The Genealogic Tree of Mycobacteria Reveals a Long-Standing Sympatric Life into Free-Living Protozoa

Otmane Lamrabet[1], Vicky Merhej[1], Pierre Pontarotti[2], Didier Raoult[1], Michel Drancourt[1]*

1 URMITE CNRS-IRD UMR 6236, IFR48, Méditerranée Infection, Aix-Marseille Université, Marseille, France, 2 Equipe Evolution Biologique et Modélisation UMR 6632, IRF48, Aix-Marseille Université/CNRS, Marseille, France

Abstract

Free-living protozoa allow horizontal gene transfer with and between the microorganisms that they host. They host mycobacteria for which the sources of transferred genes remain unknown. Using BLASTp, we searched within the genomes of 15 mycobacteria for homologous genes with 34 amoeba-resistant bacteria and the free-living protozoa *Dictyostelium discoideum*. Subsequent phylogenetic analysis of these sequences revealed that eight mycobacterial open-reading frames (ORFs) were probably acquired via horizontal transfer from beta- and gamma-Proteobacteria and from Firmicutes, but the transfer histories could not be reliably established in details. One further ORF encoding a pyridine nucleotide disulfide oxidoreductase (pyr-redox) placed non-tuberculous mycobacteria in a clade with *Legionella* spp., *Francisella* spp., *Coxiella burnetii*, the ciliate *Tetrahymena thermophila* and *D. discoideum* with a high reliability. Co-culturing *Mycobacterium avium* and *Legionella pneumophila* with the amoeba *Acanthamoeba polyphaga* demonstrated that these two bacteria could live together in amoebae for five days, indicating the biological relevance of intra-amoebal transfer of the pyr-redox gene. In conclusion, the results of this study support the hypothesis that protists can serve as a source and a place for gene transfer in mycobacteria.

Introduction

Massive sequencing revealed that bacterial genomes have undergone a mosaic evolution, combining variable proportions of vertically acquired DNA from previous generations and horizontally acquired DNA from other organisms present in their environment [1]. Therefore, the evolution of bacterial genomes cannot be represented by trees alone but rather must be represented by more complex structures such as rhizomes illustrating the various, multiple sources of DNA that have been combined in one particular bacterial species [2]. Therefore, to a certain extent, a bacterial genome sheds light on the particular environment in which that bacterium's ancestors used to live and on the amount of DNA exchange with neighbor organisms [3]. Accordingly, genome sequencing revealed that contrary to previous conjecture, current *Mycobacterium* organisms are the result, in part, of horizontal genetic transfer from unidentified Eukarya and from environmental alpha- and gamma-Proteobacteria and Actinobacteria, as demonstrated for *Mycobacterium tuberculosis* [4–7]. However, the places in which *Mycobacterium* ancestors came in contact with other organisms for these genetic transfer events remained unknown.

Recent studies have shown that free-living protozoa, amoebae in particular, are indeed places in which horizontal genetic transfer occurs [8]. Free-living amoebae host numerous amoeba-resistant bacteria [3,9–13], fungi [14], giant DNA viruses [15] and virophages [16], all of which live in sympatry in the free-living protozoa. Moreover, free-living protozoa are "melting pots" in which microorganisms exchange DNA including genes by horizontal gene transfer (HGT) [3,17–19], as illustrated for *Rickettsia bellii* [20], *Candidatus Amoebophilus* asiaticus [21] and the recently found transfer of a *Acanthamoeba polyphaga* Mimivirus protein to *Legionella pneumophila* [22]. DNA can also be transferred from the protozoa themselves to the microorganisms, as in the cases of the *A. polyphaga* Mimivirus [15,23], *Legionella drancourtii* [22,24] and *Chloroflexus aurantiacus* [25]. Genetic transfers can also occur in the reverse direction, from the microorganisms to free-living protozoa, as in the case of *Tetrahymena thermophila*, which acquired bacterial genes involved in the catabolism of complex carbohydrates, contributing largely to its capacity to colonize the rumen [26]. There have also been documented transfers from bacteria to animals [27].

Non-tuberculous mycobacteria share aquatic and terrestrial ecological niches with free-living protozoa including ciliates, flagellates and amoebae [19,28–30]. Co-culture experiments further showed that non-tuberculous mycobacteria could be phagocytosed by the ciliate *Tetrahymena pyriformis* [28], the social amoeba *Dictyostelium discoideum* and the free-living amoeba (FLA) *Acanthamoeba polyphaga* [19,31–33] and further reside in amoebal cysts, which act as a "Trojan horse" for such amoeba-resistant mycobacteria [29,33,34]. *M. tuberculosis* complex organisms can also be phagocytosed by amoebae [35–37], and it was recently observed that, except for *Mycobacterium canetti*, *M. tuberculosis* complex members can also reside within amoebal cysts [37].

We speculated that free-living protozoa may have been places in which gene transfers into mycobacteria occurred. We performed

extensive bioinformatics comparisons of available mycobacteria genomes with those of amoeba-resistant bacteria and free-living protozoa to test this hypothesis, and we used co-culture experiment to confirm its biological relevance.

Materials and Methods

Bacterial genome sequences and homologous gene determination

The protein complement of *M. tuberculosis* H37Rv (NC_000962), *M. tuberculosis* CDC1551 (NC_002755), *Mycobacterium bovis* (NC_002945), *Mycobacterium avium* subsp. *hominissuis* 104 (NC_008595), *M. avium* subsp. *paratuberculosis* K10 (NC_002944), *M. avium* subsp. *avium* (NZ_ACFI00000000), *Mycobacterium intracellulare* (NZ_ABIN00000000), *Mycobacterium abscessus* (NC_010397), *Mycobacterium smegmatis* mc^2 155 (NC_008596), *Mycobacterium marinum* (NC_010612), *Mycobacterium ulcerans* Agy99 (NC_008611), *Mycobacterium gilvum* PYR-GCK (NC_009338), *Mycobacterium sp.* JLS (NC_009077), *Mycobacterium vanbaalenii* PYR-1 (NC_008726) and *Mycobacterium leprae* TN (NC_002677) was downloaded from the National Center for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov) (Table 1).

Each mycobacterial open reading frame (ORF) was then compared with the complete genomes of *D. discoideum* (NC_007087-92) and 34 amoeba-resistant bacteria [38] (Table S1) using the BLASTp program. The 100 hit sequences exhibiting a significant alignment (E-value$<1.10^{-4}$) and a hit sequence with coverage $\geq 80\%$ and similarity $\geq 30\%$ were selected for further phylogenetic analyses. The conserved domains of selected ORFs were searched with InterProScan (http://www.ebi.ac.uk/Tools/InterProScan).

Phylogenetic analysis and molecular data

For each set of 100 hits, the amino acid sequences were aligned using MUSCLE algorithm [39]. The alignments produced were then manually refined in order to remove regions that contain gaps or are highly divergent with the BioEdit program v7.0.9 [40].

The corrected alignments were then used for maximum likelihood (ML) and Bayesian inference (BI). ML was constructed

using PHYML [41] in the PHYLIP package version 3.5c with 100 and 1,000 randomizations of input order. The substitution model was set to WAG and enabled the optimization options for tree topology, branch lengths, and rate parameters. To test the robustness of inferred topologies, posterior probabilities were determined by a Bayesian Markov chain Monte Carlo (MCMC) method implemented in the program MR BAYES V3.0 [42]. One million generations were run using the WAG matrix and model parameters (gamma shape and proportion invariant), and the trees were sampled every 100 generations. The posterior probability stabilized after 100,000 generations, so all parameter estimates before generation 100,000 were omitted. The tree with maximum posterior probability was assessed using a consensus of the final 100 000 trees. Bootstrap support of >75% and posterior probability of >90% were considered to identify supported nodes.

Substitution rates were calculated by dating the nodes in the 16S rRNA gene sequence-based phylogeny. Distances or numbers of substitutions per site separating pairs of species were estimated from the absolute numbers of differences between pairs of nucleotide sequences. We converted these data into measures of time divergence using the constant rate of 16S rRNA divergence of 0.01–0.02 per 50 million years found by Moran *et al.* [43]. All distance calculations were based on the same 1,440 sites, for which there were no missing data.

The species tree of mycobacteria was constructed based on the 16S rRNA gene sequences. The 16S rRNA sequences from the 15 studied *Mycobacterium* spp. were retrieved from NCBI database and aligned using MUSCLE. The phylogenetic relationships were inferred using the Neighbor-joining method.

Co-culture experiments

The *A. polyphaga* Linc-AP1 strain (a gift from T. J. Rowbotham, Public Health Laboratory, Leeds, United Kingdom) was grown at 28°C for 3 days in 150-cm^3 culture flasks (Corning, New York, USA) containing 30 ml of peptone-yeast extract-glucose (PYG) broth [44–46]. When the average amoeba concentration reached 5×10^5 cells/ml, amoebae were centrifuged at 500 g for 10 min, and the pellet was suspended twice in 30 ml of Page's modified Neff's amoeba saline (PAS) (solution A-NaCl 1.20 g; MgSO$_4$.7H$_2$0 0.04 g; Na$_2$HPO$_4$ 1.42 g; KH$_2$PO$_4$ 1.36 g/100 ml of glass distilled water; solution B-CaCl$_2$.2H$_2$O 0.04 g/100 ml of distilled water; amoeba saline, 10 ml of solution A+10 ml of solution B+980 ml distilled water) [44,46,47]. Liquid medium-cultured *M. avium* subsp. *avium* CIP104244^T [33] and *L. pneumophila* strain Lens [12] organisms were washed two times with sterile phosphate-buffered saline (PBS), and the pellet was suspended in PAS. This inoculum was vortexed to minimize mycobacterial clumping. Ten milliliters of the amoebal suspension in PAS ($\sim 10^5$ amoeba/ml) was inoculated with $\sim 10^6$ *L. pneumophila*/ml or $\sim 10^6$ *M. avium*/ml (MOI = 10) or co-infected with both bacteria. As controls, *A. polyphaga*, *L. pneumophila* and *M. avium* were cultured separately in PAS. After a 3-h incubation at 32°C, the coculture was washed three times with PAS to remove any remaining extracellular or adherent mycobacteria, and it was incubated in 10 ml PAS for 5 days at 32°C. At 0, 3 and 5 days of co-culture, *A. polyphaga* monolayers were lysed with 0.1% sodium dodecyl sulfate (SDS) (Sigma-Aldrich Logistic Gmbh, Lyon, France) for 30 min and passed through a 26-gauge needle to ensure complete lysis of the amoebae. The lysate (100 μl) was plated onto 7H10 agar for *M. avium* or Buffered Charcoal Yeast Extract (BCYE) agar plates for *L. pneumophila* and incubated for 5 to 15 days at 35 or 37°C to determine the number of colonies (CFU) of intracellular *M. avium* and *L. pneumophila*. All experiments were performed in triplicate.

Table 1. Workflow summarizing the steps followed in the identification of HGT genes in mycobacteria.

1. The proteomes of 15 *Mycobacterium* spp., *D. discoideum* and 34 amoeba-resistant bacteria (Table S1) were downloaded from the National Center for Biotechnology Information (NCBI).

2. Search for homologous genes of mycobacterial open reading frames (ORFs) in the genomes of *D. discoideum* and 34 amoeba-resistant bacteria using the BLASTp program from NCBI (E-value<1.10–4, similarity >30% and coverage >80%).

3. Search for the homologous sequences of the mycobacterial ORFs found in step 2 in the NR database using BLASTp (E-value<1.10–4, similarity >30% and coverage >80%).

4. Selection of ORFs from mycobacteria found in step 3 presenting significant homology with *D. discoideum* and 34 amoeba-resistant bacteria in the first 100 hits.

5. Infer phylogenetic relationships between the protein sequences found in steps 4 using MUSCLE for alignment and two construction methods (Maximum Likelihood within the PHYML program and M. Bayes).

6. Analysis of the trees generated in step 5, looking for possible HGT between *Mycobacterium* spp. and amoebae and/or amoeba-resistant bacteria.

More details can be found in the materials and methods.
doi:10.1371/journal.pone.0034754.t001

Statistical analyses

All statistical analyses mentioned in this study were performed using the chi2-square test with a significance level of $p = 0.05$.

Results

Identification of genes homologous to amoeba and amoeba-resisting bacteria in mycobacterial genomes

We searched for homologous sequences for the 65,812 ORFs of the 15 studied mycobacterial genomes in a database of free living protozoa and amoeba-resisting bacteria using a BLASTp. We found a total of 11,783 that have homologous sequences in the free living protozoa *D. discoideum* and/or amoeba-resisting bacteria (E-value<1.10–4, similarity >30% and coverage >80%). We found a total of 88 mycobacterial ORFs (0.13%) that present significant homology in the genome of the free-living protozoa *D. discoideum*. The number of ORFs with significant homology ranged from 4 genes in *M. leprae* to 29 genes in *M. smegmatis*. When comparing the 15 genomes of *Mycobacterium* spp. with the 34 available genomes of amoeba-resisting bacteria we could identify a total of 11,695 ORFs (17.8%) with significant homology in amoeba-resisting bacterial genomes. The number of mycobacterial ORFs with significant homology in the amoeba-resisting bacteria ranged from 365 for *M. leprae* to 1,208 for *M. smegmatis*. The closely related homologous genes were found in beta-Proteobacteria (30.5% ORFs), gamma-Proteobacteria (18.3% ORFs), Firmicutes (17.6% ORFs), Bacteroidetes (10.8% ORFs), delta-Proteobacteria (7.8% ORFs), Chlamydiae (6.7% ORFs) and alpha-Proteobacteria (8.3% ORFs) (Figure S1).

Phylogenetic analyses and horizontal transfer history

We searched for homologous sequences for the 11,783 ORFs in the NR database. We selected the only queries that contain free living protozoa *D. discoideum* and/or amoeba-resisting bacteria in the 100 first hits. This analysis yielded 151 sets of 100 homologous genes including sequences from free living protozoa *D. discoideum* and/or amoeba-resisting. We made 151 phylogenetic trees on the basis of these 151 gene sequences. Eight out of the 151 gene-trees showed *Mycobacterium* species in a clade with amoeba-resisting bacteria (Fig. S2, S3, S4, S5, S6, S7) and one gene (encoding for pyr-redox) showed *Mycobacterium* species in a clade with *D. discoideum* and amoeba-resisting bacteria (Fig. 1) (Table S2). Mycobacterial sequences clustered with gamma-Proteobacteria in 2/9 trees; with Archaea, gamma-Proteobacteria and Plancto-myces in 1/9 trees; with Bacteroidetes and gamma-Proteobacteria in 1/9 trees; with Firmicutes spp. in 2/9 trees; with beta-Proteobacteria in 2/9 trees; and with Eukarya in 1/9 trees (Fig. S2, S3, S4, S5, S6, S7).

The gene encoding for hypothetical hydrolase placed *M. marinum* and *M. ulcerans* in a clade with *Methanosarcina acetivorans*, *Desulfovibrio salexigens*, *Planctomyces limnophilus* and *Vibrio cholerae* (Fig. S2). The gene encoding for hypothetical protein MT3512 placed *M. tuberculosis* H37Rv in a clade with *Gramella forsetii* and *Francisella tularensis* (Fig. S3). The gene encoding for amidase placed *M. marinum* in a clade with *Legionella* spp. (Fig. S4). The gene encoding for Two ORFs encoding for Acetyl CoA hydrolase in *M. marinum*, *M. ulcerans* and transcriptional regulator in *M. smegmatis*, placed these mycobacteria in clade with *Burkholderia* spp. (Fig. S5). Two ORFs encoding for sulphate transporter in *tuberculosis*, *M. bovis* and betalactamase in *M. abscessus*, placed these mycobacteria in clade with *Bacillus* spp. (Fig. S6). Finally, the gene encoding for amino acid permease placed *M. smegmatis* in a clade with *Pseudomonas putida* (Fig. S7). Further phylogenetic analyses of an ORF encoding a pyridine nucleotide disulfide oxidoreductase (pyr-redox) placed

M. marinum, *M. ulcerans*, *M. avium*, *M. intracellulare*, *M. abscessus*, *Mycobacterium parascrofulaceum* and *M. smegmatis* in a clade with *Legionella* spp., *Francisella* spp., *Coxiella burnetii*, *T. thermophila* and *D. discoideum* with a high reliability (Fig. 1). The different construction methods showed that *Mycobacterium* spp. formed a highly supported group (bootstrap values, 94–95%) with gamma-Proteobacteria (*Legionella* spp., *C. burnetii* and *Francisella* spp.), *D. discoideum* and *T. thermophila*. In addition, we observed that *Legionella* spp. did not cluster with the other gamma-Proteobacteria but rather with *Mycobacterium* spp., *D. discoideum* (amoeba) and *T. thermophila* (ciliates) (Fig. 1). The phylogenetic construction using M. Bayes gave the same topology. The tree topology is the same when carrying out with 100 or 1,000 bootstrap replicas in what concerns the place of mycobacteria in a highly supported clade with amoeba and amoeba-resistant bacteria. Interestingly, the pyr-redox sequences matched with genes encoding for a monooxygenase with coverage of 60% and identity 25% in *Rhodococcus* and coverage of 58% and identity of 24% in *Nocardia*. These results suggest that the HGT event of pyr-redox concerns only the mycobacteria genus.

### Characteristics and functions of the horizontally transferred genes

Our findings showed that environmental mycobacteria and mycobacteria from *M. tuberculosis* complex are all affected by HGT. However, the source organisms differ between the 2 groups of mycobacteria: *M. tuberculosis* complex underwent HGT from Firmicutes, Bacteroidetes and gamma-Proteobacteria spp. while the environmental mycobacteria acquired their 7 ORFs from Firmicutes, beta- and gamma-Proteobacteria, Archaea and Eukarya (Table S2).

The nine transferred genes identified here account for 0.02–0.09% of the mycobacterial genomic content. From the nine HGT, four candidates encode for proteins involved in metabolism and five genes encode for proteins involved in information storage and processing (Table S2). Among the five genes encoding for information storage and processing, two genes encode for amidase proteins that hydrolyse the CO-NH$_2$ bond with production of NH$_3$, one gene encodes for a betalactamase implicated in the bacterial resistance to beta-lactam antibiotics, one gene encodes for one transcriptional regulator and one gene encodes for a hypothetical protein, characterized by the presence of a formyl_trans_N domain and belonging to the transferase family. Among the genes encoding for metabolic proteins, two genes are implicated in transporter of different substrates across the membrane including the sulfate transporter and the amino acid permease and two genes encode for the Acetyl-CoA hydrolase and a pyridine nucleotide disulfide oxidoreductase.

The gene length of these ORFs varies from 702 to 1,485 pb. These ORFs are widely distributed across the genomes of *Mycobacterium* spp. The detailed observation of the regions surrounding these HGT candidates, i.e. 10 genes upstream and downstream, revealed the presence of 4 transposases in 3 mycobacterial genomes *M. ulcerans*, *M. smegmatis* and *M. avium* (Table S3). The GC content of 5 transferred genes significantly differ from the GC content of the genome in *M. tuberculosis*, *M. bovis*, *M. smegmatis*, *M. ulcerans* and *M. marinum* ($p<0.05$) (Table S3). Only 2 out of 9 transfered genes present both a GC% significantly differing from that of the mycobacterial host genome and transposase gene in the close vicinity.

Co-culture experiments

We co-cultured the amoeba *A. polyphaga* with both *L. pneumophila* and *M. avium*, and we observed that *L. pneumophila and M. avium*
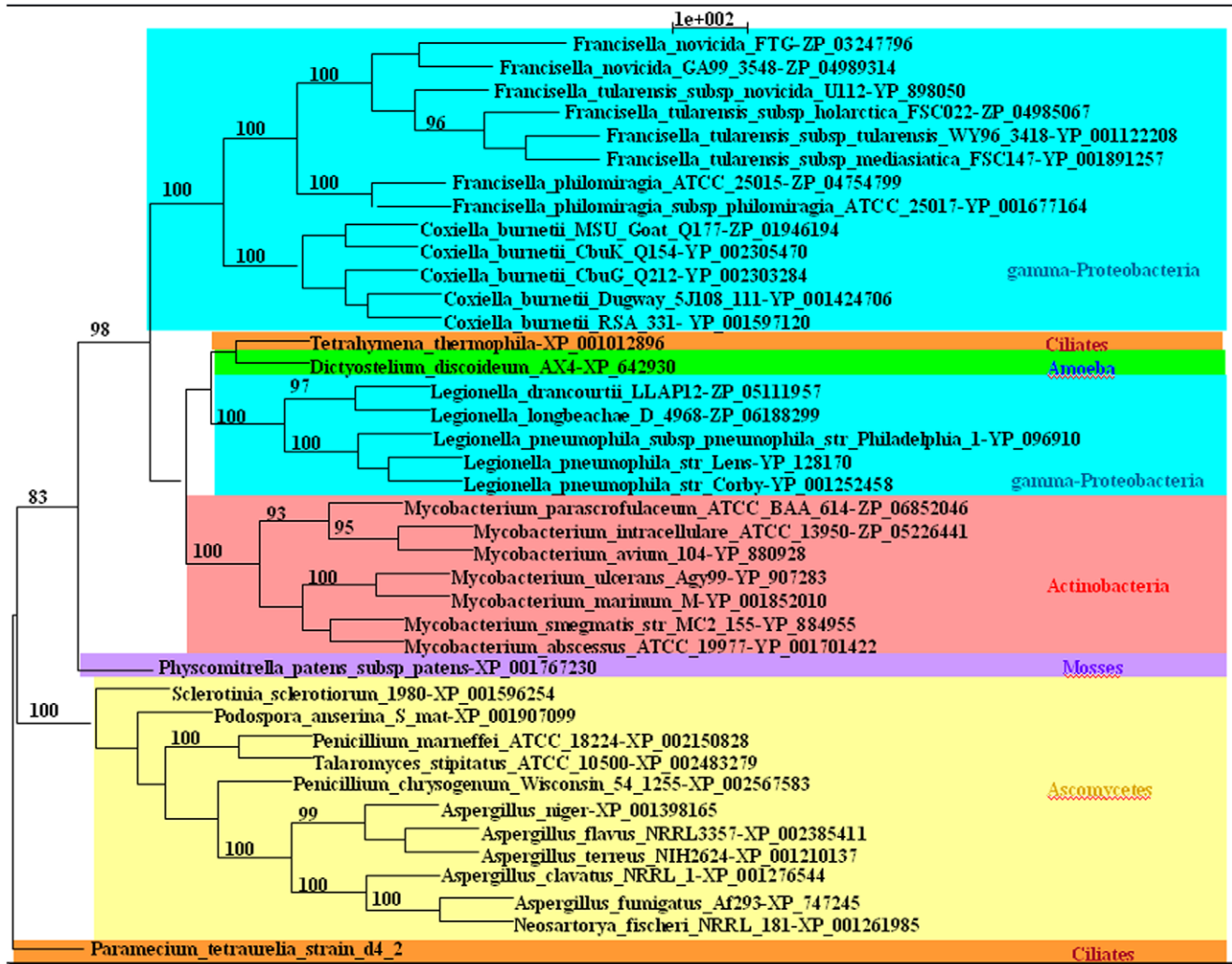
Figure 1. Phylogeny as inferred from the pyr-redox gene. The phylogenetic tree was obtained using maximum likelihood with the amino acid dataset. Numbers at the nodes represent bootstrap percentages. Only high bootstraps (>75) are indicated.
doi:10.1371/journal.pone.0034754.g001

could indeed live together in amoebae for at least five days. We first observed that the number of *A. polyphaga* trophozoites infected with *M. avium*, *L. pneumophila* or both strains increased significantly (p≤0.05) over the course of the experiments. The quantification of the colony forming units (CFU) of *M. avium* and *L. pneumophila* when co-cultured with amoebae yielded $1.66 \times 10^6 \pm 1.68 \times 10^5$ CFU/mL at day 0, $1.52 \times 10^9 \pm 2.5 \times 10^8$ CFU/mL at day 3 and $1.65 \times 10^9 \pm 2.76 \times 10^8$ CFU/mL at day 5 for *L. pneumophila* and $1.99 \times 10^5 \pm 1.63 \times 10^4$ CFU/mL at day 0, $5.2 \times 10^6 \pm 7.07 \times 10^5$ CFU/mL at day 3 and $2.05 \times 10^7 \pm 1.48 \times 10^7$ CFU/mL at day 5 for *M. avium* (Fig. 2).

Discussion

Our phylogenetic analyses identified eight mycobacterial genes that have close phylogenetic relationships with bacteria other than *Actinobacteria* spp. Given that most of these species are amoeba-resistant, the phylogenies were highly suggestive of possible HGT within amoeba. Nonetheless, the lack of information about the direction of the transfer hampered the elucidation of the HGT history. Furthermore, we found one gene encoding for pyr-redox

that gave insight into the history of the HGT events in relation with the mycobacterial lifestyle within free-living protozoa.

It has been previously shown that the *Mycobacterium* spp. and gamma-Proteobacteria studied herein are able to live alone in amoebae [9,29,38,48] as well as in ciliates [12] or together in *Acanthamoeba castellanii* [49]. We therefore co-cultured the amoeba *A. polyphaga* with both *L. pneumophila* and *M. avium*, and we observed that *L. pneumophila* and *M. avium* could indeed live together in amoebae for at least five days. Thus, our data expand the previous demonstration of intra-amoebal surviving of both *Legionella* and mycobacteria in amoeba *A. castellanii* to another species of amoeba, *A. polyphaga*. This sympatric lifestyle, *i.e.*, various microorganisms living together, provides opportunities for DNA exchange and gene transfer within amoebae [3,15,23]. This hypothesis agrees with the current model for the evolution of mycobacteria, which postulates that the ancestor of mycobacteria was an environmental organism living in an aquatic habitat [50]. Recent genome analysis of the environmental *Mycobacterium indicus pranii*, a member of the *M. avium* complex, further supports this hypothesis in which the most recent common ancestor of mycobacteria gave rise to waterborne *M. marinum* and *M. ulcerans* on one branch, the *M.*
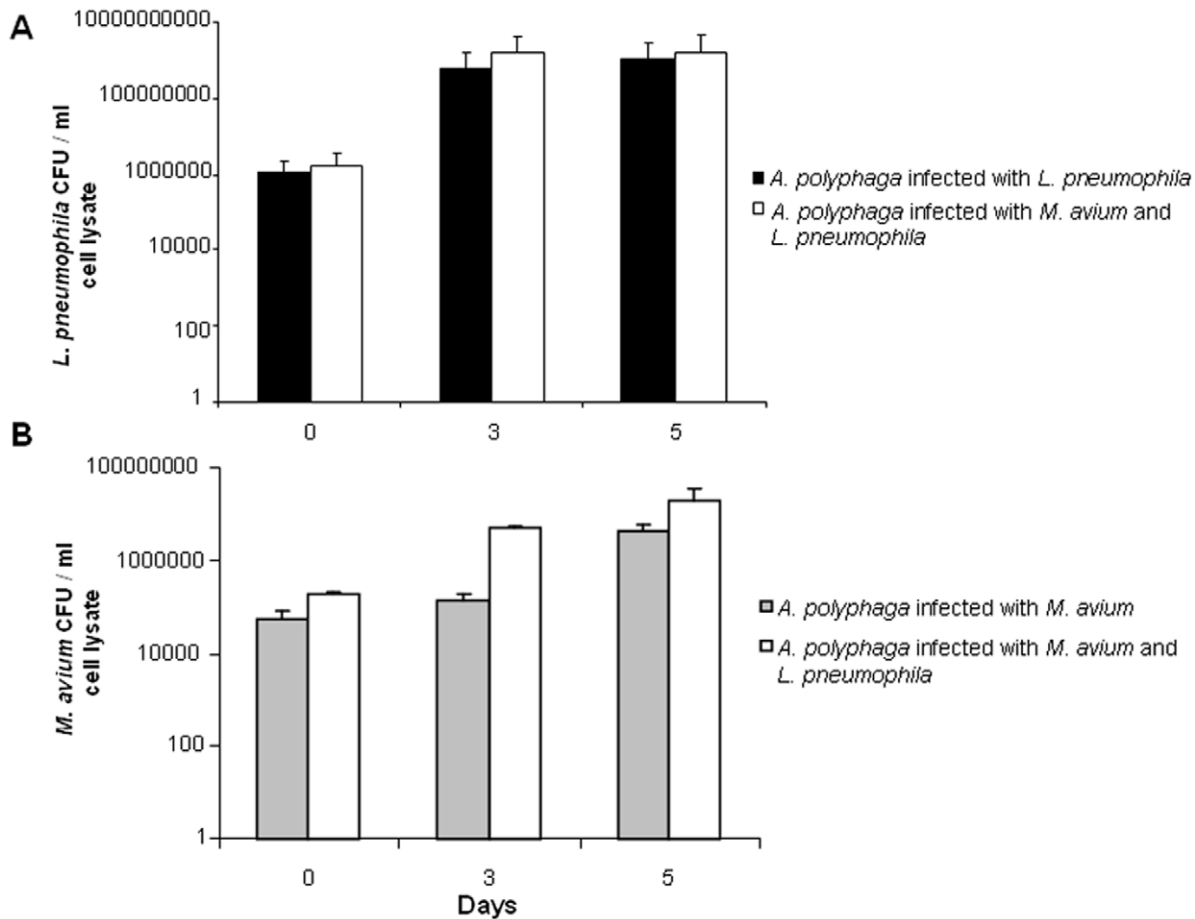
Figure 2. *A. polyphaga* co-cultured with *M. avium* and *L. pneumophila* for 5 days. A) The number of *L. pneumophila* colonies was obtained after plating the lysate of *L. pneumophila* and *A. polyphaga* culture or *L. pneumophila*, *M. avium* and *A. polyphaga* co-culture in BCYE agar medium. B) The number of *M. avium* colonies was obtained after plating the lysate of *M. avium* and *A. polyphaga* culture or *M. avium*, *L. pneumophila* and A. polyphaga co-culture in 7H10 agar medium. Data points are the means of triplicate wells, and the standard errors are represented by error bars.
doi:10.1371/journal.pone.0034754.g002

avium complex on a second branch and the *M. tuberculosis* complex on a third branch [50].

Life in free-living amoebae has been demonstrated to protect amoeba-resistant organisms, such as environmental mycobacteria and *Legionella*, against adverse environmental conditions [3,9,38,48], to increase their resistance to some antibiotics [48,51,52] and to enhance their virulence [36,48,51]. The pyr-redox gene studied herein is present in *Mycobacterium* spp. that have been shown to survive in amoebal cysts. The significant association between the presence of pyr-redox and the survival in amoebal cysts (chi2-square, p = 0.002) highlights the possible role of this protein in the intraamoebal lifestyle and life inside macrophages [19,48]. During phagocytosis, amoebae and macrophages produce the oxygen metabolites nitric oxide and hydrogen peroxide, which generate a toxic environment that can kill phagocytized bacteria [53–55]. *Mycobacterium* spp. deploy multiple strategies to resist to this oxidative stress, including the expression of catalase/peroxidase [56] and superoxide dismutase [57], a thiol-based detoxification response [58] and the pyr-redox response [59,60]. Pyr-redox complements the anti-oxidative arsenal of mycobacteria during their survival in amoebae and macrophages.

Phylogenetic trees have indicated that phylogenetically distant organisms have acquired the pyr-redox via HGT, but the source of

this transfer is ambiguous (Fig. 3). According to one scenario, pyr-redox was acquired by *Mycobacterium* spp., *D. discoideum* and *T. thermophila* from gamma-Proteobacteria, specifically from *Legionella* spp. (Fig. 3). This result agrees with previously published observations that genes acquired by HGT in Actinobacteria mostly originated from beta- and gamma-Proteobacteria [6]. The genes acquired by HGT in fungi [61] and HGT in ciliates such as *T. thermophila* [26,62] mostly originated from bacteria. *D. discoideum* may have transferred the pyr-redox gene to *Mycobacterium* spp. or may have been the place for transfer. According to a second scenario, there were multiple gene losses in *Legionella* spp., *D. discoideum* and *T. thermophila* and recent acquisitions from *Mycobacterium* spp. (Fig. 3). Whereas several studies have demonstrated HGT between mycobacteria [63], HGT originating from *Mycobacterium* spp. has never been reported. Considering the paraphyly of *Tetrahymena* spp. and *Dictyostelium* spp., we have to postulate a minimum of two independent HGT events in both scenari: the first one event from *Legionella* spp. or mycobacteria (ancestors) into amoebae and amoeba-resistant bacteria and the second event from *Legionella* spp. or mycobacteria (ancestors) into *T. thermophila*. Alternatively, the scenarii might have required a single ancient HGT event in a certain common ancestor of eukaryotes and subsequent multiple losses from organisms except *T. thermophila* and *D. discoideum*.
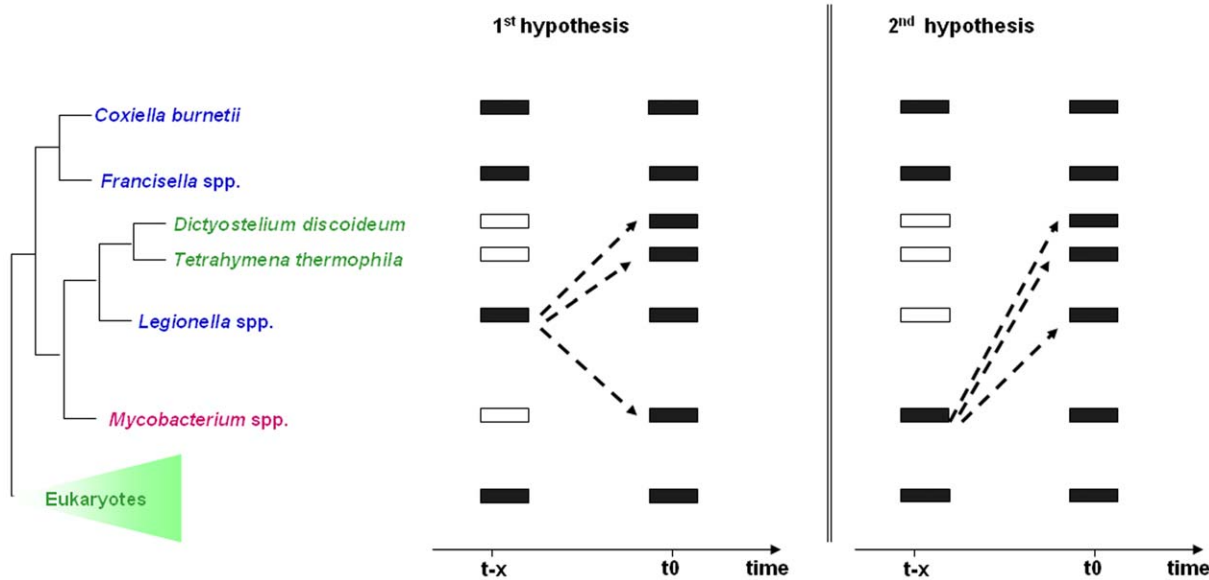
Figure 3. Schematic representation of two alternative explanations of the evolutionary history of the pyridine nucleotide disulfide oxidoreductase gene. Dashed arrows indicate the possible lateral transfer of the gene. First hypothesis: The gene encoding pyr-redox exists in the gamma-Proteobacteria species and is absent in *D. discoideum* and *T. thermophila*. This gene was acquired from *Legionella* spp. by *Mycobacterium* spp., *D. discoideum* and *T. thermophila*. Second hypothesis: The gene encoding pyr-redox was lost from *Legionella* spp., *D. discoideum* and *T. thermophila* and was acquired later from *Mycobacterium* spp. T0 corresponds to the time of observation and T-x to the time when the event occurred.
doi:10.1371/journal.pone.0034754.g003

Both scenarii involve HGT from a bacterium to a eukaryote (*Tetrahymena* spp. and *Dictyostelium*) following the loss of the eukaryotic pyr-redox gene from these genomes. Indeed, the gene encoding pyr-redox might have become disused or lost its functional importance, allowing the loss of the gene. The second scenario requires additional losses from *Legionella* genomes that occurred before the HGT and is less parsimonious than the first one. Thus, the scenario that postulates HGT from *Legionella* or *Dictyostelium* into mycobacteria seems to be more likely. The molecular clock showed that mycobacteria and amoeba-resistant gamma-Proteobacteria exchanged the pyr-redox gene between 33 and 267 Million Years Ago, after the separation of gamma-Proteobacteria spp. and before the radiation of *Legionella* spp. (Fig. 4). This range provides an estimated time-frame for the intracellular association of mycobacteria within amoebae and subsequent horizontal gene transfers. The pyr-redox gene has been found in 5/21 annotated *Mycobacterium* genomes, with the notable exception of the *M. tuberculosis* complex members. The genome of *M. tuberculosis* has been shown to exhibit the highest ratio of eukaryotic-prokaryotic gene fusion [64], but this observation was made before protist genomes were sequenced. The most parsimonious scenario suggests that the pyr-redox gene was acquired by an ancestor of all mycobacterial species, followed by a loss by the *M. tuberculosis* complex members *M. leprae*, *M. vanbaalenii*, *M. gilvum* and *Mycobacterium* sp. JLS.

In conclusion, our phylogenetic analyses found 8 ORFs most likely acquired through HGT in *Mycobacterium* spp. and one further pyr-redox ORF that elucidates the history of the HGT events in relation with the mycobacterial lifestyle within free-living protozoa. The experimental data reported herein support these genome-based analyses. Amoebae or other phagocytic organisms may have been the places in which the gene exchanges occurred. Thus, *Mycobacterium* spp. have followed an evolutionary strategy similar to that of other intracellular bacteria: they interfere with host cellular processes through the expression of genes horizontally acquired from the host. HGT may have contributed to the adaptation of mycobacteria to an intracellular lifestyle.

Supporting Information

Figure S1 Putative sources of homologous ORFs from bacteria other than Actinobacteria in the mycobacterial genome.
(PDF)

Figure S2 Extended phylogenetic tree showing representatives of the conserved hypothetical hydrolase. Phylogenetic trees showing HGT events as generated by the Maximum Likelihood method. Numbers at nodes are bootstrap percentages based on 100 resamplings. The scale bar represents the number of estimated changes per position for a unit of branch length. *Mycobacterium* spp. are colored in red.
(PDF)

Figure S3 Extended phylogenetic tree showing representatives of hypothetical protein MT3512. Phylogenetic trees showing HGT events as generated by the Maximum Likelihood method. Numbers at nodes are bootstrap percentages based on 100 resamplings. The scale bar represents the number of estimated changes per position for a unit of branch length. *Mycobacterium* spp. are colored in red.
(PDF)

Figure S4 Extended phylogenetic tree showing representatives of amidase. Phylogenetic trees showing HGT events as generated by the Maximum Likelihood method. Numbers at nodes are bootstrap percentages based on 100 resamplings. The scale bar represents the number of estimated changes per position for a unit of branch length. *Mycobacterium* spp. are colored in red.
(PDF)
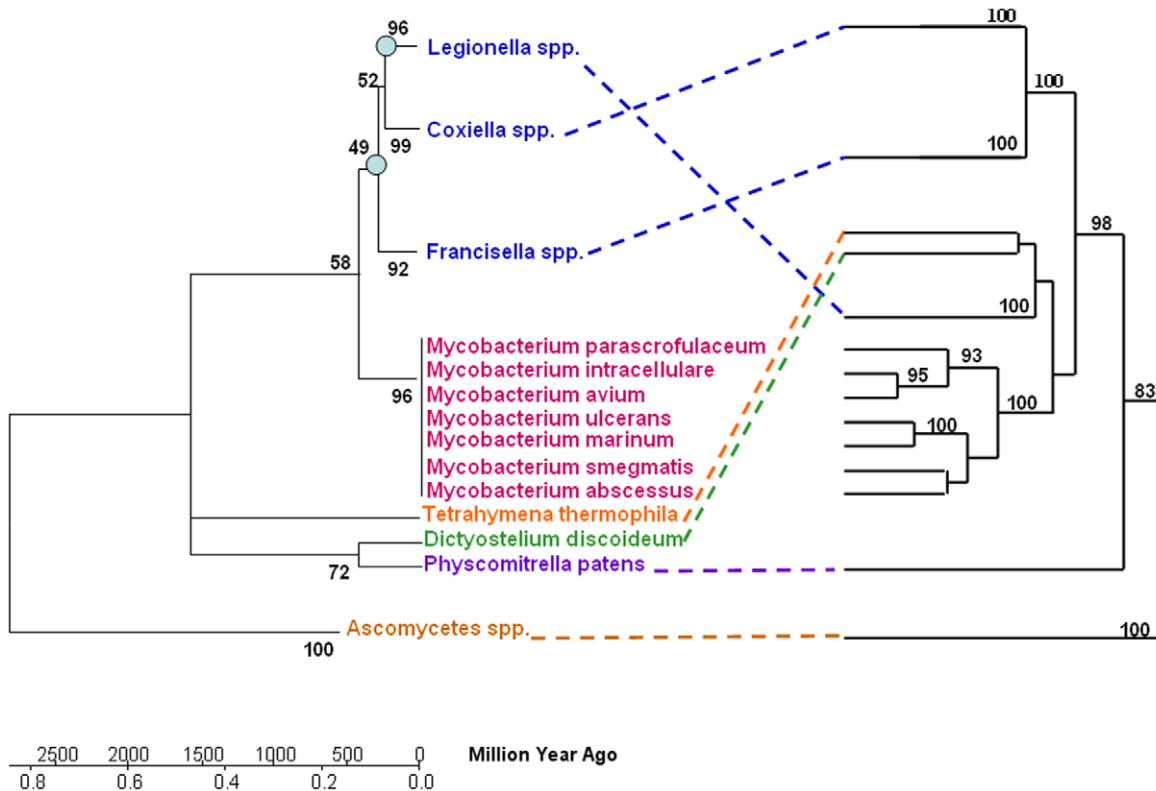
Figure 4. Schematic representation of HGT and molecular clock. The left tree shows the relationships and approximate dates of divergence of the species. The phylogeny was reconstructed based on 16S and 18S rDNA sequences, and the expected divergence was calculated as a function of time, 1–2% per 50 MYA. The right tree shows the relationships among species based on pyr redox protein. The circles on the nodes indicate the radiation of *Legionella* spp. (33 MYA) and the separation of gamma-Proteobacteria spp. (267 MYA).
doi:10.1371/journal.pone.0034754.g004

Figure S5 Extended phylogenetic tree showing representatives of A) acetyl CoA hydrolase and B) transcriptional regulator. Phylogenetic trees showing HGT events as generated by the Maximum Likelihood method. Numbers at nodes are bootstrap percentages based on 100 resamplings. The scale bar represents the number of estimated changes per position for a unit of branch length. *Mycobacterium* spp. are colored in red.
(PDF)

Figure S6 Extended phylogenetic tree showing representatives of A) sulfate transporter and B) beta-lactamase. Phylogenetic trees showing HGT events as generated by the Maximum Likelihood method. Numbers at nodes are bootstrap percentages based on 100 resamplings. The scale bar represents the number of estimated changes per position for a unit of branch length. *Mycobacterium* spp. are colored in red.
(PDF)

Figure S7 Extended phylogenetic tree showing representatives of amino acid permease. Phylogenetic trees showing HGT events as generated by the Maximum Likelihood method. Numbers at nodes are bootstrap percentages based on

100 resamplings. The scale bar represents the number of estimated changes per position for a unit of branch length. Mycobacterium spp. are colored in red.
(PDF)

Table S1 Genome sequence of amoeba-resistant bacteria utilized in this study.
(DOC)

Table S2 Genes probably transferred by horizontal gene transfer.
(DOC)

Table S3 Description of the nine probably transfered genes.
(DOC)

Author Contributions

Conceived and designed the experiments: DR MD. Performed the experiments: OL VM PP. Analyzed the data: OL VM PP MD DR. Contributed reagents/materials/analysis tools: DR PP. Wrote the paper: OL PP MD.

References

1. Audic S, Robert C, Campagna B, Parinello H, Claverie JM, et al. (2007) Genome analysis of *Minibacterium massiliensis* highlights the convergent evolution of water-living bacteria. PLoS Genet 3: e138.
2. Raoult D (2010) The post-Darwinist rhizome of life. Lancet 375: 104–105.
3. Moliner C, Fournier PE, Raoult D (2010) Genome analysis of microorganisms living in amoebae reveals a melting pot of evolution. FEMS Microbiol Rev 34: 281–294.
4. Kinsella RJ, Fitzpatrick DA, Creevey CJ, McInerney JO (2003) Fatty acid biosynthesis in *Mycobacterium tuberculosis*: lateral gene transfer, adaptive evolution, and gene duplication. Proc Natl Acad Sci U S A 100: 10320–10325.
5. Marri PR, Bannantine JP, Paustian ML, Golding GB (2006) Lateral gene transfer in *Mycobacterium avium* subspecies *paratuberculosis*. Can J Microbiol 52: 560–569.

6. Becq J, Gutierrez MC, Rosas-Magallanes V, Rauzier J, Gicquel B, et al. (2007) Contribution of horizontally acquired genomic islands to the evolution of the tubercle bacilli. Mol Biol Evol 24: 1861–1871.
7. Veyrier F, Pletzer D, Turenne C, Behr MA (2009) Phylogenetic detection of horizontal gene transfer during the step-wise genesis of *Mycobacterium tuberculosis*. BMC Evol Biol 9: 196.
8. Saisongkorh W, Robert C, La Scola B, Raoult D, Rolain JM (2010) Evidence of transfer by conjugation of type IV secretion system genes between *Bartonella* species and *Rhizobium radiobacter* in amoeba. PLoS One 5: e12666.
9. Greub G, Raoult D (2004) Microorganisms resistant to free-living amoebae. Clin Microbiol Rev 17: 413–433.
10. Brandl MT, Rosenthal BM, Haxo AF Berk SG (2005) Enhanced survival of *Salmonella enterica* in vesicles released by a soilborne *Tetrahymena* species. Appl Environ Microbiol 71: 1562–1569.
11. Snelling WJ, McKenna JP, Hack CJ, Moore JE, Dooley JS (2006) An examination of the diversity of a novel *Campylobacter* reservoir. Arch Microbiol 186: 31–40.
12. Pagnier I, Raoult D, La Scola B (2008) Isolation and identification of amoeba-resisting bacteria from water in human environment by using an *Acanthamoeba polyphaga* co-culture procedure. Environ Microbiol 10: 1135–1144.
13. Pagnier I, Merchat M, La Scola B (2009) Potentially pathogenic amoeba-associated microorganisms in cooling towers and their control. Future Microbiol 4: 615–629.
14. Chrisman CJ, Alvarez M, Casadevall A (2010) Phagocytosis of *Cryptococcus neoformans* by, and nonlytic exocytosis from, *Acanthamoeba castellanii*. Appl Environ Microbiol 76: 6056–6062.
15. Raoult D, Boyer M (2010) Amoebae as genitors and reservoirs of giant viruses. Intervirology 53: 321–329.
16. La Scola B, Desnues C, Pagnier I, Robert C, Barrassi L, et al. (2008) The virophage as a unique parasite of the giant mimivirus. Nature 455: 100–104.
17. Hotopp JC, Clark ME, Oliveira DC, Foster JM, Fischer P, et al. (2007) Widespread lateral gene transfer from intracellular bacteria to multicellular eukaryotes. Science 317: 1753–1756.
18. Boyer M, Yutin N, Pagnier I, Barrassi L, Fournous I, et al. (2009) Giant Marseillevirus highlights the role of amoebae as a melting pot in emergence of chimeric microorganisms. Proc Natl Acad Sci U S A 106: 21848–21853.
19. Thomas V, Greub G (2010) Amoeba/amoebal symbiont genetic transfers: lessons from giant virus neighbours. Intervirology 53: 254–267.
20. Ogata H, La Scola B, Audic S, Renesto P, Blanc G, et al. (2006) Genome sequence of *Rickettsia bellii* illuminates the role of amoebae in gene exchanges between intracellular pathogens. PLoS Genet 2: e76.
21. Schmitz-Esser S, Tischler P, Arnold R, Montanaro J, Wanger M, et al. (2010) The genome of the amoeba symbiont "Candidatus Amoebophilus asiaticus" reveals common mechanisms for host cell interaction among amoeba-associated bacteria. J Bacteriol 192: 1045–1057.
22. Lurie-Weinberger MN, Gomez-Valero L, Merault N, Glockner G, Buchrieser C, et al. (2010) The origins of eukaryotic-like proteins in *Legionella pneumophila*. Int J Med Microbiol 300: 470–481.
23. Moreira D, Brochier-Armanet C (2008) Giant viruses, giant chimeras: the multiple evolutionary histories of Mimivirus genes. BMC Evol Biol 8: 12.
24. Moliner C, Raoult D, Fournier PE (2009) Evidence that the intra-amoebal *Legionella drancourtii* acquired a sterol reductase gene from eukaryotes. BMC Res Notes 2: 51.
25. Da Lage JL, Feller G, Janecek S (2004) Horizontal gene transfer from Eukarya to bacteria and domain shuffling: the alpha-amylase model. Cell Mol Life Sci 61: 97–109.
26. Ricard G, McEwan NR, Dutilh BE, Jouany JP, Macheboeuf D, et al. (2006) Horizontal gene transfer from Bacteria to rumen Ciliates indicates adaptation to their anaerobic, carbohydrates-rich environment. BMC Genomics 7: 22.
27. Hotopp JCD (2011) Horizontal gene transfer between bacteria and animals. Trends in Genetics 27: 157–163.
28. Stahl C, Kunetzko S, Kaps I, Seeber S, Engelhardt H, et al. (2001) MspA provides the main hydrophilic pathway through the cell wall of *Mycobacterium smegmatis*. Mol Microbiol 40: 451–464.
29. Adekambi T, Ben Salah S, Khlif M, Raoult D, Drancourt M (2006) Survival of environmental mycobacteria in Acanthamoeba polyphaga. Appl Environ Microbiol 72: 5974–5981.
30. Vaerewijck MJ, Sabbe K, Van Hende J, Bare J, Houf K (2010) Sampling strategy, occurrence and diversity of free-living protozoa in domestic refrigerators. J Appl Microbiol 109: 1566–1578.
31. Greub G, La Scola B, Raoult D (2004) Amoebae-resisting bacteria isolated from human nasal swabs by amoebal coculture. Emerg Infect Dis 10: 470–477.
32. Steinert M, Heuner K (2005) *Dictyostelium* as host model for pathogenesis. Cell Microbiol 7: 307–314.
33. Ben Salah I, Drancourt M (2010) Surviving within the amoebal exocyst: the *Mycobacterium avium* complex paradigm. BMC Microbiology 10: 99.
34. Barker J, Brown MR (1994) Trojan horses of the microbial world: protozoa and the survival of bacterial pathogens in the environment. Microbiology 140: 1253–1259.
35. Taylor SJ, Ahonen LJ, de Leij FA, Dale JW (2003) Infection of *Acanthamoeba castellanii* with *Mycobacterium bovis* and *M. bovis* BCG and survival of *M. bovis* within the amoebae. Appl Environ Microbiol 69: 4316–4319.
36. Hagedorn M, Rohde KH, Russell DG, Soldati T (2009) Infection by tubercular mycobacteria is spread by nonlytic ejection from their amoeba hosts. Science 323: 1729–1733.
37. Mba Medie F, Ben Salah I, Henrissat B, Raoult R, Drancourt M (2011) *Mycobacterium tuberculosis* complex mycobacteria as amoeba-resistant organisms. PLoS One 6: e20499.
38. Thomas V, McDonnell G, Denyer SP, Maillard JY (2009) Free-living amoebae and their intracellular pathogenic microorganisms: risks for water quality. FEMS Microbiol Rev 34: 231–259.
39. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, et al. (2007) Clustal W and Clustal X version 2.0. Bioinformatics 53: 2947–8.
40. Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symposium Series 41: 95–98.
41. Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol 52: 696–704.
42. Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19: 1572–1574.
43. Moran NA, Munson MA, Baumann P, Ishikawa H (1993) A Molecular Clock in Endosymbiotic Bacteria is Calibrated Using the Insect Hosts. Proc R Soc Lond B 253: 167–171.
44. La Scola B, Mezi L, Weiller PJ, Raoult D (2001) Isolation of *Legionella anisa* using an amoebic coculture procedure. J Clin Microbiol 39: 365–366.
45. Greub G, Raoult D (2002) Crescent bodies of *Parachlamydia acanthamoeba* and its life cycle within *Acanthamoeba polyphaga*: an electron micrograph study. Appl Environ Microbiol 68: 3076–3084.
46. Greub G, La Scola B, Raoult D (2004) Amoebae-resisting bacteria isolated from human nasal swabs by amoebal coculture. Emerg Infect Dis 10: 470–477.
47. Rowbotham TJ (1983) Isolation of *Legionella pneumophila* from clinical specimens via amoebae, and the interaction of those and other isolates with amoebae. J Clin Pathol 36: 978–986.
48. Ben Salah I, Ghigo E, Drancourt M (2009) Free-living amoeba, a training field for macrophage resistance of mycobacteria. Clin Microbiol Infect 15: 894–905.
49. Boyer M, Azza S, Barrassi L, Klose T, Campocasso A, et al. (2001) Mimivirus shows dramatic genome reduction after intraamoebal culture. Proc Natl Acad Sci U S A 108: 10296–301.
50. Ahmed N, Saini V, Raghuvanshi S, Khurana JP, Tyagi AK, et al. (2007) Molecular analysis of a leprosy immunotherapeutic bacillus provides insights into *Mycobacterium* evolution. PLoS One 2: e968.
51. Miltner EC, Bermudez LE (2000) *Mycobacterium avium* grown in *Acanthamoeba castellanii* is protected from the effects of antimicrobials. Antimicrob Agents Chemother 44: 1990–1994.
52. Thomas V, Loret JF, Jousset M, Greub G (2008) Biodiversity of amoebae and amoebae-resisting bacteria in a drinking water treatment plant. Environ Microbiol 10: 2728–2745.
53. Chan J, Xing Y, Magliozzo RS, Bloom BR (1992) Killing of virulent *Mycobacterium tuberculosis* by reactive nitrogen intermediates produced by activated murine macrophages. J Exp Med 175: 1111–1122.
54. Fabrino DL, Bleck CK, Anes E, Hasilik A, Niederweis M, et al. (2009) Porins facilitate nitric oxide-mediated killing of mycobacteria. Microbes Infect 11: 868–875.
55. Ghigo E, Pretat L, Desnues B, Capo C, Raoult D, et al. (2009) Intracellular life of *Coxiella burnetii* in macrophages. Ann N Y Acad Sci 1166: 55–66.
56. Faguy DM, Doolittle WF (2000) Horizontal transfer of catalase-peroxidase genes between archaea and pathogenic bacteria. Trends Genet 16: 196–197.
57. Zhang Y, Lathigra R, Garbe T, Catty D, Young D (1991) Genetic analysis of superoxide dismutase, the 23 kilodalton antigen of *Mycobacterium tuberculosis*. Mol Microbiol 5: 381–391.
58. Ung KS, Av-Gay Y (2006) Mycothiol-dependent mycobacterial response to oxidative stress. FEBS Lett 580: 2712–2716.
59. Zahrt TC, Deretic V (2002) Reactive nitrogen and oxygen intermediates and bacterial defenses: unusual adaptations in *Mycobacterium tuberculosis*. Antioxid Redox Signal 4: 141–159.
60. Venketaraman V, Dayaram YK, Talaue MT, Connell ND (2005) Glutathione and nitrosoglutathione in macrophage defense against *Mycobacterium tuberculosis*. Infect Immun 73: 1886–1889.
61. Garcia-Vallve S, Romeu A, Palau J (2000) Horizontal gene transfer of glycosyl hydrolases of the rumen fungi. Mol Biol Evol 17: 352–361.
62. Devillard E, Newbold CJ, Scott KP, Forano E, Wallace RJ, et al. (1999) A xylanase produced by the rumen anaerobic protozoan Polyplastron multi-vesiculatum shows close sequence similarity to family 11 xylanases from gram-positive bacteria. FEMS Microbiol Lett 181: 145–152.
63. Coros A, DeConno E, Derbyshire KM (2008) IS6110, a *Mycobacterium tuberculosis* complex-specific insertion sequence, is also present in the genome of *Mycobacterium smegmatis*, suggestive of lateral gene transfer among mycobacterial species. J Bacteriol 190: 3408–3410.
64. Gamieldien J, Ptitsyn A, Hide W (2002) Eukaryotic genes in *Mycobacterium tuberculosis* could have a role in pathogenesis and immunomodulation. Trends Genet 18: 5–8.