Title: Functional genomics of intracellular bacteria.

Authors: de Barsy M, Greub G

Journal: Briefings in functional genomics

Year: 2013 Jul

Volume: 12

Issue: 4

Pages: 341-53

DOI: 10.1093/bfgp/elt012

Functional genomics of intracellular bacteria

Marie de Barsy[1] and Gilbert Greub[1]

[1] Institute of Microbiology, University Hospital Center and University of Lausanne, Lausanne, Switzerland

Corresponding author:
Prof. Gilbert Greub, MD PhD
Center for Research on Intracellular Bacteria (CRIB), Institute of Microbiology
University Hospital Center and University of Lausanne,
Bugnon 48, 1011 Lausanne, SWITZERLAND
Phone : +41 21 314 49 79, Fax : +41 21 341 40 60
e-mail : gilbert.greub@chuv.ch

Abstract

During the genomic era, a large amount of whole genome sequences accumulated, which identified many hypothetical proteins of unknown function. Rapidly, functional genomics, which is the research domain that assign a function to a given gene product, has thus been developed. Functional genomics of intracellular pathogenic bacteria exhibit specific peculiarities due to the fastidious growth of most of these intracellular micro-organisms, due to the close interaction with the host cell, due to the risk of contamination of experiments with host cell proteins and, for some strict intracellular bacteria such as *Chlamydia*, due to the absence of simple genetic system to manipulate the bacterial genome. In order to identify virulence factors of intracellular pathogenic bacteria, functional genomics often rely on bioinformatic analyses compared to model organisms such *as E. coli* and *B. subtilis.* The use of heterologous expression is another common approach. Given the intracellular lifestyle and the many effectors that are used by the intracellular bacteria to corrupt host cell functions, functional genomics is also often targeting the identification of new effectors such as those of the T4SS of *Brucella* and *Legionella*.

1. Introduction

In the late 1990's, the development of automated-DNA sequencing revolutionized microbiology through the availability of complete genomes of bacteria. The first bacterial genome sequence was reported for the bacterial pathogen *Haemophilus influenzae* in 1995 [1]. The availability of a complete genome sequence rapidly appeared to be insufficient to understand the complexity of the bacterial world. The genome sequence data is not an end by itself but rather the starting point to raise testable functional hypotheses. Thus, the transition to the 21st century gave rise to a large increase of research in functional genomics, often referred as the post-genomic era (Figure 1). Functional genomics consist to assign a function to a protein encoded by a given gene. A key feature of functional genomics is the "genome-wide" approach that requires adapting the experimental design to large-scale research. Basically, functional genomics is based onto two main approaches: the sequence-based and the experiment-based function assignment. Sequence-based functional genomics relies mainly on similarity of sequence at nucleotide and/or protein levels as well as, the overall structure and composition of a genome. This approach provides clues but does not establish gene product function and needs experimental verification. On the other hand, experience-based functional genomics relies on experiments performed at different levels such as DNA (transpositional screen, random mutagenesis), RNA (microarrays, RNAseq) and protein (2D gel followed by mass spectrometry, ORFeome, heterologous expression,...). Studying the function at the protein level helps characterizing directly the molecular actors in the cell and also includes most functional

63 screens such as protein-protein interactions, subcellular localization as well as,
64 secretion/translocation screens.
65 Functional genomics was first applied to model organisms such as *E. coli* and *B. subtilis*
66 and then to pathogenic bacteria in order to gain insight on their virulence factors.
67 Functional genomics of intracellular bacteria cannot be directly extrapoled from
68 functional genomics data of *E. coli* because gene content and function largely reflects
69 the ecology of a given bacterium and sustained host-pathogen interactions shape the
70 bacterial genomes of intracellular bacteria. This review is devoted to functional
71 genomics of facultative and obligate intracellular pathogenic bacteria. Obligate
72 intracellular pathogenic bacteria only proliferate inside host cells and no defined media
73 are yet available that sustain their bacterial growth whereas, facultative intracellular
74 bacteria may be grown axenically on synthetic media. Functional genomics of
75 intracellular bacteria implies two major challenges: (i) to distinguish the bacterial
76 components of interest from the cellular fraction of the host eukaryotic cell, (ii) the low
77 number of intracellular bacteria at early time point post-infection (p.i.) (before bacterial
78 proliferation) forces the experimenter to deal with low amount of bacterial components.
79 Moreover, the absence of simple genetic system for some strict intracellular bacteria,
80 such as *Chlamydia* spp., is another challenge that explains the large variety of
81 functional genomics approaches that have been developed by chlamydologists. In
82 addition to *Chlamydia*, this review will also focus on functional genomics applied to
83 study two facultative intracellular pathogenic bacteria, *Brucella* spp. and *Legionella* spp..
84 and will present the different strategies used to identify the effectors translocated into
85 the host cell thanks to their Type IV secretion system (T4SS).

Functional genomics of *Brucella*

Generalities

Brucella spp. are the ethiological agent of brucellosis, a widespread worldwide zoonosis affecting a large range of mammals including humans and responsible of dramatic economical losses in endemic countries. *Brucella* spp. are Gram-negative bacteria belonging to alpha-2 subclass of *Proteobacteria* [2]. The *Brucella* genus is divided into 10 species according to their hosts [3, 4]. *Brucella* spp. are able to infect professional and non-professional phagocytes. Once internalized, *Brucella* resides in a vacuole called "*Brucella* containing vacuole" (BCV), which successively interacts with endocytic compartments. Then, *Brucella* reaches the endoplasmic reticulum (ER) at particular sites, the ER exit sites (ERES), where it extensively proliferates [5-7].

Brucella *genome*

The first complete *Brucella* genomes were reported in 2002 for *B. melitensis* strain 16M and *B. suis* strain 1330 [8, 9]. The complete genome of *Brucella abortus* was published in 2005 [10]. Before the availability of complete genome sequence, several groups used transposon mutagenesis and signature-tagged transposon mutagenesis in order to identify virulence factors [11-15]. These experiments were essentially performed *in vitro* on eukaryotic cells and led to the identification of genes essential for intracellular survival. These gene products were involved in various pathways such as amino acid and DNA metabolism, LPS biosynthesis and the T4SS, which suggested that *Brucella* genome does not contain toxins and other canonical virulence factors used by other pathogens.

***ORFeome of* Brucella**

In 2004, the first *B. melitensis* ORFeome was constructed [16]. The *B. melitensis* ORFeome is a library of all protein-encoding open reading frame (ORF) cloned in an entry vector compatible with the Gateway cloning technology, easily transferrable in expression vector by recombination. The ORFeome is a convenient resource for high-throughput functional genomics and can be used for various purposes such as protein over-expression, mutant construction, interaction mapping... Various screens based on the ORFeome availability were performed such as (i) functional screen in yeast for anti-apoptotic effector candidate, (ii) screen for proteins with a polar localization in *B. abortus* or (iii) yeast-two hybrid between all *B. melitensis* proteins and human phagosomal protein or endoplasmic reticulum exit site associated proteins [17-20].

Thanks to the presence of the same flanking sequences for each coding sequence, the ORFeome is also useful for the construction of a PCR product microarrays for the global analysis of gene expression of *Brucella* in laboratory conditions [21].

***Trancriptional analysis of* Brucella**

Transcriptional analysis by using microarrays was essentially performed to study the global regulation of a transcriptional regulator. This was the case of BvrR/S, the two component system essential for *Brucella* virulence [22]. The genes regulated by BvrR/S were determined by parallel whole genome microarray analyses of the wild type and the *bvrR* mutant strains grown under the same conditions [23]. A similar approach was used for VjbR and BabR, two transcriptional factors belonging to LuxR family responding to quorum sensing autoinducer, by combining proteomic study to genome microarray analyses [24-27]. The aim of this work was to identify the quorum sensing regulon by

132 focusing on the most likely targets. The putative targets included those identified by
133 combining proteomic and microarray analysis and those by the microarray analysis
134 alone confirmed by qRT-PCR or chromatin immunoprecipitation (chIP) [27].
135 Other transcriptomic studies were performed in order to identify differences in global
136 gene expression level between bacteria subjected to different stimuli or in different
137 stages of axenic growth or intracellular growth. Rosetti and colleagues performed a
138 microarray analysis between *Brucella melitensis* at the late logarithmic phase of growth
139 (the most invasive culture) and stationary phase (the least invasive). The majority of up-
140 regulated genes in late-log growth phase were associated with growth, including DNA
141 replication, transcription, translation, intermediate metabolism, energy production and
142 conversion, membrane transport, and biogenesis of the cell envelope and outer
143 membrane [28]. Another study characterized the transcriptional profile at 4h (non-
144 proliferative phase) and 12h (proliferative phase) after infection of HeLa cells with *B.*
145 *melitensis*. As many as 151 and 115 genes were differentially expressed at 4 and 12h
146 p.i., compared to the inoculum (a culture at late-log phase of growth). These genes
147 mainly involved in growth and metabolism were down-regulated at 4h p.i. and up-
148 regulated at 12h p.i. [29]. The aim of these two studies was first to identify genes
149 encoding proteins (i) involved in invasion by comparing transcriptional profile between
150 the most and the least invasive growth phase of *B. melitensis* cultures and (ii) involved
151 in survival and proliferation (by comparing transcriptional profile at 4 and 12h p.i). The
152 results presented in these studies reflect a global adaptation. It is therefore difficult to
153 identify a single gene or a set of gene involved in the invasion, survival or proliferation.

The recent advance in RNAseq provides a novel approach for transcriptomic studies where host and pathogen can be both analyzed in parallel. This technique also allows the identification of the transcription start site (TSS), alternative TSS and operon organization as well as non-coding RNAs, antisense RNAs, and 5'-/3'-untranslated regions. Small regulatory RNAs are involved in post-transcriptional regulation and even in modification of protein activity. Hfq protein binds RNA and is usually required for the function and/or stability of this family of sRNAs, in Gram-negative bacteria [30]. The presence of *hfq* gene in *B. abortus* genome suggests that such a sRNA regulation exist. Until now, only two sRNAs were identified which are both orthologous to AbcR1 and 2 of *A. tumefaciens* [31]. Moreover the small regulatory RNAs of *Brucella* still remain poorly known and RNAseq approach should help to study this domain.

Proteomics investigations

The first proteomic study of *Brucella* was performed on bacterial cells grown on blood agar in aerobic condition [32]. A total of 883 proteins spots were detected on 2D gel among which 440 proteins were identified by mass spectrometry. These proteins represent 187 genes that correspond to 6% of the predicted genes present in the genome. Later, various proteomic studies were performed on infected cells or on bacteria grown under microaerobic or anaerobic conditions [33-35]. It was reported that the basal metabolism is reduced under microaerobic and anaerobic conditions, which is expected with low or absence of growth. Under these both conditions, glycolysis and denitrification were favored. When oxygen became limiting, basic metabolism processes were maintained and various respiratory pathways were observed. This flexibility confers to *Brucella* an advantage to survive in low oxygen environments such as

177 damaged host tissues [34]. Lamontagne and colleagues showed that *Brucella* prepare
178 for cell division soon after their internalization in mouse macrophages by overproducing
179 several proteins involved in division and DNA metabolism, such as PleC and XseA [35].
180 In addition to provide such functional information, proteomic studies can also be useful
181 to confirm that proteins were correctly annotated in term of length and that they are
182 produced [36].

183 ***Brucella* VirB T4SS**

184 The T4SS of *Brucella* was discovered in 1999, by a transposon mutagenesis screen for
185 mutants attenuated during infection of HeLa cells. The involvement of the T4SS VirB in
186 *Brucella* virulence and its regulation were extensively studied [5, 11, 15, 25, 37-40].
187 Briefly, *virB* expression is regulated by BvR/S and VjbR regulators [25, 39]. It was also
188 reported that VirB T4SS is required to sustain interaction with the ER and generate a
189 proliferative organelle, probably through the action of translocated effectors into the host
190 cell or the vacuolar membrane [5]. A common strategy to identify effectors relies on
191 bioinformatic approaches. One study hypothesized that proteins translocated by VirB
192 must be co-regulated with the *virB* operon by the VjbR regulator [41]. A conserved motif,
193 in the *virB* promoter, required for VjbR activation was determined and then 144
194 promoters containing this motif were identified. For interesting candidates, translocation
195 into macrophages using TEM-β-lactamase reporter were tested. Thus, two proteins,
196 VceA and VceC, were reported as the first T4SS substrates. A second genome-wide
197 bioinformatics screen was initiated to identify additional effectors. This screen was
198 based on different criteria such as the homology to known effectors and the occurrence
199 of eukaryotic-like domain or motif. Using this bioinformatic approach, 84 *B. abortus*

putative effectors (BPEs) were identified [42]. Translocation of these putative effectors was tested using adenylate cyclase reporter and six were translocated into the eukaryotic cytoplasm.

Another strategy to identify effectors, consist to focus on a particular feature of these effectors such as their interactions with host proteins, as well as their translocation into host cell cytoplasm. In this prospect, a genome-wide yeast-two hybrid between all *B. melitensis* and human phagosomal protein was performed [18]. This approach was possible thanks to the availability of both human and *B. melitensis* ORFeome. A specific interaction was identified between the human Rab GTPase Rab2 and a *Brucella* protein called RicA. This interaction was confirmed by GST pull down and RicA was shown to be translocated into host cell cytoplasm of macrophages using the TEM-β-lactamase reporter. Functional screen in yeast to identify anti-apoptotic effector candidates and translocation screen using the Yersinia YopP as a reporter system [17, 44] are alternative strategies used to identify T4SS effectors.

Despite overlapping many different approaches, *Brucella* effectors still remain poorly characterized possibly due to the difficulties to adapt tools and experiments to this biosafety level 3 bacterial pathogen.

Functional genomics of *Legionella*

Generalities

Legionella pneumophila is a Gram-negative bacterium commonly found in aquatic environment where it replicates inside protozoan hosts [45, 46]. *L. pneumophila* is the

causing agent of a severe pneumonia, called Legionnaires' disease [47-49]. The alveolar macrophages are the primary sites of bacterial proliferation. The *L. pneumophila* virulence seems to rely on its ability to avoid phagosome-lysosome fusion, since mutants defective for this particular phenotype are unable to proliferate inside the host cell and thus to cause the disease [50, 51]. Once inside the cell, *L. pneumophila* reside in a vacuole called *Legionella* containing vacuoles (LCV). The LCV rapidly acquire the characteristics of an ER-like compartment by recruiting vesicles from the early secretory pathway [52]. This is essential for bacterial proliferation and require a functional Dot/Icm type IV secretion system (Dot/Icm T4SS) that translocates effector proteins and represents a major virulence factor (see below) [53].

***Genomics* of Legionella**

The 3 first complete genomes of *L. pneumophila* were published in 2004 and a fourth sequence from the same species was reported in 2007 [54-56]. Many important factors involved in internalization and intracellular proliferation have been identified during the pre-genomic era such as Dot/Icm Type IV secretion system, the Type II secretion system Lsp and the Mip (macrophage infectivity potentiator) [57, 58]

The availability of complete *L. pneumophila* genomes is an open window to better understand the *L. pneumophila* biology. For example, sequence analysis allowed the identification of a putative type I secretion system (Lss) encoded by the lssXYZABD locus [59]. *L. pneumophila* genome was also screened to identify patatin-like proteins (PLPs) and 11 PLPs were identified designated PatA to PatK. These PLPs form a new family of phospholipases. Four of these PLPs (PatA/VipD, PatC/VpdA, PatG/VpdB and PatF/VpdC) have been identified and characterized previously [60, 61].

***Trancriptional analyses of* Legionella**

The first transcriptional study of *L. pneumophila* using microarrays was performed during infection of its natural host, i.e. *Acanthamoeba castellanii* [62]. Virulence traits such as Dot/Icm substrates, factors associated to invasion, virulence and motility as well as more than 90 proteins without characterized function, were overexpressed during the transmissive phase (>10h post-infection, p.i.) compared to the proliferative phase (<10h p.i.). Another transcriptional analysis was performed on *L. pneumophila* biofilms cells by comparing transcriptional profile of sessile cells with two distinct populations of planktonic cells [63]. The results showed that sessile cells have a similar gene expression profile to proliferative phase *L. pneumophila*. Recently, to detect putative virulence factors involved in resistance to macrophages, the transcriptional response of *L. pneumophila* once internalized by human macrophages, was analysed at 0, 6 and 18h p.i. and was compared to exponential and post-exponential axenic growth. Interestingly, 8 of the 10 most highly induced genes were of unknown function. These genes could represent virulence traits. Three new translocated effectors were identified by scanning the genome in order to detect regions enriched in genes without assigned function and showing a similar expression patterns to their neighbouring effector genes [64].

In addition many putative sRNA molecules were identified by both bioinformatic analyses and deep RNA-sequencing on *L. pneumophila* grown in broth and inside *A. castellanii* [65-68]. Thus, deep RNA-sequencing gave new insights on the global transcriptional regulation and response to particular conditions and also helped identifying small non-coding RNA involved in the post-transcriptional regulation.

Proteomics investigations

The first proteomic study performed on *L. pneumophila* was reported in 2005. This work on total cell extracts provided a reference map for further investigations [69]. A proteomic approach was also used to identify the T2SS secretome [70, 71], as well as the whole secretome [72] and the membranome and surfaceome [73]. Proteomic analysis was also performed at both exponential phase and post-exponential phase (virulent) of *L. pneumophila*, to confirm differences observed at the transcription level and to identify proteins possibly associated to virulence [74]. This led to the identification of 68 proteins among which 64 were overproduced at the post-exponential phase. Of these, nine proteins of unknown function were found, among which 6 were demonstrated specific for *L. pneumophila* by southern blot analysis. Two of them were associated to haemolysis by conducting contact dependent hemolysis assay using SRBC (sheep red blood cells) and two were translocated into macrophages by the Dot/Icm T4SS demonstrating the usefulness of proteomics to decipher protein functions.

***Legionella* Dot/Icm T4SS**

The T4SS Dot/Icm was identified by various selections and screens for *L. pneumophila* mutants defective for intracellular growth and/or defective for macrophage killing. These mutants were called *dot* for defect in organelle trafficking, in Isberg lab, and *icm* for deficient in intracellular multiplication in Shuman lab [75-80]. Up to now, 26 *dot/icm* genes have been identified and are essential for intracellular growth, in particular to prevent phagosome-lysosome fusion.

More than 300 *L. pneumophila* effectors have been identified using an arsenal of methods such as bioinformatic screens for genes encoding eukaryotic-like domain,

genetic screens for particular phenotypes, yeast screens and translocation screens. Sequencing and analysis of the *Legionella* genome identified a wide variety of proteins exhibiting eukaryotic-like domains such as ankyrin repeat, Sel-1, SET, Sec7 motifs, U-box and F-box domains. Using these bioinformatic approaches, a *L. pneumophila* protein, called RalF, containing a Sec7 homology domain was identified [81]. Sec7 homology domains are found in a family of eukaryotic ARF-GEF, which stimulates exchange of GDP for GTP. Arf1 is a small GTPase involved in the regulation of the vesicle traffic between the ER and the Golgi. It was reported that RalF is required for the localization of Arf1 on phagosomes containing *L. pneumophila*. Moreover RalF is translocated through the phagosomal membrane by the T4SS Dot/Icm. RalF is the first translocated substrate of the T4SS with an identified function.

One interesting genetic screen relied on the identification of mutant strains, obtained by transposon mutagenesis, exhibiting a similar phenotype than that of the *dotL* mutant, i.e. lethality [82]. This lethality is likely due to the assembly of a poison Dot/Icm complex, caused by the dysregulation of the molecular flow through the translocator. This study allowed the identification of LidA (lowered viability in the presence of dot). LidA function was then extensively studied [43, 83-86].

The yeast was also exploited to identify and characterized *L. pneumophila* effectors. A yeast lethality screen was performed to identifiy *L. pneumophila* proteins which interfere with yeast growth. This screen led to the identification of YlfA for yeast lethal factor. YlfA was also shown to be translocated by the Dot/Icm apparatus and is associated with vesicles of the early secretory pathway including ER [87]. The yeast can be also used to identify proteins which cause a membrane trafficking (vacuole protein sorting, VPS)

defect in yeast. Using this screen, 3 proteins which inhibit vacuolar traffic were identified called VipA (VPS inhibitory protein), VipD and VipF [60]. These 3 proteins are also translocated into host macrophages through the Dot/Icm T4SS. VipD possesses a patatin domain and is thus also called PatA.

All these methods considerably improved our knowledge on *L. pneumophila* intracellular life cycle and subversion of host cell processes to its own advantage. Moreover, yeasts represent a useful tool to define the function of putative effectors.

Functional genomics of *Chlamydia*

Generalities

Chlamydia trachomatis and *C. pneumoniae* are important human pathogens causing ocular infection and respiratory diseases such as pneumonia, respectively. Until now, we are still unable to genetically manipulate these organisms by targeted mutagenesis or transposon mutagenesis, prompting the development of other approaches. One major reason for the absence of genetic system to manipulate the genome of *Chlamydia*, resides in the obligate intracellular life cycle of these bacteria. The chlamydial development cycle is indeed characterized by two distinct developmental stages, which are morphologically and functionally different. Elementary bodies (EBs) are the infectious form that may survive extracellularly whereas reticulate bodies (RBs) are non-infectious and proliferate inside the host cell. Basically, EBs are internalized and differentiate into RBs which replicate by binary fission. RBs then redifferentiate into EBs which are released after cell lysis. Finally, EBs initiate a new infection cycle.

Chlamydia *genome and genetic approaches*

The complete genome sequence of *C. trachomatis* and *C. pneumonaie* were published in 1998 and 1999, respectively [88, 89].

One interesting approach to tackle the intractability of *Chlamydia* is the use of a combinatorial approach coupling a chemical mutagenesis with whole genome sequencing and a system of DNA exchange within infected cells [90]. Practically, *Chlamydia* infected Vero cells were treated with the alkylating agent ethylmethylsulfonate (EMS). Mutated *Chlamydia* were then used to reinfect a monolayer of Vero cells, overlaid with agar to observe plaque formation. Mutants were classified according to their plaque morphotypes and the whole genome of mutants sequenced to identify mutated genes sharing the same phenotype. Finally, co-infection between the wild type and mutant strains were performed to obtain recombinants where particular mutated genes could be linked to a phenotype. This method led to the identification of 4 mutants which form large granular plaques (Gnr). Three were mutated in the *glgB* gene encoding a glycogen-branching enzyme. Microscopic analysis of HeLa cells infected with Gnr mutants showed an accumulation of large precipitates in the lumen of inclusions, likely glycogen. Recombinant strains were obtained to address the link between genotype and phenotype and showed that all strains with a mutated *glgB* (even single mutation) were accumulating glycogen inside inclusion. Altogether, a loss-of-function of *glgB* is leading to the accumulation of glycogen. Similarly, in the Gnr4 mutant, a mutation was identified in the *gspE* gene which is homologous to ATPases of the Type II secretion system. One hypothesis is that *gspE* mutant accumulates glycogen because a key glycogen hydrolase is not secreted. Interestingly, *gspE* mutant is attenuated during HeLa cells infection compared to wild type and *glgB* mutant strains.

This suggests that the T2SS is involved in the secretion of other factors essential for bacterial survival. This combinatorial approach should be applicable to other genetically intractable pathogenic bacteria.

***Trancriptional analyses of* Chlamydia**

Transcriptomic and proteomic analyses are important to study the global adaptation of *Chlamydia* to their host. Microarray analyses of the temporal gene expression during the developmental cycle have been performed, repectively on *C. trachomatis* serovar L2 and D [91, 92] and on *C. pneumoniae* [93]. Belland and colleagues identified 29 early genes expressed as early as 1h p.i. [91]. Analysis of these genes suggests that *Chlamydia* established systems for nutrient acquisition and modify its inclusion by expressing particular genes during the early stage of infection. Transcriptomic studies allowed the definition of a new class of genes called the “very late” (or “tardy” genes), in addition to the “late” genes. Basically, “late” genes encode early proteins required for EBs infection and “tardy” correspond to genes which mRNA transcripts are present in EBs. A total of 26 and 70 “late” genes were respectively, identified by Belland *et al.* and Nicholson *et al.* [91, 92]. Among these genes, *omcAB* and *hctAB* were previously characterized as “late” genes. *HctAB* encodes HctA and HctB, two chlamydial histone-like proteins, which mediate chromosomal condensation during the differentiation of RBs to EBs. *OmcAB* encodes for OmcA and OmcB are two cysteine-rich outer membrane proteins interacting with OmpA, the major outer membrane protein, to form a highly disulfide crosslinked complex. This complex is considered essential for the resistance of EBs to osmotic stress when outside host cells, since *Chlamydia* do not have a classical peptidoglycan layer.

Recently, RNAseq was performed on the *C. trachomatis* L2b and *C. pneumoniae* CWL-029 on purified elementary bodies and reticulate bodies [94, 95]. For *C. trachomatis*, 363 transcription start sites have been mapped and 43 non-coding RNA identified. As many as 83 genes showed differential expression level between RBs and EBs [95]. For *C. pneumoniae*, 565 transcriptional start sites of annotated genes and novel transcripts were mapped. Semi-quantitatvive analysis showed significant differences in genes expression between EBs and RBs for 288 genes. Moreover, 75 non-coding RNA were identified [94]. By intergenic tiling microarray on RNA of *C. trachomatis* D at 40h p.i., 34 non-coding RNAs were identified, 16 being confirmed by northern blot [96]. One of the non-coding RNA regulated *ftsI* expression by inducing degradation of *ftsI* mRNA. This is especially interesting since in other bacterial lineages, FtsI is involved in peptidoglycan synthesis. The true role of FtsI in *Chlamydia* remains to be determined [97]. Given the susceptibility of *Chlamydia* to penicillin derivatives, the FtsI and others genes of the *fts* operon are likely important in chlamydial multiplication and/or in the biosynthesis of the chlamydial peptidoglycan-like layer.

Proteomics investigations

Different proteomic studies using 2D gel and mass spectrometry were performed during the past 10 years in order to better understand events such as differentiation of EBs to RBs. Several studies were performed on purified EBs or on both purified EBs and RBs, in order to determine the proteome and study its temporal variation [98-100]. These studies confirmed that ORFing of *C. trachomatis* and *C. pneumoniae* genomes was correct for these hypothetical proteins. It was also showed that the entire set of glycolytic enzymes were present in the so-called metabolically inert form (EBs)

suggesting that there are metabolic flux also in EBs. These results were recently confirmed by a quantitative study showing that proteins of the central metabolism and glucose catabolism were more abundant in EBs, whereas in RBs, proteins involved in ATP generation, proteins synthesis, and nutrient transport were predominant [101].

Proteomic approaches were also used to identify translocated chlamydial proteins into host cell cytoplasm [102, 103]. The first study allowed the identification of CPAF (chlamydial protease or proteasome-like activity factor), a factor previously characterized by Zhong and colleagues [104]. A second study allowed the identification of CT621, which localized to the host cell cytoplasm and nucleus and whose translocation is dependent of the T3SS [103].

An ORFeome was recently constructed for *C. pneumoniae* [105]. The ORFeome is an essential tool for functional genomics of such intractable bacteria allowing functional screens in yeast such as two-hybrid, lethality screen, membrane traffic defect, adhesion assays, as well as screen to identify T3SS translocated proteins, for instance by heterologous expression in *Shigella* using reporter gene fusions.

Advances in* Chlamydia *transformation

In one of the pioneering works in this field, Binet and colleagues [106] constructed pUC plasmid derivatives carrying different lengths of rRNA regions containing 4 nucleotide substitutions. Three substitutions located in the 16Sr RNA gene conferred resistance to kasugamycin and spectinomycin, and caused a loss of one HpaI restriction site. *C. psittaci* 6BC was then electroporated with various concentrations of circular or linearized plasmids. Allelic replacements of the endogenous rRNA operon were selected by incubation of the infected cells with the two antibiotics. Allelic replacements were

observed at frequencies greater than 10^{-6}. This showed that genetic manipulations are feasible in *Chlamydia*. Recently, *C. trachomatis* transformation was reported using penicillin selection and calcium chloride treatment of EBs to render them competent [107]. A GFP plasmid was constructed based on the plasmid of the Swedish new variant strain (a strain with a deletion of a 400bp region in the canonical 7'500kb plasmid of *C. trachomatis*). This plasmid was used to obtain penicillin resistant *C. trachomatis* strains expressing *gfp*. These recent advances in *C. trachomatis* transformation and mutagenesis by allelic recombination in *C. psittaci* open the window to future development of genetic tools in order to perform targeted and random mutagenesis, which could considerably improve our knowledge on the biology of these bacteria [106, 107].

Discussion

During the last two decades, complete genomes were obtained for many bacteria. All these data are however not sufficient to understand the bacterial biology. This led to the development of functional genomics whose main feature is its genome-wide approach (Figure 2). One challenge of the functional genomics on intracellular pathogen is to discriminate bacterial material from host cell material. The early times post-infection are also critical since there is not yet bacterial proliferation and we have to deal with very small amount of bacteria. Functional genomics approaches are useful to better understand host-pathogen interaction which is tightly regulated by the two interacting partners and leads to accumulating data at the DNA, RNA and protein level. These data

are quite difficult to interpret and need to conciliate all 3 levels. Each approach presents advantages and limitations which are summarized in the Table 1. One major disadvantage of microarray analysis is that results may be different than results obtained at proteomic level due to post-transcriptional modifications. It is therefore interesting to perform both analyses in parallel on a same sample. The main disadvantage of proteomic studies, using 2D gel and mass spectrometry, is that they require supplementary steps to separate/distinguish host cell proteins from bacterial proteins. RNAseq compared to microarrays, has several advantages, and especially in term of cost and quality of data, but this new technology has not yet been much applied to intracellular pathogenic bacteria. RNAseq also require differentiating bacterial RNA and host cell RNA. This is possible by mapping reads to reference bacterial genome sequence or, before the RNAseq, by a physical separation of bacteria from host cells or by sequential purification steps to isolate bacterial RNA.

Genome-wide studies give a global view of the bacterial response to a particular environment. They may be combined with specific functional screens in order to determine the role of a particular protein.

Extracellular pathogenic bacteria secrete toxins/enzymes which are considered as virulence factors. For intracellular bacteria, toxins/enzymes counterparts are effector proteins. These proteins are translocated by a secretion system and are involved in the host cell hijacking. Intracellular pathogenic bacteria possess a battery of non-canonical effectors with redundant functions, rendering their identification extremely difficult. In this context, functional genomics approaches are very useful tools.

In the future, functional genomics will increasingly use RNAseq and will likely also investigate bacterial metabolism using metabolomics. The major interest of RNAseq is the identification of non-coding RNAs which are new actors in genes regulation. Non-coding RNA role is still likely underestimated. By giving a snapshot of the metabolites present at a define time, metabolomics will provide important insights into bacterial physiology. The next challenge in the functional genomics field is to integrate data from transcriptomic, proteomic and metabolomic studies in order to obtain a global picture of the bacterial state in a defined condition and to better characterize host-pathogen interaction.

References

[1] R.D. Fleischmann, M.D. Adams, O. White, R.A. Clayton, E.F. Kirkness, A.R. Kerlavage, C.J. Bult, J.F. Tomb, B.A. Dougherty, J.M. Merrick, et al., Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd, Science 269 (1995) 496-512.

[2] E. Moreno, E. Stackebrandt, M. Dorsch, J. Wolters, M. Busch, H. Mayer, *Brucella abortus* 16S rRNA and lipid A reveal a phylogenetic relationship with members of the alpha-2 subdivision of the class Proteobacteria, Journal of bacteriology 172 (1990) 3569-3576.

[3] E. Moreno, A. Cloeckaert, I. Moriyon, *Brucella* evolution and taxonomy, Veterinary microbiology 90 (2002) 209-227.

[4] S. Audic, M. Lescot, J.M. Claverie, H.C. Scholz, *Brucella microti*: the genome sequence of an emerging pathogen, BMC genomics 10 (2009) 352.

[5] J. Celli, C. de Chastellier, D.M. Franchini, J. Pizarro-Cerda, E. Moreno, J.P. Gorvel, Brucella evades macrophage killing via VirB-dependent sustained interactions with the endoplasmic reticulum, The Journal of experimental medicine 198 (2003) 545-556.

[6] T. Starr, T.W. Ng, T.D. Wehrly, L.A. Knodler, J. Celli, *Brucella* intracellular replication requires trafficking through the late endosomal/lysosomal compartment, Traffic 9 (2008) 678-694.

[7] J. Pizarro-Cerda, S. Meresse, R.G. Parton, G. van der Goot, A. Sola-Landa, I. Lopez-Goni, E. Moreno, J.P. Gorvel, *Brucella abortus* transits through the autophagic pathway and replicates in the endoplasmic reticulum of nonprofessional phagocytes, Infection and immunity 66 (1998) 5711-5724.

[8] V.G. DelVecchio, V. Kapatral, R.J. Redkar, G. Patra, C. Mujer, T. Los, N. Ivanova, I. Anderson, A. Bhattacharyya, A. Lykidis, G. Reznik, L. Jablonski, N. Larsen, M. D'Souza, A. Bernal, M. Mazur, E. Goltsman, E. Selkov, P.H. Elzer, S. Hagius, D. O'Callaghan, J.J. Letesson, R. Haselkorn, N. Kyrpides, R.

Overbeek, The genome sequence of the facultative intracellular pathogen *Brucella melitensis*, Proceedings of the National Academy of Sciences of the United States of America 99 (2002) 443-448.
[9] I.T. Paulsen, R. Seshadri, K.E. Nelson, J.A. Eisen, J.F. Heidelberg, T.D. Read, R.J. Dodson, L. Umayam, L.M. Brinkac, M.J. Beanan, S.C. Daugherty, R.T. Deboy, A.S. Durkin, J.F. Kolonay, R. Madupu, W.C. Nelson, B. Ayodeji, M. Kraul, J. Shetty, J. Malek, S.E. Van Aken, S. Riedmuller, H. Tettelin, S.R. Gill, O. White, S.L. Salzberg, D.L. Hoover, L.E. Lindler, S.M. Halling, S.M. Boyle, C.M. Fraser, The *Brucella suis* genome reveals fundamental similarities between animal and plant pathogens and symbionts, Proceedings of the National Academy of Sciences of the United States of America 99 (2002) 13148-13153.
[10] S.M. Halling, B.D. Peterson-Burch, B.J. Bricker, R.L. Zuerner, Z. Qing, L.L. Li, V. Kapur, D.P. Alt, S.C. Olsen, Completion of the genome sequence of *Brucella abortus* and comparison to the highly similar genomes of *Brucella melitensis* and *Brucella suis*, Journal of bacteriology 187 (2005) 2715-2726.
[11] D. O'Callaghan, C. Cazevieille, A. Allardet-Servent, M.L. Boschiroli, G. Bourg, V. Foulongne, P. Frutos, Y. Kulakov, M. Ramuz, A homologue of the *Agrobacterium tumefaciens* VirB and *Bordetella pertussis* Ptl type IV secretion systems is essential for intracellular survival of *Brucella suis*, Molecular microbiology 33 (1999) 1210-1220.
[12] P. Lestrate, A. Dricot, R.M. Delrue, C. Lambert, V. Martinelli, X. De Bolle, J.J. Letesson, A. Tibor, Attenuated signature-tagged mutagenesis mutants of Brucella melitensis identified during the acute phase of infection in mice, Infection and immunity 71 (2003) 7053-7060.
[13] S. Kohler, V. Foulongne, S. Ouahrani-Bettache, G. Bourg, J. Teyssier, M. Ramuz, J.P. Liautard, The analysis of the intramacrophagic virulome of *Brucella suis* deciphers the environment encountered by the pathogen inside the macrophage host cell, Proceedings of the National Academy of Sciences of the United States of America 99 (2002) 15711-15716.
[14] V. Foulongne, G. Bourg, C. Cazevieille, S. Michaux-Charachon, D. O'Callaghan, Identification of *Brucella suis* genes affecting intracellular survival in an in vitro human macrophage infection model by signature-tagged transposon mutagenesis, Infection and immunity 68 (2000) 1297-1303.
[15] R.M. Delrue, M. Martinez-Lorenzo, P. Lestrate, I. Danese, V. Bielarz, P. Mertens, X. De Bolle, A. Tibor, J.P. Gorvel, J.J. Letesson, Identification of Brucella spp. genes involved in intracellular trafficking, Cellular microbiology 3 (2001) 487-497.
[16] A. Dricot, J.F. Rual, P. Lamesch, N. Bertin, D. Dupuy, T. Hao, C. Lambert, R. Hallez, J.M. Delroisse, J. Vandenhaute, I. Lopez-Goni, I. Moriyon, J.M. Garcia-Lobo, F.J. Sangari, A.P. Macmillan, S.J. Cutler, A.M. Whatmore, S. Bozak, R. Sequerra, L. Doucette-Stamm, M. Vidal, D.E. Hill, J.J. Letesson, X. De Bolle, Generation of the *Brucella melitensis* ORFeome version 1.1, Genome research 14 (2004) 2201-2206.
[17] G. Laloux, M. Deghelt, M. de Barsy, J.J. Letesson, X. De Bolle, Identification of the essential *Brucella melitensis* porin Omp2b as a suppressor of Bax-induced cell death in yeast in a genome-wide screening, PloS one 5 (2010) e13274.
[18] M. de Barsy, A. Jamet, D. Filopon, C. Nicolas, G. Laloux, J.F. Rual, A. Muller, J.C. Twizere, B. Nkengfac, J. Vandenhaute, D.E. Hill, S.P. Salcedo, J.P. Gorvel, J.J. Letesson, X. De Bolle, Identification of a *Brucella* spp. secreted effector specifically interacting with human small GTPase Rab2, Cellular microbiology 13 (2011) 1044-1058.
[19] M. de Barsy, A. Mirabella, J.J. Letesson, X. De Bolle, A *Brucella abortus cstA* mutant is defective for association with endoplasmic reticulum exit sites and displays altered trafficking in HeLa cells, Microbiology 158 (2012) 2610-2618.
[20] D. Dotreppe, C. Mullier, J.J. Letesson, X. De Bolle, The alkylation response protein AidB is localized at the new poles and constriction sites in *Brucella abortus*, BMC microbiology 11 (2011) 257.
[21] C. Viadas, M.C. Rodriguez, J.M. Garcia-Lobo, F.J. Sangari, I. Lopez-Goni, Construction and evaluation of an ORFeome-based *Brucella* whole-genome DNA microarray, Microbial pathogenesis 47 (2009) 189-195.

[22] A. Sola-Landa, J. Pizarro-Cerda, M.J. Grillo, E. Moreno, I. Moriyon, J.M. Blasco, J.P. Gorvel, I. Lopez-Goni, A two-component regulatory system playing a critical role in plant pathogens and endosymbionts is present in *Brucella abortus* and controls cell invasion and virulence, Molecular microbiology 29 (1998) 125-138.
[23] C. Viadas, M.C. Rodriguez, F.J. Sangari, J.P. Gorvel, J.M. Garcia-Lobo, I. Lopez-Goni, Transcriptome analysis of the *Brucella abortus* BvrR/BvrS two-component regulatory system, PloS one 5 (2010) e10216.
[24] J.J. Letesson, P. Lestrate, R.M. Delrue, I. Danese, F. Bellefontaine, D. Fretin, B. Taminiau, A. Tibor, A. Dricot, C. Deschamps, V. Haine, S. Leonard, T. Laurent, P. Mertens, J. Vandenhaute, X. De Bolle, Fun stories about *Brucella*: the "furtive nasty bug", Veterinary microbiology 90 (2002) 317-328.
[25] R.M. Delrue, C. Deschamps, S. Leonard, C. Nijskens, I. Danese, J.M. Schaus, S. Bonnot, J. Ferooz, A. Tibor, X. De Bolle, J.J. Letesson, A quorum-sensing regulator controls expression of both the type IV secretion system and the flagellar apparatus of *Brucella melitensis*, Cellular microbiology 7 (2005) 1151-1161.
[26] S. Uzureau, M. Godefroid, C. Deschamps, J. Lemaire, X. De Bolle, J.J. Letesson, Mutations of the quorum sensing-dependent regulator VjbR lead to drastic surface modifications in *Brucella melitensis*, Journal of bacteriology 189 (2007) 6035-6047.
[27] S. Uzureau, J. Lemaire, E. Delaive, M. Dieu, A. Gaigneaux, M. Raes, X. De Bolle, J.J. Letesson, Global analysis of quorum sensing targets in the intracellular pathogen *Brucella melitensis* 16 M, Journal of proteome research 9 (2010) 3200-3217.
[28] C.A. Rossetti, C.L. Galindo, S.D. Lawhon, H.R. Garner, L.G. Adams, *Brucella melitensis* global gene expression study provides novel information on growth phase-specific gene regulation with potential insights for understanding Brucella:host initial interactions, BMC microbiology 9 (2009) 81.
[29] C.A. Rossetti, C.L. Galindo, H.R. Garner, L.G. Adams, Transcriptional profile of the intracellular pathogen *Brucella melitensis* following HeLa cells infection, Microbial pathogenesis 51 (2011) 338-344.
[30] J. Vogel, B.F. Luisi, Hfq and its constellation of RNA, Nature reviews. Microbiology 9 (2011) 578-589.
[31] C.C. Caswell, J.M. Gaines, P. Ciborowski, D. Smith, C.H. Borchers, C.M. Roux, K. Sayood, P.M. Dunman, R.M. Roop Ii, Identification of two small regulatory RNAs linked to virulence in *Brucella abortus* 2308, Molecular microbiology 85 (2012) 345-360.
[32] M.A. Wagner, M. Eschenbrenner, T.A. Horn, J.A. Kraycer, C.V. Mujer, S. Hagius, P. Elzer, V.G. DelVecchio, Global analysis of the *Brucella melitensis* proteome: Identification of proteins expressed in laboratory-grown culture, Proteomics 2 (2002) 1047-1060.
[33] S. Al Dahouk, V. Jubier-Maurin, H.C. Scholz, H. Tomaso, W. Karges, H. Neubauer, S. Kohler, Quantitative analysis of the intramacrophagic *Brucella suis* proteome reveals metabolic adaptation to late stage of cellular infection, Proteomics 8 (2008) 3862-3870.
[34] S. Al Dahouk, S. Loisel-Meyer, H.C. Scholz, H. Tomaso, M. Kersten, A. Harder, H. Neubauer, S. Kohler, V. Jubier-Maurin, Proteomic analysis of *Brucella suis* under oxygen deficiency reveals flexibility in adaptive expression of various pathways, Proteomics 9 (2009) 3011-3021.
[35] J. Lamontagne, A. Forest, E. Marazzo, F. Denis, H. Butler, J.F. Michaud, L. Boucher, I. Pedro, A. Villeneuve, D. Sitnikov, K. Trudel, N. Nassif, D. Boudjelti, F. Tomaki, E. Chaves-Olarte, C. Guzman-Verri, S. Brunet, A. Cote-Martin, J. Hunter, E. Moreno, E. Paramithiotis, Intracellular adaptation of *Brucella abortus*, Journal of proteome research 8 (2009) 1594-1609.
[36] J. Lamontagne, M. Beland, A. Forest, A. Cote-Martin, N. Nassif, F. Tomaki, I. Moriyon, E. Moreno, E. Paramithiotis, Proteomics-based confirmation of protein expression and correction of annotation errors in the *Brucella abortus* genome, BMC genomics 11 (2010) 300.

[37] D.J. Comerci, M.J. Martinez-Lorenzo, R. Sieira, J.P. Gorvel, R.A. Ugalde, Essential role of the VirB machinery in the maturation of the *Brucella abortus*-containing vacuole, Cellular microbiology 3 (2001) 159-168.
[38] A.B. den Hartigh, H.G. Rolan, M.F. de Jong, R.M. Tsolis, VirB3 to VirB6 and VirB8 to VirB11, but not VirB7, are essential for mediating persistence of *Brucella* in the reticuloendothelial system, Journal of bacteriology 190 (2008) 4427-4436.
[39] C. Martinez-Nunez, P. Altamirano-Silva, F. Alvarado-Guillen, E. Moreno, C. Guzman-Verri, E. Chaves-Olarte, The two-component system BvrR/BvrS regulates the expression of the type IV secretion system VirB in *Brucella abortus*, Journal of bacteriology 192 (2010) 5603-5608.
[40] R. Sieira, D.J. Comerci, D.O. Sanchez, R.A. Ugalde, A homologue of an operon required for DNA transfer in *Agrobacterium* is required in *Brucella abortus* for virulence and intracellular multiplication, Journal of bacteriology 182 (2000) 4849-4855.
[41] M.F. de Jong, Y.H. Sun, A.B. den Hartigh, J.M. van Dijl, R.M. Tsolis, Identification of VceA and VceC, two members of the VjbR regulon that are translocated into macrophages by the Brucella type IV secretion system, Molecular microbiology 70 (2008) 1378-1396.
[42] M.I. Marchesini, C.K. Herrmann, S.P. Salcedo, J.P. Gorvel, D.J. Comerci, In search of *Brucella abortus* type IV secretion substrates: screening and identification of four proteins translocated into host cells through VirB system, Cellular microbiology 13 (2011) 1261-1274.
[43] M.P. Machner, R.R. Isberg, Targeting of host Rab GTPase function by the intravacuolar pathogen *Legionella pneumophila*, Developmental cell 11 (2006) 47-56.
[44] J.P. Lavigne, G. Patey, F.J. Sangari, G. Bourg, M. Ramuz, D. O'Callaghan, S. Michaux-Charachon, Identification of a new virulence factor, BvfA, in *Brucella suis*, Infection and immunity 73 (2005) 5524-5529.
[45] B.S. Fields, The molecular ecology of *legionellae*, Trends in microbiology 4 (1996) 286-290.
[46] G. Greub, D. Raoult, Microorganisms resistant to free-living amoebae, Clinical microbiology reviews 17 (2004) 413-433.
[47] J.E. McDade, C.C. Shepard, D.W. Fraser, T.R. Tsai, M.A. Redus, W.R. Dowdle, Legionnaires' disease: isolation of a bacterium and demonstration of its role in other respiratory disease, The New England journal of medicine 297 (1977) 1197-1203.
[48] F. Lamoth, G. Greub, Fastidious intracellular bacteria as causal agents of community-acquired pneumonia, Expert review of anti-infective therapy 8 (2010) 775-790.
[49] F. Lamoth, G. Greub, Amoebal pathogens as emerging causal agents of pneumonia, FEMS microbiology reviews 34 (2010) 260-280.
[50] M.A. Horwitz, The Legionnaires' disease bacterium (*Legionella pneumophila*) inhibits phagosome-lysosome fusion in human monocytes, The Journal of experimental medicine 158 (1983) 2108-2126.
[51] M.A. Horwitz, Characterization of avirulent mutant *Legionella pneumophila* that survive but do not multiply within human monocytes, The Journal of experimental medicine 166 (1987) 1310-1328.
[52] J.C. Kagan, C.R. Roy, *Legionella* phagosomes intercept vesicular traffic from endoplasmic reticulum exit sites, Nature cell biology 4 (2002) 945-954.
[53] C.R. Roy, K.H. Berger, R.R. Isberg, *Legionella pneumophila* DotA protein is required for early phagosome trafficking decisions that occur within minutes of bacterial uptake, Molecular microbiology 28 (1998) 663-674.
[54] C. Cazalet, C. Rusniok, H. Bruggemann, N. Zidane, A. Magnier, L. Ma, M. Tichit, S. Jarraud, C. Bouchier, F. Vandenesch, F. Kunst, J. Etienne, P. Glaser, C. Buchrieser, Evidence in the *Legionella pneumophila* genome for exploitation of host cell functions and high genome plasticity, Nature genetics 36 (2004) 1165-1173.

[55] M. Chien, I. Morozova, S. Shi, H. Sheng, J. Chen, S.M. Gomez, G. Asamani, K. Hill, J. Nuara, M. Feder, J. Rineer, J.J. Greenberg, V. Steshenko, S.H. Park, B. Zhao, E. Teplitskaya, J.R. Edwards, S. Pampou, A. Georghiou, I.C. Chou, W. Iannuccilli, M.E. Ulz, D.H. Kim, A. Geringer-Sameth, C. Goldsberry, P. Morozov, S.G. Fischer, G. Segal, X. Qu, A. Rzhetsky, P. Zhang, E. Cayanis, P.J. De Jong, J. Ju, S. Kalachikov, H.A. Shuman, J.J. Russo, The genomic sequence of the accidental pathogen *Legionella pneumophila*, Science 305 (2004) 1966-1968.

[56] M. Steinert, K. Heuner, C. Buchrieser, C. Albert-Weissenberger, G. Glockner, *Legionella* pathogenicity: genome structure, regulatory networks and the host cell response, International journal of medical microbiology : IJMM 297 (2007) 577-587.

[57] N.P. Cianciotto, B.I. Eisenstein, C.H. Mody, G.B. Toews, N.C. Engleberg, A *Legionella pneumophila* gene encoding a species-specific surface protein potentiates initiation of intracellular infection, Infection and immunity 57 (1989) 1255-1262.

[58] M.R. Liles, V.K. Viswanathan, N.P. Cianciotto, Identification and temperature regulation of *Legionella pneumophila* genes involved in type IV pilus biogenesis and type II protein secretion, Infection and immunity 66 (1998) 1776-1782.

[59] S. Jacobi, K. Heuner, Description of a putative type I secretion system in *Legionella pneumophila*, International journal of medical microbiology : IJMM 293 (2003) 349-358.

[60] N. Shohdy, J.A. Efe, S.D. Emr, H.A. Shuman, Pathogen effector protein screening in yeast identifies *Legionella* factors that interfere with membrane trafficking, Proceedings of the National Academy of Sciences of the United States of America 102 (2005) 4866-4871.

[61] S.M. VanRheenen, Z.Q. Luo, T. O'Connor, R.R. Isberg, Members of a *Legionella pneumophila* family of proteins with ExoU (phospholipase A) active sites are translocated to target cells, Infection and immunity 74 (2006) 3597-3606.

[62] H. Bruggemann, A. Hagman, M. Jules, O. Sismeiro, M.A. Dillies, C. Gouyette, F. Kunst, M. Steinert, K. Heuner, J.Y. Coppee, C. Buchrieser, Virulence strategies for infecting phagocytes deduced from the *in vivo* transcriptional program of *Legionella pneumophila*, Cellular microbiology 8 (2006) 1228-1240.

[63] T. Hindre, H. Bruggemann, C. Buchrieser, Y. Hechard, Transcriptional profiling of *Legionella pneumophila* biofilm cells and the influence of iron on biofilm formation, Microbiology 154 (2008) 30-41.

[64] S.P. Faucher, C.A. Mueller, H.A. Shuman, *Legionella Pneumophila* Transcriptome during Intracellular Multiplication in Human Macrophages, Frontiers in microbiology 2 (2011) 60.

[65] S.P. Faucher, G. Friedlander, J. Livny, H. Margalit, H.A. Shuman, *Legionella pneumophila* 6S RNA optimizes intracellular multiplication, Proceedings of the National Academy of Sciences of the United States of America 107 (2010) 7533-7538.

[66] T. Sahr, H. Bruggemann, M. Jules, M. Lomma, C. Albert-Weissenberger, C. Cazalet, C. Buchrieser, Two small ncRNAs jointly govern virulence and transmission in *Legionella pneumophila*, Molecular microbiology 72 (2009) 741-762.

[67] T. Sahr, C. Rusniok, D. Dervins-Ravault, O. Sismeiro, J.Y. Coppee, C. Buchrieser, Deep sequencing defines the transcriptional map of *L. pneumophila* and identifies growth phase-dependent regulated ncRNAs implicated in virulence, RNA biology 9 (2012) 503-519.

[68] B.A. Weissenmayer, J.G. Prendergast, A.J. Lohan, B.J. Loftus, Sequencing illustrates the transcriptional response of *Legionella pneumophila* during infection and identifies seventy novel small non-coding RNAs, PloS one 6 (2011) e17570.

[69] I. Lebeau, E. Lammertyn, E. De Buck, L. Maes, N. Geukens, L. Van Mellaert, L. Arckens, J. Anne, S. Clerens, First proteomic analysis of *Legionella pneumophila* based on its developing genome sequence, Research in microbiology 156 (2005) 119-129.

[70] S. DebRoy, J. Dao, M. Soderberg, O. Rossier, N.P. Cianciotto, *Legionella pneumophila* type II secretome reveals unique exoproteins and a chitinase that promotes bacterial persistence in the lung,

698 Proceedings of the National Academy of Sciences of the United States of America 103 (2006) 19146-
699 19151.
700 [71] O. Rossier, J. Dao, N.P. Cianciotto, The type II secretion system of *Legionella pneumophila*
701 elaborates two aminopeptidases, as well as a metalloprotease that contributes to differential infection
702 among protozoan hosts, Applied and environmental microbiology 74 (2008) 753-761.
703 [72] F. Galka, S.N. Wai, H. Kusch, S. Engelmann, M. Hecker, B. Schmeck, S. Hippenstiel, B.E. Uhlin, M.
704 Steinert, Proteomic characterization of the whole secretome of *Legionella pneumophila* and functional
705 analysis of outer membrane vesicles, Infection and immunity 76 (2008) 1825-1836.
706 [73] A. Khemiri, A. Galland, D. Vaudry, P. Chan Tchi Song, H. Vaudry, T. Jouenne, P. Cosette, Outer-
707 membrane proteomic maps and surface-exposed proteins of *Legionella pneumophila* using cellular
708 fractionation and fluorescent labelling, Analytical and bioanalytical chemistry 390 (2008) 1861-1871.
709 [74] T. Hayashi, M. Nakamichi, H. Naitou, N. Ohashi, Y. Imai, M. Miyake, Proteomic analysis of growth
710 phase-dependent expression of *Legionella pneumophila* proteins which involves regulation of bacterial
711 virulence traits, PloS one 5 (2010) e11718.
712 [75] H.L. Andrews, J.P. Vogel, R.R. Isberg, Identification of linked *Legionella pneumophila* genes
713 essential for intracellular growth and evasion of the endocytic pathway, Infection and immunity 66
714 (1998) 950-958.
715 [76] G. Segal, M. Purcell, H.A. Shuman, Host cell killing and bacterial conjugation require overlapping
716 sets of genes within a 22-kb region of the *Legionella pneumophila* genome, Proceedings of the National
717 Academy of Sciences of the United States of America 95 (1998) 1669-1674.
718 [77] J.P. Vogel, H.L. Andrews, S.K. Wong, R.R. Isberg, Conjugative transfer by the virulence system of
719 *Legionella pneumophila*, Science 279 (1998) 873-876.
720 [78] B.C. Brand, A.B. Sadosky, H.A. Shuman, The *Legionella pneumophila* icm locus: a set of genes
721 required for intracellular multiplication in human macrophages, Molecular microbiology 14 (1994) 797-
722 808.
723 [79] K.H. Berger, R.R. Isberg, Two distinct defects in intracellular growth complemented by a single
724 genetic locus in *Legionella pneumophila*, Molecular microbiology 7 (1993) 7-19.
725 [80] A. Marra, S.J. Blander, M.A. Horwitz, H.A. Shuman, Identification of a *Legionella pneumophila*
726 locus required for intracellular multiplication in human macrophages, Proceedings of the National
727 Academy of Sciences of the United States of America 89 (1992) 9607-9611.
728 [81] H. Nagai, J.C. Kagan, X. Zhu, R.A. Kahn, C.R. Roy, A bacterial guanine nucleotide exchange factor
729 activates ARF on *Legionella* phagosomes, Science 295 (2002) 679-682.
730 [82] G.M. Conover, I. Derre, J.P. Vogel, R.R. Isberg, The *Legionella pneumophila* LidA protein: a
731 translocated substrate of the Dot/Icm system associated with maintenance of bacterial integrity,
732 Molecular microbiology 48 (2003) 305-321.
733 [83] W. Cheng, K. Yin, D. Lu, B. Li, D. Zhu, Y. Chen, H. Zhang, S. Xu, J. Chai, L. Gu, Structural insights
734 into a unique *Legionella pneumophila* effector LidA recognizing both GDP and GTP bound Rab1 in their
735 active state, PLoS pathogens 8 (2012) e1002528.
736 [84] I. Derre, R.R. Isberg, LidA, a translocated substrate of the *Legionella pneumophila* type IV
737 secretion system, interferes with the early secretory pathway, Infection and immunity 73 (2005) 4370-
738 4380.
739 [85] M.R. Neunuebel, Y. Chen, A.H. Gaspar, P.S. Backlund, Jr., A. Yergey, M.P. Machner, De-
740 AMPylation of the small GTPase Rab1 by the pathogen *Legionella pneumophila*, Science 333 (2011) 453-
741 456.
742 [86] S. Schoebel, A.L. Cichy, R.S. Goody, A. Itzen, Protein LidA from *Legionella* is a Rab GTPase
743 supereffector, Proceedings of the National Academy of Sciences of the United States of America 108
744 (2011) 17945-17950.

[87] E.M. Campodonico, L. Chesnel, C.R. Roy, A yeast genetic system for the identification and characterization of substrate proteins transferred into host cells by the *Legionella pneumophila* Dot/Icm system, Molecular microbiology 56 (2005) 918-933.
[88] S. Kalman, W. Mitchell, R. Marathe, C. Lammel, J. Fan, R.W. Hyman, L. Olinger, J. Grimwood, R.W. Davis, R.S. Stephens, Comparative genomes of *Chlamydia pneumoniae* and *C. trachomatis*, Nature genetics 21 (1999) 385-389.
[89] R.S. Stephens, S. Kalman, C. Lammel, J. Fan, R. Marathe, L. Aravind, W. Mitchell, L. Olinger, R.L. Tatusov, Q. Zhao, E.V. Koonin, R.W. Davis, Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis, Science 282 (1998) 754-759.
[90] B.D. Nguyen, R.H. Valdivia, Virulence determinants in the obligate intracellular pathogen *Chlamydia trachomatis* revealed by forward genetic approaches, Proceedings of the National Academy of Sciences of the United States of America 109 (2012) 1263-1268.
[91] R.J. Belland, G. Zhong, D.D. Crane, D. Hogan, D. Sturdevant, J. Sharma, W.L. Beatty, H.D. Caldwell, Genomic transcriptional profiling of the developmental cycle of *Chlamydia trachomatis*, Proceedings of the National Academy of Sciences of the United States of America 100 (2003) 8478-8483.
[92] T.L. Nicholson, L. Olinger, K. Chong, G. Schoolnik, R.S. Stephens, Global stage-specific gene regulation during the developmental cycle of *Chlamydia trachomatis*, Journal of bacteriology 185 (2003) 3179-3189.
[93] A.P. Maurer, A. Mehlitz, H.J. Mollenkopf, T.F. Meyer, Gene expression profiles of *Chlamydophila pneumoniae* during the developmental cycle and iron depletion-mediated persistence, PLoS pathogens 3 (2007) e83.
[94] M. Albrecht, C.M. Sharma, M.T. Dittrich, T. Muller, R. Reinhardt, J. Vogel, T. Rudel, The transcriptional landscape of *Chlamydia pneumoniae*, Genome biology 12 (2011) R98.
[95] M. Albrecht, C.M. Sharma, R. Reinhardt, J. Vogel, T. Rudel, Deep sequencing-based discovery of the *Chlamydia trachomatis* transcriptome, Nucleic acids research 38 (2010) 868-877.
[96] Y.M. Abdelrahman, L.A. Rose, R.J. Belland, Developmental expression of non-coding RNAs in *Chlamydia trachomatis* during normal and persistent growth, Nucleic acids research 39 (2011) 1843-1854.
[97] A.J. McCoy, A.T. Maurelli, Building the invisible wall: updating the chlamydial peptidoglycan anomaly, Trends in microbiology 14 (2006) 70-77.
[98] A.C. Shaw, K. Gevaert, H. Demol, B. Hoorelbeke, J. Vandekerckhove, M.R. Larsen, P. Roepstorff, A. Holm, G. Christiansen, S. Birkelund, Comparative proteome analysis of *Chlamydia trachomatis* serovar A, D and L2, Proteomics 2 (2002) 164-186.
[99] B.B. Vandahl, S. Birkelund, H. Demol, B. Hoorelbeke, G. Christiansen, J. Vandekerckhove, K. Gevaert, Proteome analysis of the *Chlamydia pneumoniae* elementary body, Electrophoresis 22 (2001) 1204-1223.
[100] P. Skipp, J. Robinson, C.D. O'Connor, I.N. Clarke, Shotgun proteomic analysis of *Chlamydia trachomatis*, Proteomics 5 (2005) 1558-1573.
[101] H.A. Saka, J.W. Thompson, Y.S. Chen, Y. Kumar, L.G. Dubois, M.A. Moseley, R.H. Valdivia, Quantitative proteomics reveals metabolic and pathogenic properties of *Chlamydia trachomatis* developmental forms, Molecular microbiology 82 (2011) 1185-1203.
[102] A.C. Shaw, B.B. Vandahl, M.R. Larsen, P. Roepstorff, K. Gevaert, J. Vandekerckhove, G. Christiansen, S. Birkelund, Characterization of a secreted *Chlamydia* protease, Cellular microbiology 4 (2002) 411-424.
[103] A.S. Hobolt-Pedersen, G. Christiansen, E. Timmerman, K. Gevaert, S. Birkelund, Identification of *Chlamydia trachomatis* CT621, a protein delivered through the type III secretion system to the host cell cytoplasm and nucleus, FEMS immunology and medical microbiology 57 (2009) 46-58.

[104] G. Zhong, P. Fan, H. Ji, F. Dong, Y. Huang, Identification of a chlamydial protease-like activity factor responsible for the degradation of host transcription factors, The Journal of experimental medicine 193 (2001) 935-942.
[105] C.J. Maier, R.H. Maier, D.P. Virok, M. Maass, H. Hintner, J.W. Bauer, K. Onder, Construction of a highly flexible and comprehensive gene collection representing the ORFeome of the human pathogen *Chlamydia pneumoniae*, BMC genomics 13 (2012) 632.
[106] R. Binet, A.T. Maurelli, Transformation and isolation of allelic exchange mutants of *Chlamydia psittaci* using recombinant DNA introduced by electroporation, Proceedings of the National Academy of Sciences of the United States of America 106 (2009) 292-297.
[107] Y. Wang, S. Kahane, L.T. Cutcliffe, R.J. Skilton, P.R. Lambden, I.N. Clarke, Development of a transformation system for *Chlamydia trachomatis*: restoration of glycogen biosynthesis by acquisition of a plasmid shuttle vector, PLoS pathogens 7 (2011) e1002258.
[108] Z.Q. Luo, R.R. Isberg, Multiple substrates of the *Legionella pneumophila* Dot/Icm system identified by interbacterial protein transfer, Proceedings of the National Academy of Sciences of the United States of America 101 (2004) 841-846.
[109] A. Subtil, C. Delevoye, M.E. Balana, L. Tastevin, S. Perrinet, A. Dautry-Varsat, A directed screen for chlamydial proteins secreted by a type III mechanism identifies a translocated protein and numerous other new candidates, Molecular microbiology 56 (2005) 1636-1647.
[110] K. Moelleken, J.H. Hegemann, The *Chlamydia* outer membrane protein OmcB is required for adhesion and exhibits biovar-specific differences in glycosaminoglycan binding, Molecular microbiology 67 (2008) 403-419.

Figures

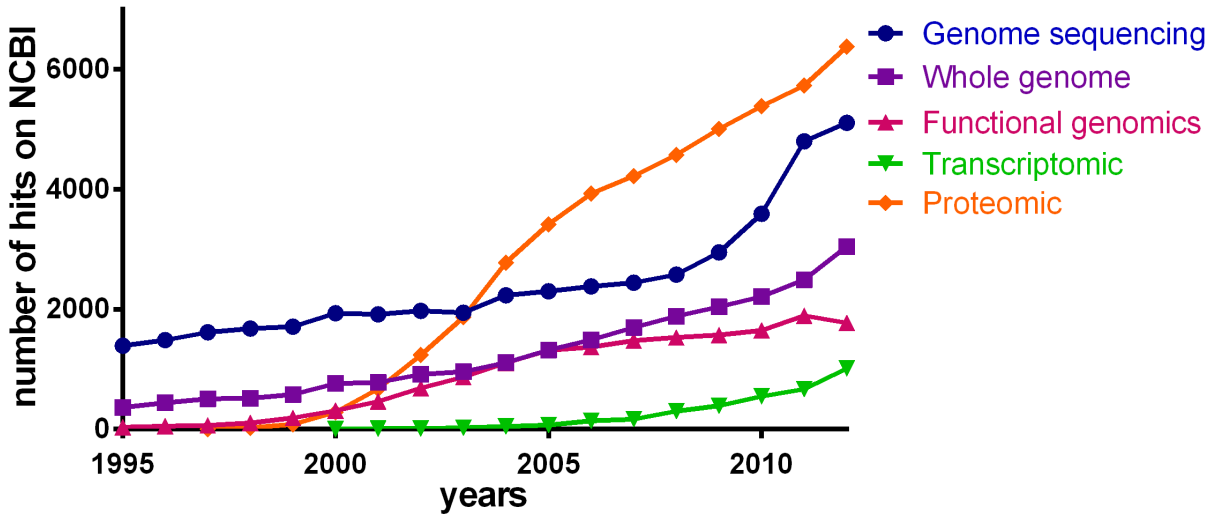

Figure 1. Graph showing the increasing number of hits when we made a NCBI research with particular key words. This reflects the expansion of the functional genomics during the last decade.

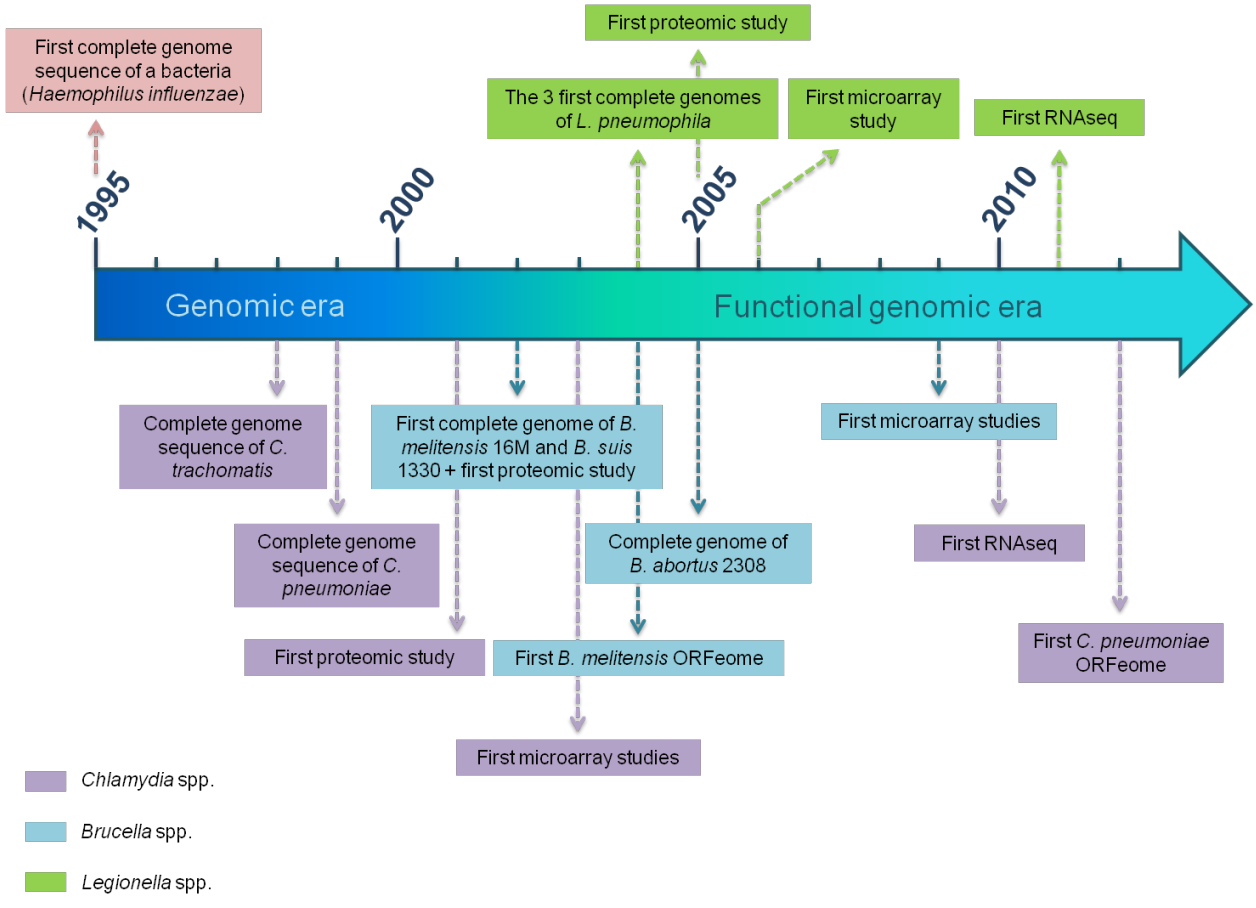

Figure 2. Timeline for *Chlamydia* spp., *Brucella* spp. and *Legionella* spp. during genomic and post-genomic era (functional genomics).

Table1. Advantages and limitations of various genomewide approaches to define genes function of *Legionella, Chlamydia* and *Brucella* spp.

	Advantages	Disavantages	*Brucella* references	*Legionella* references	*Chlamydia* references
Transcriptomics (microarrays)	Global transcriptional response	Not reflected in the proteomic level and based on genome annotation	[23, 27-29]	[62-64]	[91-93]
Transcriptomics (RNAseq)	Identification of RNA present in particular condition (mRNA, rRNA, ncRNA)	Need to distinguish host cell RNA (during analysis or before)		[67]	[94, 95]
Proteomics (2D gels+MS)	Identification of gene products present in particular conditions	Sensitivity issues, need to isolate bacteria from the host cells	[32-36]	[69-72, 74]	[98-103]
Translocation screens	Identification of putative effectors	No information on the putative function of the translocated proteins	[41, 42] *	[108]	[109]*
Yeast screens	Give clues on the function	Need confirmation in models of cellular infection	[17-19]	[60, 87]	[110]*

* not performed at the genome wide level