Acta Tropica 197 (2019) 105041

Contents lists available at ScienceDirect

Acta Tropica

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

Mycobacterium leprae's evolution and environmental adaptation

Bibiana Chavarro-Portillo[a,b,c], Carlos Yesid Soto[b], Martha Inírida Guerrero[a,*]

[a] Hospital Universitario Centro Dermatológico Federico Lleras Acosta, Bogotá, Colombia
[b] Departamento de Química, Facultad de Ciencias, Universidad Nacional de Colombia, Bogotá, Colombia
[c] Doctorado en Biotecnología - Instituto de Biotecnología, Facultad de ciencias, Universidad Nacional de Colombia, Bogotá, Colombia

ARTICLE INFO

Keywords:
Mycobacterium leprae
Leprosy
Genome
Reductive evolution
Pseudogene
Intracellular pathogen

ABSTRACT

Leprosy is an ancient disease caused by the acid-fast bacillus *Mycobacterium leprae*, also known as Hansen's bacillus. *M. leprae* is an obligate intracellular microorganism with a marked Schwann cell tropism and is the only human pathogen capable of invading the superficial peripheral nerves. The transmission mechanism of *M. leprae* is not fully understood; however, the nasal mucosa is accepted as main route of *M. leprae* entry to the human host. The complete sequencing and the comparative genome analysis show that *M. leprae* underwent a genome reductive evolution process, as result of lifestyle change and adaptation to different environments; some of lost genes are homologous to those of host cells. Thus, *M. leprae* reduced its genome size to 3.3 Mbp, contributing to obtain the lowest GC content (approximately 58%) among mycobacteria. The *M. leprae* genome contains 1614 open reading frames coding for functional proteins, and 1310 pseudogenes corresponding to 41% of the genome, approximately. Comparative analyses to different microorganisms showed that *M. leprae* possesses the highest content of pseudogenes among pathogenic and non-pathogenic bacteria and archaea. The pathogen adaptation into host cells, as the Schwann cells, brought about the reduction of the genome and induced multiple gene inactivation. The present review highlights the characteristics of genome's reductive evolution that *M. leprae* experiences in the genetic aspects compared with other pathogens. The possible mechanisms of pseudogenes formation are discussed.

1. Introduction

Mycobacterium leprae, the causative agent of leprosy or Hansen's disease, is an intracellular pathogen with a pronounced tropism for both myelinated and nonmyelinated Schwann cells (SC) (Scollard et al., 2006). Inside host cells, *M. leprae* induces the cellular proliferation to ensure a niche for long-term survival (Rambukkana, 2010). *M. leprae* is a straight or slightly curved bacillus with rounded ends, ranging from 1 to 8 μm in length and 0.2 to 0.5 μm in diameter. When viewed by optical microscopy, *M. leprae* is observed as individual cells or as clumped bacilli known as *globias* (Scollard et al., 2006). *M. leprae* is non-motile, cannot form spores, divides by binary fission, and is a very slow growing mycobacteria; the generation time of *M. leprae* ranges from 12 to 14 days, approximately, the longest ever reported among bacteria (Scollard et al., 2006). Thus, the success of *M. lepraae* infection in the SC is pivotal for its survival (Rambukkana, 2000).

The SC lack of antimicrobial mechanisms to defeat an infection, therefore they can tolerate the bacterial load produced by *M. leprae* and allow releasing bacilli to the environment.

M. leprae is capable of invading SC through binding PGL-1 of the cell surface to the G domain of Laminin α2 chain (LNα2 G) on the basal lamina that surrounds SC (Ng et al., 2000; Rambukkana, 2004). As the infection progress, the bacilli replicate without restriction. Therefore, the availability of non-myelinating SC represents a limitation for the bacterial growth; to overcome this limitation, *M. leprae* activates quiescent non-myelinating SC inducing them to re-enter the cell cycle (Rambukkana, 2010). *M. leprae* also invade macrophages (primary host cells), where the bacilli replicate in the absence of an effective adaptive immune response (Misch et al., 2010). *M. leprae* is the only known bacterial human pathogen that invades superficial peripheral nerves. The peripheral nerves protects *M. leprae* from the immune response due to the blood-brain barrier (Rambukkana, 2000). The infection of the SC and the nerve colonization by *M. leprae* produce complex biological and pathological alterations such as axonal degeneration, demyelination, and fibrosis that induce nerve damage, loss of sensitivity and tissue damage, which are the distinctive clinical characteristics of leprosy (Rambukkana, 2000, 2004).

Leprosy presents a wide spectrum of clinical manifestations,

* Corresponding author at: Av. Calle 1 # 13ª-61, Hospital Universitario Centro Dermatológico Federico Lleras Acosta, Oficina de Docencia e Investigación, Bogotá, Colombia.
E-mail address: marthainiridag@yahoo.com (M.I. Guerrero).

https://doi.org/10.1016/j.actatropica.2019.105041
Received 25 September 2018; Received in revised form 28 May 2019; Accepted 28 May 2019
Available online 30 May 2019
0001-706X/

including tuberculoid pole (TT) or paucibacillary (PB) leprosy, and lepromatous pole (LL) or multibacillary (MB) leprosy, among them, there are intermediate forms (Ridley and Jopling, 1966). Clinically, patients with LL display a high bacterial load in skin biopsy samples, multiple skin lesions and anesthesia. In contrast, patients with TT show a low bacilli load in skin biopsy samples and can display a single anesthetic injury (Misch et al., 2010). The treatment of leprosy recommended by the World Health Organization (WHO) is based on the disease classification after a rigorous examination and skin smear results (http://www.who.int/lep/disease/en/). The Multidrug therapy (MDT) for PB patients includes dapsone and rifampicin for at least 6 months, or dapsone, rifampicin and clofazimine for at least 12 months for MB patients (WHO, 2013). The duration of the treatment against leprosy, as well as tuberculosis, has been a controversial topic. It is known that after long treatments for months or even years, there are still residual bacilli; in the case of *M. leprae,* the duration of treatment is a consequence of the low replication rate of the bacilli and the challenges that systemic antibiotics must overcome to reach the target site, and get into the bacterial cell (Penna et al., 2017).

It is noteworthy that *M. leprae* cannot be grown in artificial culture medium, so its culture is restricted to animal models, low growth rates of bacilli can be obtained in mice foot pads (Scollard et al., 2006), and high growth rates in immunocompromised rodents or nine-banded armadillos susceptible to leprosy (Scollard et al., 2006). For the above reasons, experimental analyses to answer fundamental questions about the genetics, metabolism, pathogenicity and antimicrobial susceptibility of *M. leprae* are difficult to perform.

Currently, whole-genome sequencing (WSG) and comparative genomic analyses have permitted the evaluation of the *M. leprae* genome size, gene content, and essential features for environmental adaptation, helping to undertand genotypical and phenotypical *M. leprae* behavior. In this context, this review aims to discuss the genetic characteristics of *M. leprae* associated with the reductive evolution process and pseudogene formation incurred by the leprosy bacilli through the course of its evolution.

2. Mechanism of leprosy transmission

Although the actual mechanisms of leprosy transmission are not well understood, the nasal mucosa is considered the main route of bacillus entry (Smith et al., 2014; Smith and Aerts, 2014). Leprosy transmission among particular populations depends on the natural susceptibility of the healthy individuals, which is modulated by cellular immunity (Scollard et al., 2006), as well as the contact frequency between healthy and infected individuals (Fischer et al., 2010; Moet et al., 2006). Therefore, leprosy patients are usually in contact with family members and cohabitants (Bratschi et al., 2015); thus, the household contact is the highest infection risk (Moet et al., 2006). In this context, there are alternate modes of leprosy transmission, which are described in the following sections.

2.1. Dasypus novemcinctus as a model of M. leprae infection and leprosy transmission

Several studies have shown that the nine-banded armadillo is a natural reservoir of *M. leprae*. Epidemiological studies, where the same disease-causing *M. leprae* strains are observed in both humans and the nine-banded armadillo, show that the Nine-banded are implicated in zoonotic transmission of leprosy (Sharma et al., 2015; Truman et al., 2011). Inflammation and loss of hair around the snout, lips, eyelids, ears, and legs is typical of leprosy in nine-banded armadillos (*Dasypus novemcinctus*). The low body temperature of the nine-banded armadillo, among others, has promoted this wildlife animal as an appropriate model to evaluate the *M. leprae* proliferation (Storrs, 1971). *D. novemcinctus* is susceptible to *M. leprae* infection not only on superficial tissues but also through dissemination to the heart, liver, spleen and lymph nodes (Kirchheimer, 1975; Kirchheimer et al., 1972). Therefore, the armadillo model provides sufficient amounts of bacilli to carry out extensive biochemical and genetic studies (Kirchheimer, 1975).

2.2. Leprosy in red-tailed squirrels

Recently, Avanzi et al., 2016 reported that the nine-banded armadillo is not the only wildlife species that is a reservoir of the leprosy bacillus (Avanzi et al., 2016). Leprosy also has been detected in red squirrels from the British Isles (Avanzi et al., 2016). Red squirrels from Brownsea Island in Dorset exhibited symptoms of the disease caused by *M. leprae*. By contrast, leprosy in red squirrels from the Isle of Wight (England), Scotland and Ireland, is caused by strains of *M. lepromatosis*, which was initialy isolated in 2008 from Mexican patients (Han et al., 2008). The complete genome sequence of *M. lepromatosis* isolated from human biopsies (Singh et al., 2015) shows that strains from Mexico and Europe diverged from a common ancestor 27,000 years ago (Avanzi et al., 2016).

2.3. Arthropods as M. leprae reservoirs

Multiple studies suggest arthropods as a source of leprosy transmission (Kirchheimer, 1976). It could explain the *M. leprae* transmission between the nine-banded armadillo and humans mediated by arthropods such as *Culex fatigans* and *Aedes aegypti* in leprosy-endemic agricultural regions of Brazil (Banerjee et al., 1991; Neumann et al., 2016). Insects bite exposed skin from infected people, acquiring viable bacilli and transmit them to healthy people (Banerjee et al., 1991). In addition, Neumann et al. reported that *M. leprae* is not viable after 20 days within the mosquito digestive tract, where the bacilli viability decreases due to there is no bacilli proliferation in the insect gut (Neumann et al., 2016). Similarly the Buruli ulcer, a dermal infection caused by *M. ulcerans,* is also transmitted by arthropods such as the Hemiptera and Diptera orders (Lavender et al., 2011).

2.4. Free-living amoebae as a reservoir and source of the leprosy bacillus

In general, mycobacteria have developed mechanisms to replicate within free-living amoebae (FLA) (Salah et al., 2009). Thus, *Mycobacterium* species have been found sharing habitats with FLA genera such as *Hartmannella, Echinamoeba, Acanthamoeba* and *Naegleria* (Kennedy et al., 2012; Lamrabet et al., 2012; Salah et al., 2009). Although *M. leprae* is not able to multiply within amoebas, FLA is capable of providing conditions for maintaining the viability of the leprosy bacillus (Wheat et al., 2014). As obligate intracellular mycobacteria, *M. leprae* can interact with FLA, which have physiological characteristics similar to human phagocytes, such as their ability to form phagocytic vacuoles. Once ingested, *M. leprae* can survive and remain viable within FLA for at least 72 h, suggesting that FLA are putative environmental reservoir of the leprosy bacillus (Lahiri and Krahenbuhl, 2008). Based on previous reports, *M. leprae* can be transmitted from FLA to humans (Wheat et al., 2014); however, further studies are needed to determine whether FLA actually facilitates the transmission of leprosy to susceptible hosts.

3. *The M. leprae* genome

3.1. The M. leprae genome size and content

Apparently, 50% of the *M. leprae* genome has no function; indeed, *M. leprae* has the largest pseudogene content among pathogenic or nonpathogenic bacteria or archaea (Liu et al., 2004). To date, four different *M. leprae* strains—TN (India), Tahi-53 (Thailand), NHDP63 (USA) and Br4923 (Brazil)—have been sequenced (Table 1) (Cole and Singh, 2012; Singh and Cole, 2012). Genomic comparisons show that *M. leprae* genomes have 99.995% of identity and their genome size

Table 1
General features of the four sequenced *Mycobacterium leprae* genomes.

Strain	SNPs InDels	SNP subtype	Additional pseudogene(s) – (function)	Sequencing Method	Source/origin
TN (India) (3,268,212 bp)	21	1A	no data	ABI	Patient from Tamil Nadu, India
Thai-53 (Thailand)	23		ML0472 – (unknown) ML2472c – (TyrA) ML2687c – (unknown)	Illumina	Patient from Thailand
NHDP63 (USA)	65	3I	ML2472c – (TyrA) ML2678c – (unknown)	Illumina	Patient from Texas (without foreign residence history)
Br4923 (Brazil) (3,268,071 bp)	65	4P	ML0825c – (ArsR regulator)	ABI	Patient from Brazil

SNPs: single nucleotide polymorphism, InDel: Insertion-deletion.
Modified from Singh and Cole, 2011.

differences are less than 0.005%, which comprises 289 polymorphic sites including single nucleotide polymorphisms (SNPs) and small insertion-deletion events (InDels) (Monot et al., 2009; Singh and Cole, 2012).

The strain TN strain, originally isolated from skin lesions of an Indian leprosy patient (Tamil Nadu, India), was the first reported complete *M. leprae* genome, in 2001 (Cole et al., 2001a). The *M. leprae* TN genome comprises 3,268,203 bp, with only 1614 ORFs (49.6%), meanwhile 1,310 are pseudogenes and noncoding regions (41%) (Cole et al., 2001a; Singh and Cole, 2012). The remaining *M. leprae* TN genome (9.4%) does not appear to encode genes; it corresponds to residual genes and unrecognizable regulatory sequences having a large amount of accumulated mutations (Eiglmeier et al., 2001). In addition, VNTRs structures (Variable Number Tandem Repetitions), which are frequently used for mycobacterial typing, are also common in the TN strain (Groathouse et al., 2004; Truman et al., 2004).

The genome sequence of the Br4923 strain contains 3,268,071 bp (141 bp less than the TN strain). The Br4934 and TN genomes show differences in 241 SNPs and 48 InDels, including 58 variations in the dispersed repeats (Monot et al., 2009; Singh and Cole, 2012). Of note, Br4923 displays a lower number of VNTR compared with the TN strain. (Monot et al., 2005, 2009; Singh and Cole, 2012)

3.2. The Mycobacterium leprae genome compared with other mycobacteria

3.2.1. M. leprae vs. M. ulcerans

M. ulcerans is the causative agent of the Buruli ulcer, an infectious disease characterized by a necrotizing infection of the subcutaneous tissue (Johnson et al., 2005). As observed in *M. leprae,* pseudogenes accumulation is relevant in *M. ulcerans,* as consequence of adaptation processes inside host cells (Cole et al., 2001a; Stinear et al., 2007). For example, the genome size of the *M. ulcerans* Agy99 strain (isolated from a Ghanan patient in 1999) is 5,631,606 bp, having a GC content of 62.5%, and containing 4160 coding genes and 771 pseudogenes (Fig. 1). The *M. ulcerans* Agy99 strain also contains the virulence plasmid pMUM001, which is 174,155 bp in size (Stinear et al., 2007). Apparently, genome differences between *M. ulcerans* and *M. leprae* are associated to their different lifestyles, as known *M. ulcerans* is not an obligate intracellular pathogen (Silva et al., 2009).

3.2.2. M. leprae vs. M. avium subsp. paratuberculosis

Mycobacterium avium subsp. *paratuberculosis* (MAP) is a mycobacterial pathogen causing chronic intestinal inflammation in humans that is known as the Johne's disease (Uzoigwe et al., 2007). MAP has the highest GC content (69.3%) within the *Mycobacterium* genus. The MAP genome is 4,829,781 bp in size (Li et al., 2005) displaying the highest percentage (91.5%) of genome encoding for proteins among the mycobacterial species (Fig. 1), such as *M. tuberculosis* (90.9%), and *M. leprae* (49.5%) (Cole et al., 1998, 2001a; Marri et al., 2006).

3.2.3. M. leprae vs. M. tuberculosis

M. tuberculosis H37Rv was the first WGS obtained for mycobacteria. This strain was isolated in 1905 from a patient with pulmonary tuberculosis; currently, H37Rv is widely used as reference strain because it has retained its virulence over time. The *M. tuberculosis* genome has 4,411,532 bp, 4000 coding genes and a 65.6% GC content (Fig. 1) (Cole et al., 1998). *M. tuberculosis* and *M. leprae* are phylogenetically close; most of the essential *M. tuberculosis* genes survived to the pseudogenization process in *M. leprae* (Muro et al., 2011).

3.2.4. M. leprae vs. M. marinum

The *M. marinum* genome is the largest reported mycobacterial genome. The *M. marinum* genome contains 6,636,827 bp with a GC content of 62.5%, 5424 coding genes and 65 pseudogenes (Fig. 1). In addition, *M. marinum* contains a 23 kb plasmid (Stinear et al., 2008). *M. marinum* infects a wider range of hosts compared to *M. lepra*e, which has a specific host cell (Eiglmeier et al., 2001; Stinear et al., 2008).

4. Reductive genome evolution in *Mycobacterium leprae* involves massive gene inactivation that leaves pseudogenes as hallmarks

4.1. Reductive genome evolution of bacterial pathogens

When becoming as an obligate intracellular parasite, bacterial pathogens including *M. leprae* overcame biological adversities such as i) how to get to the host cell, ii) how to avoid being killed, iii) how to avoid interfering with the host cell functions, iv) how to maintain the bacterial replication, v) how to obtain new bacterial generations inside host cells, and vi) how to infect new cells (Moulder, 1985). To thrive as intracellular obligate organism, bacterial pathogens have undergone processes of gene disintegration and deletions resulting in a genome contraction (Andersson and Kurland, 1998; Moran, 2002), which known as reductive genome evolution. This phenomenon is mainly produced by excision and insertion of genomic sequences. Therefore, intrachromosomic recombination in repeated sequences, a high proportion of pseudogenes and the accumulation of insertion sequences can reflect gen eliminations. Some bacterial genes and their corresponding functions are lost during the adaptation process to specific environmental conditions (Moran, 2002). This process frequently occurs in most bacterial lineages having close interrelationship with host cells, such as in commensal symbiosis or parasitic relationships (Cole et al., 2001a; Silva et al., 2001; Wernegreen, 2002). Sometimes, genes that code for DNA reparation are lost suggesting that it could be necessary steps for reductive genome evolution (Gomez-Valero et al.,

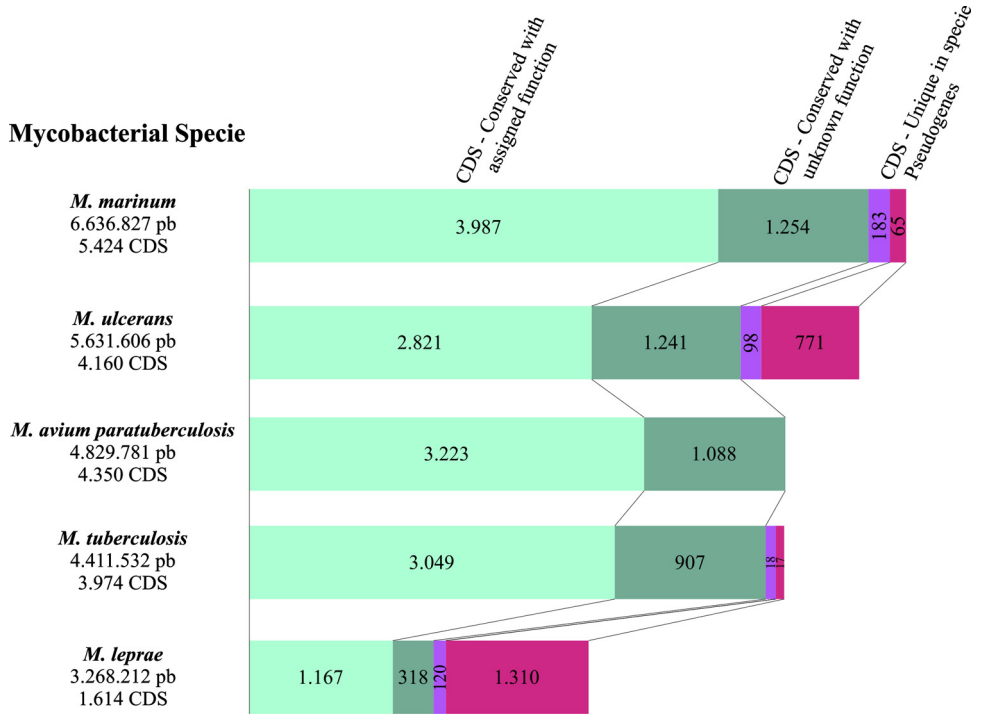

Fig. 1. Open Reading Frames and their predictive functionality in some *Mycobacterium* species, including *M. leprae*. Protein-coding genes (CDS).

2007; Moran, 2002).

Reductive genome evolution leads to biological implication such as i) change from a free-living to a strictly intracellular or host-associated life. On an early stage of this process, for example, facultative intracellular pathogens such as *Legionella* spp. were capable of growing in the cytoplasm of the host cell retaining their capacity to thrive as a free-living organism. Among others, this growth characteristic were obtained by changes in the genes that encode for external surface structures, including genes associated with horizontal gene transfer. ii) restriction from multiple to specific hosts, and/or iii) a change from several to specific host tissues (Toft and Andersson, 2010). In the lattest, the capability of growing on artificial media could be lost; this is the case of intracellular pathogens such as *Rickettsia spp* and *Chlamydia* spp. These bacterial pathogens are highly selective in their host cell (Toft and Andersson, 2010).

Bacterial adaptations are associted with an extensive loss of genes and pseudogenization (Gomez-Valero et al., 2007; Hybiske and Stephens, 2007). After the lifestyle changes occur, the natural pressure to generate and accumulate mutations diminishes, becoming constant among the bacterial population. Over time, some bacterial DNA repair functions are lost, increasing the rate at which deleterious mutations occur (Moran and Plague, 2004).

This genomic reduction is not an aleatory process; thus, when genes are removed from the bacterial genome, usually the same genes are present in the host cell to avoid negative effects on the pathogen, as the case of some genes that encode for amino acid biosynthesis (Andersson and Kurland, 1998). As a consequence of the reduction of genes involved in metabolic functions, the pathogen focuses on basic gene functions for its survival such as replication, transcription, and translation, which are essential fucntions but not enough for growth and survival outside the host cell (Andersson and Kurland, 1998; Moran, 2002). The cytoplasm of the host cell is an isolated environment; thus, the horizontal gene transfer to other bacteria is difficult when pathogenic bacteria are inside the host cell (Andersson and Andersson, 1999b; Andersson and Kurland, 1998).

4.2. Evolutionary reduction in some endosymbiotic bacteria and human pathogens

Many intracellular bacterial species having reductive evolution display small and compact genomes in contrast to relative species with a free-living lifestyle (Andersson et al., 1998). For example, bacterial pathogens such as *Rickettsia prowazekii*, *Chlamydia trachomatis* and *M. leprae*, and in the endosybiotic proteobacteria *Buchnera aphidicola* are only able to replicate inside host cells and have suffered reductive genome evolution (Andersson and Andersson, 1999a; Klasson and Andersson, 2004; Moran, 2002; Moran and Plague, 2004; Stephens et al., 1998; Wixon, 2001).

4.2.1. Buchnera aphidicola

Some bacterial species have established essential relationships with their eukaryotic hosts during their co-evolutionary process allowing them to survive and replicate. This is the case of the endosybiotic bacterial species *B. aphidicola*, which resides inside bacteriocytes present in the aphid (*Homoptera, Insects*), and are transmitted exclusively to the insect offspring. This host-microorganism mutualism is so strict, in consequence, the independent bacterial reproduction is not possible avoiding the bacterial interaction with other bacterial lineages (Moran et al., 2008).

B. aphidicola is phylogenetically close to *E. coli* and *H. influenzae* (Wixon, 2001). The GC content of *B. aphidicola* is low (26.3%)

Table 2
Comparison of genomic characteristics of obligate intracellular bacteria.

Strain	*Buchnera aphidicola*	*Rickettsia prowazekii*	*Chlamydia trachomatis*	*Mycobacterium leprae*
Size	640,681 bp	1,111,523 bp	1,042,519 bp	3,268,203 bp
% GC content	26.3	29.1	41.3	58.0
ORFs	7	834	894	1614
% encoding genome		75.4	89.5	49.6
% Pseudogenes		24		41
Plasmids	7.786 kb		7493 bp	

(Table 2), as seen in other intracellular bacteria displaying reductive genome evolution (Shigenobu et al., 2000). The presence in *B. aphidicola* of genes involved in the biosynthesis of essential amino acids but not those required for non-essential amino acids biosynthesis, is also a relevant genome feature of this proteobacteria. Noteworthy, *B. aphidicola* lacks genes involved in the synthesis of some cell surface components, such as lipopolysaccharides, phospholipids, and genes involved in cell defense. As consequence, *Buchnera* is completely symbiotic and is only viable in the bacteriocyte, its host cell (Gil et al., 2002; Shigenobu et al., 2000). The genetic reduction in *Buchnera* has continued, although its symbiosis with the aphid there has been for more than 150 million years because the ancestral *Buchnera* already had a reduced genome (Gil et al., 2002). The loss of genes causes the production of new pseudogenes and intergenic spacers, which implies that this microorganism can support and compensate for the continuous loss of genes (Gil et al., 2002).

4.2.2. Rickettsia prowazekii

R. prowazekii is an obligate intracellular pathogen that is the etiological agent of typhoid fever in humans, and is transmitted by the human body louse, *Pediculus humanus*, whose target cells are the vascular endothelial cells. Without proper antibiotic treatment, *R. prowazekii* infections have a mortality ranging from 10 to 60% (Bechah et al., 2008). In spite of patients being symptom-free for years, *R. prowazekii* may produce recurrent subclinical typhus after the successful treatment of the primary infection (Mahajan, 2012). *R. prowazekii* has a genome size of 1,111,523 bp, with low GC content (29.1%), and 834 ORFs occupying 75.4% of the whole genome. The average length of the *R. prowazekii* ORF is 1.005 bp (Table 2). Twenty-four percent of the *R. prowazekii* genome consists of pseudogenes and non-coding DNA characterized by the presence of stop codons within ORFs encoding different amino acids and other genome insertions and small deletions. As a consequence, *R. prowazekii* has incurred the loss or degradation of genes involved in the biosynthesis of amino acids and nucleotides, in addition to those involved in anaerobic glycolysis (Zomorodipour and Andersson, 1999). This genomic reduction led to the obligate intracellular lifestyle of this pathogen, which is enforced with the degradation of genes involved in metabolic pathways, whose functions are assumed by the host cell (Andersson and Andersson, 1999a). Finally, there is a bias in the genome deletions and mutations increasing the A/T content, which indicates that the genomes in “ricketsians” have not reached the size and minimum G/C content to be able to survive outside a host cell (Blanc et al., 2007).

4.2.3. Chlamydia trachomatis

Chlamydia trachomatis is an obligate intracellular bacterium having a rigid cell wall and needs eukaryotic cells to multiply due to the bacterial inability to produce essential nutrients, making *C. trachomatis* strictly dependent on the biosynthetic pathways of the host (Wyrick, 2000). The *C. trachomatis* genome sequence contains 1,042,519 bp, with a GC content of 41.3%, 894 ORFs (89.5% coding DNA) and a 7493 bp plasmid (Stephens et al., 1998). *C. trachomatis* depends on the host's cytoplasmic amino acid pool because its genome is significantly lacking in the genes involved in amino acid biosynthesis (Table 2) (Zomorodipour and Andersson, 1999). There is an absence of genes encoding the enzymes involved in purine and pyrimidine biosynthesis in *C. trachomatis* (Stephens et al., 1998). This suggests that nucleosides can be absorbed directly from the host cell cytoplasm and subsequently converted to essential nucleotides and deoxynucleotides, with biosynthetic genes involved in those interconversions in *C. trachomatis* (Stephens et al., 1998). The transition from a free-living lifestyle to an intracellular environment in *C. trachomatis* is associated with a massive loss of genetic information affecting the biosynthetic pathways (Borges et al., 2012).

4.2.4. Mycobacterium leprae

Genomic analyses show that the contraction of the *M. leprae* genome was probably due to a strong alteration in the lifestyle (Gomez-Valero et al., 2007; Liu et al., 2004). It leads to which 50% of genes became pseudogenes, approximately (Table 2) (Singh and Cole, 2011). An study performed by Marri et al. in 2006 showed that *M. leprae* evolved retaining a minimum set of genes representing essential gene families, eliminating several metabolic pathways, and leaving *M. leprae* with markedly specific growth requirements (Cole et al., 2001a; Eiglmeier et al., 2001; Marri et al., 2006; Vissa and Brennan, 2001). This evidence strongly suggests that massive gene inactivation and reductive genome evolution have occurred in the *M. leprae* genome (Akinola et al., 2016; Gomez-Valero et al., 2007).

Different analyses have shown that *M. leprae* has clonal patterns that agree with the migration routes of humans over the past 100,000 years, which suggests that leprosy originated in Africa (Monot et al., 2005, 2009). Therefore, its genome reduction occurred well before the global *M. leprae* expansion and its interaction with different human populations (Taylor et al., 2013). Recombination phenomena of some interspersed repetitive sequences such as RLEP, REPLEP, LEPREP and LEPRPT (37, 15, 8 and 5 copies, respectively) in *M. leprae* genome (TN strain) have occurred as adaptation mechanisms against gene suppression, translocation, and inversions; which are the most relevant observed features of *M. leprae* genome variability (Cole et al., 2001b; Taylor et al., 2013). The presence of RLEP repeats in the 3´ end of some genes, often within pseudogenes, suggests that these sequences are transposon residues that have lost their transposition capacity. Moreover, chromosomal rearrangements (inversions, translocations and gene deletions) that are mediated by recombination events of these repetitive elements could also explain the loss of synteny (Cole et al., 2001b). Another starting point for the massive gene inactivation leading to reductive genome evolution in *M. leprae*, as suggested by several authors, could be a change in the occupied niche and adaptation to surviving in highly specialized cells, such as Schwann cells. Once *M. leprae* hosted in Schwann cells, bacilli were free from selective pressure from other microorganisms (Marri et al., 2006).

5. Pseudogenization in *M. leprae*

Pseudogenes are non-functional genes as consequence of frequent nucleotide replacements that usually produce mutations. Elimination of typical start codons and/or vital gene segments, and accumulation of stop codons within ORFs are frequently found in pseudogenes (Mira and Pushker, 2005). These types of mutations sometimes inactivates promoters, regulatory sequences or ribosome-binding sites, resulting in the silencing of transcriptional and/or translational functions, and the premature termination of protein synthesis (Andersson and Andersson, 1999a, b; Lawrence et al., 2001; Mira and Pushker, 2005)

Pseudogenization in *M. leprae* occurred through a gene-by-gene process (Gomez-Valero et al., 2007). This hypothesis is supported by the chromosomal distribution of the lost genes in the ancestral *M. leprae* genome. On the other hand, there are *M. leprae* genes, and pseudogenes that do not have orthologs in other mycobacterial genomes, suggesting that *M. leprae* pseudogenization could be also promoted by horizontal gene transfer after the *M. leprae* and *M. tuberculosis* divergence (Muro et al., 2011). At least 200 pseudogenes were added to *M. leprae* after its divergence from *M. tuberculosis* (Gomez-Valero et al., 2007). Many of the insertion events that produced new pseudogenes apparently did not provide advantages to the *M. leprae* pathogenicity (Liu et al., 2004).

M. leprae retained essential genes during pseudogenization. For example, among the cell-structure components conserved through evolution, *M. leprae* expresses phenolic glycolipid (PGL-1) on the cell wall, a relevant *M. leprae* antigen (Scollard et al., 2006). Despite suffering an evolutionary reduction, there are not pseudogenes associated with PGL-1 biosynthesis (Callegaro-Filho et al., 2010), suggesting that PGL-1 could be an essential cell-wall component for the *M. leprae*

survival within host cells. Nowadays is known that PGL-1 is pivotal for the *M. leprae* recognition and entry into Schwann cells (Rambukkana, 2000). Intriguingly, PGL-1 is not essential for the cellular response to *M. leprae*, which in turn is responsible for controlling the infection progress (Scollard et al., 2006; Spencer and Brennan, 2011).

5.1. Mycobacterium leprae pseudogenes retain functional gene structures

Both the number and proportion of pseudogenes in the *M. leprae* genome are unique among prokaryotes and eukaryotes (Harrison and Gerstein, 2002; Liu et al., 2004). Surprisingly, 82.3% of pseudogenes sequences in *M. leprae* retain characteristics of functional genes. Although some ORFs in *M. leprae* pseudogenes contain several stop codons (ranging from 1 to 40), many of those ORFs are transcribed and/or translated (Mira and Pushker, 2005; Williams et al., 2009). Most pseudogenes are highly conserved among *M. leprae* strains; however, recent studies have shown that s of them are strain-specific (Table 1) (Monot et al., 2009; Singh and Cole, 2011). On the other hand, Muro et al. (2011) observed that bacterial intergenic regions of approximately 750 bp in size usually display a linkage map distance of zero in the *M. leprae* genome. This finding suggests that contiguous genes possibly belong to the same operon in *M. leprae* (Muro et al., 2011). This work also found that essential *M. leprae* genes tend to be positioned in the 5′ half of operons; however, the pseudogenes location in *M. leprae* is biased for the 3′ half of operons. These findings suggest that selective pressure has caused rearrangements leading to functionally important genes being located upstream of less important genes in operons (Madan Babu, 2003; Muro et al., 2011).

5.2. Possible mechanisms involved in Mycobacterium leprae pseudogenization

The lack of DNA polymerase III proofreading activity, which greatly contributes to an accumulation of gene mutations, is one of the most likely pseudogenization mechanisms proposed in bacteria (Liu et al., 2004). Different *M. leprae* pseudogenization mechanisms have been proposed including the loss of sigma (σ) factors and two-component systems (Madan Babu, 2003; Tyagi and Saini, 2004); however, the lack of proofreading DNA polymerase III activity mediated by DnaQ promoted pseudogenization in *M. leprae* (Cole et al., 2001a; Mizrahi et al., 2000). The impaired proofreading activity increases genomic mutation rates, restriction *M. leprae* to confined environments (see Fig. 2B) (Liu et al., 2004). In addition, the elimination of intergenic regulatory regions, or the accumulation of mutations in the ribosome binding site and promoter regions, are also possible mechanisms for the production of *M. leprae* pseudogenes (Williams et al., 2009).

Alteration or loss of functional Shine-Dalgarno (SD) sequences results in a lack of efficient translation, which in turn diminishes or eliminates the proteins production (Fig. 2C) (Mira and Pushker, 2005). In *M. leprae*, where pseudogenes have only slightly reduced length (average length is 804 bp versus 977 bp for functional genes), the SD sequence in pseudogenes is highly degraded suggesting an altered protein translation. The *M. leprae* metabolic rate is potentially reduced by the production of defective proteins, and although the pseudogenes can persist for a long time in the genome, they are effectively "silenced" by a decreased binding of the ribosome to the SD sequences (Mira and Pushker, 2005; Williams et al., 2009).

In the *M. leprae* pseudogenization process, multiple mutations in the genome produced accumulation of mutations and stop codons (Madan Babu, 2003). It occurred in two different phases: first, the loss of σ factors avoided the expression of genes associated with particular stresses that were absent in the specific niche (Schwann cells) hosting *M. leprae* (Fig. 2A); and second, once *M. leprae* was hosted in Schwann cells, it lacked selective pressure, leading to the induction of genome mutations and gene inactivations (Madan Babu, 2003). During a successful infection process, mycobacteria face the host immune response through a coordinated regulation of gene transcription and expression, in which σ factors plays a pivotal role (Tyagi and Saini, 2004). The mycobacterial RNA polymerase is a holoenzyme composed of the RNA polymerase core and a small interchangeable subunit known as the sigma factor σ required for the initiation of transcription and specific recognition of the -35 and -10 promoter sequences (Smith, 2003). The σ factor binding and dissociation to the RNA polymerase core mediates specific cellular responses (Manganelli et al., 2004, 2001). It is known that mycobacterial σ factors regulate the temporal expression of specific regulons in response to several physiological and/or environmental stimuli. In addition, σ factors are also important for: adhesion, invasion, intracellular replication, dissemination via host-pathogen interactions, transcription of housekeeping genes, survival during stationary phase, and the stress response (Gruber and Gross, 2003; Smith, 2003).

Typical mycobacterial genomes, such as *M. tuberculosis,* encodes 13 different σ factors (See Table 3) essential for gene transcription (Cole et al., 1998; Manganelli et al., 2004). The specific lack of the σ factors *sigE, sigH,* and *sigF* makes tubercle bacilli sensitive to different environmental stresses, such as heat shock, oxidative stress, and growth in macrophages. In addition, they are also involved in the attenuation of *M. tuberculosis* (Manganelli, 2014; Manganelli et al., 2002, 2001; Raman et al., 2001). *M. leprae* lost 9 of the typical 13 σ factors present in the *Mycobacterium* genus, including *M. tuberculosis* (Manganelli et al., 2004). Some of those σ factors evolved into pseudogenes that display several functions in *M. leprae*. The σ factors still retained in *M. leprae* are: *sigA,* which is associated with intracellular growth; *sigB,* which is involved in the heat shock response and oxidative stress; *sigC*, which is downregulated during stationary phase and under several stress conditions, and *sigE,* whose expression is correlated with the response to low phosphate levels, alkaline pH, and the response against vancomycin (Manganelli, 2014). Genes such as *sigD, sigF, sigG, sigH, sigI, sigJ, sigK y sigM* were lost in *M. leprae.* It is believed that the suppression of sigF, which is induced when the bacteria is exposed to multiple environments including anaerobiosis, and other sigma factors, makes difficult the *M. leprae* adaptation to different environmental conditions (Cole et al., 2001a; Manganelli et al., 2004).

Another possible mechanism for *M. leprae* pseudogenization is the loss of specific signal transduction systems, such as two-component systems compromising the ability of *M. leprae* to adapt to particular environments (Tyagi and Saini, 2004). For instance, the latency global regulator DevR/DosR, which regulates the expression of genes involved in the stress response to hypoxia in mycobacteria, evolved into pseudogenes and producing an impaired response to hypoxia in *M. leprae*, (O'Toole et al., 2003; Park et al., 2003). Tyagi and Saini (2004) looked hortologuous *M. tuberculosis* genes activated by hypoxia in *M. leprae* (Cole et al., 2001a; Tyagi and Saini, 2004), and observed that forty-eight of the sixty-eight genes involved in the *M. tuberculosis* response to hypoxia evolved into pseudogenes or were absent in *M. leprae*, including the two-component system DevR/DosR. This finding indicates that *M. leprae* lost over 68% of the genes associated with the hypoxic response, which is proportional to the σ factor lost in *M. leprae* (69%) (Tyagi and Saini, 2004). Thus, the loss of DevR/DosR could have induced the accumulation of gene mutations involved in the *M. leprae* response to low oxygen tension. Six of the eleven orthologous *M. tuberculosis* two-component systems evolved to pseudogene in *M. leprae*.

Although it is unknown whether the suppression of two-component systems preceded the loss of σ factors or vice versa, the question is why so many systems that regulate adaptation processes were susceptible to mutations or were lost in *M. leprae* genome and only retaining essential genes (Tyagi and Saini, 2004). For example, the central carbon and energy pathways are still complete in *M. leprae*, whereas the alternative pathways genes are degenerated; thus, *M. leprae* is able of biosinthesize most of these pathway components, except for methionine, which must be obtained from host cells (Wheeler, 2001). The distribution of genes of anabolic pathways are virtually complete evidencing lost of genetic redundancy instead of functionality. The energy metabolism of *M.*

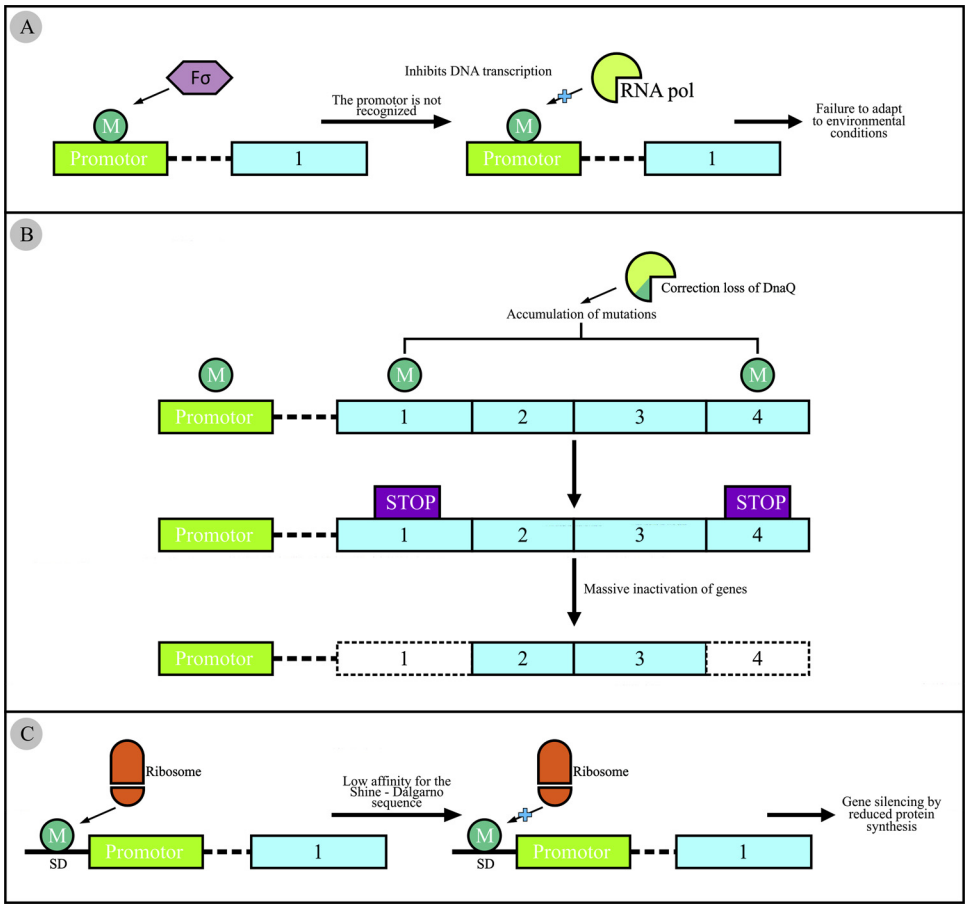

Fig. 2. Possible mechanisms involved in the pseudogene formation. A) The coordinated regulation of the mycobacterial gene expression mainly occurs during the initiation of transcription mediated by the RNA polymerase holoenzyme, which contains the σ, for the promoter recognition. When mutations in the *M. leprae* regulatory sequences occur, the σ may not recognize the promoter inhibiting the DNA transcription and in which is turn could avoid the adaptation to different environmental conditions. B) The loss of activity mediated DNAQ correction in DNA polymerase III could induce mutations in different parts of the *M. leprae* genome and accumulation of stop codons, pseudogenes production and therefore massive gene inactivation. C. The Shine-Dalgarno (SD) region is a sequence complementary to a highly conserved region in the 16S rRNA subunit, which is involved in the mRNA binding to the ribosome; when functional SD sequences are lost, translation is inefficient leading to reduction or loss of protein synthesis.
Fσ, sigma factor; SD, Shine-Dalgarno sequence; M, Mutation; DNAQ, DNA polymerase III subunit epsilon; STOP, Stop codon.

Table 3
Genes encoding sigma factors in mycobacteria.

Gen Presence	*M. tuberculosis*	*M. avium*	*M. paratuberculosis*	*M. marinum*	*M. leprae*	Inducing conditions [a]
sigA	+	+	+	+	+	Growth in macrophages
sigB	+	+	+	+	+	Heat shock, oxidative stress, surface stress, growth in macrophages
sigC	+	+	+	+	+	Downregulated in stationary phase and in response to several stress conditions
sigC like	–	–	–	+	–	
sigD	+	+	+	+	P	Nutrient depletion; downregulated in response to several stress conditions
sigE	+	+	+	+	+	Surface stress, oxidative stress, heat shock, low phosphate, alkaline pH, vancomycin, growth in macrophages
sigF	+	+	+	+	P	Nutrient depletion
sigF like	–	+	+	–	–	
sigG	+	+	+	+	P	Growth in macrophages, DNA-damaging agents
sigH	+	+	+	+	P	Heat shock, oxidative stress, growth in macrophages
sigI	+	–	–	–	P	Stationary phase, heat or cold shock
sigJ	+	+	+	+	P	Stationary phase, growth in macrophages
sigK	+	+	–	+	P	
sigL	+	P?	+	+	–	
sigM	+	+	+	+	P	Heat shock, late stationary phase

+ Presence of the gene; -, Absence of the gene; P, pseudogene. [a] functions established in *M. tuberculosis.*
Modified from Manganelli et al., 2004.

leprae has less than a third of the functional genes than *M. tuberculosis* and has lost alternative pathways giving a result of limited anaerobic metabolism. (Wheeler, 2001).

5.3. Differential transcription of pseudogenes affects M. leprae physiology

Although the functional roles of pseudogenes and other noncoding regions are not-well-understood in *M. leprae*, it is known that some of them are highly transcribed (Akama et al., 2009; Williams et al., 2009). Levels of pseudogene transcription in *M. leprae* varies after macrophage infection both *in vitro* and *in vivo* (Nakamura et al., 2009; Suzuki et al., 2006). Williams et al. (2009) found that as many as 36% of the total *M. leprae* transcripts belong to pseudogenes, and that 43% of all pseudogenes were transcriptionally active (Williams et al., 2009). Independent microarray analyses demonstrated that only half of the *M. leprae* transcripts are derived from coding genes, while the remaining ones are produced by pseudogenes and noncoding regions, including *M. leprae*-specific RLEP sequences. Most transcribed pseudogenes (75%) lack conventional start codons, while 67% of expressed pseudogenes contain more than five stop codons (Williams et al., 2009).

The metabolic cost of transcribing apparently nonfunctional sequences could be detrimental to cell growth and physiology, or may

Table 4
General comparative characteristics of microorganisms revised.

Process	Compared microorganisms			
	Buchnera aphidicola	*Rickettsia prowazekii*	*Chlamydia trachomatis*	*Mycobacterium leprae*
Extreme genome reduction model	1 Multiple disintegration events from dispersed genes, through out the genome. 2 Contraction of the highly dynamic genome during the early stages of symbiosis 3 Large deletions causing massive losses of genetic material 4 Relatively minor losses with elimination of large genomic areas of redundant and useless genes in the stable intracellular context, generating pseudogenes	1 Gradual loss in time of genetic material by cascade events initiated for a point mutation inactivating a dispensable gene and subsequent mutations that transform the pseudogene into an unrecognizable sequence, passing to small fragments and finally there is extinction. 2 In this genome, more than 20 pseudogenes have been identified, with short pieces of a gene showing strong similarities with the *spoT* and *relA* genes.	1 No fragments of *metK* genes have been identified and there are no *spoT* or *relA* homologs either. 2 In *R. prowazekii spoT* and *relA* are pseudogenes and in *C. trachomatis*, they were eliminated, possibly this is more efficient eliminating sequences of non-functional genes. 3 *C. trachomatis* is an obligate intracellular parasite older than *R. prowazekii* and their population structures are slightly different.	1 Genome reduction by recombination events between repetitive sequences that occupy ≈2% of the TN genome. 2 At the 3' end of genes or within pseudogenes there are RLEPs that suggest that they are traces of transposons no longer capable of transposing. 3 *M. leprae* has > 26 IS inactive elements, in addition to the 4 families of scattered repetitions: 4 RLEP (37 copies); 5 REPLEP (15 copies); 6 LEPREP (8 copies); 7 • LEPRPT (5 copies).
Loss of repair genes	1 There is a limited capacity for DNA repair and recombination 2 It lacks to. a the genes essential for recombination in bacteria, b the *recA* and *uvrABC* genes required for the excision repair system c the genes of the SOS system. 3 It lacks the genes that encode the methylation proteins and the restriction enzymes.	1 Their increased mutation rate correlates with the loss of genes involved in DNA repair pathways. 2 Loss of six genes belonging to DNA repair processes including: phrB, radC, MutM and MutT, these last two protect the cell against the effects of the oxidative stress product. 3 Likewise, two putative DNA repair protein genes are rented	1 In general, the repair systems of *Chlamydia sp.*, seem to be more complete 2 Is notable the absence of *MutT* or MutM orthologs coding for proteins that repair the oxidative damage in all other bacteria.	1 Most *M. leprae* genes have orthologs in *M. tuberculosis*, but some appear to be unique and may have new functions. 2 Hypothetical proteins with unknown function and a series of potential enzymes, such as inorganic pyrophosphatase encoded by ppa, are included. 3 Enzymes have been acquired by horizontal gene transfer as ProS, for the only aminoacyl-tRNA synthetase of M. leprae without counterpart in M. tuberculosis.
rRNA genes disposition	1 The rRNA genes are present as unique copies in *Buchnera,* while in *E. coli* they are present in 7 copies 2 In *Buchnera,* a single copy is considered a characteristic of slow-growing organisms.	1 The rRNA genes in all available *Rickettsia* genomes have an unusual disposition 2 The 16S rRNA gene is separated from the group of 23S and 5S rRNA genes, 3 The upstream spacer of 23S rRNA gene is rearranged in some *Rickettsia* species 4 The rearrangement of the rRNA genes was produced by intra-chromosomal recombination before speciation in *Rickettsia spp*, since it contains short repetitive sequences, eliminated in other related species,		1 The synthesis of macromolecules such as ribosomes, aminoacyl tRNA, RNA and proteins is reasonably intact.
Nucleotide *de novo* synthesis	1 It is not clear if *Buchnera* has retained the ability to synthesize all its nucleotides. 2 Unlike *E. coli*, it does not possess all enzymes that are required for the inter-conversion of deoxy, mono, di and tri-phosphorylated nucleotides 3 In *Buchnera,* the enzymes show more specificity. 4 Essential precursors can be imported directly from the host cell.	1 Did not find genes for *de novo* synthesis of nucleosides in R. prowazekii genome. 2 Four genes are present for the conversion of mono to nucleoside diphosphates (*adk, gmk, cmk* and *pyrH*) 3 There are two genes that encode ribonucleotide reductase, which converts ribonucleoside diphosphate to deoxyribonucleoside diphosphate 4 The *ndk* gene of the nucleoside diphosphate kinase that converts NDP / dNDP into NTP / dNTP is present in *R. prowazekii.* 5 There is a complete set of genes for conversion of dCTP and dUTP into TTP, including thyA, which encodes thymidylate synthase.	1 The ability' lacks, to synthesize purines and pyrimidines de novo, except for the synthesis of cytosine-triphosphate. 2 The genes of enzymes required for the synthesis and interconversion of each of the deoxyribonucleotides are present, except for the thymidylate synthase gene, for the synthesis of deoxythymidine, monophosphate obtained from the host cell.	1 The nucleic acids and bioproducts of purine synthesis pathway are essential components of energy metabolism and protein synthesis. 2 *M leprae* can't synthesize *de novo* purines, which partly explains its inability to grow in artificial media

(*continued on next page*)

Table 4 (*continued*)

Process	Compared microorganisms			
	Buchnera aphidicola	*Rickettsia prowazekii*	*Chlamydia trachomatis*	*Mycobacterium leprae*
		6 *R. prowazekii* genome encodes all enzymes for the interconversion of nucleoside monophosphates to all other required nucleotides. 7 The nucleoside monophosphate is imported from the eukaryotic host.		

have some biological relevance in *M. leprae*. The rate of pseudogene deletion in *M. leprae* appears to be slower than that in other bacteria, suggesting that these sequences could provide backup functions, or be activated by events such as gene conversion in unusual circumstances (Mira et al., 2006). Pseudogenes might represent a reservoir for genetic diversity (Singh and Cole, 2011). Many authors have speculated that some of these transcribed noncoding regions and pseudogenes play roles in the regulation of infection, intracellular parasitism, and replication of *M. leprae* (Nakamura et al., 2009); however, further research still needs to be done with regards to these postulates.

6. Conclusions

In addition to the human's nasal mucosa, there are other possible transmission sources and reservoirs of the leprosy bacillus, such as FLA, the nine-banded armadillo, arthropods and red-tailed squirrels. The changes in the *M. leprae* genome that affected it lifestyle, led it to become an intracellular pathogen with an exquisite tropism to Schwann cells that provides an adequate environment for the bacilli's grow. During this process, *M. leprae* underwent an evolutionary genome reduction in which a minimum set of genes essential for an obligate intracellular life style were retained, leaving *M. leprae* with very specific growth requirements that preclude its growth or multiplication in artificial media. Other intracellular pathogens, such as *C. trachomatis* and *R. prowazekii* also eliminated some biosynthetic pathways that making them dependent on their specific host cells (Table 4). This phenomenon of specialization observed in many pathogens that become obligate intracellular after having been environmental, is demonstrated in the loss of some specific biosynthetic pathways that are essential for them to survive as environmental microorganisms but that after these molecular changes are converted into depending on a specific host, after responding to the environmental pressure they have suffered and subsequently trying to adapt to it, they lose the genetic machinery that is left over or redundant. All the above mainly occurs in groups of genes of the genomic repertoire necessary for environmental life and are mainly the pathways of synthesis of purines and pyrimidines, the synthesis of proteins with enzymatic character and the repair pathways of DNA damaged by stress conditions, among others. The evolutionary process gave the *M. leprae* genome the highest percentage of pseudogenes found in any prokaryotic or eukaryotic microorganism. Comparative genomic studies have shown that a large number of pseudogenes are conserved among different reference strains of *M. leprae* isolated in distant geographic regions; however, some of the pseudogenes are strain specific suggesting that the pseudogenization processes may occur after the strains spreading around the world. Although approximately 41% of the *M. leprae* genome consists of pseudogenes, many of them transcriptionally active, the processes leading to their formation occurred in different ways, including the lost of σ factors and the DnaQ polymerase proofreading activity, accumulation of stop codons, and the development of mutations in translation-promoting regions. The role of pseudogenes in the biology of *M. leprae* remains elusive; thus, more deeply studies are necessary to gain a better understanding of their impact on the *M. leprae* biology.

References

Akama, T., Suzuki, K., Tanigawa, K., Kawashima, A., Wu, H., et al., 2009. Whole-genome tiling array analysis of *Mycobacterium leprae* RNA reveals high expression of pseudogenes and noncoding regions. J. Bacteriol. 191, 3321–3327.

Akinola, R.O., Mazandu, G.K., Mulder, N.J., 2016. A quantitative approach to analyzing genome reductive evolution using protein-protein interaction networks: a case study of *Mycobacterium leprae*. Front. Genet. 7, 39.

Andersson, J.O., Andersson, S.G., 1999a. Genome degradation is an ongoing process in *Rickettsia*. Mol. Biol. Evol. 16, 1178–1191.

Andersson, J.O., Andersson, S.G., 1999b. Insights into the evolutionary process of genome degradation. Curr. Opin. Genet. Dev. 9, 664–671.

Andersson, S.G., Zomorodipour, A., Andersson, J.O., Sicheritz-Ponten, T., Alsmark, U.C., et al., 1998. The genome sequence of *Rickettsia prowazekii* and the origin of

mitochondria. Nature 396, 133–140.
Andersson, S.G.E., Kurland, C.G., 1998. Reductive evolution of resident genomes. Trends Microbiol. 6, 263–268.
Avanzi, C., Del-Pozo, J., Benjak, A., Stevenson, K., Simpson, V.R., et al., 2016. Red squirrels in the British Isles are infected with leprosy bacilli. Science 354, 744–747.
Banerjee, R., Banerjee, B.D., Chaudhury, S., Hati, A.K., 1991. Transmission of viable *Mycobacterium leprae* by *Aedes aegypti* from lepromatous leprosy patients to the skin of mice through interrupted feeding. Lepr. Rev. 62, 21–26.
Bechah, Y., Capo, C., Mege, J.L., Raoult, D., 2008. Epidemic typhus. Lancet Infect. Dis. 8, 417–426.
Blanc, G., Ogata, H., Robert, C., Audic, S., Suhre, K., et al., 2007. Reductive genome evolution from the mother of Rickettsia. PLoS Genet. 3, e14.
Borges, V., Nunes, A., Ferreira, R., Borrego, M.J., Gomes, J.P., 2012. Directional evolution of *Chlamydia trachomatis* towards niche-specific adaptation. J. Bacteriol. 194, 6143–6153.
Bratschi, M.W., Steinmann, P., Wickenden, A., Gillis, T.P., 2015. Current knowledge on *Mycobacterium leprae* transmission: a systematic literature review. Lepr. Rev. 86, 142–155.
Callegaro-Filho, D., Shrestha, N., Burdick, A.E., Haslett, P.A., 2010. A potential role for complement in immune evasion by *Mycobacterium leprae*. J. Drugs Dermatol. 9, 1373–1382.
Cole, S.T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., et al., 1998. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. Nature 393, 537–544.
Cole, S.T., Eiglmeier, K., Parkhill, J., James, K.D., Thomson, N.R., et al., 2001a. Massive gene decay in the leprosy bacillus. Nature 409, 1007–1011.
Cole, S.T., Singh, P., 2012. History and phylogeography of leprosy. In: Nunzi, E., Massone, C. (Eds.), Leprosy: A Practical Guid. Springer, Milan, Milano, pp. 3–13.
Cole, S.T., Supply, P., Honore, N., 2001b. Repetitive sequences in *Mycobacterium leprae* and their impact on genome plasticity. Lepr. Rev. 72, 449–461.
Eiglmeier, K., Simon, S., Garnier, T., Cole, S.T., 2001. The integrated genome map of *Mycobacterium leprae*. Lepr. Rev. 72, 462–469.
Fischer, E., De Vlas, S., Meima, A., Habbema, D., Richardus, J., 2010. Different mechanisms for heterogeneity in leprosy susceptibility can explain disease clustering within households. PLoS One 5, e14061.
Gil, R., Sabater-Munoz, B., Latorre, A., Silva, F.J., Moya, A., 2002. Extreme genome reduction in *Buchnera spp.*: toward the minimal genome needed for symbiotic life. Proc. Natl. Acad. Sci. U. S. A. 99, 4454–4458.
Gomez-Valero, L., Rocha, E.P., Latorre, A., Silva, F.J., 2007. Reconstructing the ancestor of *Mycobacterium leprae*: the dynamics of gene loss and genome reduction. Genome Res. 17, 1178–1185.
Groathouse, N.A., Rivoire, B., Kim, H., Lee, H., Cho, S.N., et al., 2004. Multiple polymorphic loci for molecular typing of strains of *Mycobacterium leprae*. J. Clin. Microbiol. 42, 1666–1672.
Gruber, T.M., Gross, C.A., 2003. Multiple sigma subunits and the partitioning of bacterial transcription space. Annu. Rev. Microbiol. 57, 441–466.
Han, X.Y., Seo, Y.H., Sizer, K.C., Schoberle, T., May, G.S., et al., 2008. A new *Mycobacterium* species causing diffuse lepromatous leprosy. Am. J. Clin. Pathol. 130, 856–864.
Harrison, P.M., Gerstein, M., 2002. Studying genomes through the aeons: protein families, pseudogenes and proteome evolution. J. Mol. Biol. 318, 1155–1174.
Hybiske, K., Stephens, R.S., 2007. Mechanisms of host cell exit by the intracellular bacterium *Chlamydia*. Proc. Natl. Acad. Sci. U. S. A. 104, 11430–11435.
Johnson, P.D., Stinear, T., Small, P.L., Pluschke, G., Merritt, R.W., et al., 2005. Buruli ulcer (*M. ulcerans*infection): new insights, new hope for disease control. PLoS Med. 2, e108.
Kennedy, G.M., Morisaki, J.H., Champion, P.A., 2012. Conserved mechanisms of *Mycobacterium marinum* pathogenesis within the environmental amoeba *Acanthamoeba castellanii*. Appl. Environ. Microbiol. 78, 2049–2052.
Kirchheimer, W.F., 1975. Experimental leprosy in the nine-banded armadillo. Public Health Rep. 90, 483–485.
Kirchheimer, W.F., 1976. The role of arthropods in the transmission of leprosy. Int. J. Leprosy Mycobacterial Dis. 44, 104–107.
Kirchheimer, W.F., Storrs, E.E., Binford, C.H., 1972. Attempts to establish the Armadillo (*Dasypus novemcinctus linn.*) as a model for the study of leprosy. II. Histopathologic and bacteriologic post-mortem findings in lepromatoid leprosy in the Armadillo. Int. J. Leprosy Mycobacterial Dis. 40, 229–242.
Klasson, L., Andersson, S.G., 2004. Evolution of minimal-gene-sets in host-dependent bacteria. Trends Microbiol. 12, 37–43.
Lahiri, R., Krahenbuhl, J.L., 2008. The role of free-living pathogenic amoeba in the transmission of leprosy: a proof of principle. Lepr. Rev. 79, 401–409.
Lamrabet, O., Mba Medie, F., Drancourt, M., 2012. *Acanthamoeba polyphaga*-enhanced growth of *Mycobacterium smegmatis*. PLoS One 7, e29833.
Lavender, C.J., Fyfe, J.A.M., Azuolas, J., Brown, K., Evans, R.N., et al., 2011. Risk of Buruli ulcer and detection of *Mycobacterium ulcerans* in mosquitoes in Southeastern Australia. PLoS Negl. Trop. Dis. 5, e1305.
Lawrence, J.G., Hendrix, R.W., Casjens, S., 2001. Where are the pseudogenes in bacterial genomes? Trends Microbiol. 9, 535–540.
Li, L., Bannantine, J.P., Zhang, Q., Amonsin, A., May, B.J., et al., 2005. The complete genome sequence of *Mycobacterium avium subspecies paratuberculosis*. Proc. Natl. Acad. Sci. U. S. A. 102, 12344–12349.
Liu, Y., Harrison, P.M., Kunin, V., Gerstein, M., 2004. Comprehensive analysis of pseudogenes in prokaryotes: widespread gene decay and failure of putative horizontally transferred genes. Genome Biol. 5, R64.
Madan Babu, M., 2003. Did the loss of sigma factors initiate pseudogene accumulation in M. leprae? Trends Microbiol. 11, 59–61.
Mahajan, S.K., 2012. Rickettsial diseases. J. Assoc. Physicians India 60, 37–44.
Manganelli, R., 2014. Sigma Factors: Key Molecules in *Mycobacterium tuberculosis* Physiology and Virulence. Microbiology Spectrum 2, MGM2-0007-2013.
Manganelli, R., Provvedi, R., Rodrigue, S., Beaucher, J., Gaudreau, L., et al., 2004. Sigma factors and global gene regulation in *Mycobacterium tuberculosis*. J. Bacteriol. 186, 895–902.
Manganelli, R., Voskuil, M.I., Schoolnik, G.K., Dubnau, E., Gomez, et al., 2002. Role of the extracytoplasmic-function sigma factor sigma (H) in *Mycobacterium tuberculosis* global gene expression. Mol. Microbiol. 45, 365–374.
Manganelli, R., Voskuil, M.I., Schoolnik, G.K., Smith, I., 2001. The *Mycobacterium tuberculosis* ECF sigma factor sigmaE: role in global gene expression and survival in macrophages. Mol. Microbiol. 41, 423–437.
Marri, P.R., Bannantine, J.P., Golding, G.B., 2006. Comparative genomics of metabolic pathways in *Mycobacterium* species: gene duplication, gene decay and lateral gene transfer. FEMS Microbiol. Rev. 30, 906–925.
Mira, A., Pushker, R., 2005. The silencing of pseudogenes. Mol. Biol. Evol. 22, 2135–2138.
Mira, A., Pushker, R., Rodriguez-Valera, F., 2006. The Neolithic revolution of bacterial genomes. Trends Microbiol. 14, 200–206.
Misch, E.A., Berrington, W.R., Vary Jr., J.C., Hawn, T.R., 2010. Leprosy and the human genome. Microbiol. Mol. Biol. Rev. 74, 589–620.
Mizrahi, V., Dawes, S.S., Rubin, H., 2000. Molecular Genetics of *Mycobacteria*. American Society for Microbiology, Washington, DC.
Moet, F.J., Pahan, D., Schuring, R.P., Oskam, L., Richardus, J.H., 2006. Physical distance, genetic relationship, age, and leprosy classification are independent risk factors for leprosy in contacts of patients with leprosy. J. Infect. Dis. 193, 346–353.
Monot, M., Honore, N., Garnier, T., Araoz, R., Coppee, J.Y., et al., 2005. On the origin of leprosy. Science 308, 1040–1042.
Monot, M., Honore, N., Garnier, T., Zidane, N., Sherafi, D., et al., 2009. Comparative genomic and phylogeographic analysis of *Mycobacterium leprae*. Nat. Genet. 41, 1282–1289.
Moran, N.A., 2002. Microbial minimalism: genome reduction in bacterial pathogens. Cell 108, 583–586.
Moran, N.A., McCutcheon, J.P., Nakabachi, A., 2008. Genomics and evolution of heritable bacterial symbionts. Annu. Rev. Genet. 42, 165–190.
Moran, N.A., Plague, G.R., 2004. Genomic changes following host restriction in bacteria. Curr. Opin. Genet. Dev. 14, 627–633.
Moulder, J.W., 1985. Comparative biology of intracellular parasitism. Microbiol. Rev. 49, 298–337.
Muro, E.M., Mah, N., Moreno-Hagelsieb, G., Andrade-Navarro, M.A., 2011. The pseudogenes of *Mycobacterium leprae* reveal the functional relevance of gene order within operons. Nucleic Acids Res. 39, 1732–1738.
Nakamura, K., Akama, T., Bang, P.D., Sekimura, S., Tanigawa, K., et al., 2009. Detection of RNA expression from pseudogenes and non-coding genomic regions of *Mycobacterium leprae*. Microb. Pathog. 47, 183–187.
Neumann, Ad.S., Dias, Fd.A., Ferreira, Jd.S., Fontes, A.N.B., Rosa, P.S., et al., 2016. Experimental infection of rhodnius prolixus (*Hemiptera, Triatominae*) with *Mycobacterium leprae* indicates potential for leprosy transmission. PLoS One 11, e0156037.
Ng, V., Zanazzi, G., Timpl, R., Talts, J.F., Salzer, J.L., et al., 2000. Role of the cell wall phenolic glycolipid-1 in the peripheral nerve predilection of *Mycobacterium leprae*. Cell 103, 511–524.
O'Toole, R., Smeulders, M.J., Blokpoel, M.C., Kay, E.J., Lougheed, K., et al., 2003. A two-component regulator of universal stress protein expression and adaptation to oxygen starvation in *Mycobacterium smegmatis*. J. Bacteriol. 185, 1543–1554.
Park, H.D., Guinn, K.M., Harrell, M.I., Liao, R., Voskuil, M.I., et al., 2003. Rv3133c/dosR is a transcription factor that mediates the hypoxic response of *Mycobacterium tuberculosis*. Mol. Microbiol. 48, 833–843.
Penna, G.O., Bührer-Sékula, S., Kerr, L.R.S., Stefani, M.Md.A., Rodrigues, L.C., et al., 2017. Uniform multidrug therapy for leprosy patients in Brazil (U-MDT/CT-BR): results of an open label, randomized and controlled clinical trial, among multibacillary patients. PLoS Negl. Trop. Dis. 11, e0005725.
Raman, S., Song, T., Puyang, X., Bardarov, S., Jacobs, W.R., et al., 2001. The alternative sigma factor SigH regulates major components of oxidative and heat stress responses in *Mycobacterium tuberculosis*. J. Bacteriol. 183, 6119–6125.
Rambukkana, A., 2000. How does *Mycobacterium leprae* target the peripheral nervous system? Trends Microbiol. 8, 23–28.
Rambukkana, A., 2004. *Mycobacterium leprae*-induced demyelination: a model for early nerve degeneration. Curr. Opin. Immunol. 16, 511–518.
Rambukkana, A., 2010. Usage of signaling in neurodegeneration and regeneration of peripheral nerves by leprosy bacteria. Prog. Neurobiol. 91, 102–107.
Ridley, D.S., Jopling, W.H., 1966. Classification of leprosy according to immunity. A five-group system. Int. J. Leprosy Mycobacterial Dis. 34, 255–273.
Salah, I.B., Ghigo, E., Drancourt, M., 2009. Free-living amoebae, a training field for macrophage resistance of mycobacteria. Clin. Microbiol. Infect. 15, 894–905.
Scollard, D.M., Adams, L.B., Gillis, T.P., Krahenbuhl, J.L., Truman, R.W., et al., 2006. The continuing challenges of leprosy. Clin. Microbiol. Rev. 19, 338–381.
Sharma, R., Singh, P., Loughry, W.J., Lockhart, J.M., Inman, W.B., et al., 2015. Zoonotic leprosy in the Southeastern United States. Emerging Infect. Dis. 21, 2127–2134.
Shigenobu, S., Watanabe, H., Hattori, M., Sakaki, Y., Ishikawa, H., 2000. Genome sequence of the endocellular bacterial symbiont of aphids *Buchnera* sp. APS. Nature 407, 81–86.
Silva, F.J., Latorre, A., Moya, A., 2001. Genome size reduction through multiple events of gene disintegration in *Buchnera* APS. Trends Genet. 17, 615–618.
Silva, M.T., Portaels, F., Pedrosa, J., 2009. Pathogenetic mechanisms of the intracellular parasite *Mycobacterium ulcerans* leading to Buruli ulcer. Lancet Infect. Dis. 9,

699–710.
Singh, P., Benjak, A., Schuenemann, V.J., Herbig, A., Avanzi, C., et al., 2015. Insight into the evolution and origin of leprosy bacilli from the genome sequence of *Mycobacterium lepromatosis*. Proc. Natl. Acad. Sci. U. S. A. 112, 4459–4464.
Singh, P., Cole, S., 2012. The genomics of leprosy. In: Nelson, K.E., Jones-Nelson, B. (Eds.), Genomics Applications for the Developing World. Springer, New York, pp. 39–49.
Singh, P., Cole, S.T., 2011. *Mycobacterium leprae:* genes, pseudogenes and genetic diversity. Future Microbiol. 6, 57–71.
Smith, C.S., Noordeen, S.K., Richardus, J.H., Sansarricq, H., Cole, S.T., et al., 2014. A strategy to halt leprosy transmission. Lancet Infect. Dis. 14, 96–98.
Smith, I., 2003. *Mycobacterium tuberculosis* pathogenesis and molecular determinants of virulence. Clin. Microbiol. Rev. 16, 463–496.
Smith, W.C., Aerts, A., 2014. Role of contact tracing and prevention strategies in the interruption of leprosy transmission. Lepr. Rev. 85, 2–17.
Spencer, J.S., Brennan, P.J., 2011. The role of *Mycobacterium leprae* phenolic glycolipid I (PGL-I) in serodiagnosis and in the pathogenesis of leprosy. Lepr. Rev. 82, 344–357.
Stephens, R.S., Kalman, S., Lammel, C., Fan, J., Marathe, R., et al., 1998. Genome sequence of an obligate intracellular pathogen of humans: chlamydia trachomatis. Science 282, 754–759.
Stinear, T.P., Seemann, T., Harrison, P.F., Jenkin, G.A., Davies, J.K., et al., 2008. Insights from the complete genome sequence of *Mycobacterium marinum* on the evolution of *Mycobacterium tuberculosis*. Genome Res. 18, 729–741.
Stinear, T.P., Seemann, T., Pidot, S., Frigui, W., Reysset, G., et al., 2007. Reductive evolution and niche adaptation inferred from the genome of *Mycobacterium ulcerans*, the causative agent of Buruli ulcer. Genome Res. 17, 192–200.
Storrs, E.E., 1971. The nine-banded armadillo: a model for leprosy and other biomedical research. Int. J. Leprosy Mycobacterial Dis. 39, 703–714.
Suzuki, K., Nakata, N., Bang, P.D., Ishii, N., Makino, M., 2006. High-level expression of pseudogenes in *Mycobacterium leprae*. FEMS Microbiol. Lett. 259, 208–214.
Taylor, G.M., Tucker, K., Butler, R., Pike, A.W., Lewis, J., et al., 2013. Detection and strain typing of ancient *Mycobacterium leprae* from a medieval leprosy hospital. PLoS One 8, e62406.
Toft, C., Andersson, S.G.E., 2010. Evolutionary microbial genomics: insights into bacterial host adaptation. Nat. Rev. Genet. 11, 465.
Truman, R., Fontes, A.B., De Miranda, A.B., Suffys, P., Gillis, T., 2004. Genotypic variation and stability of four variable-number tandem repeats and their suitability for discriminating strains of *Mycobacterium leprae*. J. Clin. Microbiol. 42, 2558–2565.
Truman, R.W., Singh, P., Sharma, R., Busso, P., Rougemont, J., et al., 2011. Probable zoonotic leprosy in the southern United States. N. Engl. J. Med. 364, 1626–1633.
Tyagi, J.S., Saini, D.K., 2004. Did the loss of two-component systems initiate pseudogene accumulation in *Mycobacterium leprae*? Microbiology 150, 4–7.
Uzoigwe, J.C., Khaitsa, M.L., Gibbs, P.S., 2007. Epidemiological evidence for *Mycobacterium avium subspecies paratuberculosis* as a cause of Crohn's disease. Epidemiol. Infect. 135, 1057–1068.
Vissa, V.D., Brennan, P.J., 2001. The genome of *Mycobacterium leprae*: a minimal mycobacterial gene set. Genome Biol. 2 REVIEWS1023.
Wernegreen, J.J., 2002. Genome evolution in bacterial endosymbionts of insects. Nat. Rev. Genet. 3, 850–861.
Wheat, W.H., Casali, A.L., Thomas, V., Spencer, J.S., Lahiri, R., et al., 2014. Long-term survival and virulence of *Mycobacterium leprae* in amoebal cysts. PLoS Negl. Trop. Dis. 8, e3405.
Wheeler, P.R., 2001. The microbial physiologist's guide to the leprosy genome. Lepr. Rev. 72, 399–407.
WHO, 2013. In: Organization, W.H. (Ed.), WHO Multidrug Therapy (MDT), Geneva, Switzerland.
Williams, D.L., Slayden, R.A., Amin, A., Martinez, A.N., Pittman, T.L., et al., 2009. Implications of high level pseudogene transcription in *Mycobacterium leprae*. BMC Genomics 10, 397.
Wixon, J., 2001. Featured organism: reductive evolution in bacteria: *Buchnera* sp., *rickettsia prowazekii* and *Mycobacterium leprae*. Comp. Funct. Genomics 2, 44–48.
Wyrick, P.B., 2000. Intracellular survival by Chlamydia. Cell. Microbiol. 2, 275–282.
Zomorodipour, A., Andersson, S.G., 1999. Obligate intracellular parasites: *Rickettsia prowazekii* and *Chlamydia trachomatis*. FEBS Lett. 452, 11–15.