**Methods**

**Molecular epidemiology of Q fever: a review of Coxiella burnetii genotyping methods and main achievements**

K. Sidi-Boumedine, E. Rousset

ANSES, Laboratory of Sophia-Antipolis, Ruminant Pathology Unit, France

The strictly intracellular bacterium, *C. burnetii*, is the causative agent of Q fever, a worldwide zoonosis. Rapid identification of the spread of a given *C. burnetii* strain within and between hosts population is a prerequisite for thorough epidemiological investigations (human outbreaks, surveillance, follow-up of prophylactic measures efficiency). Many genotyping methods have been explored for their ability to discriminate between *C. burnetii* isolates deriving from various geographic areas and hosts, primarily from patients exhibiting different symptoms. However, these methods have not yet been used extensively for molecular epidemiology studies of Q fever. This is thought to be mainly linked to the biologic particularities and fastidious nature of this pathogen. This review aims at describing the methods explored for the molecular typing of *C. burnetii* and emphasises on the latest approaches and developments. Therefore, the inventory of the methods is made to allow future users to select the most interesting for epidemiological investigations of Q fever, and thereby determine their interpretation criteria.

**Introduction**

The obligate intracellular Gram negative bacterium *Coxiella burnetii* is the causative agent of Q fever, a widespread zoonosis. Q fever has been reported worldwide, even in Artic areas, with the possible exception of New Zealand (Hilbink et al., 1993; Maurin and Raoult, 1999; Koch et al., 2010). In certain epidemiological circumstances, Q fever could result in explosive human outbreaks (van der Hoek et al., 2010). In addition, for particular risk groups, the public health impact and thereby also the impact on society and/or the economy can be significant (EFSA, 2010).

The reservoir of *C. burnetii* is vast, ranging from ruminants, domestic carnivores, wildlife mammals, birds to arthropods, mainly ticks (Cutler et al., 2007). This bacterium appeared to easily cross the species barrier. The different species may play a role in the dissemination or maintenance of the disease as pathogen-carriers or as vectors. However, uncertainties remain on the roles of most of them. Notwithstanding, domestic ruminants constitute the main reservoir of the disease, as major human outbreaks are associated with exposure to these animals. Indeed, animals may shed organisms through vaginal mucus, milk, urine and faeces. Large numbers of bacteria are found in the placenta and foetal fluids, particularly at kidding during an abortion wave due to Q fever. Moreover, the environmental resistance of *C. burnetii*, as a spore-like form, increases considerably the exposure risks because bacteria may persist for long periods and be spread. Inhalation of contaminated aerosols or dusts constitutes the primary mode of transmission to humans (Maurin and Raoult, 1999; Woldehiwet, 2004; Arnicou-Bouvery and Rodolakis, 2005).

Although sheep and goats are more frequently related to human Q fever outbreaks than other animal species, the sources of human infections remain often unidentified. The genotyping of *C. burnetii* isolates could contribute to explain the scenario leading to spread as well as outbreaks of Q fever, and thereby to find the best options for its control. Obviously, a systematic genotyping would provide a descriptive database to determine the temporal and geographical evolution of Q fever. A considerable range of methods have been explored for the molecular typing and characterisation of *C. burnetii* strains. However, investigations on the genotyping methods are still essential for gaining insight into the molecular epidemiology of *C. burnetii*.

The first complete genome sequence was achieved in 2003, although the Q fever agent was discovered in the thirtys. Identification of *C. burnetii* virulence factors has been thwarted because of the lack of genetic manipulation system. Besides, most of the investigators have primarily employed molecular typing methods to study the association between genetic polymorphisms of strains and the different symptoms of Q fever in humans. The molecular works on *C. burnetii* have been impaired by its intracellular lifestyle and fastidious nature. Moreover, the high infectious capacity of *C. burnetii* limits the exchange of strains, and thus in many typing studies are based only on local or regional isolates. It appears clearly that the majority of the studies carried out included a restricted number of strains. The comparisons of the methods were based only on the presence of few common strains, but not by performing the different methods in a same study. Thus, no typing method is validated or considered as a reference. The efforts to better understand the epidemiology of this important pathogen using molecular typing methods are more recent. The large-scale human outbreak in the Netherlands increased the emergency of this need (EFSA, 2010).

This review aims at describing the methods explored for *C. burnetii* typing, emphasising on the current approaches and developments. The Table 1 provides a helpful overview of the various methods characteristics, according to the literature. The promising potential in epidemiological investigations of Q fever of some of these methods is discussed and some insight for their use in particular epidemiologic situations is also given (Figure 1).
### Methods

#### Table 1. Characteristics of *C. burnetii* molecular typing methods already in use and/or that might be useful

<table>
<thead>
<tr>
<th>Typing Method</th>
<th>Discriminatory power</th>
<th>Profiles number (total tested strains)</th>
<th>Diversity index*</th>
<th>Repeatability</th>
<th>Reproducibility</th>
<th>Time to result (days)</th>
<th>Relative cost</th>
<th>Comments</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sequencing of entire genome</td>
<td>High</td>
<td>4 (4)</td>
<td>1</td>
<td>High</td>
<td>High</td>
<td>Months to years</td>
<td>Very high</td>
<td>Not all genes will be present in the sequenced strain</td>
<td>Beare et al., 2009</td>
</tr>
<tr>
<td>Microarray based genotyping</td>
<td>High</td>
<td>8 (23)</td>
<td>0.7154</td>
<td>Medium &gt; High</td>
<td>Medium &gt; High</td>
<td>Weeks to months</td>
<td>High</td>
<td>Initial selection of target genes might be time consuming</td>
<td>Beare et al., 2006</td>
</tr>
<tr>
<td>MLST and variants</td>
<td>Moderate &gt; High</td>
<td>34 (176)</td>
<td>0.9273</td>
<td>High</td>
<td>High</td>
<td>3+</td>
<td>High</td>
<td>Reliability dependent on DNA yield and purity</td>
<td>Glazunova et al., 2005</td>
</tr>
<tr>
<td>VNTR typing</td>
<td>Moderate &gt; High</td>
<td>9 (16)</td>
<td>0.9083</td>
<td>High</td>
<td>Potentially High</td>
<td>2–3</td>
<td>Medium &gt; High</td>
<td>Equipment: Medium to High. Labor &amp; Supplies: Medium.</td>
<td>Svraka et al., 2006</td>
</tr>
<tr>
<td>PFGE</td>
<td>Moderate &gt; High</td>
<td>20 (82)</td>
<td>0.8585</td>
<td>Medium &gt; High</td>
<td>Medium</td>
<td>3 to weeks</td>
<td>Equipment: High. Labor &amp; Supplies: High.</td>
<td>Jager et al., 1998</td>
<td></td>
</tr>
<tr>
<td>REA of entire genome</td>
<td>Moderate &gt; High</td>
<td>6 (32)</td>
<td>0.7944</td>
<td>Medium &gt; High</td>
<td>Medium</td>
<td>1–3</td>
<td>Medium</td>
<td>Discrimination depends on type and number of enzymes selected</td>
<td>Hendrix et al., 1991</td>
</tr>
<tr>
<td>Direct sequencing of one genetic region</td>
<td>Moderate &gt; High</td>
<td>4 (21)</td>
<td>0.5429</td>
<td>High</td>
<td>High</td>
<td>2–3</td>
<td>Equipment: Low to High. Labor &amp; Supplies: Low to High.</td>
<td>Zhang et al., 1997</td>
<td></td>
</tr>
<tr>
<td>Restriction endonuclease on a single amplified product</td>
<td>Low &gt; Moderate</td>
<td>5 (37)</td>
<td>0.6351</td>
<td>High</td>
<td>High</td>
<td>1–2</td>
<td>Medium</td>
<td>Patterns can vary with equipment used</td>
<td>Sekeyova et al., 1999</td>
</tr>
<tr>
<td>REP and RAPD PCR</td>
<td>Low &gt; Moderate</td>
<td>4 (5)</td>
<td>0.9</td>
<td>Medium</td>
<td>Low</td>
<td>1</td>
<td>Equipment: Low to Medium. Labor &amp; Supplies: Low.</td>
<td>Rustscheff et al., 2000</td>
<td></td>
</tr>
<tr>
<td>IS1111 PCR</td>
<td>Moderate</td>
<td>5 (21)</td>
<td>0.7286</td>
<td>Medium &gt; High</td>
<td>High</td>
<td>1</td>
<td>Low to Medium</td>
<td>IS absent from the test strains can be detected but those absent in the test strains cannot be detected</td>
<td>Denison et al., 2007</td>
</tr>
<tr>
<td>IRS-PCR</td>
<td>Moderate &gt; High</td>
<td>7 (14)</td>
<td>0.8791</td>
<td>Medium &gt; High</td>
<td>Medium &gt; High</td>
<td>1</td>
<td>Medium to High</td>
<td>Discrimination depends on type of enzymes selected</td>
<td>Arricau-Bouvery et al., 2006</td>
</tr>
<tr>
<td>Real time PCR SNP typing</td>
<td>Moderate &gt; High</td>
<td>9 (65)</td>
<td>0.7438</td>
<td>High</td>
<td>High</td>
<td>1</td>
<td>Medium to High</td>
<td>Discrimination depends on type and number of SNPs selected</td>
<td>Huijsmans et al., 2011</td>
</tr>
<tr>
<td>Plasmid profiles</td>
<td>Low</td>
<td>NA</td>
<td>NA</td>
<td>High</td>
<td>Medium</td>
<td>1</td>
<td>Equipment: Low</td>
<td>Discrimination could be enhanced by the use of RE digestion of the plasmids</td>
<td>Aarts et al., 2001</td>
</tr>
</tbody>
</table>

NA: Not available.

* Diversity Index is calculated according to Hunter and Gaston, 1988.
### Methods

#### Table 1. Characteristics of C. burnetii molecular typing methods already in use and/or that might be useful

<table>
<thead>
<tr>
<th>Typing Method</th>
<th>Discriminatory power</th>
<th>Profiles number (total tested strains)</th>
<th>Diversity Index*</th>
<th>Repeatability</th>
<th>Reproducibility</th>
<th>Time to result (days)</th>
<th>Relative cost</th>
<th>Comments</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>RFLP based on Southern blotting</td>
<td>Moderate &gt; High (depends on number of bands observed)</td>
<td>NA</td>
<td>NA</td>
<td>Medium &gt; High</td>
<td>Medium</td>
<td>1–3</td>
<td>Medium</td>
<td>Could help resolving the number and positions of IS elements</td>
<td>van Soolingen, 2001</td>
</tr>
<tr>
<td>AFLP</td>
<td>Moderate &gt; High</td>
<td>NA</td>
<td>NA</td>
<td>High</td>
<td>Medium&gt;High</td>
<td>2</td>
<td>Equipment: Low to Medium Labor &amp; Supplies: Low</td>
<td>Proved to be useful for very homogeneous intracellular bacteria</td>
<td>Boumedine and Rodolakis, 1998</td>
</tr>
<tr>
<td>AP-PCR (using random primers)</td>
<td>Low &gt; Moderate</td>
<td>NA</td>
<td>NA</td>
<td>Low</td>
<td>Low</td>
<td>1</td>
<td>Equipment: Low to Medium Labor &amp; Supplies: Low</td>
<td>Patterns can vary with equipment used but when applied under well controlled conditions, provide a rapid production of results at very low cost</td>
<td>Sidi-Boumedine et al., 2009</td>
</tr>
</tbody>
</table>

NA: Not available.
* Diversity Index is calculated according to Hunter and Gaston, 1988.

#### Local outbreak investigation
- Isolate comparisons, determine:
  - Origin of an outbreak;
  - Inter species transmission;
  - Presence of double infection or re-infection;
  - Laboratory cross contamination.

**REA, REP-PCR, RAPD, VNTR**

#### Large scale national and international studies
- Study of collected time-space surveillance isolates, determine:
  - Presence of dominant patterns in a country or region;
  - If outbreak is propagated;
  - Chain of transmission;
  - Relation of clinical outcomes to strain types.

**ECP, AFLP, VNTR, MLST**

---

**Better comprehension of the epidemiology of the studied organism**
- Predict emergence and spread of disease.
- Implement actions to stop the spread.
- Improve public health intervention.

**Study of bacterial population genetics**
- Phylogenetic and historical studies, determine:
  - Clonal spread versus independent origin of a particular strain over disparate areas;
  - Flow of infection from one group to another;
  - Identification of pathogenic factors.

**PFGE, MLST, CGH micro-array, Whole Genome Sequencing**

---

Figure 1. Purposes and epidemiologic applications of bacterial typing, with indication of the required techniques for each (adapted from Foxman et al., 2005)
Methods

Molecular typing methods explored for the characterisation of Coxiella burnetii

Though, members of the single species belonging to the genus Coxiella exhibit a high level of genetic homogeneity; as shown by DNA-DNA hybridisation and 16S rRNA gene sequencing (Stein and Raoult, 1993; Vodkin et al., 1986). A genetic polymorphism amongst C. burnetii strains, as described above in this manuscript, has been shown using different molecular techniques with or without DNA amplification (Table 1).

Currently, five full genome and plasmid sequences are available (Table 2). In 2003, the first complete genome sequence was obtained for the reference strain: Nine Mile RSA 493 (Seshadri et al., 2003). Whole genome sequences from two representatives of human chronic endocarditis and a naturally attenuated rodent isolates were recently released (Beare et al., 2009). Genome and plasmid sizes ranged from 2.03 to 2.22 Mb and 37 to 54 kb respectively. Cross-genome comparisons revealed the presence of numerous pseudogenes, a relatively higher number of insertion sequences dispersed in the chromosome than other obligate intracellular bacteria and extensive transposon-mediated genomic plasticity within the genus Coxiella (Seshadri et al., 2003; Beare et al., 2009). Several molecular typing methods, developed for molecular epidemiology of various bacterial species, extensively reviewed elsewhere (Olive and Bean, 1999; Foxman et al., 2003; Beare et al., 2009), have been applied with, more or less, success to the discrimination of C. burnetii isolates.

Plasmid profiling

Plasmid characterization led to the identification of 4 C. burnetii plasmid types: QpH1 (36 kb)/QpDG, QpRS (39 kb), QpDV (33 kb) and a plasmid without designation derived from a Chinese isolate (Stein and Raoult, 1993; Vodkin et al., 1991). It is well known that epidemiologically unrelated strains harbour large QpRS like homologous sequences integrated to the chromosome (Jager et al., 2002). Moreover, strains without plasmids which have been applied with, more or less, success to the discrimination of C. burnetii isolates.

Methods based on Restriction Endonuclease Analysis of chromosomal DNA

Restriction endonuclease analysis (REA) of chromosomal DNA

C. burnetii isolates from geographic and/or host origins can be differentiated by REA, using different combinations of restriction endonucleases, the best resolution being obtained using BamHI digestion (Vodkin et al., 1986; Hendrix et al., 1991). Interestingly, the acute disease isolates were clustered into three similar but distinct patterns forming genomic groups I, II, and III, whereas chronic disease isolates were clustered into groups IV and V. Three additional isolates of unknown pathogenicity obtained from feral rodents in Dugway (Utah, USA), exhibited a unique REA digestion pattern and were assigned to group VI (Hendrix et al., 1991). The advantages of REA are its wide applicability and universal use, whilst the disadvantages include the need for pure, good quality DNA for analysis and the difficulty of comparing the complex profiles generated, which consist of hundreds of fragments.

Pulsed Field Gel Electrophoresis analysis (PFGE)

Once optimal conditions had been reached, using rare cutting restriction enzymes, such as NotI, in combination with PFGE specific electrophoretic conditions and image analysis, C. burnetii isolates could be separated into 20 different patterns related to their geographic origin (Heinzen et al., 1990; Thiele et al., 1993; Jager et al., 1998). The main advantages of PFGE are its discriminatory power, relatively simple banding patterns and reproducibility. Furthermore, PFGE can be used for the determination of the genome size by summing up the resulting fragments (Willems et al., 1998). However, the necessity to culture C. burnetii, the use of long and laborious DNA isolation procedures and digestion of the samples mean it may take from few days up to weeks to obtain data.

Table 2. List of C. burnetii fully sequenced genomes

<table>
<thead>
<tr>
<th>Strain</th>
<th>Gen bank accession number</th>
<th>Genotype*</th>
<th>Plasmid type (accession number)</th>
<th>Host</th>
<th>Origin</th>
<th>Date of isolation</th>
<th>Clinical pattern</th>
<th>Publication</th>
</tr>
</thead>
<tbody>
<tr>
<td>C. burnetii</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RSA 493</td>
<td>NC_002971</td>
<td>I</td>
<td>QpH1 (NC_002118 and NC_004704)</td>
<td>Tick</td>
<td>Montana, USA</td>
<td>1935</td>
<td>Acute</td>
<td>Seshadri et al, 2003</td>
</tr>
<tr>
<td>RSA 331</td>
<td>NC_010117</td>
<td>II</td>
<td>QpH1 (not sequenced from this strain)</td>
<td>Human blood</td>
<td>Italy</td>
<td>1945</td>
<td>Acute</td>
<td>None</td>
</tr>
<tr>
<td>C. burnetii</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CbuK_Q154</td>
<td>NC_011528</td>
<td>IV</td>
<td>QpRS (NC_011526)</td>
<td>Human endocarditis</td>
<td>Oregon, USA</td>
<td>1976</td>
<td>Chronic</td>
<td>Beare et al., 2009</td>
</tr>
<tr>
<td>C. burnetii</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CbuG_Q212</td>
<td>NC_011527</td>
<td>V</td>
<td>Plasmidless</td>
<td>Human endocarditis</td>
<td>Nova Scotia, Canada</td>
<td>1982</td>
<td>Chronic</td>
<td>Beare et al., 2009</td>
</tr>
<tr>
<td>C. burnetii</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dugway 7E-12</td>
<td>NC_009727</td>
<td>VI</td>
<td>QpDG (NC_009726)</td>
<td>Rodent</td>
<td>Dugway, Utah, USA</td>
<td>1957</td>
<td>Attenuated</td>
<td>Beare et al., 2009</td>
</tr>
</tbody>
</table>

* As defined by restriction enzyme banding patterns (Hendrix et al., 1991).
Methods

Amplification-based methods
Repetitive-element based PCR (Rep-PCR) and Random Amplified Polymorphic DNA (RAPD) fingerprinting methods

Even though very few C. burnetii strains were tested, Rep-PCR has been successfully used. Different C. burnetii genetic variants were found between strains isolated from aborted goats, the environment (hay) in Sweden and from reference strains (Sjöstedt et al., 1998; Rustscheff et al., 2000). Recently, RAPD has been used as a rapid and cheap mean to assess the genetic variability of isolates from ruminants (Sidi-Boumedine et al., 2009a). These methods suffer essentially from weak reproducibility, and so from the lack of portability as well as difficulties in interlaboratory interpretations.

Infrequent Restriction Site-PCR (IRS-PCR)

Recently, this method has been applied to the typing of 13 C. burnetii isolates from ruminants and of the Nine Mile reference strain (Arriacu-Bouvery et al., 2006). Seven distinct IRS-PCR patterns were identified. A good performance was observed, although MLVA could discriminate up to 11 clusters for the same batch of isolates. To our knowledge, IRS-PCR could be improved by selecting a combination of restriction enzymes and adapters suitable for the typing of C. burnetii. Indeed, the combination of enzymes which was previously used for IRS-PCR of C. burnetii was originally designed for the discrimination of Brucella species (Cloeckaert et al., 2003). This combination was likely not the best appropriate for C. burnetii typing.

PCR-based IS1111 fingerprinting

A high degree of variability in the number of IS1111 elements, between different C. burnetii isolates, has been recently described (Klee et al., 2006). However, this feature has only recently been used for genotyping, when an IS1111 PCR-based typing method was described (Denison et al., 2007). This PCR test, which uses four primer pairs and a two-step PCR allowed the differentiation of C. burnetii isolates into 5 genomic groups. The later were correlated to the ones previously described by using REA analysis (Hendrix et al., 1991).

Variable Number Tandem Repeats-PCR (VNTR-PCR)

Two multiple locus VNTR analysis (MLVA) systems for the characterisation of C. burnetii, have been described (Svraka et al., 2006; Arricau-Bouvery et al., 2006). The MLVA scheme proposed by Svraka et al. is composed of 7 VNTR loci and is included within the scheme from Arricau-Bouvery et al. Using 17 VNTR loci and 42 isolates of C. burnetii, MLVA led to the identification of 36 different genotypes. Thus, these data give evidence that MLVA is a suitable tool for epidemiological studies for C. burnetii (Arricau-Bouvery et al., 2006). Interestingly, electronic resources over the internet are freely available to facilitate the setting-up of new MLVA assays (http://bacterial-genotyping.igmors.u-psud.fr) or to query existing data (http://minisatellites.u-psud.fr).

So far MLVA data for C. burnetii derive from a small number of strains, especially those that have served for the elaboration of the method. Currently, many laboratories are willing to implement MLVA method in order to identify the type of C. burnetii strains, prevailing in a given country, in a normal context as well as in the case of declared outbreaks (Chmielewski et al., 2009; Klaassen et al., 2009; Roest et al., 2011). However, a number of inconsistencies were confirmed in routine practice, as shown by the results obtained from an interlaboratory comparison with 7 European participants (Sidi-Boumedine et al., 2009b). A first revision and guidelines for C. burnetii MLVA genotyping will, therefore, be soon made available (to be published elsewhere). Briefly, significant improvements were described with regards to discrepancies due to clerical errors, to different definitions or nomenclature used, to overlapping marker sequences and to the choice of the repeated motifs.

Real time PCR SNP typing

A novel real time PCR SNP-genotyping test was recently designed (Huijsmans et al., 2011). A panel of 10 SNPs, 7 located in single copy genes and 3 located in the multicopy IS1111, allowed the distinction of 9 SNP-genotypes among 28 strains of C. burnetii. A good performance of SNP typing was observed, although 14 MLVA types were previously described for the same batch of isolates.

The analysis of Dutch outbreak samples revealed the presence of 3 and 4 distinct genotypes for 14 human patients and various veterinary samples tested, respectively. Two genotypes identified in human samples were also present in 9 farms from the outbreak area. Thus, this method was helpful to provide epidemiological linkage, showing that multiple C. burnetii strains have infected humans in the Dutch Q-fever outbreak. Moreover, this method is advantageous as it is rapid and suitable for direct typing of C. burnetii from clinical veterinary samples. Further research will have to resolve international standardization and comparison of genotyping data from different users.
Thrusa multispacer sequence typing (MST) method has been described for C. burnetii (Glazunova et al., 2005). MST is a variant of MLST, based on sequence investigation of parts of the genome located between 2 open reading frames (ORFs). Given that sequences of spacers are considered potentially more variable than those of coding gene since they are subject to lower selection pressure, MST focuses on the sequences of such noncoding zones rather than of housekeeping genes. In this work, the authors screened 173 isolates of C. burnetii and 10 variable spacers were sequenced allowing the characterisation of 34 different sequence types (STs). An up to date website hosting MLST schemes, databases, and analysis software have been developed for a range of microorganisms (http://pubmlst.org/). In a similar way, a database for C. burnetii MST exists, where an ST type assignment can be performed through comparison of a query sequence against wellknown, validated MST sequences, on a dedicated website (http://ifr48.timone.univ-mrs.fr). The inferred phylogenetic analysis led to the characterisation of 3 monophyletic groups, which could be subdivided into different clusters. Worth mentioning, results are concordant between MST, com1 and djlA sequence comparisons and plasmid type. A link between plasmid type, some STs and disease type was also observed. A geographic distribution of isolates was observed, but some types were not delineated on the basis of geographic origin. This is thought to be due to movements of infected patients, animals, or ticks, complicating an establishment of the geographical relationship.

Hybridisation-based methods
With the availability of complete genome sequences of a large variety of microbial species, whole genome analyses become possible by microarray or DNA chip technology screening between different strains of the same species. An array consists of a collection of DNA probes, derived from the available genome sequence and spotted on a solid support, then labelled nucleic acids are hybridised against them. The presence or absence of DNA regions of the sequenced genome in query isolates is analysed, in one single hybridisation experiment. Recently, a comparative genome hybridisation (CGH) method using DNA microarrays was described for C. burnetii (Beare et al., 2006). Genomic DNA of 24 isolates of C. burnetii of diverse geographical and environmental origins were hybridised to microarray chip containing probe sets corresponding to all ORFs of the Nine Mile phase I isolate. It was shown that the genomes of the 24 C. burnetii isolates were highly conserved, however 139 ORFs were polymorphic, with the majority being complete ORF deletions. Twenty-five polymorphisms were partial ORF deletions, point mutations, or small insertions within ORFs that also resulted in negative hybridisation signals, with a small number of these predicted to be functional. These ORF polymorphisms may contribute to the biological properties of the isolates such as their virulence potential. With regards to genotyping, CGH analysis divided the 24 isolates into 8 groups. Genomic groups I to VI showed good agreement with the grouping obtained by REA analysis. In addition the inferred phylogenetic trees suggested a divergent evolution whereby genomic group I is the ancestor of genomic group III, and genomic group III is the ancestor of genomic group II. Four isolates previously ungrouped by REA analysis formed two new genomic groups (VII and VIII, respectively). Finally, genomic groups IV, V, VI, VII, and VIII constitute a distinct clade from genomic groups I, II, III.

Methods

Sequence-based methods

Single-locus sequence typing (SLST)
Several targets for locus specific analysis have been evaluated for their variability and their suitability for C. burnetii typing. Sequence comparison and/or PCR-restriction fragment length polymorphism (PCR-RFLP) analysis of com1 (encoding an outer membrane-associated immunoreactive protein) and mucZ (known to confer a mucoid property to bacteria), which was renamed djlA when the whole genome of C. burnetii was sequenced (Seshadri et al., 2003), led to the separation of 37 strains of C. burnetii in 5 different mucZ/com1 groups (Sekeyova et al., 1999). This study showed parallels with established division of C. burnetii by PFGE. However, the strains were not delineated on the basis of geographic origin, which could be due to the fact that they were essentially from Europe and mostly from France. In another study, based on com1 sequencing, 21 strains were separated into 4 different groups, and all isolates from Japan belonged to the same group (Zhang et al., 1997).

In addition, sequencing of the isocitrate dehydrogenase (icd) gene could differentiate between isolates with various geographical origins and phenotypic properties (Nguyen and Hirai, 1999). Nineteen isolates could be divided into 3 groups: isolates from acute cases of Q fever, ticks and cows were in group 1, whereas isolates from chronic Q fever patients and a prototype strain from an aborted goat were in group 2 and 3, respectively. In the same study, PCR-RFLP analysis was developed for rapid differentiation of C. burnetii isolates. In a larger study, including 72 isolates from various sources in Japan (human, cattle, cats, dogs and ticks), sequence determination and PCR-RFLP of both icd and com1 genes suggested the presence of C. burnetii isolates specific to Japan (Andoh et al., 2004).

Multilocus sequence typing (MLST) and variants
MLST is based upon the sequences of short internal fragments of multiple housekeeping genes in order to define sequence types corresponding to alleles at each locus (Maiden et al., 1998). MLST targets coding regions and uses variations that accumulate very slowly in the population. That is why MLST is used for long-term epidemiology and for the identification of lineages that have an increased propensity to cause disease. A low variability was found in C. burnetii housekeeping genes.
deciphering the genomic plasticity of *C. burnetii* isolates and their role in the complex epidemiology of Q fever. These complementary knowledges will enhance the assessment of the epidemiological information values for the genetic markers and methods.

**Acknowledgement**

This work was supported by the Network of Excellence Med-Vet-Net within the WorkPackage25 on Q fever (www.medvetnet.org). A part of work aimed to devise and evaluate *C. burnetii* molecular typing tools was presented at the 5th Annual Scientific Conference, Med-Vet-Net Association, Madrid, Spain, 3–6 June 2009 (Sidi-Boumedine et al., 2009b).

**Reference List**


G., 2006. Molecular characterization of *C. burnetii* isolates by infrequent restriction site-PCR and MLVA typing. BMC. Microbiol. 6, 38.

G., 2006. Molecular characterization of *C. burnetii* isolates by infrequent restriction site-PCR and MLVA typing. BMC. Microbiol. 6, 38.


