	1	0	0	France	1	1994	32	1	0	1									0
129	20	Raoult, Arch Intern Med, 2002 - case 5	1	0	0	France	1	1994	39	1	0	1									0
130	20	Raoult, Arch Intern Med, 2002 - case 6	1	0	0	France	1	1995	22	0	0	1									0
131	20	Raoult, Arch Intern Med, 2002 - case 7	1	0	0	France	1	1995	34	1	0	1									0
132	20	Raoult, Arch Intern Med, 2002 - case 8	1	0	0	France	1	1996	29	0	1	1		1							1
133	20	Raoult, Arch Intern Med, 2002 - case 9	1	0	0	France	1	1997	24	1	0	1									0
134	20	Raoult, Arch Intern Med, 2002 - case 10	1	0	0	France	1	1997	26	1	0	1									0
135	20	Raoult, Arch Intern Med, 2002 - case 11	1	0	0	France	1	1997	35	1	0	1									0
136	20	Raoult, Arch Intern Med, 2002 - case 12	1	0	0	France	1	1999	23	0	1	1									0
137	20	Raoult, Arch Intern Med, 2002 - case 13	1	0	0	France	1	1999	36	0	1	1									0
138	20	Raoult, Arch Intern Med, 2002 - case 14	1	0	0	France	1	2000	35	0	0	1									0
139	20	Raoult, Arch Intern Med, 2002 - case 15	1	0	0	France	1	2000	23	0	1	1		1							1
140	20	Raoult, Arch Intern Med, 2002 - case	1	0	0	France	1	2000	36	1	0	1									0

		16																		
141	20	Raoult, Arch Intern Med, 2002 - case 17	1	0	0	France	1	2000	42	0	1	1								0
142	21	Riechman, Am J Med, 1988	1	0	0	Israel	0	1988	29	0	1	0			1					1
143	22	Shinar, Obstet Gynecol, 2012	1	0	0	Israel	0	2012	34	0	0	0								0
144	23	Stein, Clin Infect Dis, 1998 - case 1 (case also published in Raoult, N Engl J Med, 1994)	1	0	0	France	1	1997	26	1	1	1	Carcopino, Clin Infect Dis, 2007, Case 14		1					1
145	23	Stein, Clin Infect Dis, 1998 - case 2	1	0	0	France	1	1997	28	0	0	1								0
146	23	Stein, Clin Infect Dis, 1998 - case 3	1	0	0	France	1	1997	18	0	0	1	Carcopino, Clin Infect Dis, 2007, Case 3	1						1
147	23	Stein, Clin Infect Dis, 1998 - case 4	1	0	0	France	1	1997	31	1	0	1			1					1
148	23	Stein, Clin Infect Dis, 1998 - case 5	1	0	0	France	1	1997	33	1	0	1								0
149	24	Syrucek, J Hyg Epidemiol Microbiol Immunol, 1958 - case 1	1	0	0	Czech	0	1958	21	0	NA	1								0
150	24	Syrucek, J Hyg Epidemiol Microbiol Immunol, 1958 - case 2	1	0	0	Czech	0	1958	20	0	NA	1								0
151	24	Syrucek, J Hyg Epidemiol Microbiol Immunol, 1958 - case 3	1	0	0	Czech	0	1958	24	0	NA	1								0
152	24	Syrucek, J Hyg Epidemiol Microbiol	1	0	0	Czech	0	1958	27	0	NA	1								0

		Immunol, 1958 - case 4																			
153	24	Syrucek, J Hyg Epidemiol Microbiol Immunol, 1958 - case 5	0	0	1	Czech	4	1958	32	ND (rubella)	NA	1									0
154	25	Tellez, J Infect, 1998	1	0	0	Spain	0	1998	28	0	0	0									0

French NRC: French Nationtal Referral center for Q fever. NA: not available, ND: not determined

Table S4

Number of included case in our review	Number of article with at least one case included	Reference	Country	Case from the French NRC database	Case series	Case series number	Year of diagnosis	Age of the mother	Age class	Fetal Loss related to Coxiella burnetii infection	Fetal malformation	Fetal malformation detail	Treatment YESNO	Treatment with sufficient duration YES/NO (TMPSMX>5wks)	Treatment duration (weeks)	Cotrimoxazole	Tetracycline	Erythromycin	Ciprofloxacin	Rifampin	Azithromycin	Clarithromycin	Sum treatment	Remarks
1	1	Angelakis, Eur J Clin Microbiol Infect Dis, 2012 - case 1	France	1	1	1	2006	27	2	0	0		1	1	12	1							1	
2	1	Angelakis, Eur J Clin Microbiol Infect Dis, 2012 - case 2	France	1	1	1	2006	27	2	0	0		1	1	24	1							1	
3	1	Angelakis, Eur J Clin Microbiol Infect Dis, 2012 - case 3	France	1	1	1	2006	27	2	0	0		1	1	24	1							1	
4	1	Angelakis, Eur J Clin Microbiol Infect Dis, 2012 - case 4	France	1	1	1	2006	28	2	1	0		0	0										
5	1	Angelakis, Eur J Clin Microbiol Infect Dis, 2012 - case 5	France	1	1	1	2006	28	2	1	0		0	0										

6	1	Angelakis, Eur J Clin Microbiol Infect Dis, 2012 - case 6	France	1	1	1	2006	26	2	1	0		1	1	1			1					1	
7	1	Angelakis, Eur J Clin Microbiol Infect Dis, 2012 - case 7	France	1	1	1	2006	35	3	0	0		1	1	12	1							1	
8	1	Angelakis, Eur J Clin Microbiol Infect Dis, 2012 - case 8	France	1	1	1	2006	28	2	1	0		0	0										
9	1	Angelakis, Eur J Clin Microbiol Infect Dis, 2012 - case 9	France	1	1	1	2006	31	2	1	0		1	0	2	1							1	
10	1	Angelakis, Eur J Clin Microbiol Infect Dis, 2012 - case 10	France	1	1	1	2006	35	3	1	0		1	0	4	1							1	
11	1	Angelakis, Eur J Clin Microbiol Infect Dis, 2012 - case 11	France	1	1	1	2007	30	2	0	0		1	1	16	1							1	
12	1	Angelakis, Eur J Clin Microbiol Infect Dis, 2012 - case 12	France	1	1	1	2007	36	3	1	0		0	0										
13	1	Angelakis, Eur J Clin Microbiol Infect Dis, 2012 - case 13	France	1	1	1	2007	25	2	1	0		0	0										
14	1	Angelakis, Eur J Clin Microbiol Infect Dis,	France	1	1	1	2007	30	2	0	0		1	1	24	1							1	

		2012 - case 14																					
15	1	Angelakis, Eur J Clin Microbiol Infect Dis, 2012 - case 15	Fra nce	1	1	1	2007	24	1	1	0		1	0	4	1						1	
16	1	Angelakis, Eur J Clin Microbiol Infect Dis, 2012 - case 16	Fra nce	1	1	1	2007	30	2	0	0		1	1	24	1						1	
17	1	Angelakis, Eur J Clin Microbiol Infect Dis, 2012 - case 17	Fra nce	1	1	1	2007	31	2	0	0		0	0									
18	1	Angelakis, Eur J Clin Microbiol Infect Dis, 2012 - case 18	Fra nce	1	1	1	2008	36	3	0	0		1	1	12	1						1	
19	1	Angelakis, Eur J Clin Microbiol Infect Dis, 2012 - case 19	Fra nce	1	1	1	2008	28	2	0	0		1	1	12	1						1	
20	1	Angelakis, Eur J Clin Microbiol Infect Dis, 2012 - case 20	Fra nce	1	1	1	2008	28	2	1	0		0	0									
21	1	Angelakis, Eur J Clin Microbiol Infect Dis, 2012 - case 21	Fra nce	1	1	1	2008	36	3	0	0		1	1	24	1						1	
22	1	Angelakis, Eur J Clin Microbiol Infect Dis,	Fra nce	1	1	1	2008	27	2	0	0		1	1	12	1						1	

		2012 - case 22																					
23	1	Angelakis, Eur J Clin Microbiol Infect Dis, 2012 - case 23	Fra nce	1	1	1	2009	34	2	0	0		1	1	8	1						1	
24	1	Angelakis, Eur J Clin Microbiol Infect Dis, 2012 - case 24	Fra nce	1	1	1	2010	34	2	1	0		0	0									
25	1	Angelakis, Eur J Clin Microbiol Infect Dis, 2012 - case 25	Fra nce	1	1	1	2010	42	3	1	0		0	0									
26	1	Angelakis, Eur J Clin Microbiol Infect Dis, 2012 - case 26	Fra nce	1	1	1	2010	29	2	0	0		1	1	24	1						1	
27	1	Angelakis, Eur J Clin Microbiol Infect Dis, 2012 - case 27	Fra nce	1	1	1	2010	40	3	1	0		0	0									
28	1	Angelakis, Eur J Clin Microbiol Infect Dis, 2012 - case 28	Fra nce	1	1	1	2011	42	3	1	0		1	0	1						1		1	
29	1	Angelakis, Eur J Clin Microbiol Infect Dis, 2012 - case 29	Fra nce	1	1	1	2011	30	2	1	0		1	1	20	1						1	
30	1	Angelakis, Eur J Clin Microbiol Infect Dis,	Fra nce	1	1	1	2011	26	2	1	0		0	0									

		2012 - case 30																					
31	2	Bental, Clin Infect Dis, 1995	Israel	0	0		1995	28	2	0	0		1	1	8			1		1			2	
32	3	Carcopino, Clin Infect Dis, 2007 - case 4	France	1	1	2	1994	22	1	0	0		0	0										
33	3	Carcopino, Clin Infect Dis, 2007 - case 7	France	1	1	2	1995	37	3	1	0		0	0										
34	3	Carcopino, Clin Infect Dis, 2007 - case 8	France	1	1	2	1995	42	3	1	0		0	0										
35	3	Carcopino, Clin Infect Dis, 2007 - case 10	France	1	1	2	1996	25	2	0	0		1	0	3	1							1	
36	3	Carcopino, Clin Infect Dis, 2007 - case 11	France	1	1	2	1996	25	2	1	0		0	0										
37	3	Carcopino, Clin Infect Dis, 2007 - case 13	France	1	1	2	1997	34	2	0	0		0	0										
38	3	Carcopino, Clin Infect Dis, 2007 - case 15	France	1	1	2	1998	33	2	1	0		1	0	4	1							1	
39	3	Carcopino, Clin Infect Dis, 2007 - case 20	France	1	1	2	1999	24	1	0	0		0	0										
40	3	Carcopino, Clin Infect Dis, 2007 - case 21	France	1	1	2	1999	23	1	0	0		0	0										
41	3	Carcopino, Clin Infect Dis, 2007 - case 22	France	1	1	2	1999	30	2	0	0		1	0	4	1							1	
42	3	Carcopino, Clin Infect Dis, 2007 -	France	1	1	2	2000	25	2	0	0		0	0										

		case 24																					
43	3	Carcopino, Clin Infect Dis, 2007 - case 25	Germany	1	1	2	2001	28	2	0	0		0	0									
44	3	Carcopino, Clin Infect Dis, 2007 - case 26	France	1	1	2	2001	25	2	0	0		0	0									
45	3	Carcopino, Clin Infect Dis, 2007 - case 27	France	1	1	2	2001	31	2	1	0		0	0									
46	3	Carcopino, Clin Infect Dis, 2007 - case 28	France	1	1	2	2001	29	2	0	0		1	1	23	1						1	
47	3	Carcopino, Clin Infect Dis, 2007 - case 29	France	1	1	2	2002	21	1	0	0		0	0									
48	3	Carcopino, Clin Infect Dis, 2007 - case 30	France	1	1	2	2002	24	1	1	0		0	0									
49	3	Carcopino, Clin Infect Dis, 2007 - case 31	France	1	1	2	2002	37	3	0	0		0	0									
50	3	Carcopino, Clin Infect Dis, 2007 - case 32	France	1	1	2	2002	29	2	0	0		1	1	11	1						1	
51	3	Carcopino, Clin Infect Dis, 2007 - case 33	France	1	1	2	2002	22	1	0	0		1	0	2	1						1	
52	3	Carcopino, Clin Infect Dis, 2007 - case 34	France	1	1	2	2002	27	2	0	0		1	1	12	1						1	
53	3	Carcopino, Clin Infect Dis, 2007 - case 35	France	1	1	2	2002	34	2	0	0		1	1	15	1						1	
54	3	Carcopino, Clin Infect Dis, 2007 -	France	1	1	2	2002	24	1	0	0		1	1	25	1						1	

		case 36																					
55	3	Carcopino, Clin Infect Dis, 2007 - case 37	Fra nce	1	1	2	2003	28	2	0	0		1	1	13	1							1	
56	3	Carcopino, Clin Infect Dis, 2007 - case 38	Fra nce	1	1	2	2003	34	2	0	0		1	1	21	1							1	
57	3	Carcopino, Clin Infect Dis, 2007 - case 39	Fra nce	1	1	2	2003	23	1	1	1	hypospa dias	0	0										
58	3	Carcopino, Clin Infect Dis, 2007 - case 40	Fra nce	1	1	2	2003	26	2	0	0		1	1	11	1							1	
59	3	Carcopino, Clin Infect Dis, 2007 - case 41	Fra nce	1	1	2	2003	34	2	0	0		1	1	29	1							1	
60	3	Carcopino, Clin Infect Dis, 2007 - case 42	Fra nce	1	1	2	2003	32	2	0	0		1	1	21	1							1	
61	3	Carcopino, Clin Infect Dis, 2007 - case 43	Fra nce	1	1	2	2003	29	2	0	0		1	1	28	1							1	
62	3	Carcopino, Clin Infect Dis, 2007 - case 44	Fra nce	1	1	2	2004	28	2	0	0		0	0										
63	3	Carcopino, Clin Infect Dis, 2007 - case 45	Fra nce	1	1	2	2004	32	2	0	0		0	0										
64	3	Carcopino, Clin Infect Dis, 2007 - case 46	Fra nce	1	1	2	2004	36	3	0	0		0	0										
65	3	Carcopino, Clin Infect Dis, 2007 - case 47	Fra nce	1	1	2	2004	33	2	1	0		1	0	2	1							1	
66	3	Carcopino, Clin Infect Dis, 2007 -	Fra nce	1	1	2	2005	27	2	1	0		0	0										

		case 48																					
67	3	Carcopino, Clin Infect Dis, 2007 - case 49	France	1	1	2	2005	27	2	1	0		0	0									
68	3	Carcopino, Clin Infect Dis, 2007 - case 50	France	1	1	2	2005	26	2	0	1	Potter syndrome with bilateral agenesis	1	1	13	1						1	
69	3	Carcopino, Clin Infect Dis, 2007 - case 51	France	1	1	2	2005	30	2	0	0		1	1	10	1						1	
70	3	Carcopino, Clin Infect Dis, 2007 - case 52	France	1	1	2	2005	35	3	1	0		0	0									
71	3	Carcopino, Clin Infect Dis, 2007 - case 53	France	1	1	2	2005	26	2	0	0		0	0									
72	4	Cerar, Wien Klin Wochenschr, 2009	Slovenia	0	0		2008	29	2	0	0		1	1	1						1		1	
73	5	Denman, Int Med J, 2009	Australia	0	0		2009	22	1	0	0		1	1	12	1							1	
74	6	Dindinaud, J Gynecol Obstet Biol Reprod, 1991 - case 1	France	0	1	3	1991	22	1	1	0		0	0									
75	6	Dindinaud, J Gynecol Obstet Biol Reprod, 1991 - case 2	France	0	1	3	1991	34	2	1	0		0	0									
76	7	Ellis, QJ Med, 1983	UK	0	0		1983	19	1	1	0		NA	NA									
77	8	Friedland, Lancet, 1994	UK	0	0		1994	26	2	1	0		0	0									
78	9	Hellmeyer, Z Geburtshilfe Neonatol, 2002	Germany	0	0		2002	27	2	0	0		1	1	5			1		1		1	3	

79	10	Jover-Diaz, Infect Dis Obstet Gynecol, 2001	Spain	0	0		2001	18	1	0	0		1	1	1							1	1	
80	11	Kaplan, Acta Obstet Gynecol Scand, 1995	Israel	0	0		1995	24	1	1	0		0	0										Less than 1 week of treatment before fetal loss, *Coxiella burnetii* organisms revealed in the placenta by immunofluorescence staining.
81	12	Ludlam, J Infect, 1997	UK	0	0		1997	26	2	0	0		1	1	3				1				1	
82	13	Marrie, Infect Dis Clin Pract, 1993	Canada	0	1	4	1993	38	3	0	0		0	0										
83	13	Marrie, Infect Dis Clin Pract, 1993	Canada	0	1	4	1993	27	2	0	0		1	1	19		1						1	
84	14	Million, Clin Infect Dis, New case 1	France	1	1	5	2011	29	2	0	0		1	1	20	1							1	
85	14	Million, Clin Infect Dis, New case 2	France	1	1	5	2011	42	3	1	0		1	1	2							1	1	
86	14	Million, Clin	Fra	1	1	5	2012	20	1	1	1	Omphal	0	0										

		Infect Dis, New case 3 - case detailed in this MS	nce								ocele												
87	14	Million, Clin Infect Dis, New case 4	Fra nce	1	1	5	2013	41	3	1	0		0	0									
88	15	Munster, Ned Tijdschr Geneeskd, 2011	Holl and	0	0		2011	42	3	0	0		1	1	NA			1				1	
89	16	Munster, Placenta, 2012 - case 1	Holl and	0	1	6	2009	31	2	0	0		1	1	NA			1				1	
90	16	Munster, Placenta, 2012 - case 2	Holl and	0	1	6	2009	32	2	0	0		1	1	NA			1				1	
91	16	Munster, Placenta, 2012 - case 3	Holl and	0	1	6	2009	31	2	0	0		1	1	NA			1				1	
92	16	Munster, Placenta, 2012 - case 4	Holl and	0	1	6	2009	34	2	0	0		0	0									
93	17	Munster, Placenta, 2012 - case 5	Holl and	0	1	6	2009	33	2	0	1	Congeni tal hydrone phrosis	0	0									
94	18	Nielsen, Emerging Infectious Diseases, 2014 - patient 1 pregnancy 1	Den mar k	0	1	7	2007-2011	33	2	0	0		1	1	23	1						1	1
95	18	Nielsen, Emerging Infectious Diseases, 2014 - patient 2 pregnancy 1	Den mar k	0	1	7	2007-2011	40	3	1	0		0	0									0
96	18	Nielsen, Emerging Infectious Diseases, 2014 - patient 3	Den mar k	0	1	7	2007-2011	30	2	0	0		1	1	29	1						1	1

		pregnancy 1																					
97	18	Nielsen, Emerging Infectious Diseases, 2014 - patient 5 pregnancy 1	Denmark	0	1	7	2007-2011	32	2	0	0		0	0									0
98	18	Nielsen, Emerging Infectious Diseases, 2014 - patient 5 pregnancy 2	Denmark	0	1	7	2007-2011	33	2	0	0		0	0									0
99	18	Nielsen, Emerging Infectious Diseases, 2014 - patient 6	Denmark	0	1	7	2007-2011	26	2	0	0		0	0									0
100	18	Nielsen, Emerging Infectious Diseases, 2014 - patient 7	Denmark	0	1	7	2007-2011	32	2	0	0		1	1	31	1						1	1
101	18	Nielsen, Emerging Infectious Diseases, 2014 - patient 8 pregnancy 1	Denmark	0	1	7	2007-2011	24	1	1	1	Undetailed	0	0									0
102	18	Nielsen, Emerging Infectious Diseases, 2014 - patient 8 pregnancy 2	Denmark	0	1	7	2007-2011	26	2	0	0		1	1	8	1						1	1
103	18	Nielsen, Emerging Infectious Diseases, 2014 -	Denmark	0	1	7	2007-2011	30	2	0	0		1	1	19	1						1	1

		patient 9 pregnancy 1																					
104	18	Nielsen, Emerging Infectious Diseases, 2014 - patient 9 pregnancy 2	Denmark	0	1	7	2007-2011	33	2	0	0		0	0									0
105	18	Nielsen, Emerging Infectious Diseases, 2014 - patient 10 pregnancy 1	Denmark	0	1	7	2007-2011	30	2	0	0		1	1	29	1						1	
106	18	Nielsen, Emerging Infectious Diseases, 2014 - patient 10 pregnancy 2	Denmark	0	1	7	2007-2011	33	2	0	0		0	0									
107	18	Nielsen, Emerging Infectious Diseases, 2014 - patient 11	Denmark	0	1	7	2007-2011	30	2	0	0		0	0									
108	18	Nielsen, Emerging Infectious Diseases, 2014 - patient 12	Denmark	0	1	7	2007-2011	31	2	0	0		1	1	17	1						1	
109	19	Raoult, Arch Intern Med, 2002 - case 1	France	1	1	8	1991	30	2	0	0		0	0									0
110	19	Raoult, Arch Intern Med, 2002 - case 3	France	1	1	8	1993	32	2	0	0		0	0									0
111	19	Raoult, Arch Intern Med, 2002 - case 4	France	1	1	8	1994	32	2	1	0		0	0									0
112	19	Raoult, Arch Intern Med, 2002 - case 5	France	1	1	8	1994	39	3	1	0		0	0									0

113	19	Raoult, Arch Intern Med, 2002 - case 6	France	1	1	8	1995	22	1	0	0		0	0									0
114	19	Raoult, Arch Intern Med, 2002 - case 7	France	1	1	8	1995	34	2	1	0		0	0									0
115	19	Raoult, Arch Intern Med, 2002 - case 8	France	1	1	8	1996	29	2	0	0		1	1	24	1						1	1
116	19	Raoult, Arch Intern Med, 2002 - case 9	France	1	1	8	1997	24	1	1	0		0	0									0
117	19	Raoult, Arch Intern Med, 2002 - case 10	France	1	1	8	1997	26	2	1	0		0	0									0
118	19	Raoult, Arch Intern Med, 2002 - case 11	France	1	1	8	1997	35	3	1	0		0	0									0
119	19	Raoult, Arch Intern Med, 2002 - case 12	France	1	1	8	1999	23	1	0	0		1	1	24	1						1	0
120	19	Raoult, Arch Intern Med, 2002 - case 13	France	1	1	8	1999	36	3	0	0		1	1	24	1						1	0
121	19	Raoult, Arch Intern Med, 2002 - case 14	France	1	1	8	2000	35	3	0	0		0	0									0
122	19	Raoult, Arch Intern Med, 2002 - case 15	France	1	1	8	2000	23	1	0	0		1	1	24	1						1	1
123	19	Raoult, Arch Intern Med, 2002 - case 16	France	1	1	8	2000	36	3	1	0		0	0									0
124	19	Raoult, Arch Intern Med, 2002 - case 17	France	1	1	8	2000	42	3	0	0		1	1	24	1						1	0
125	20	Riechman, Am J Med, 1988	Israel	0	0		1988	29	2	0	0		1	1	7		1					1	
126	21	Shinar,	Isra	0	0		2012	34	2	0	0		0	0									

		Obstet Gynecol, 2012	el																				
127	22	Stein, Clin Infect Dis, 1998 - case 1	Fra nce	1	1	9	1992	26	2	1	0		1	0	3	1						1	
128	22	Stein, Clin Infect Dis, 1998 - case 2	Fra nce	1	1	9	1997	28	2	0	0		0	0									
129	22	Stein, Clin Infect Dis, 1998 - case 3	Fra nce	1	1	9	1993	18	1	0	0		0	0									
130	22	Stein, Clin Infect Dis, 1998 - case 4	Fra nce	1	1	9	Befor e 1998	31	2	1	0		0	0									
131	22	Stein, Clin Infect Dis, 1998 - case 5	Fra nce	1	1	9	1997	33	2	1	0		0	0									
132	23	Syrucek, J Hyg Epidemiol Microbiol Immunol, 1958 - case 1	Cze ch	0	1	10	1958	21	1	0	0		NA	NA									
133	23	Syrucek, J Hyg Epidemiol Microbiol Immunol, 1958 - case 2	Cze ch	0	1	10	1958	20	1	0	0		NA	NA									
134	23	Syrucek, J Hyg Epidemiol Microbiol Immunol, 1958 - case 3	Cze ch	0	1	10	1958	24	1	0	0		NA	NA									
135	23	Syrucek, J Hyg Epidemiol Microbiol Immunol, 1958 - case 4	Cze ch	0	1	10	1958	27	2	0	1	hypospa dias, adhesio n of the prepuce and hydrocel e	NA	NA									
136	24	Tellez, J Infect, 1998	Spai n	0	0		1998	28	2	0	0		0	0									

French NRC: French National Referral center for Q fever. NA: Not available

Table S5

Data extracted from Nybo Andersen, BMJ, 2000

Maternal age	Live births	Spontaneous abortions	Induced abortions	Ectopic pregnancies	Stillbirths	All pregnancy outcomes
12-19	44674	5427	49884	808	223	101016
20-24	246038	24465	74683	4163	1046	350395
25-29	312904	33728	59014	7233	1270	414149
30-34	157457	22391	48641	5861	699	235049
35-39	43471	11369	36195	2679	226	93940
40-44	5101	3962	15421	614	34	25132
>44	117	509	1184	54	1	1865

Data calculated based on data extracted from Nybo Andersen, BMJ, 2000

Maternal age	Live births	Spontaneous abortions	Stillbirths
12-24	290712	29892	1269
25-34	470361	56119	1969
35-44	48572	15331	260
Total	809645	101342	3498

Maternal age	Total fetal death	Total pregnancy (without ectopic pregnancies and induced abortions)	Rate (%) of fetal death (without ectopic pregnancies and induced abortions)
12-24	31161	321873	10
25-34	58088	528449	11
35-44	15591	64163	24
Total	104840	914485	11

Supplementary Table 6. Fetal death according to case series after exclusion of possible duplicated cases

	Treated		Not treated		Data on treatment not available	Total
	Fetal death	Total	Fetal death	Total		
Outside the French NRC for Q Fever						
13 cases reports	0	8	2	4	1	13
Syrucek 1958 (Czech rep)	-	-	-	-	4	4
Dindinaud 1991 (France)	0	0	2	2	0	2
Marrie 1993 (Canada)	0	1	0	1	0	2
Munster 2012 (Netherlands)	0	3	0	2	0	5
Nielsen 2014 (Denmark)	0	7	2	8	0	15
French NRC for Q Fever						
Stein 1998 (France)[a]	1	1	2	4	0	5
Raoult 2002 (France)	0	5	7	11	0	16
Carcopino 2007 (France)[b]	2	18	9	22	0	40
Angelakis 2012 (France)	6	19	10	11	0	30
Million 2014 (France)[c]	1	2	2	2	0	4

[a]Case published by Raoult in 1986 was included in this case series [24]. [b]13 cases were considered as duplicates of previous series published by Stein [4] and Raoult [3] and excluded before analysis. [c]Present series.

Supplementary Table 7. Fetal death according to age in Q fever patients and the general population

	General population[a] (n=914485)	Total Q fever patients (n=136)	Untreated Q fever patients (n=67)	Treated Q fever patients (n=64)	Q fever patients with unknown treatment (n=5)
12-24 yrs	31161/321873 (10%)	9/24 (37%)	7/13 (54%)	1/7 (14%)	1/4
25-34 yrs	58088/528449 (11%)	24/88 (27%)	18/39 (46%)	6/48 (12.5%)	0/1
35-44 yrs	15591/64163 (24%)	14/24 (58%)	11/15 (73%)	3/9 (33%)	-
Total	104840/914485 (11%)	47/136 (35%)	36/67 (54%)	10/64 (16%)	1/5

[a] Based on the largest population based study available to date [8].

Supplementary Table 8. Fetal death age-specific and combined (pooled) cumulative incidence ratios in Q fever patients versus general population

		PR[1] [95% CI]	PR[2] [95% CI]	p-value[3]
Total Q fever patients	12-24 yrs	3.9 [2.3-6.5]		
	25-34 yrs	2.5 [1.8-3.5]		
	35-44 yrs	2.4 [1.7-3.4]		
	Total		2.6 [2.1-3.3]	$<10^{-3}$
Untreated Q fever patients	12-24 yrs	5.6 [3.4-9.2]		
	25-34 yrs	4.2 [3.0-5.9]		
	35-44 yrs	3.0 [2.2-4.1]		
	Total		3.9 [3.2-4.8]	$<10^{-3}$
Treated Q fever patients	12-24 yrs	1.5 [0.2-9.0]		
	25-34 yrs	1.1 [0.5-2.4]		
	35-44 yrs	1.4 [0.5-3.4]		
	Total		1.2 [0.7-2.1]	0.48

[1]Mantel-Haenszel age-specific cumulative incidence ratios. [2]Mantel-Haenszel combined (pooled) cumulative incidence ratio. [3]Cochran-Mantel-Haenszel chi-2 test

Supplementary Table 9. Fetal malformation in 122 Q fever patients and two French registries

	Q fever patients (n=122[a])	Paris [18]	Central East France register of congenital malformation (REMERA) [18]	Incidence in the French general population estimated based on two large scale registries (Paris and REMERA [18])
Hypospadias	2/122	35/26563 (13.2 for 10000)	69/58853 (11.7 for 10000)	104/85416 (12.2 for 10000)
Renal agenesis	1/122	4/26563 (1.5 for 10000)	32/58853 (5.4 for 10000)	36/85416 (4.2 for 10000)
Omphalocele	1/122	13/26563 (4.9 for 10000)	21/58853 (3.5 for 10000)	34/85416 (4.0 for 10000)
Total	4/122	52/26563 (20.0 for 10000)	122/58853 (20.1 for 10000)	174/85416 (20.4 for 10000)

The cumulative incidence ratio of fetal malformations for a Q fever patient, compared to a person in the general population, is estimated at 16.1 (95% CI: 6.1-42.7). [a]The total number of Q fever patients assessable for malformations did not included the 14 included pregnancies published in

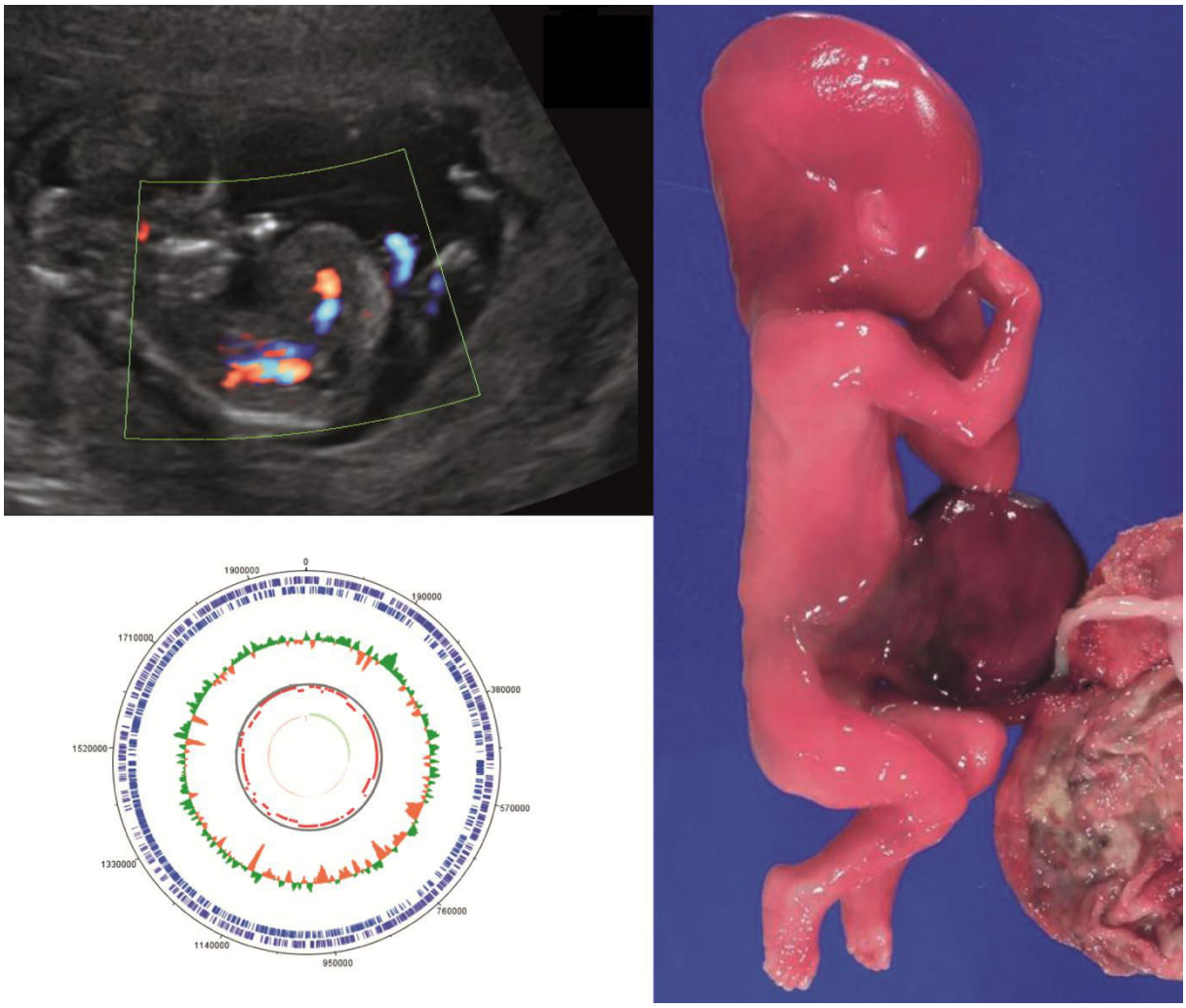
Nielsen *et al.* [14]. In their series, a fetus was severely malformed and died few hours after delivery, but there was no detail about the malformation so we could not compare to the Paris and REMERA registries.

Supplementary Table 10. Population-based serological studies included in the meta-analysis

Study name	Positive serology		Negative serology	
	Fetal deaths	Total	Fetal deaths	Total
Outside the French NRC for Q Fever				
Langley 2003 (Canada)	4	104	41	4483
Baud 2009 (England)	12	20	318	418
Quijada 2011 (Spain)	70	108	203	392
Van der Hoek 2011 (Netherlands)	0	49	24	1125
Nielsen 2013 (Denmark)	2	169	6	687
Eyigor 2013 (Turkey)	10	12	26	46
Munster 2013 (Netherlands)	0	183	6	1046
French NRC				
Rey 2000 (France)	2	16	736	9493

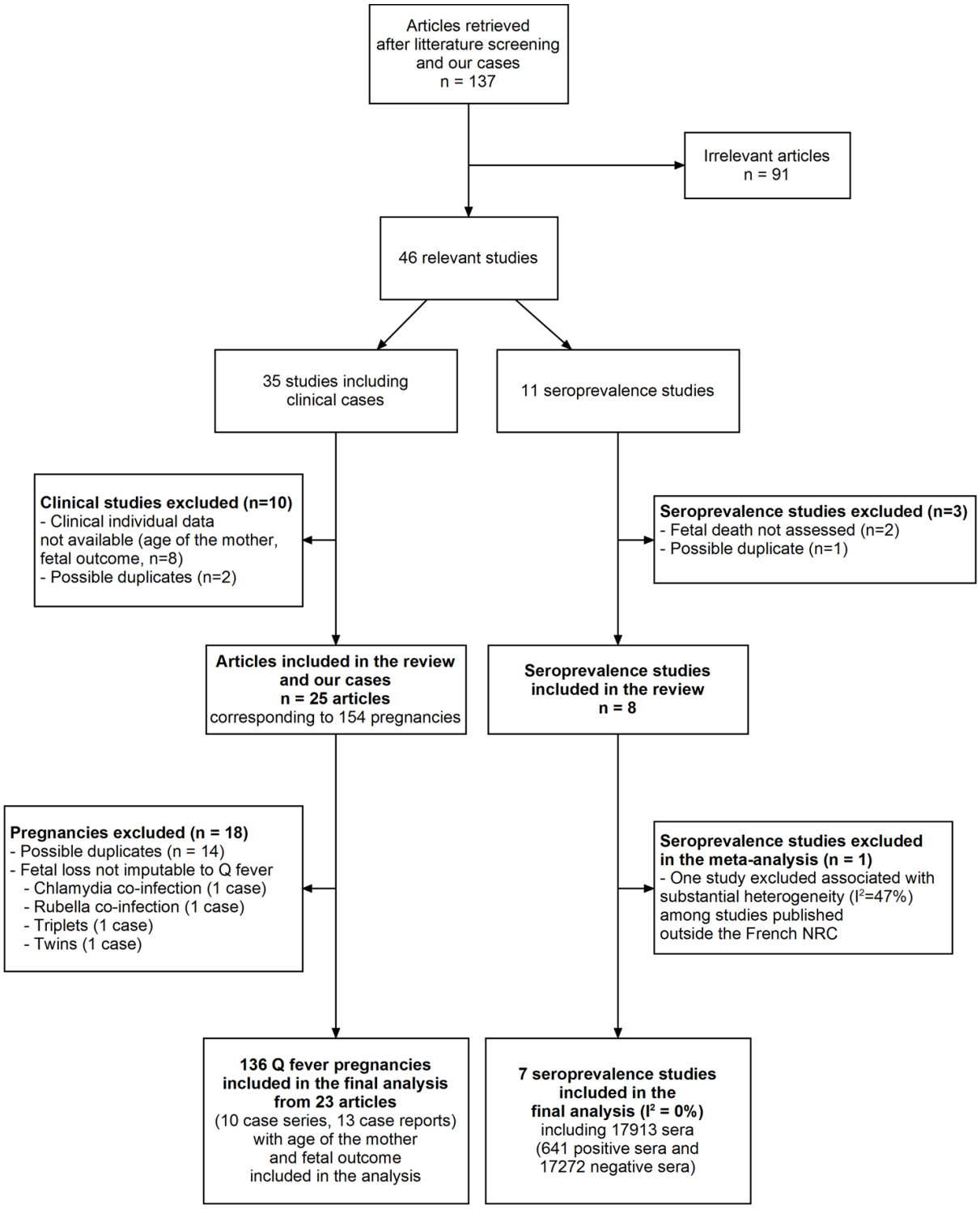
Supplementary Figures

Supplementary Figure 1. Fetus infected with *Coxiella burnetii*

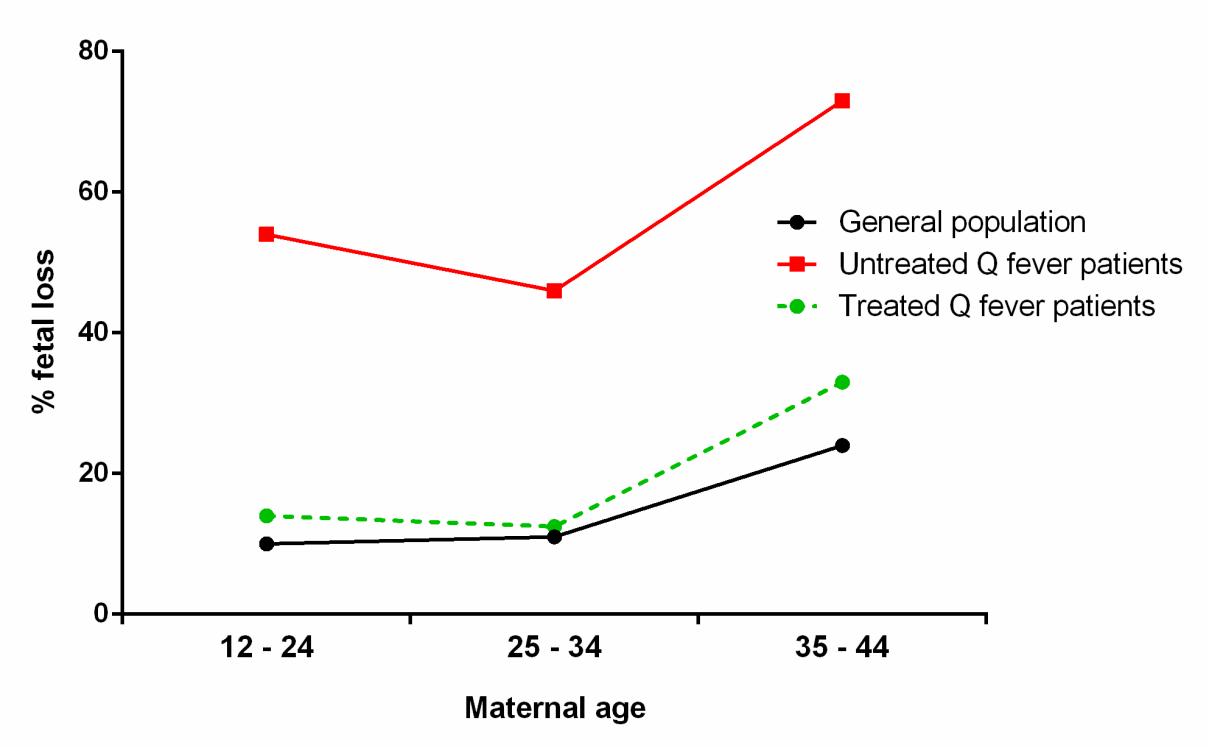


Left up: an anterior wall closing defect was detected by echocardiography at 3 months of pregnancy. Right: examination of the fetus confirmed a large omphalocele containing the liver and the small bowel associated with a discrete adrenal hypoplasia. Left bottom: *Coxiella burnetii* was isolated from the placenta, and the genome was sequenced. Blue: chromosome. Red: plasmid (QpH1). Orange: GC content below average. Green: GC content above average.

Supplementary Figure 2. Flowchart of included studies

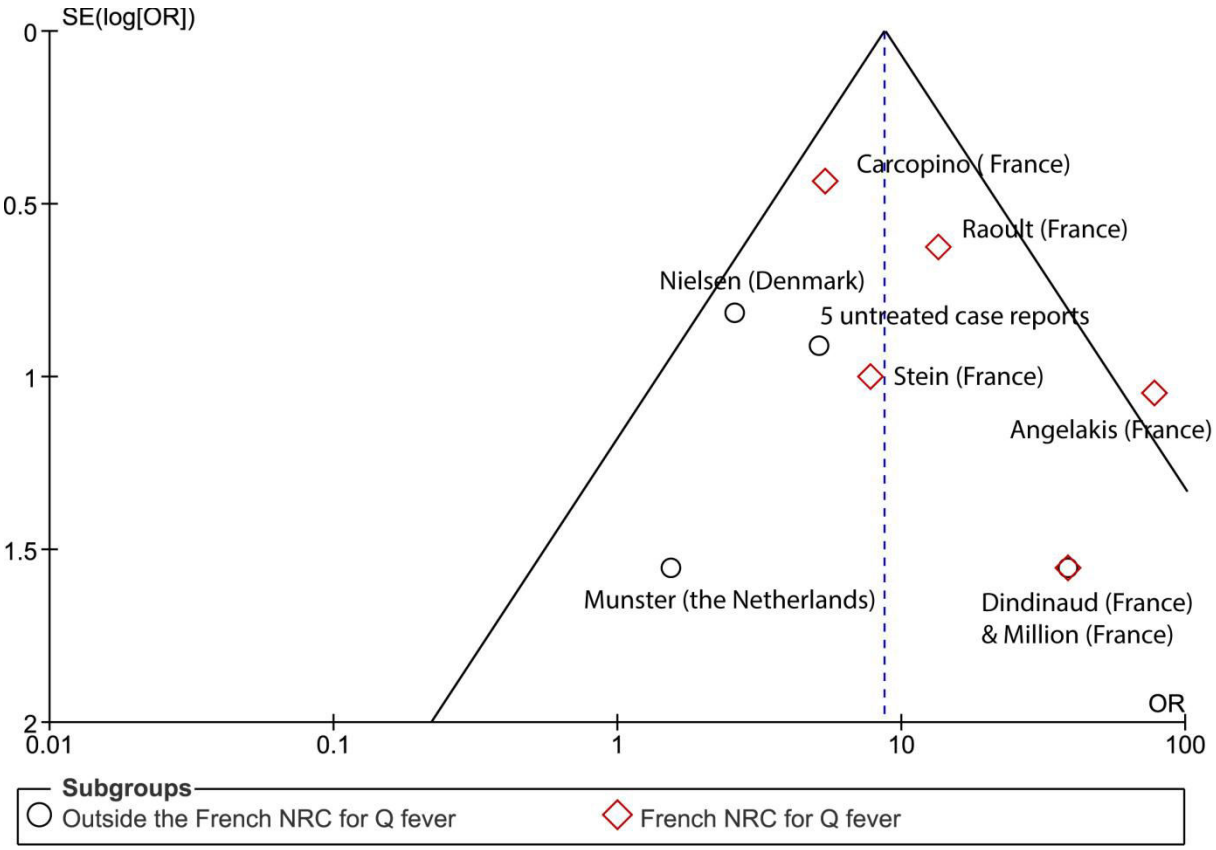


Supplementary Figure 3. Comparison of cumulative incidence ratio of fetal death according to age, Q fever and treatment



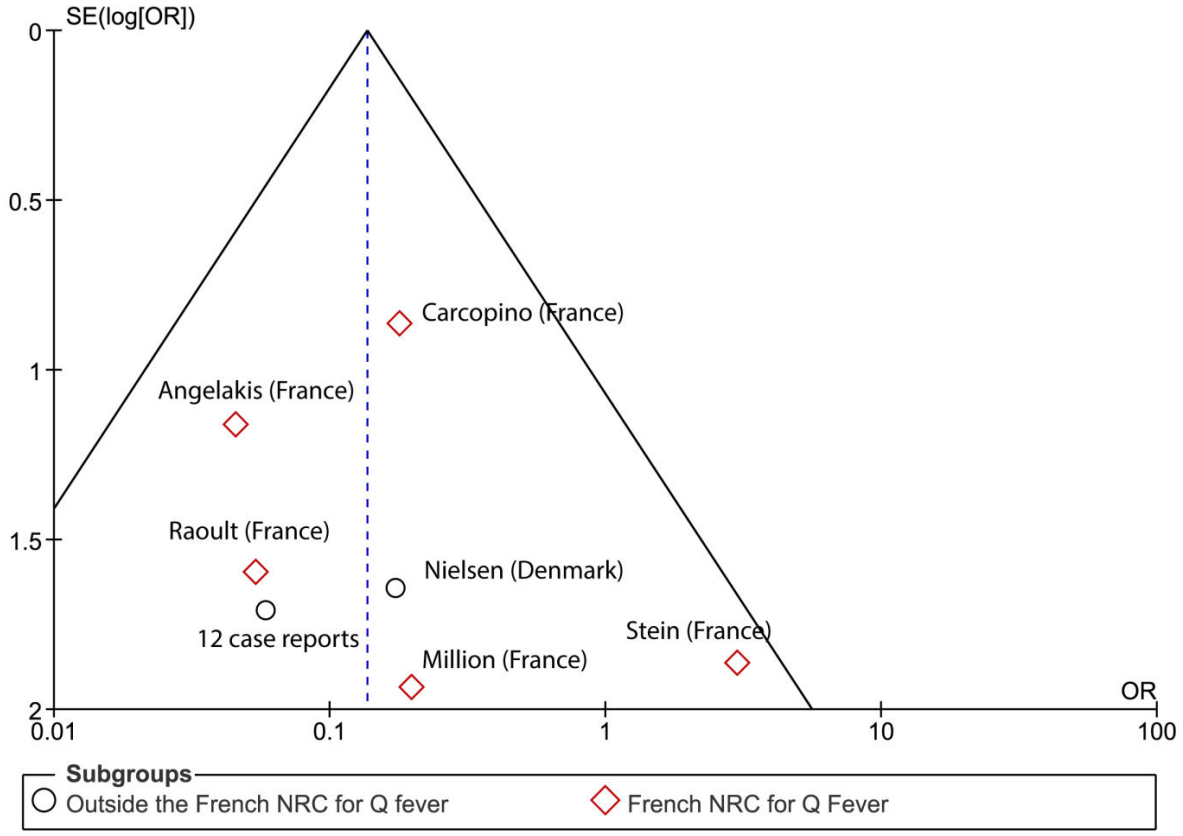
An increased cumulative incidence ratio of fetal death was found in the untreated Q fever population versus the general population (Cumulative incidence ratio 3.9; 95% confidence interval [CI], 3.2.-4.8; $P < .001$), but no difference was found for the treated Q fever population (1.2; 95% CI, 0.7-2.1; $P = .48$).

Supplementary Figure 4. Funnel plot of studies included in the meta-analysis assessing the fetal death rate in untreated Q fever pregnancies compared to the general population



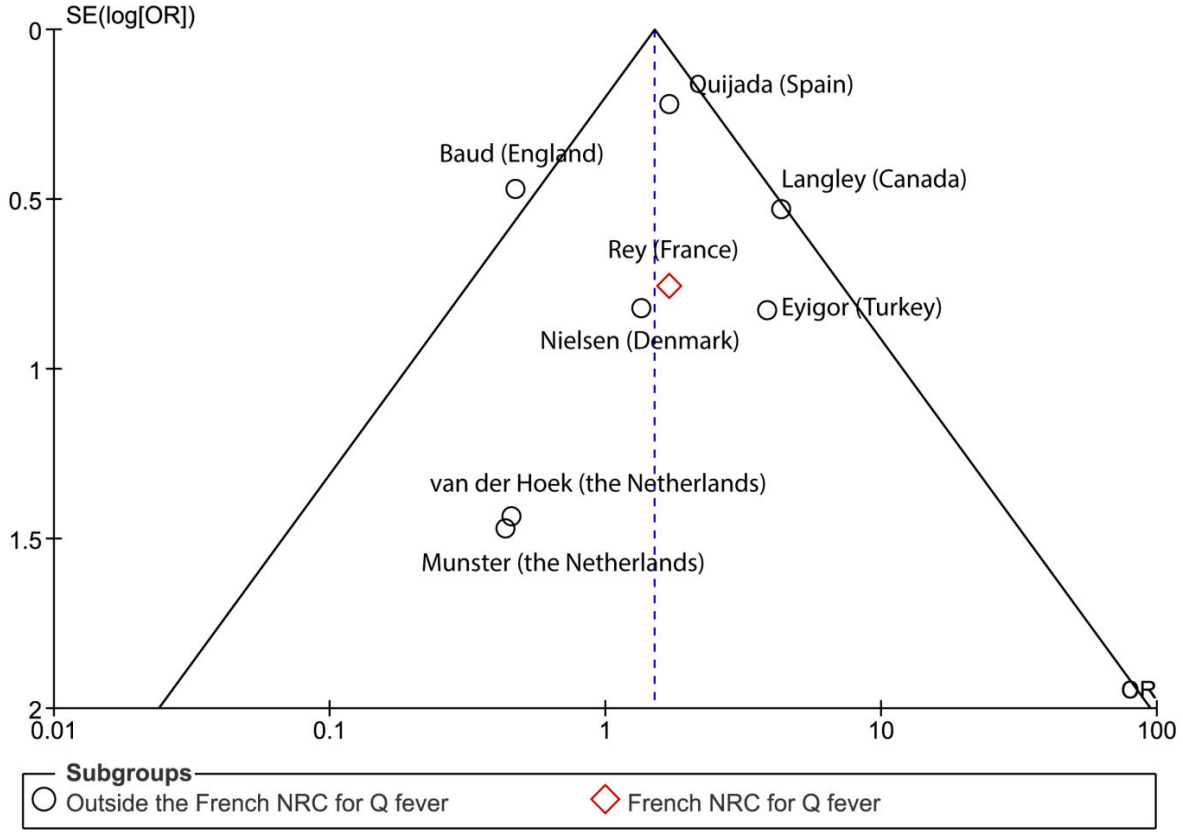
Standard difference in means was plotted against the standard error of the difference in means of each study to identify asymmetry in the distribution of trials (random effect model). The overall effect estimate is indicated by the vertical line. The two diagonal lines represent the 95 % confidence interval. Each comparison is represented by a single dot. Examination of the funnel plot revealed an outlier on the right part corresponding to increased fetal death rate (Angelakis, Eur J Clin Microbiol Infect Dis, 2013 [6]), and this outlier was subsequently excluded.

Supplementary Figure 5. Funnel plot of studies included in the meta-analysis assessing the effectiveness of antibiotics for the prevention of fetal death associated with Q fever during pregnancy



Standard difference in means was plotted against the standard error of the difference in means of each study to identify asymmetry in the distribution of trials (random effect model). The overall effect estimate is indicated by the vertical line. The two diagonal lines represent the 95 % confidence interval. Each comparison is represented by a single dot. No outlier was identified.

Supplementary Figure 6. Funnel plot of published studies on the association between positive *Coxiella burnetii* serology and fetal death



Standard difference in means was plotted against the standard error of the difference in means of each study to identify asymmetry in the distribution of trials (random effect model). The overall effect estimate is indicated by the vertical line. The two diagonal lines represent the 95 % confidence interval. Each comparison is represented by a single dot. Examination of the funnel plot revealed an outlier to the left of the mean (Baud, Clin Microbiol Infect, 2009 [23]), and this outlier was excluded. Strikingly, the two studies from the Netherlands have very similar results.

References

(1) Million M, Thuny F, Richet H, Raoult D. Long-term outcome of Q fever endocarditis: a 26-year personal survey. Lancet Infect Dis **2010 Aug**; 10(8):527-35.

(2) Rouli L, Rolain JM, El FA, Robert C, Raoult D. Genome sequence of *Coxiella burnetii* 109, a doxycycline-resistant clinical isolate. J Bacteriol **2012 Dec**; 194(24):6939.

(3) Raoult D, Fenollar F, Stein A. Q fever during pregnancy: diagnosis, treatment, and follow-up. Arch Intern Med **2002 Mar 25**; 162(6):701-4.

(4) Stein A, Raoult D. Q fever during pregnancy: a public health problem in southern France. Clin Infect Dis **1998 Sep**; 27(3):592-6.

(5) Carcopino X, Raoult D, Bretelle F, Boubli L, Stein A. Managing Q fever during pregnancy: the benefits of long-term cotrimoxazole therapy. Clin Infect Dis **2007 Sep 1**; 45(5):548-55.

(6) Angelakis E, Million M, D'Amato F, et al. Q fever and pregnancy: disease, prevention, and strain specificity. Eur J Clin Microbiol Infect Dis **2013 Mar**; 32(3):361-8.

(7) International clearinghouse for birth defects surveillance and research. Annual report 2010. Roma, Italy: The centre of the international clearinghouse for birth defects surveillance and research; **2010**.

(8) Nybo Andersen AM, Wohlfahrt J, Christens P, Olsen J, Melbye M. Maternal age and fetal loss: population based register linkage study. BMJ **2000 Jun 24**; 320(7251):1708-12.

(9) Jewell NP. Statistics for Epidemiology. New York: **2013**.

(10) Stata Statistical Software: Release 12. College Station, TX: StataCorp LP.: **2011**.

(11) Review manager (RevMan) Version 5.2. Copenhagen: The Nordic Cochrane centre: **2012**.

(12) McGivern D, White R, Paul ID, Caul EO, Roome AP, Westmoreland D. Concomitant zoonotic infections with ovine Chlamydia and 'Q' fever in pregnancy: clinical features, diagnosis, management and public health implications. Case report. Br J Obstet Gynaecol **1988 Mar**; 95(3):294-8.

(13) Syrucek L, Sobeslavsky O, Gutvirth I. Isolation of *Coxiella burnetii* from human placentas. J Hyg Epidemiol Microbiol Immunol **1958**; 2(1):29-35.

(14) Nielsen S, Molbak K, Henriksen T, Krogfelt K, Larsen C, Villumsen S. Q fever during pregnancy and maternofoetal consequences: a case series from Denmark. Emerg Infect Dis. In press. 2014.

(15) Gaburro D, Del Campo A. Considerazioni epidemiologiche e cliniche su un'infezione da *Coxiella burnetii* in tre gemelle immature. 8 ed. **1956**.

(16) Bental T, Fejgin M, Keysary A, et al. Chronic Q fever of pregnancy presenting as *Coxiella burnetii* placentitis: successful outcome following therapy with erythromycin and rifampin. Clin Infect Dis **1995 Nov**; 21(5):1318-21.

(17) Hellmeyer L, Schmitz-Ziegler G, Slenczka W, Schmidt S. Q Fever in pregnancy: a case report and review of the literature. Z Geburtshilfe Neonatol **2002 Sep**; 206(5):193-8.

(18) Ellis ME, Smith CC, Moffat MA. Chronic or fatal Q-fever infection: a review of 16 patients seen in North-East Scotland (1967-80). Q J Med **1983**; 52(205):54-66.

(19) Numazaki K, Ueno H, Yokoo K, Muramatsu Y, Chiba S, Morita C. Detection of serum antibodies to *Bartonella henselae* and *Coxiella burnetii* from Japanese children and pregnant women. Microbes Infect **2000 Oct**; 2(12):1431-4.

(20) Fiset P, Wisseman CL, Batawi YE. Immunologic evidence of human fetal infection with *Coxiella burnetii.* Am J Epidemiol **1975 Jan**; 101(1):65-9.

(21) Nielsen SY, Andersen AM, Molbak K, et al. No excess risk of adverse pregnancy outcomes among women with serological markers of previous infection with *Coxiella burnetii*: evidence from the Danish National Birth Cohort. BMC Infect Dis **2013**; 13:87.

(22) Nielsen SY, Hjollund NH, Andersen AM, et al. Presence of antibodies against *Coxiella burnetii* and risk of spontaneous abortion: a nested case-control study. PLoS ONE **2012**; 7(2):e31909.

(23) Baud D, Peter O, Langel C, Regan L, Greub G. Seroprevalence of *Coxiella burnetii* and Brucella abortus among pregnant women. Clin Microbiol Infect **2009 May**; 15(5):499-501.

(24) Raoult D, Stein D. Q fever during pregnancy - a risk for women, fetuses, and obstetricians (letter). N Engl J Med **1994**;(330):371.

CONCLUSIONS ET PERSPECTIVES

La génomique a représenté un outil important pour progresser rapidement dans la connaissance de *Coxiella burnetii* et améliorer la gestion de la fièvre Q. Cependant, de nombreux aspects de cette bactérie doivent encore être étudiés. Premièrement, bien que le séquençage de l'ARN ribosomique 16S ait permis d'établir la taxonomie correcte de *C. burnetii*, son origine évolutive n'est toujours pas claire. Récemment il a été émis l'hypothèse selon laquelle un ancêtre commun serait issu d'une Coxiella endosymbionte des tiques molles . De plus, la physiopathologie reste encore incomplète et la corrélation entre les souches et les effets cliniques sont à élucider. Plusieurs hypothèses ont été proposées telles que le lien avec le type de plasmide ou le groupe génomique et plusieurs études sur des modèles animaux ont été effectuées. Cependant, les résultats ne reflètent pas souvent ces hypothèses ou ne sont pas cohérents avec ce qui est observé chez l'homme. Par exemple, le plasmide QpH1, qui initialement a été associé à des formes aiguës de la maladie, a été retrouvé plus tard dans les symptômes les plus graves de formes non aiguës, comme l'endocardite. De même, les observations sur des modèles expérimentaux montrant un niveau inférieur de pathogénicité pour la souche Q212 ne laissaient pas penser que le génotype MST21 aurait été le seul clone détecté au Canada et responsable de plusieurs épidémies.

Au cours de ces trois années de thèse, nous avons cherché les propriétés génétiques et les particularités des souches épidémiques afin d'arriver à distinguer différents pathotypes et d'établir la relation avec leurs effets cliniques. Nos résultats ont été cohérents avec les études phylogénétiques précédentes qui ont montré que *C.burnetii* n'a pas évolué par l'acquisition d'un matériel génétique étranger mais plutôt, par la modification et la perte de gènes existants. Après l'analyse du clone de la souche épidémique aux Pays-Bas nous

avons pu exclure tout cas d'acquisition ou de perte de gènes par rapport à la référence NMI. Les différences trouvées entre les deux souches étaient uniquement basées sur des mutations ponctuelles. Notre possible explication au mécanisme qui a permis au génotype 33 de déterminer cette large épidémie est basée sur le changement des antigènes de la surface associé à une surpopulation de chèvres dans les zones périurbaines ayant le système immunitaire naïf contre ce nouveau clone émergent.

Concernant le clone épidémique en Guyane française, suite à la découverte de la délétion d'une région de 6105pb, nous avons confirmé ce qui a été précédemment décrit pour les bactéries les plus pathogènes: la virulence des bactéries intracellulaires est souvent liée à la perte des gènes impliqués dans la communication avec l'hôte plutôt qu'à l'acquisition de gènes externes. Comme décrit pour les « bad bugs », *Rickettsia prowazekii*, *Mycobacterium tuberculosis, Shigella dysenteriae* ou *Yersinia pestis*, l'hyper virulence est entraînée par la réduction du génome avec une diminution significative du contenu en gênes .

Nous avons démontré que cette large délétion est spécifique au génotype "Cayenne". La transformation génétique de Nine Mile, par « knock out », entraînant la suppression de ces possibles gènes de non-virulence, permettra de vérifier sur les modèles expérimentaux et cellulaires si la délétion est l'explication de la virulence du clone circulant en Guyane.

Comme précédemment discuté, la délétion contient l'opéron hlyCABD du système de sécrétion de type I (T1SS). Bien que la stratégie de virulence de *C.burnetii* soit basée sur le système de sécrétion de type IV , certains composants des systèmes de types I et II sont présents dans son génome. La fonctionnalité de ces systèmes n'a pas encore été établie. Dans *C.burnetii* NMI, le gène HlyB a été muté en deux pseudogènes, HlyA et HlyC sont absents, tandis que HlyD est toujours complet et peut être fonctionnel. Récemment, dans

L. pneumophila, il a été démontré que, malgré que T1SS ne soit pas complet, il peut encore être fonctionnel et impliqué dans le mécanisme d'internalisation dans les cellules hôtes .

Il n'y a aucune preuve expérimentale dans la littérature qui montre que T1SS ne soit pas actif chez *C.burnetii*. Les modèles expérimentaux pourront y répondre et aideront à comprendre si la virulence du clone est corrélée à la réduction génomique avec une éventuelle inactivation de T1SS. Il serait également intéressant d'étudier la réponse des macrophages et de définir la modulation de la transcription induite par l'infection par les souches épidémiques comme Cb175 et/ou Z3055.

Concernant les mutations non-synonymes trouvées dans Z3055, il serait intéressant d'analyser les effets sur les produits des gènes et confirmer le rôle de la dérive génétique dans l'émergence d'un nouveau clone. De plus, il faut remarquer que dans Cb175 et Z3055 nous avons trouvé un taux similaire de mutations non synonymes. Une analyse comparative avec toutes les souches de la collection de notre laboratoire, la plus grande dans le monde, permettra d'observer si ces catégories sont les plus souvent intéressées par les mutations non silencieuses et si celles-ci sont spécifiquement liées à l'hyper-pathogénicité.

Nos résultats préliminaires suggèrent que le pangénome de *Coxiella burnetii* est fermé, avec un ratio core génome/ pangénome de 96% calculé lors de l'étude avec 7 génomes entiers. Une étude plus complète avec plus de 40 souches permettra de confirmer cette hypothèse.

REMERCIEMENTS

Je tiens à remercier :

Mon directeur de thèse, Monsieur le **Professeur Didier RAOULT** pour m'avoir accueillie au sein de son laboratoire et m'avoir encadré pendant ces trois années.

Monsieur le **Professeur Jean-Louis MEGE** pour m'avoir fait l'honneur d'accepter la présidence de ce jury de thèse.

Monsieur le **Docteur Olivier DURON** et Monsieur **le Professeur Max MAURIN** pour avoir accepté de faire partie du jury et d'évaluer mon travail de thèse

Je remercie : **Aurélia**, **Carole**, **Micheline** et **Samad** pour avoir consacré du temps à la correction de mon français.

Je remercie pour leur travail précieux et essentiel pour que cette de thèse ait lieu :

Madame **Nathalie Duclos** et toute l'équipe de Monsieur le Professeur La Scola pour la mise en culture des souches de *Coxiella burnetii*

Madame **Catherine Robert**, **Thi Tien Nguyen**, **Noémie Labas** et **Caroline Blanc-Tailleur** de s'être occupées du séquençage

Je remercie les personnes qui ont travaillé avec moi : **Carole Eldin**, **Laetitia Rouli**, **Jeremy Delerce**, **Sophie Edouard**, **Matthieu Million**

et les autres co-auteurs : **Kalliopi Georgiades**, **Azhar EI**, **Pierre-Edouard Fournier**, **Emmanuel Angelakis**, **Tom Marrie**, **Judith Tyczka**.

Je remercie et je m'excuse avec les gens qui m'ont écoutée et aidée pendant les moments de déception et découragement.

REFERENCES des avant-propos

1. Seshadri R, Paulsen IT, Eisen JA *et al*. Complete genome sequence of the q-fever pathogen coxiella burnetii. *Proceedings of the National Academy of Sciences of the United States of America* 100(9), 5455-5460 (2003).
2. Glazunova O, Roux V, Freylikman O *et al*. Coxiella burnetii genotyping. *Emerging infectious diseases* 11(8), 1211-1217 (2005).
3. De Bruin A, Van Alphen PT, Van Der Plaats RQ *et al*. Molecular typing of coxiella burnetii from animal and environmental matrices during q fever epidemics in the netherlands. *BMC veterinary research* 8, 165 (2012).
4. Klaassen CH, Nabuurs-Franssen MH, Tilburg JJ, Hamans MA, Horrevorts AM. Multigenotype q fever outbreak, the netherlands. *Emerging infectious diseases* 15(4), 613-614 (2009).
5. Roest HI, Ruuls RC, Tilburg JJ *et al*. Molecular epidemiology of coxiella burnetii from ruminants in q fever outbreak, the netherlands. *Emerging infectious diseases* 17(4), 668-675 (2011).
6. Tilburg JJ, Roest HJ, Buffet S *et al*. Epidemic genotype of coxiella burnetii among goats, sheep, and humans in the netherlands. *Emerging infectious diseases* 18(5), 887-889 (2012).
7. Tilburg JJ, Roest HJ, Nabuurs-Franssen MH, Horrevorts AM, Klaassen CH. Genotyping reveals the presence of a predominant genotype of coxiella burnetii in consumer milk products. *Journal of clinical microbiology* 50(6), 2156-2158 (2012).
8. Mahamat A, Edouard S, Demar M *et al*. Unique clone of coxiella burnetii causing severe q fever, french guiana. *Emerging infectious diseases* 19(7), 1102-1104 (2013).
9. Angelakis E, Johani S, Ahsan A, Memish Z, Raoult D. Q fever endocarditis and new coxiella burnetii genotype, saudi arabia. *Emerging infectious diseases* 20(4), 726-728 (2014).
10. Rasko DA, Myers GS, Ravel J. Visualization of comparative genomic analyses by blast score ratio. *BMC bioinformatics* 6, 2 (2005).
11. Hackert VH, Van Der Hoek W, Dukers-Muijrers N *et al*. Q fever: Single-point source outbreak with high attack rates and massive numbers of undetected infections across an entire region. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* 55(12), 1591-1599 (2012).
12. Speelman P. The largest q fever outbreak ever reported. *The Netherlands journal of medicine* 68(12), 380-381 (2010).
13. Pfaff F, Francois A, Hommel D *et al*. Q fever in french guiana: New trends. *Emerging infectious diseases* 4(1), 131-132 (1998).
14. Grangier C DM, Ardillon V, Mahamat a, Fournier P, Simmonnet C, Et Al. . Epidemiologie de la fièvre q en guyanne, 1990–2006. . *Le Bulletin de Veille Sanitaire.CIRE Antilles Guyane* 10 (10), 2– 4. (2009).
15. Epelboin L, Chesnais C, Boulle C *et al*. Q fever pneumonia in french guiana: Prevalence, risk factors, and prognostic score. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* 55(1), 67-74 (2012).
16. Medini D, Donati C, Tettelin H, Masignani V, Rappuoli R. The microbial pan-genome. *Current opinion in genetics & development* 15(6), 589-594 (2005).
17. Chun J, Grim CJ, Hasan NA *et al*. Comparative genomics reveals mechanism for short-term and long-term clonal transitions in pandemic vibrio cholerae. *Proceedings of the National Academy of Sciences of the United States of America* 106(36), 15442-15447 (2009).
18. Rasko DA, Webster DR, Sahl JW *et al*. Origins of the e. Coli strain causing an outbreak of hemolytic-uremic syndrome in germany. *The New England journal of medicine* 365(8), 709-717 (2011).
19. Georgiades K, Raoult D. Defining pathogenic bacterial species in the genomic era. *Frontiers in microbiology* 1, 151 (2010).

20. Diene SM, Merhej V, Henry M *et al*. The rhizome of the multidrug-resistant enterobacter aerogenes genome reveals how new "killer bugs" are created because of a sympatric lifestyle. *Molecular biology and evolution* 30(2), 369-383 (2013).
21. Lefebure T, Bitar PD, Suzuki H, Stanhope MJ. Evolutionary dynamics of complete campylobacter pan-genomes and the bacterial species concept. *Genome biology and evolution* 2, 646-655 (2010).
22. D'auria G, Jimenez-Hernandez N, Peris-Bondia F, Moya A, Latorre A. Legionella pneumophila pangenome reveals strain-specific virulence factors. *BMC genomics* 11, 181 (2010).
23. Imperi F, Antunes LC, Blom J *et al*. The genomics of acinetobacter baumannii: Insights into genome plasticity, antimicrobial resistance and pathogenicity. *IUBMB life* 63(12), 1068-1074 (2011).
24. Rolain JM, Vayssier-Taussat M, Saisongkorh W *et al*. Partial disruption of translational and posttranslational machinery reshapes growth rates of bartonella birtlesii. *mBio* 4(2), e00115-00113 (2013).
25. Barrick JE, Lenski RE. Genome-wide mutational diversity in an evolving population of escherichia coli. *Cold Spring Harbor symposia on quantitative biology* 74, 119-129 (2009).
26. Bruggemann H, Cazalet C, Buchrieser C. Adaptation of legionella pneumophila to the host environment: Role of protein secretion, effectors and eukaryotic-like proteins. *Current opinion in microbiology* 9(1), 86-94 (2006).
27. Zusman T, Yerushalmi G, Segal G. Functional similarities between the icm/dot pathogenesis systems of coxiella burnetii and legionella pneumophila. *Infection and immunity* 71(7), 3714-3723 (2003).
28. Grodeland G, Mjaaland S, Roux KH, Fredriksen AB, Bogen B. DNA vaccine that targets hemagglutinin to mhc class ii molecules rapidly induces antibody-mediated protection against influenza. *Journal of immunology* 191(6), 3221-3231 (2013).
29. Edouard S, Mahamat A, Demar M, Abboud P, Djossou F, Raoult D. Comparison between emerging q fever in french guiana and endemic q fever in marseille, france. *The American journal of tropical medicine and hygiene* 90(5), 915-919 (2014).
30. Cole ST, Eiglmeier K, Parkhill J *et al*. Massive gene decay in the leprosy bacillus. *Nature* 409(6823), 1007-1011 (2001).
31. Sakharkar KR, Dhar PK, Chow VT. Genome reduction in prokaryotic obligatory intracellular parasites of humans: A comparative analysis. *International journal of systematic and evolutionary microbiology* 54(Pt 6), 1937-1941 (2004).
32. Andersson JO, Andersson SG. Insights into the evolutionary process of genome degradation. *Current opinion in genetics & development* 9(6), 664-671 (1999).
33. Wixon J. Featured organism: Reductive evolution in bacteria: Buchnera sp., rickettsia prowazekii and mycobacterium leprae. *Comparative and functional genomics* 2(1), 44-48 (2001).
34. Merhej V, Raoult D. Rickettsial evolution in the light of comparative genomics. *Biological reviews of the Cambridge Philosophical Society* 86(2), 379-405 (2011).
35. Fournier PE, El Karkouri K, Leroy Q *et al*. Analysis of the rickettsia africae genome reveals that virulence acquisition in rickettsia species may be explained by genome reduction. *BMC genomics* 10, 166 (2009).
36. Georgiades K, Raoult D. Genomes of the most dangerous epidemic bacteria have a virulence repertoire characterized by fewer genes but more toxin-antitoxin modules. *PloS one* 6(3), e17962 (2011).
37. Cazalet C, Gomez-Valero L, Rusniok C *et al*. Analysis of the legionella longbeachae genome and transcriptome uncovers unique strategies to cause legionnaires' disease. *PLoS genetics* 6(2), e1000851 (2010).

38. Kozak NA, Buss M, Lucas CE *et al*. Virulence factors encoded by legionella longbeachae identified on the basis of the genome sequence analysis of clinical isolate d-4968. *Journal of bacteriology* 192(4), 1030-1044 (2010).
39. Weisburg WG, Dobson ME, Samuel JE *et al*. Phylogenetic diversity of the rickettsiae. *Journal of bacteriology* 171(8), 4202-4206 (1989).
40. Duron O, Noel V, Mccoy KD *et al*. The recent evolution of a maternally-inherited endosymbiont of ticks led to the emergence of the q fever pathogen, coxiella burnetii. *PLoS pathogens* 11(5), e1004892 (2015).
41. Samuel JE, Frazier ME, Mallavia LP. Correlation of plasmid type and disease caused by coxiella burnetii. *Infection and immunity* 49(3), 775-779 (1985).
42. Hendrix LR, Samuel JE, Mallavia LP. Differentiation of coxiella burnetii isolates by analysis of restriction-endonuclease-digested DNA separated by sds-page. *Journal of general microbiology* 137(2), 269-276 (1991).
43. Minnick MF, Heinzen RA, Reschke DK, Frazier ME, Mallavia LP. A plasmid-encoded surface protein found in chronic-disease isolates of coxiella burnetti. *Infection and immunity* 59(12), 4735-4739 (1991).
44. Minnick MF, Small CL, Frazier ME, Mallavia LP. Analysis of the cbhe' plasmid gene from acute disease-causing isolates of coxiella burnetii. *Gene* 103(1), 113-118 (1991).
45. Russell-Lodrigue KE, Andoh M, Poels MW *et al*. Coxiella burnetii isolates cause genogroup-specific virulence in mouse and guinea pig models of acute q fever. *Infection and immunity* 77(12), 5640-5650 (2009).
46. Russell-Lodrigue KE, Zhang GQ, Mcmurray DN, Samuel JE. Clinical and pathologic changes in a guinea pig aerosol challenge model of acute q fever. *Infection and immunity* 74(11), 6085-6091 (2006).
47. Stein A, Louveau C, Lepidi H *et al*. Q fever pneumonia: Virulence of coxiella burnetii pathovars in a murine model of aerosol infection. *Infection and immunity* 73(4), 2469-2477 (2005).
48. Thiele D, Willems H. Is plasmid based differentiation of coxiella burnetii in 'acute' and 'chronic' isolates still valid? *European journal of epidemiology* 10(4), 427-434 (1994).
49. Voth DE, Heinzen RA. Coxiella type iv secretion and cellular microbiology. *Current opinion in microbiology* 12(1), 74-80 (2009).
50. Fuche F, Vianney A, Andrea C, Doublet P, Gilbert C. Functional type 1 secretion system involved in legionella pneumophila virulence. *Journal of bacteriology* 197(3), 563-571 (2015).

RÉSUMÉ

Coxiella burnetii est l'agent pathogène responsable de la fièvre Q. Ce coccobacille gram négatif appartient à la famille des γ-protéobactéries. Grâce à de nombreuses études effectuées, nos connaissances sur sa taxonomie, son mode de vie et sa résistance aux antibiotiques se sont approfondies. Cependant, plusieurs aspects nécessitent encore plus d'investigations afin d'améliorer notre compréhension notamment des liens entre les différentes souches ainsi que les relations entre le génome et la virulence.

Dans le cadre de cette thèse nous nous sommes intéressés à l'étude de souches de *C.burnetii* responsables d'événements épidémiques. Nous avons séquencé une souche de génotype MST33 (Z3055), proche de la souche responsable de l'épidémie de fièvre Q aux Pays-Bas, et une souche de génotype MST17 (Cb175) clone provoquant l'une des formes les plus virulentes de fièvre Q aiguë jamais décrite auparavant et retrouvée à ce jour uniquement en Guyane Française. Ensuite, nous avons effectué une analyse génomique comparative de ces souches avec d'autres souches disponibles dans GenBank. Nous avons effectué un génotypage en utilisant la méthode de « multispacer typing (MST) » et, afin d'explorer les niveaux de conservation et les différences génétiques entre les souches étudiées, nous avons utilisé une approche basée sur l'algorithme de Blast Ratio Score (BSR).

Les résultats de ces analyses montrent que le génome de la souche Z3055 était très similaire à celui de la souche de référence Nine Mile I. Les différences observées sont liées à la présence de mutations non synonymes dans le génome de Z3055. Ces mutations sont plus fréquentes dans les gènes codant pour des protéines membranaires, des protéines contenant des domaines d'ankyrine, des facteurs de transcription et de traduction. Cette grande proportion de protéines membranaires mutées pourrait expliquer l'ampleur de cette épidémie en Hollande. En effet, le changement de profil antigénique pourrait être à l'origine de la formation d'un nouveau sérotype capable d'échapper à la réponse immunitaire de l'hôte et de diffuser facilement dans une population au système immunitaire naïf.

Nous avons d'ailleurs montré que la souche responsable de la fièvre Q en Guyane (Cb175) présente des différences chromosomiques importantes par rapport à NMI, bien au contraire de la souche Z3055. Ces différences se manifestent principalement par la présence d'une délétion d'une région de 6105pb contenant l'opéron hlyCABD du système de sécrétion de type 1 (T1SS). Ce résultat est cohérent avec ce qui a été observé chez les bactéries épidémiques les plus dangereuses comparées à leurs espèces non-épidémiques plus proches qui ont un génome réduit et contiennent moins de protéines du système de sécrétion.

D'autre part, le génotypage des souches de notre laboratoire révèle la corrélation entre le génome et leur répartition géographique, confirmant la notion de « geotyping ». En effet, un seul génotype MST21, a été détecté au Canada au cours des récentes épidémies de fièvre Q rencontrées dans ce pays. Nous avons séquencé la souche canadienne épidémique DOG UTAD (MST21) et comparé le « draft génome » correspondant avec une souche de même génotype. Les résultats indiquent que la souche DOG UTAD est très similaire à CbuG_Q212 en partageant le même génotype MST21 et la même origine géographique (Canada), ce qui suggère une évolution clonale récente. Nous avons, aussi, séquencé la souche Cb196 de génotype MST51 (uniquement décrit en Arabie Saoudite) et avons observé une forte similarité de séquence avec CbuK_Q154.

Enfin, l'analyse par BSR indique un ratio core génome/pangenome de 96%, ce qui suggère que le pangénome de *C. burnetii* est fermé. Une analyse pangénomique globale impliquant toutes les souches collectées dans notre laboratoire permettra de confirmer ce résultat préliminaire et de clarifier les liens entre l'origine géographique, le génotype et la virulence des souches.

ABSTRACT

Coxiella burnetii is a human pathogen that causes the zoonotic disease Q fever. This coccobacillus gram-negative belongs to the family of γ-proteobacteria. Several studies on this pathogen have improved our understanding of its taxonomy, lifestyle, virulence and the resistance to antibiotics. However, until nowadays, our knowledge of many aspects is still weak and need more highlights, such as the relationship between strains and the link between genome and virulence.

In order to answer these questions, we focused on the study of strains responsible for epidemic events. Particularly, we sequenced the clone of the strain responsible for Netherlands outbreak having genotype MST33 (Z3055), and strain having MST17 (Cb175) responsible for one of the most severe form of acute Q fever never reported in literature and uniquely described in French Guiana. We performed a comparative genomic analysis of *C. burnetii* strains sequenced in our laboratory with other strains whose genomes are available in GenBank database. The genotyping of *C. burnetii* was performed using multispacer sequence typing (MST) based on the most variable intergenic spacers of *C. burnetii* genome. In the aim to explore the conservation levels and genetic differences among the strains studied in this thesis, we performed the Blast Score Ratio (BSR).

Our findings showed that the Netherlands outbreak responsible strain (clone Z3055) was highly similar to the reference strain Nine Mile I. Only slight differences were observed, which were related to non-synonymous mutations in Z3055 genome. These mutations were significantly increased in genes encoding membrane proteins, ankyrin repeat domains containing proteins, transcription factors and translation proteins. The high proportion of mutated membrane proteins could explain this large-scale outbreak. Change of antigenic profile may have led to a new serotype, conferring to the novel clone the capacity to escape the host immune response and to disseminate easily in a immunologically naïve population.

On the contrary, the type strain responsible for Q fever in Guiana (Cb175) showed an important difference in its chromosome sequence compared to the reference NMI because of the deletion of a sequence of 6105bp containing the Type 1 secretion systems (T1SS) hlyCABD operon. This result appear consistent with previous findings that showed the most dangerous epidemic bacteria compared with their closest non-epidemic species are characterized by reduced genomes accompanied by significant decrease in ORF content and contain less secretion system proteins.

Moreover, the genotyping of our strains has shown a link between genome and geographic distribution of *C.burnetii*, confirming the concept of "geotyping". In fact, a single genotype MST21 has been detected in Canada during the multiple cases of Q fever. We sequenced the epidemic strain DOG UTAD (MST21) and we compared the corresponding draft genome to a strain of the same genotype. The results indicated a high similarity between DOG UTAD and CbuG_Q212. These two strains share the same genotype MST21 and were both isolated in Canada, suggesting a recent clonal radiation. Recently, a novel and unique genotype MST51 has been described in Saudi Arabia and we sequenced the corresponding strain Cb196 of the same genotype. The results showed a high sequence similarity with strain CbuK_Q154 with a phylogenetically close genotype (MST8).

Finally, results of the analysis based on the Blast Score Ratio considering 7 whole genomes showed an estimated ratio core genome/pangenome of 96%, suggesting that the pangenome of *C. burnetii* is closed. A larger pangenomic study involving all strains (more than 40) collected in our laboratory will allow confirming this result. It would also highlight the pathophysiology of Q fever by investigating the link between geographical origin, genotype, antibiotic susceptibility and other genetic specificities with clinical issues.