Multiple Acquisitions of Pathogen-Derived Francisella Endosymbionts in Soft Ticks

Jonathan G. Gerhart  
*Portland State University*

H. Auguste Dutcher  
*Portland State University*

Amanda E. Brenner  
*Portland State University*

Abraham S. Moses  
*Portland State University*, amoses@pdx.edu

Libor Grubhoffer  
*University of South Bohemia*

*See next page for additional authors*

Let us know how access to this document benefits you.

Follow this and additional works at: [https://pdxscholar.library.pdx.edu/bio_fac](https://pdxscholar.library.pdx.edu/bio_fac)

Part of the [Biology Commons](https://pdxscholar.library.pdx.edu/bio_fac), and the [Parasitology Commons](https://pdxscholar.library.pdx.edu/bio_fac)

Citation Details


This Article is brought to you for free and open access. It has been accepted for inclusion in Biology Faculty Publications and Presentations by an authorized administrator of PDXScholar. For more information, please contact pdxscholar@pdx.edu.
Authors
Multiple Acquisitions of Pathogen-Derived Francisella Endosymbionts in Soft Ticks

Jonathan G. Gerhart1,3, H. Auguste Dutcher1, Amanda E. Brenner1, Abraham S. Moses1, Libor Grubhoffer2, and Rahul Raghavan1,*

1Biology Department and Center for Life in Extreme Environments, Portland State University
2Biology Centre CAS, Institute of Parasitology, Faculty of Science, University of South Bohemia, Ceske Budejovice, Czech Republic
3Present address: Division of Parasitic Diseases and Malaria, Entomology Branch, Center for Global Health, Centers for Disease Control and Prevention, Atlanta, GA

*Corresponding author: E-mail: rahul.raghavan@pdx.edu.

Accepted: January 22, 2018

Data deposition: This project as been deposit at NCBI under the accession LVCE00000000.1.

Abstract

Bacterial endosymbionts of ticks are of interest due to their close evolutionary relationships with tick-vectored pathogens. For instance, whereas many ticks contain Francisella-like endosymbionts (FLEs), others transmit the mammalian pathogen Francisella tularensis. We recently sequenced the genome of an FLE present in the hard tick Amblyomma maculatum (FLE-Am) and showed that it likely evolved from a pathogenic ancestor. In order to expand our understanding of FLEs, in the current study we sequenced the genome of an FLE in the soft tick Ornithodoros moubata and compared it to the genomes of FLE-Am, Francisella persica—an FLE in the soft tick Argus (Persicargas) arboreus, Francisella sp. MA067296—a clinical isolate responsible for an opportunistic human infection, and F. tularensis, the established human pathogen. We determined that FLEs and MA067296 belonged to a sister taxon of mammalian pathogens, and contained inactivated versions of virulence genes present in F. tularensis, indicating that the most recent common ancestor shared by FLEs and F. tularensis was a potential mammalian pathogen. Our analyses also revealed that the two soft ticks (O. moubata and A. arboreus) probably acquired their FLEs separately, suggesting that the virulence attenuation observed in FLEs are not the consequence of a single acquisition event followed by speciation, but probably due to independent transitions of pathogenic francisellae into nonpathogenic FLEs within separate tick lineages. Additionally, we show that FLEs encode intact pathways for the production of several B vitamins and cofactors, denoting that they could function as nutrient-provisioning endosymbionts in ticks.

Key words: tick, endosymbiont, Coxiella, Francisella, Coxiella-like, Francisella-like.

Introduction

Ticks (order Ixodida) are ectoparasites of reptiles, birds, and mammals. Recent molecular clock estimates indicate that they originated in the Carboniferous period (~320 million years ago, Ma) and diverged in the early Permian (~300 Ma) into two major families: 1) Ixodidae, or hard ticks that possess a hardened chitinous scutum, and 2) Argasidae, or soft ticks that lack scutum (Jeyaprakash and Hoy 2009; Mans et al. 2016). Ticks depend on an imbalanced diet consisting entirely of vertebrate blood. Similar to endosymbiotic bacteria in insects with specialized diets—for example, Buchnera aphidicola in aphids that feed exclusively on phloem (Baumann 2005; Bennett and Moran 2015)—endosymbiotic bacteria present within tick cells are thought to compensate for the dearth of B vitamins and cofactors in blood (Smith et al. 2015; Gottlieb et al. 2015; Gerhart et al. 2016). However, unlike most primary insect endosymbionts, tick endosymbionts are not maintained within specialized cells called bacteriocytes, but are instead found in several tissues, including salivary glands (Klyachko et al. 2007; Budachetri et al. 2014), which could explain their propensity to be horizontally transmitted between hosts (Duron et al. 2017).
Ticks carry endosymbionts mostly from the genera Coxiella, *Rickettsia*, or *Francisella*. In addition to these putative primary endosymbionts (i.e. those necessary for host survival), secondary symbionts that are not strictly necessary, including *Midichloria*, *Arsenophonus*, *Rickettsiella*, and *Wolbachia* species, are found in varying rates in hard and soft tick populations (Ahantarig et al. 2013; Duron et al. 2017). Understanding the biology and evolution of tick endosymbionts is of special interest because the three primary endosymbionts share close evolutionary relationships with tick-borne pathogens *Coxiella burnetii*, *Rickettsia parkeri*, and *Francisella tularensis*, respectively, that cause human and animal diseases (Petersen et al. 2009; Ahantarig et al. 2013; Duron et al. 2015). Previous studies indicate that pathogen-symbiont transitions could occur in either direction: The mammalian pathogen *C. burnetii* probably originated from a tick-associated nonpathogenic ancestor (Smith et al. 2015; Duron et al. 2015; Moses et al. 2017), whereas the FLE in the hard tick *Amblyomma maculatum* (FLE-Am) likely evolved from a mammalian pathogen (Gerhart et al. 2016).

In this study, in order to better understand the evolution and functions of FLEs in ticks, we sequenced the genome of an FLE present in the soft tick *Ornithodoros moubata* (FLE-Om). While this project was in progress, the genome of *O. moubata* was first described more than four years ago using microscopy (Reinhardt et al. 1972). Recently, 16S rDNA PCR was used to confirm that the vertically transmitted bacterium was closely related to the human pathogen *F. tularensis* (Noda et al. 1997). Using Illumina high-throughput sequencing, we analyzed DNA isolated from an adult female *O. moubata* obtained from a laboratory colony (Rego et al. 2005). Out of the ~180 million sequencing reads, the vast majority were of host origin, with <5% of reads contributing to the assembly of the eight FLE-Om contigs (supplementary table S1, Supplementary Material online). In addition, low quantities of reads that correspond to *Clostridium* spp. and *Burkholderia* spp. were also present (<0.5% of reads). No other bacterium was detected at significant levels, including the previously described *Coxiella-like “Symbiont A”* (Noda et al. 1997). By screening several male and female *O. moubata* adults using both bacterial and *Coxiella*-specific 16S rDNA PCR primers, we confirmed that FLE-Om but not the *Coxiella-like “Symbiont A”* has been retained in the *O. moubata* laboratory colony. As shown previously in insects, this observation suggests that whereas FLE-Om is likely a primary endosymbiont that provides a critical function to *O. moubata*, the *Coxiella-like “Symbiont A”* is a secondary symbiont or a transient bacterium (McCutcheon and Moran 2010; Hall et al. 2016).

The ~1.56 Mb genome of FLE-Om is similar in size to that of other FLEs (FLE-Am and *F. persica*), but considerably smaller than the ~1.90 Mb genome of the human pathogen *F. tularensis* (table 1). Concordantly, the G + C content and number of protein-coding genes are also lower in the three FLEs in comparison to *F. tularensis*, signifying a more intimate host association in FLEs, which leads to small, A + T-biased genomes with fewer functional genes (McCutcheon and Moran 2010). Intriguingly, while FLE-Om and FLE-Am have similar GC%, coding density, average gene length, pseudogene content, rRNA gene content, etc., these characteristics are very distinct in *F. persica*, indicating that its evolutionary history is different from that of the other two FLEs.

**Francisella persica** Has Higher Virulence Potential than Other FLEs

All three FLEs encode several genes shown to be critical to the virulence of *F. tularensis* (Rowe and Huntley 2015; Melbom and Charbit 2010). However, in FLE-Am and FLE-Om, as expected in nutrient-provisioning endosymbionts, most virulence genes have been inactivated or appear to have been deleted (table 2, supplementary table S2, Supplementary Material online). *Francisella persica*, however, contains a considerably higher number of potentially functional virulence-associated genes, suggesting that this endosymbiont could be pathogenic to mammals. In fact, *F. persica* (previously misidentified as *Wolbachia persica*) was found to cause fever and death in guinea pigs (Suiter and Weiss 1961). Similarly, MA067296, a *Francisella* strain that caused an opportunistic human infection (Kugeler et al. 2008; Challacombe et al. 2017) also encodes a larger complement of intact virulence genes than FLE-Om and FLE-Am, suggesting that the presence of full-length virulence genes corresponds to the ability of a bacterium to cause disease in...
mammals. Accordingly, while it has not yet been tested, the inactivation of most virulence genes in FLE-Om and FLE-Am is assumed to have rendered these tick endosymbionts non-pathogenic to mammals.

Mobile elements may have played a significant role in pathogen attenuation through the disruption and removal of virulence-associated genes. Past studies have demonstrated that association with hosts leads to a proliferation of mobile elements in bacterial genomes, followed by an extensive purge as the genome is minimized to its essentials with the development of an obligate host–bacterium relationship (Moran and Plague 2004; McCutcheon and Moran 2010). In accordance with this trend, the genome of *F. tularensis*, which has a less intimate relationship with its host, is replete with insertion sequences (IS), whereas the genomes of host-restricted FLEs contain very few IS elements (table 1). Type VI secretion system gene clusters in FLEs appear to have been particularly disrupted by these selfish genetic elements, as a concentration of partial and pseudogenized IS elements are found in this gene cluster. It is likely that insertion and transposition of IS elements resulted in the loss of two virulence-associated genes from *F. persica* and six genes from FLE-Om and FLE-Am (fig. 1).

**Table 1**

<table>
<thead>
<tr>
<th>Feature</th>
<th>F. tularensis</th>
<th>FLE-Am</th>
<th>F. persica</th>
<th>FLE-Om</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length (bp)</td>
<td>1,892,772</td>
<td>1,556,255</td>
<td>1,540,154</td>
<td>1,564,197</td>
</tr>
<tr>
<td>G+C%</td>
<td>32.30</td>
<td>31.80</td>
<td>31.39</td>
<td>31.80</td>
</tr>
<tr>
<td>No. of coding genes</td>
<td>1,556</td>
<td>1,001</td>
<td>1,096</td>
<td>989</td>
</tr>
<tr>
<td>Coding density (%)</td>
<td>92</td>
<td>57</td>
<td>68</td>
<td>56</td>
</tr>
<tr>
<td>Average gene length (bp)</td>
<td>937</td>
<td>886</td>
<td>964</td>
<td>881</td>
</tr>
<tr>
<td>No. of pseudogenes</td>
<td>227</td>
<td>484</td>
<td>205</td>
<td>543</td>
</tr>
<tr>
<td>No. of single copy genes</td>
<td>106/111</td>
<td>106/111</td>
<td>106/111</td>
<td>106/111</td>
</tr>
<tr>
<td>16S rDNA</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>23S rDNA</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>5S rDNA</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>tRNAs</td>
<td>38</td>
<td>32</td>
<td>38</td>
<td>34</td>
</tr>
<tr>
<td>IS elementsb</td>
<td>72</td>
<td>3</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Accession #</td>
<td>NC_006570</td>
<td>LNCT00000000</td>
<td>NZ_CP013022</td>
<td>LVCE0000000</td>
</tr>
</tbody>
</table>

aAlbertsen et al. (2013).
bBoth coding and pseudogenized versions of IS elements have been included.

**Table 2**

<table>
<thead>
<tr>
<th>Genes</th>
<th>F. tularensis</th>
<th>MA067296</th>
<th>F. persica</th>
<th>FLE-Am</th>
<th>FLE-Om</th>
</tr>
</thead>
<tbody>
<tr>
<td>LPS and capsule</td>
<td>22</td>
<td>0</td>
<td>16</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Antigen synthesis</td>
<td>17</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>Type IV pili</td>
<td>8</td>
<td>0</td>
<td>8</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>Outer membrane</td>
<td>10</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>Inner membrane</td>
<td>17</td>
<td>0</td>
<td>16</td>
<td>1</td>
<td>16</td>
</tr>
</tbody>
</table>

aNumber of full-length genes.
bNumber of pseudogenes.

FLEs May Provide B Vitamins and Cofactors to Ticks

Vertebrate blood lacks several B vitamins and cofactors; therefore, vertically inherited intracellular bacteria that are present across tick lineages are thought to provide these essential nutrients to the host. We examined the metabolic pathways present in FLEs and found that they encode complete pathways for the synthesis of biotin (B7), folate (B9), lipoic acid, riboflavin (B2), and FAD (fig. 2). All lack phoA gene in the folate synthesis pathway and a phosphoric monoester hydrolase (EC 3.1.3.-) in the riboflavin pathway. However, previous studies have shown that phoA might not be required to synthesize folate (Bermingham and Derrick 2002), and these two genes have not been retained in most mutualists, suggesting that their functions could be compensated by other genes (Klein et al. 2013; Smith et al. 2015; Gottlieb et al. 2015; Manzano-Marín et al. 2015; Boyd et al. 2017). The three FLEs displayed somewhat different capabilities to synthesize...
Coxiella not a nutrient-providing endosymbiont. In contrast, the two biotin and folate, indicating that REIS is likely a parasite and have the ability to recycle nitrogen by incorporating ammonia, a metabolic waste product, into the synthesis of glutamine, as shown in several insect heme biosynthesis pathway, and have the ability to recycle cofactor production, FLEs—but not CLEs—have retained the cofactor-provisioning endosymbionts in ticks. In addition to those missing in FLEs, suggestive of their roles as primary cofactor-provisioning endosymbionts in ticks. In addition to cofactor production, FLEs—but not CLEs—have retained the heme biosynthesis pathway, and have the ability to recycle nitrogen by incorporating ammonia, a metabolic waste product, into the synthesis of glutamine, as shown in several insect endosymbionts (e.g. Zientz et al. 2006; López-Sánchez et al. 2009; Sabree et al. 2009; Hansen and Moran 2011). In sum, FLEs have the potential to supply several vitamins and cofactors that are present in very low concentrations in their hosts’ blood meals, thereby improving tick fitness and providing the impetus for ticks to maintain FLEs through generations via vertical transmission. Furthermore, unlike most other arthropod endosymbionts, F. persica has been cultured in laboratory media (Hinrichs et al. 2016), and it should be feasible to develop growth media for FLE-Am and FLE-Om using our genomic data, as previously done for other fastidious bacteria (Renesto et al. 2003; Omsland et al. 2009).

FLEs Evolved from a Pathogenic Ancestor and Soft Ticks Have Acquired FLEs Multiple Times

Our phylogenetic estimation using 404 genes in 49 fully sequenced Francisella genomes revealed that FLEs and the opportunistic human pathogen Francisella sp. MA067296 are closely related to each other and belong to a sister taxon of mammalian pathogens (fig. 3). This evolutionary relationship, combined with the presence of homologous virulence genes in both the FLE and mammalian-pathogen branches of the Francisella tree, indicate that the common ancestor of FLEs and F. tularensis was armed with potentially pathogenic characteristics (table 2, supplementary table S2, Supplementary Material online). The ability of MA067296 and F. persica to cause morbidity in humans and guinea pigs, respectively, further supports the conclusion that FLEs likely evolved from a pathogenic ancestor (Suitor and Weiss 1961; Kugeler et al. 2008; Challacombe et al. 2017). However, it should be noted that we cannot completely rule out the possibility that each Francisella acquired its complement of virulence genes via horizontal gene transfer or that some of the apparently pseudogenized genes in FLEs are actually functional.

Although soft ticks O. moubata and A. arboreus host FLE-Om and F. persica, respectively, FLE-Om is more closely related to FLE-Am in the hard tick A. maculatum than to F. persica (fig. 4). This incongruence between tick and FLE phylogenies is not entirely surprising because host-endosymbiont
coevolution is rare in ticks, possibly due to horizontal transfer of symbionts between unrelated tick species (Duron et al. 2015). While a few CLE and FLE strains have coevolved with their tick hosts (Duron et al. 2017; Azagi et al. 2017), a comprehensive survey of 81 species of hard and soft ticks showed that FLEs in several soft ticks were more closely related to FLEs in hard ticks than to each other, suggesting that FLEs have been exchanged between hard and soft ticks (Duron et al. 2017). On the basis of these observations, the two FLEs in soft ticks were probably acquired from independent sources: FLE-Om from a hard tick, and F. persica from an as yet unknown (possibly environmental) source. Horizontal transfer of
endosymbionts is more prevalent among ticks than among insects possibly because tick endosymbionts are not restricted to specialized host cells, but instead are usually present in several tissues including in the salivary glands (Klyachko et al. 2007; Budachetri et al. 2014). Since tick-borne pathogens are typically found in salivary glands and are spread via tick saliva while the arthropod is feeding on hosts, the presence of FLEs in salivary glands is likely a holdover from their
previous pathogenic lifestyle, and could facilitate their horizontal transfer to other ticks or even to other blood-feeding arthropods such as keds, while cofeeding on the same vertebrate host (Wright et al. 2015; Lee et al. 2016).

In conclusion, we show that while all three FLEs likely originated from pathogenic francisellae and could function as nutrient-provisioning endosymbionts, the FLEs in two soft ticks have dissimilar genome characteristics, virulence potentials, and evolutionary histories, indicative of independent acquisition events. The soft ticks either gained preexisting FLEs from other hosts, or acquired pathogenic *Francisella* strains with varying degrees of virulence loss within their respective hosts. Similar to FLEs, previous studies suggest that *Sodalis*-allied endosymbionts of insects probably evolved from a pathogenic progenitor, indicating that arthropods could gain new symbionts by domesticating potentially pathogenic bacteria (Clayton et al. 2012). Because of the potential to utilize insights gained from studying FLEs to control tularemia and other tick-borne diseases, it is critical that further studies are conducted to 1) delineate the evolutionary histories of all FLEs in hard and soft ticks, 2) define the evolutionary relationships between FLEs and environmental *Francisella* (Barns et al. 2005; Keim et al. 2007), 3) understand the functional consequences of the genetic differences observed between FLEs, and 4) validate the nutrients potentially provided by FLEs to ticks.

### Materials and Methods

#### Sequencing, Genome Assembly, and Bacterial Identification

A female *O. moubata* from a laboratory colony maintained at the Institute of Parasitology, Czech Republic was used to sequence the FLE-Om genome. The tick’s outer surface was sterilized using 70% ethanol, and midgut tissue was extracted and subjected to Protease K digestion, followed by DNA extraction using a DNeasy kit and protocol (Qiagen). Purified DNA was sequenced in a single lane of Illumina Hi-Seq 2500 (100 bp, paired-end) at Oregon Health and Science University’s Massively Parallel Sequencing Shared Resource, yielding ~180 million read pairs. Low confidence reads were removed using Trimmomatic (Bolger et al. 2014) and assembled into contigs with IDBA (Peng et al. 2012), using a k-mer range of 49:121, a step size of two, and a minimum contig length of 1 kb. All contigs were searched using BlastN against a library of *Francisella* genome sequences obtained from NCBI to identify putative FLE contigs. All trimmed reads were mapped to these contigs using bowtie2 and Samtools (Langmead and Salzberg 2012; Li et al. 2009). Reads that mapped to the preliminary FLE contigs along with 20% of original reads were reassembled into a final set of eight FLE-Om contigs, which have been submitted to NCBI (accession: LVCE00000000.1). The completeness of the assembled genome was examined using a single-copy gene database (Albertsen et al. 2013), yielding identical results for FLE-Om, FLE-Am, *F. persica*, and *F. tularensis* (106/111 genes). The identities of bacteria present in the tick microbiome were determined by binning BlastN results using MEGAN5’s taxonomy browser (Huson et al. 2007). To verify the absence of the *Coxiella*-like “Symbiont A” (Noda et al. 1997), DNA was extracted from five adult male and female ticks and PCR was performed using both bacterial and *Coxiella*-specific 16S rDNA primers (Klyachko et al. 2007; Klindworth et al. 2013).

#### Genome Annotation and Assessment of Metabolic and Virulence Potential

FLE-Om contigs were annotated using NCBI’s prokaryotic genome annotation pipeline. Cofactor synthesis pathways in FLEs, CLEs, and REIS were identified with BlastKOALA (Kanehisa et al. 2016), which maps protein sequences to the Kyoto Encyclopedia of Genes and Genomes (Manyam et al. 2016). In addition, biosynthesis genes present in

---

**Fig. 4** — FLEs and hosts have incongruent evolutionary histories. FLE phylogeny (left; same as in fig. 3) does not correspond to the host phylogeny (right; based on an 18S rDNA sequences). Branch lengths are not representative of evolutionary distances. Soft ticks and their FLEs are highlighted in green.
F. tularensis were searched against FLE genomes using tBlastN and verified with a reciprocal BlastP of the resultant top hits against F. tularensis coding sequences. Pathways in CLEs and REIS were verified using previously published studies (Gillespie et al. 2012; Smith et al. 2015; Gottlieb et al. 2015). A list of genes critical to the pathogenicity of F. tularensis was used to identify functional and pseudogenized versions of virulence genes present in FLEs (Rowe and Huntley 2015; Meibom and Charbit 2010). Presence of IS elements was determined using ISfinder (Siguier et al. 2006).

**Phylogenetic Analyses**

In addition to FLE-Om, FLE-Am and F. persica, we included 46 fully sequenced Francisella genomes to generate a robust phylogenetic tree (supplementary Data set S1, Supplementary Material online). Using reciprocal BlastP, a subset of 404 fully sequenced genomes to generate a robust phylogenetic tree (supplementary Data set S1, Supplementary Material online). Using reciprocal BlastP, a subset of 404 orthologous genes were identified. Concatenated sequences were aligned using Clustal Omega (Sievers et al. 2014) and trimmed using Gblocks (Talavera and Castresana 2007). Tick phylogeny was built using 18 S rDNA sequences (NCBI accessions: L76344.1, L76353.1, KJ133645.1, KY016614.1). jModelTest2 was used to select the GTR + I + G evolution model for the FLE tree and HKY85 + I model for the tick tree (Darriba et al. 2012). Maximum Likelihood trees with 1,000 bootstrap replicates were constructed using RAxML (Stamatakis 2014). Bayesian trees with a chain length of 500,000 and a burn-in fraction of 25% and sampling every 100 trees were constructed using MrBayes (Huelsenbeck and Ronquist 2001). The Gammaproteobacteria cladogram (supplementary fig. S1, Supplementary Material online) is based on a previously published phylogenetic tree (Williams et al. 2010).

**Supplementary Material**

Supplementary data are available at Genome Biology and Evolution online.

**Acknowledgments**

We thank Colleen Campbell, Jess Millar, Todd Smith, and Jim Archuleta for technical assistance and helpful discussions. This work was supported in part by National Institutes of Health grant AI126385 to R.R.

**Literature Cited**


Klyachko O, Stein BD, Grindle N, Clay K, Fuqua C. 2007. Localization and
Klein CC, et al. 2013. Biosynthesis of vitamins and cofactors in bacterium-
Noda H, Munderloh UG, Kurtti TJ. 1997. Endosymbionts of ticks and their
Keim P, Johansson A, Wagner DM. 2007. Molecular epidemiology,
Klein CC, et al. 2013. Biosynthesis of vitamins and cofactors in bacterium-
Noda H, Munderloh UG, Kurtti TJ. 1997. Endosymbionts of ticks and their
Keim P, Johansson A, Wagner DM. 2007. Molecular epidemiology,
Klein CC, et al. 2013. Biosynthesis of vitamins and cofactors in bacterium-
Noda H, Munderloh UG, Kurtti TJ. 1997. Endosymbionts of ticks and their
Keim P, Johansson A, Wagner DM. 2007. Molecular epidemiology,