DOCTOR OF SCIENCES

Tn-seq on the pathogen Brucella abortus uncovers essential functions for culture and critical pathways for macrophages infection including pyrimidines biosynthesis

Sternon, Jean-François

Award date:
2017

Awarding institution:
University of Namur

Tn-seq on the pathogen *Brucella abortus* uncovers essential functions for culture and critical pathways for macrophages infection including pyrimidines biosynthesis

Jean-François Sternon

Dissertation presented in preparation of the degree of PhD in Sciences
Public defence

Jury members

Dr. Clayton Caswell, Virginia Tech, USA
Pr. Stephan Köhler, University of Montpellier, France
Dr. Géraldine Laloux, Catholic Univeristy of Louvain, Belgium
Pr. Jean-Jacques Letesson, University of Namur, Belgium
Pr. Thierry Arnould, President, University of Namur, Belgium
Pr. Xavier De Bolle, Promoter, University of Namur, Belgium

2017

Remerciements/Acknowledgements

With this first part of my acknowledgements section, I would like to honestly thank all of my jury members for taking the time to read my manuscript and provide me with helpful comments and interesting discussions. Your help was well appreciated and the kind atmosphere in which the private defence occurred only yielded constructive criticisms and will remain as a pleasant memory.

Je voudrais ensuite remercier Xavier De Bolle, mon promoteur, qui a toujours su donner à la recherche un aspect stimulant et humain et sans qui mon parcours n'aurait surement pas été aussi agréable. Je voudrais aussi remercier les PI du labo (Jean-Yves, Régis, Francesco, Eric) pour les discussions intéressantes que nous avons eu au cours de ces quatre années. Je voudrais aussi adresser un merci tout particulier à Mathieu, Françoise et Aurélie, sans qui la logistique du labo serait un véritable cauchemar !

Je tiens ensuite à remercier mes parents et ma soeur pour m'avoir supporté pendant ces 4 années, certains moments n'auraient pas été été évidents sans votre soutien. De plus il était quand même temps qu'on ait un veritable docteur dans la famille ! Au même titre, je tiens également à remercier Didi, tonton Jean-Luc et tonton Chris, eux seuls savent pourquoi.

En cette fin de parcours, j'ai une pensée pour toute particulière pour Mike, mon ex-tuteur depuis longtemps parti mais pourtant jamais oublié… Lui qui m'a tout appris (comme j'aimais le lui laisser croire). Sache que toutes mes pensées ont déjà été écrites dans mon mémoire, donc tu n'auras qu'à le relire… Bon allez, pour résumer, c'est vrai que tu m'as appris quelques trucs, mais pas que des trucs catholiques (je me demande si ELLE est au courant ?). Tout celà reste de bons souvenirs, car au fond, le temps fini par guérir les blessures.

Un petit mot maintenant pour mes compagnons de galère de l'URBM. Je pense tout particulièrement à ceux qui subissent, comme cette chère Katy qui, comme moi, a connu et connait quelques bons échecs successifs. Ce qui ne nous tue pas nous rend plus fort (ou nous pousse vers une longue depression aux accents suicidaires… mais ne pensons pas à ça !). Tu verras ça va aller, *la science, c'est robuste*! Je souhaite bon vent à Phuong, bonne fin de travail à Vicky, et bonne chance aux plus jeunes (Agnès, mais bien entendu Mathilde et Pierre). Vous qui avez été mes mémorants, bravo, vous avez eu beaucoup de chance ! Sachez que je suis fier de vous, car comme je vous le demandais souvent, vous n'avez jamais foiré (attention, tout n'est pas encore fini!). N'oubliez pas qu'il y a deux maîtres mots, perseverance et addiction ! Une dernière chose, Jimmy, trouve le bouton du fun pour le FRIA… De manière plus générale, je fais une spéciale dédicasse à mes (ex-)collègues, Bruce, Mr. Thib, Eme, Nayla, Arnaud, Gwen, Mr Jé, Mathieu, Kevin, Hubert, Gautier, les mémo Béa et Gangstalita, …(d'office je sais que j'en oublie, mais personne n'a été oublié volontairement !)

Aussi, le groupe, LE groupe… Cellar Twins (on a une page facebook). Vous verrez les gars, on sera bientôt des stars avec plein de thunes et de matériels! Vivement les roadies ! De plus Carl, sache que tes tips étaient bien faits… du travail de qualité!

On arrive lentement vers la fin, et comment mieux conclure ces remerciements qu'en mentionnant celle sans qui rien n'aurait été pareil. Celle qui a supporté ma mauvaise humeur et qui a su me redonner le sourire quand j'en avais besoin, Elodie. Merci beaucoup pour tout ce que tu as fait pour moi, et pour tout ce que tu feras encore !

Et finalement, à vous aussi, lecteurs, qui vous apprêtez par plaisir ou par obligation à lire ces quelques pages, je vous souhaite une bonne lecture!

Let's be a little more original than a simple citation, for all of you who like to crack codes, enjoy !

162813922536126317271194272795941332361263376764313641639121578495363710653
126938341367263212524512617941537426590270691603666631041361313339627632121
568250062743543159647201276195372145942729558566263383565094941148739249211
52136476333950638636222271181332761 81

Abbreviations list

AICAR	5-Aminoimidazole-4-carboxamide ribonucleotide
BCV	*Brucella* containing vacuole
BMDM	Bone marrow derived macrophage(s)
Bvr	*Brucella* virulence related
chr I	Chromosome I
chr II	Chromosome II
CFU	Colony forming unit(s)
DolK	Dolichol kinase
EEA-1	Early endosomal antigen 1
ER	Endoplasmic reticulum
ERES	Endoplasmic reticulum exit sites
gDNA	Genomic DNA
IGP	imidazole glycerol phosphate
IR	Inverted repeats
LAMP1	Lysosome associated membrane protein
LPS	Lipopolysacharride
NGS	Next generation sequencing
OMP	Outer membrane protein(s)
ORF	Open reading frame
*ori*I	Origin of replication of chromosome 1
*ori*II	Origin of replication of chromosome 2
PBP	Penicillin binding protein(s)
PG	Peptidoglycan
PI	Post-infection
PRPP	phosphoribosyl pyrophosphate
Tn-seq	Transposon sequencing
TTM	Transposon tolerance map
UPR	Unfolded protein response

Table of contents

Introduction

I. Brucellosis and *Brucella*

1) Discovery

At the end of the 19th century, a British army doctor named David Bruce based on Malta Island was confronted to various cases of soldiers affected by a comparable but yet unknown illness. The most commonly reported symptoms were flue-like afflictions such as long lasting undulant fever cycles associated with chills, fatigue, muscular pain and nocturnal sweating. In addition, the mortality rate was higher than the average observed for conditions causing similar symptoms (up to 2 %), thus suggesting a previously uncharacterized disease. By gathering and analyzing samples from infected individuals, David Bruce isolated the microorganism "*Micrococcus melitensis*" (later renamed *Brucella melitensis*) and was the first to set an actual causal link between Malta fever and the presence of the microbe (Theodorides, 1996).

2) Epidemiology

Due to its ability to mostly infect animals but also, to a lesser extent, humans, brucellosis is referred to as a zooanthroponosis. However, it should be noted that the infection of humans can be considered as an evolutionary dead end since no transmission from human to human seem to happen naturally (Moreno & Moriyon, 2006). Due to its worldwide distribution and associated economical losses, brucellosis is considered a major zooanthroponosis. However, while the number of new cases per year has been estimated at 500,000 by some authors (Pappas *et al.,* 2006), cases quantification remains difficult in numerous areas of the world because of census related issues such as the lack of symptoms specific to brucellosis coupled to the absence of *bona fide* methods for brucellosis diagnosis (Moreno & Moriyon, 2006). Thanks to the systematic diagnosis and subsequent herds culling in the case of infected animals, the majority of western Europe has been declared brucellosis free (Moreno & Moriyon 2006). Interestingly, it has been shown that brucellosis can be found in endemic wildlife, as exemplified by *B. suis* biovar 2 that is commonly isolated from wild boars (Wood *et al.,* 1976) and that was reported to cause infections to pigs from neighboring farms (Barlozzari *et al.,* 2015) but also to hunters that had prepared wild boar meat (Starnes *et al.,* 2004).

In animals, symptoms of brucellosis include long term septic arthritic lesions, sterility in males, and abortion of pregnant female, the latter being one of main routes of *Brucella* spreading (Moreno & Moriyon, 2006). In fact, the instinctive behavior of licking aborted fetuses in conjunction with the fetuses high bacterial load (*i.e* up to 10^{10} bacteria per g) strongly supports brucellosis dissemination throughout the herd. Similarly, the breathing of contaminated aerosols occurring during abortion as well as the consumption of infected milk constitute major spreading routes. Likewise, humans exposed to aerosols or unpasteurized

contaminated dairy products share the same contamination risks (Moreno & Moriyon, 2006). If left untreated, brucellosis will eventually reach a potentially lifelong chronic state characterized by occasional reoccurrences of previously described undulant fever episodes in addition to severe complications such as suppurative abscesses in various organs in about 10 % of the cases (Moreno & Moriyon, 2006).

3) Human brucellosis treatments and vaccines

To this day, the recommended treatment for brucellosis consists in multiple antibiotherapy. While numerous combinations have been tested, including the use of doxycycline or tetracycline in conjunction with rifampicin or with aminoglycosides such as gentamycin or streptomycin, the best results have been obtained using strong doses of tetracycline and streptomycin (*i.e*. doxycycline 100 mg twice/day for 45 days, and streptomycin 1 g/day for 14 days). However, it should be noted that even in the case of an adequate treatment, brucellosis relapses occur at a rate 3 to 5 % (Moreno & Moriyon, 2006).

Regarding prophylaxis, no efficient vaccine is currently available for preventing human brucellosis. For animal brucellosis, several vaccinal strains have been developed, the most commonly used being the live strains S19 and RB51 for *B. abortus* and Rev.1 for *B. melitensis.* Unfortunately, while their respective efficiency varies, providing protection in 75 % of cases at maximum for S19 (Nicoletti, 1990), these vaccines suffer from several major drawbacks. Indeed, all three vaccinal strains are still infectious for humans, and can be shed in the milk of their respective vaccinated hosts (Nicoletti, 1990; Uza *et al.,* 2000 ; Moreno & Moriyon, 2006). Moreover, due to their smooth lipopolysaccharide (LPS), the S19 and Rev.1 strains elicit serological responses, which in turn interfere with the genuine diagnosis of field brucellosis (Cunningham, 1977). For this reason, the rough *B. abortus* 2308 derivative RB51 was engineered, but due to its virulence, human infections can occur but cannot be diagnosed by typical serological tests. In addition, some of these strains are resistant to antibiotics used as treatments for brucellosis, such as streptomycin resistance for Rev.1 and rifampicin resistance for RB51. Taken together, these data strongly suggest that the development of vaccines against brucellosis require improvements in terms of both efficiency and safety of use.

4) The *Brucella* genus

4.a) Species

The *Brucella* genus includes 13 species, which are mainly distinguishable based on an epidemiological point of view. Indeed, despite their respective host specificity, those species are characterized by high genome identity of more than 90 % (Whatmore, 2009). As a side note, it should be noted that *Brucella* taxonomy is still a disputed and discussed matter as examined by Ficht (Ficht, 2010).

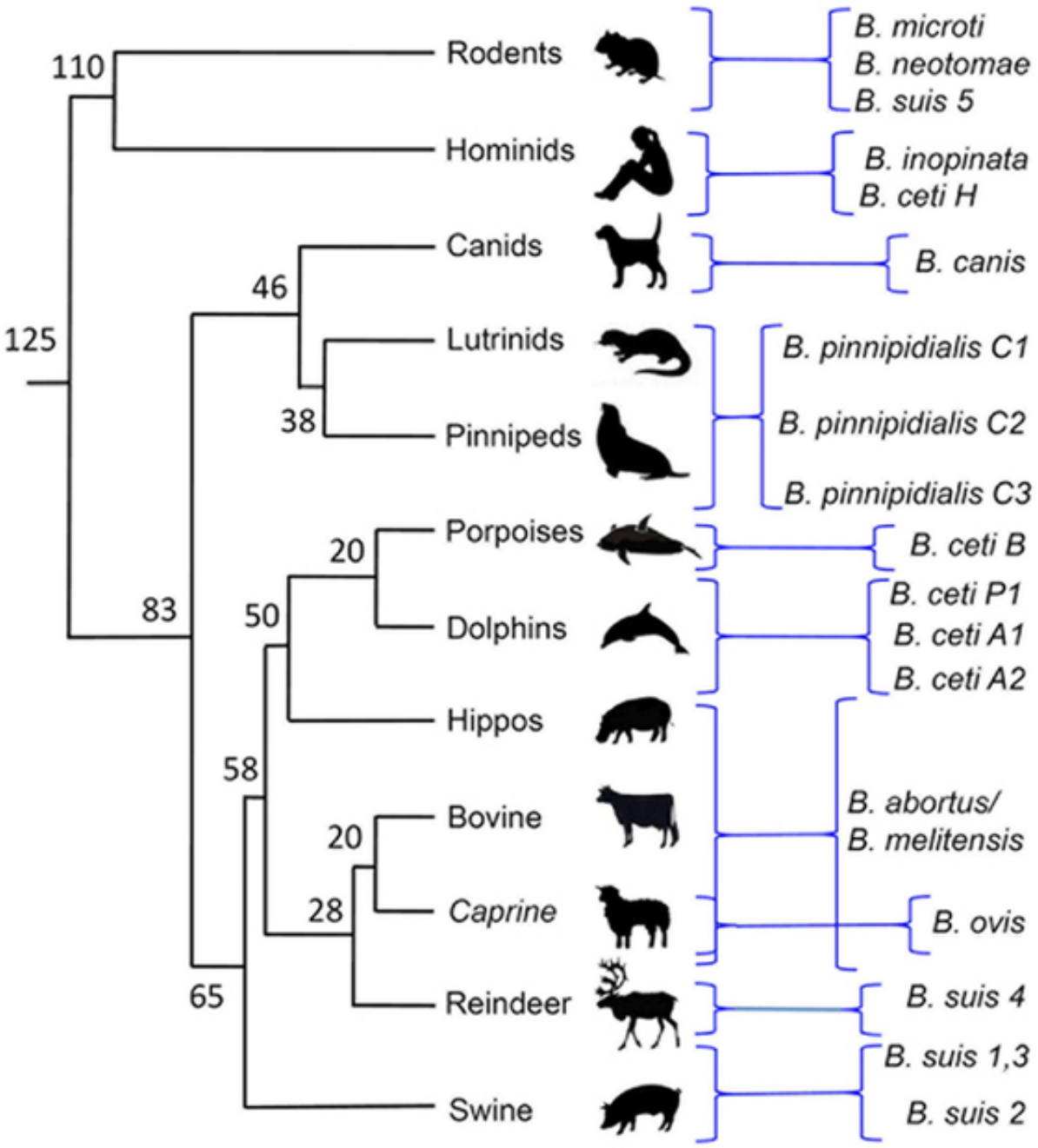


Figure 1. (Adapted from Guzmán-Verri *et al.*, 2012) Overview of *Brucella* phylogeny and host specificity. The numbers present on each tree branch represent a time estimation in millions of years of the *Brucella* genus dispersion.

The genus *Brucella* can be split in two parts. The first part includes 9 "typical" species that are assumed to have evolved as a single group (Wattam *et al.,* 2014), namely *B. melitensis, B. abortus, B. suis, Brucella ovis, Brucella canis, Brucella ceti, Brucella pinnipedialis, Brucella neotomae,* and *Brucella papionis*. Among these, only *B. melitensis, B. abortus,* and *B. suis* have been shown to be consistently able to infect humans, while such species respective hosts are caprines, bovines, and suids (but also wild boars, hares, reindeers, and rodents depending on the biovar)(Figure 1). The remaining species vary in host specificity, including several marine mammals as well as baboons, and are less pathogenic to humans (Whatmore, 2009). The second part of the *Brucella* genus composed of 4 additional "atypical" species and comprises strains displaying unusual growth (Al dahouk *et al.,* 2010) such as *Brucella microti* and *Brucella inopinata,* isolated from rodents and breast implants, respectively, as well as species for which little information is available due to their recent discovery, such as *Brucella vulpis* found in foxes (Scholz *et al.,* 2016) and unnamed *Brucella* species found in African frogs (Soler-Lloréns *et al.,* 2016). Interestingly, species isolated from frogs are motile as opposed to other *Brucella* species and show active proliferation inside macrophage-like murine J774-cells with a growth rate that is similar to fast-growing *Brucella* strains such as *B. microti*, and are able to colonize the liver and spleen of BALB/C mice (Al Dahouk *et al.,* 2017).

4.b) The bacterium

Bacteria from the *Brucella* genus are small Gram-negative cocobacilli ranging from 0.5 to 1.5 µm in size and belonging to the Rhizobiales order within the Alphaproteobacteria class. As a main feature of Alphaproteobacteria, *Brucella* shares the archetypal DivK-CtrA regulon (Curtis & Brun, 2014), which has been extensively characterized in the model Alphaproteobacteria *C. crescentus* due to its involvement in bacterial cell cycle regulation and in the asymmetrical development of *C. crescentus* daughter cells (Jacobs *et al.,* 2001 ; Matroule *et al.,* 2004). Briefly, the phosphorylation status of the response regulator DivK is sensed and transmitted through a complex phosphorelay to numerous interactants which in turn regulate the transcription, translation, post-translational modification and proteolysis of the master transcription factor CtrA (suggested reviews: Tsokos & Laub 2012; Kirkpatrick & Viollier 2011). Consequently, DivK indirectly regulates CtrA, thus dictating bacterial cell development. Interestingly, it has been shown that *Brucella* cell division also generates asymmetrical daughter cells (*i.e.* the small and the large daughter cells) (Hallez *et al.,* 2004), and that such asymmetry could also be traced to the subcellular localization of DivK-CtrA interactants such as the histidine kinase PdhS (Hallez *et al.,* 2007). However, it should be noted that no difference could be observed when comparing the small and the large daughter cells in infection (C. Van der Henst, M. Deghelt, C. Mullier, *unpublished*).

Brucella also shares characteristics commonly found among Rhizobiales, such as a multipartite genome (Slater *et al.,* 2009). In fact, the genome of *Brucella* is composed of two replicons, one of 2.1 Mb defined as chromosome I (chr I) and a second one of 1.2 Mb defined as chromosome II (chr II). Remarkably, while chr I possesses a typical prokaryotic chromosome replicon, chr II only harbours a plasmid-like repABC replicon (Paulsen *et al.,* 2002; Pinto *et al.,* 2012; Deghelt *et al.,* 2014). This observation, in conjunction with

additional genome structure and synteny analyses, suggests that chr II could have been acquired ancestrally as a megaplasmid which would have then been maintained due to selective pressure (Paulsen *et al.,* 2002; Slater *et al.,* 2009; Pinto *et al.,* 2012; Deghelt *et al.,* 2014). It is worth noting that the multipartie genome does not apply to all *Brucella* strains, as it has been shown that *B. suis* biovar 3 only possesses a single chromosome (Jumas-Bilak *et al.,* 1998). It has been demonstrated in a recent paper from our lab that the timing of chromosome replication initiation was chromosome specific (Deghelt *et al.,* 2014). Indeed, it was shown that the replication of chr I was steadily initiated prior to the replication chr II, and it was postulated that this could result in a synchronization of chromosome replication termination due to the difference in chromosome sizes (Deghelt *et al.,* 2014).

Another typical trait of Rhizobiales found in *Brucella* is its unipolar growth mode (Brown *et al.,* 2012). Indeed, it was shown in several Rhizobiales such as the plant pathogen *Agrobacterium tumefaciens,* the plant symbiont *Synorhizobium meliloti*, the soil bacterium and opportunistic pathogen *Ochrobactrum anthropi* and *B. abortus* that cell wall elongation is spatially confined to a single pole (Brown *et al.,* 2012). It should be noted that the only growth site of *Brucella* is its new pole, which is the pole generated through cell division (Brown *et al.,* 2012). This is different from the more classical models such as *Escherichia coli* and *Bacillus subtilis* where elongation relies on the incorporation of new cell wall material all along the cell (Burman *et al.,* 1983; Pooley *et al.,* 1978), or the midpoint incorporation through an equatorial ring as seen in *S. penumoniae* (Briles & Tomasz 1970). Remarkably, while most rod shaped bacteria possess two cell wall synthesis machineries dedicated to elongation and septation, respectively (*i.e.* the elongasome and the divisome), *Brucella* lacks major elongasome components such as *mreB, mreC, merD,* and *rodA* and possesses a single homolog for both PBP2 and PBP3 functions, *ftsI*. Consequently, it appears that *Brucella* solely relies on its divisome for both elongation and division. It should be noted that other pathogens are able to survive with a limited number of genes involved in peptidoglycan synthesis. For example, *S. aureus* only houses 9 genes involved in PG biosynthesis, 7 of which can be deleted without seemingly affecting growth and division *in vitro* including *mreB* and *mreC* (Reed *et al.,* 2015; Tavares *et al.,* 2015). As a side note, it is worth mentioning that due to their strong mechanistic similarities, it has been postulated that the archetypical bacterial elongasome could actually have been evolutionarily derived from the divisome, presumably by loosing functions such as inner membrane constriction (Szwedziak & Löwe 2013). Nonetheless, this hypothesis unlikely explains the absence of elongasome in *Brucella* since the related Alphaproteobacterium *C. crescentus* possesses homologs for *mreB, mreC, mreD,* and *rodA*. However, the absence of these genes in *Brucella* and other Rhizobiales such as *A. tumefaciens* and *S. meliloti* suggests that elongasome components may have been lost during Rhizobiales evolution.

5) Infectious cycle and cell cycle

5.a) Cellular infection

Thanks to infection of pregnant goats with *B. abortus* and subsequent electron microscopy imaging, it has been reported that in its natural host, *Brucella* replicates inside compartments derived from the endoplasmic reticulum (ER) (Anderson & Cheville 1986). For the sake of simplicity and convenience, infections have been typically performed using more classical infection models such as HeLa cells, bone marrow derived macrophages (BMDM), and RAW 264.7 macrophages, showing that *Brucella* was also able to reach and proliferate inside ER-like compartments (Pizarro-Cerda *et al.,* 1998; Celli *et al.,* 2003; Archambaud *et al.,* 2010), thus suggesting the suitability of such model cell lines for studying *Brucella* intracellular trafficking.

When infecting cultured host cells, *Brucella* undergoes a typical biphasic infection process composed of a first dormancy-like phase, characterized by the absence of bacterial proliferation and which varies in time depending on the host cell type, and a second phase displaying strong bacterial proliferation (Comerci *et al.,* 2001; Celli *et al.,* 2003).

During the non-proliferative phase of infection, the low number of infected cell (*i.e.* 5 to 10 % for HeLa cells and 10 to 20 % for RAW 264.7 macrophages) in conjunction with the low number of bacteria found inside host cells (*i.e.* one to two bacteria per infected cells) result in a small but constant bacterial load as shown by fluorescent microscopy imaging and stable colony forming units (CFU) values (Deghelt *et al.,* 2014). It is worth mentioning that in more aggressive cell types such as BMDM the tendencies depicted above vary. In the latter case, the non-proliferative phase also displays a strong and consistent bacterial load reduction clearing up to 90 % of the internalized bacteria (Celli *et al.,* 2003). The length of the non-proliferative phase varies depending on the host cell type and typically ranges from 5 to 6 hours in HeLa cells (Comerci *et al.,* 2001) and from 4 to 5 hours in RAW 264.7 macrophages (Celli *et al.,* 2003). Bacterial mechanisms promoting *Brucella* entry inside host cells are poorly understood. However, several adhesion mediating proteins or adhesins belonging to the autotransporter family (or type V secretion system) have been shown to promote adhesion and/or invasion of host cells. For example, in *Brucella suis*, the monomeric autotransporter protein BmaC has been demonstrated influence adhesion to HeLa and infection of A549 cells thanks to its ability to mediate adhesion to fibronectin (Posadas *et al.,* 2012). Similarly, the two trimeric autotransporters BtaE and BtaF have also been characterized as adhesion factors required for *B. suis* virulence, likely by mediating adhesion to other extracellular matrix components such as hyaluronic acid and collagen I (Ruiz-Ranwez *et al.,* 2013; Ruiz-Ranwez *et al.,* 2013), and the protein SP41 has been shown to mediate *Brucella* binding to HeLa cells, and more specifically sialic acid (Castañeda-Roldán *et al.,* 2006). Likewise, a fourth adhesin named BigA was identified as promoting adhesion of *Brucella abortus* to hepitelial cells (Czibener *et al.,* 2016).

Despite the absence of proliferation during this first phase of infection, it has been shown that the *Brucella* containing vacuole (BCV) undergoes numerous changes (for review:

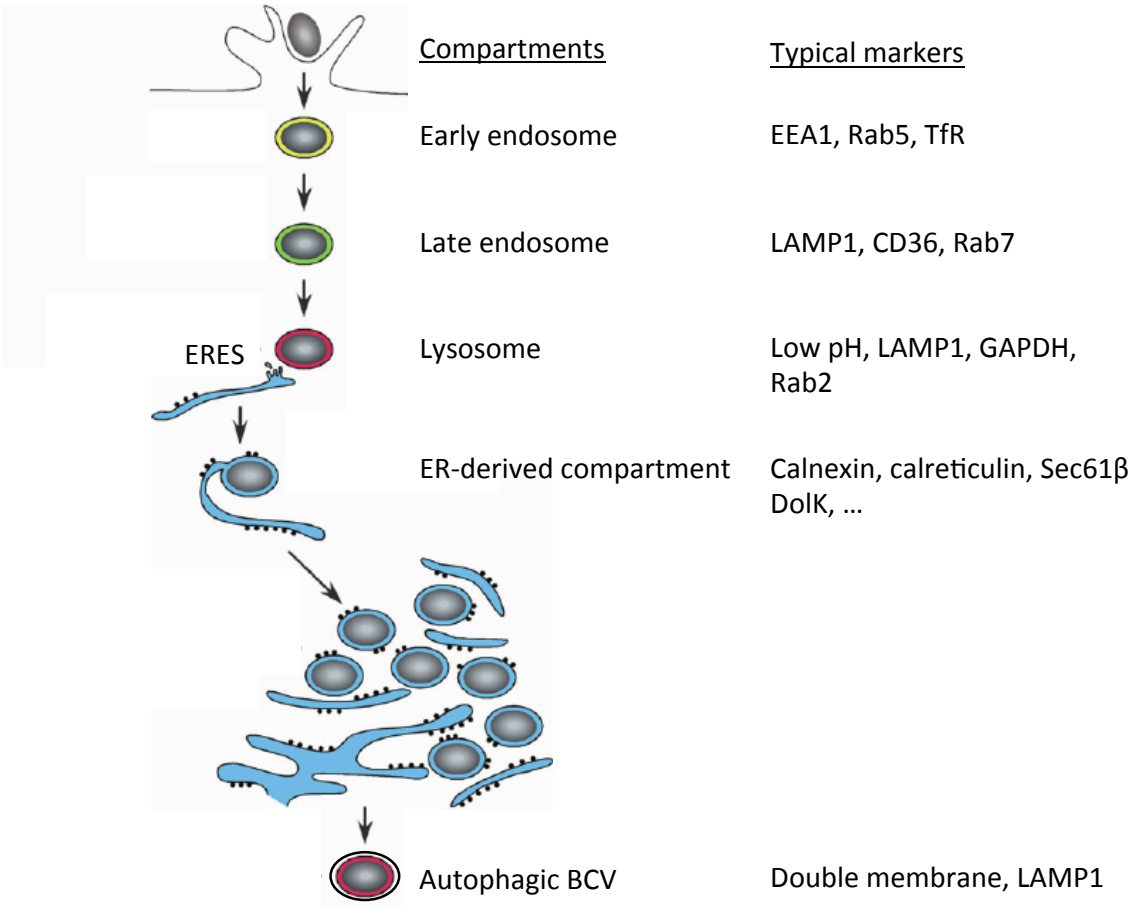


Figure 2. (Adapted from Starr *et al.* 2008) Overview of the *Brucella* intracellular trafficking displaying the different compartments interacting with the BCV along with typical markers of such compartments.

Celli, 2015). Upon bacterial internalization, which is suggested to be dependent on lipid raft (Watarai *et al.,* 2002), the BCV first interact with early endosomes and acquires typical early endosome markers such as EEA1, Rab5 and the transferrin receptor TfR prior to 1 h post-infection (PI) (Pizarro-Cerda *et al.,* 1998; Comerci *et al.,* 2001; Celli *et al.,* 2003)(Figure 2). The resulting endosomal BCV quickly maturates by acquiring markers associated to the late endosomes such as LAMP1, CD63 and Rab7 prior to 2 h PI and the intravacuolar pH then drops to 4.5 (Pizarro-Cerda *et al.,* 1998; Starr *et al.,* 2008). Interestingly, it has been shown in *B. abortus* that mutants in the two-component system encoding *bvrRS* operon were strongly attenuated at early PI time points in HeLa cells, peritoneal murine macrophages, and mice (Sola-Landa *et al.,* 1998). Moreover, the implication of this operon in *Brucella* outer membrane homeostasis and resistance to surface-targeted antimicrobial peptides was demonstrated, thus suggesting that this system could be involved in the extracellular to intracellular transition as well as early intracellular survival (Guzman-Verri *et al.,* 2002).

While the canonical fusion of the BCV to lysosomes is lethal for *Brucella*, the strict absence of lysosomal contacts prevents *Brucella* from accessing its replication niche, thus suggesting that controlled lysosomal interactions are mandatory for *bona fide* intracellular trafficking (Starr *et al.,* 2008). Concomitantly, the acidification of the BCV has been shown to induce the expression of the *virB* operon (Boschiroli *et al.,* 2002), a type IV secretion system translocating at least 15 effector proteins to the host cell (Ke *et al.,* 2015) and that is mandatory for virulence by enabling access the ER derived replication niche (Porte *et al.,* 1999*;* Sieira *et al.,* 2000; O'callaghan *et al.,* 1999; Comerci *et al.,* 2001; Celli *et al.,* 2003). Eventually, the BCV is able to reach the ER by interacting with host cell proteins associated with vesicular trafficking such as the glyceraldehyde-3-phosphate dehydrogenase and Rab2 (Fugier *et al.,* 2009), as well as the ER exit sites (ERES) through which *Brucella* is thought to enter the ER (Celli *et al.,* 2005) and become a proliferation permissive compartment also called replicative BCV. It has been suggested that the unfolded protein response (UPR) as well as autophagy could play a role in the biogenesis of the replicative BCV (de Jong *et al.,* 2013; Smith *et al.,* 2013; Taguchi *et al.,* 2015), as exemplified by the involvement of the autophagy related UPR sensor protein IRE1α (Qin *et al.,* 2008). However, the partial requirement of such pathways and the controversial nature of some of these data call for further investigations (Celli, 2015).

Once inside the ER, *Brucella* actively proliferates in a fashion that resembles *in vitro* culture. Expectedly, numerous metabolic mutants display attenuation in infection (Barbier *et al.,* 2011) examples being mutants involved in the biosynthesis of purines (Drazek *et al.,* 1995; Lestrate *et al.,* 2000; Köhler *et al.,* 2002), pyrimidines (Köhler *et al.,* 2002), histidine (Köhler *et al.,* 2002; Abdo *et al.,* 2011) as the as well as metallic ion uptake systems such as the zinc-specific *znu* system (Kim *et al.,* 2004). However, while *Brucella* access to the replication niche and further proliferation is becoming increasingly documented, the process by which the bacterium exits infected cells and spread to others is less clear (Celli, 2015). It has been recently shown that at late cellular infection time points (*i.e.* 48 h PI and beyond), some replicative BCV evolve to a multi-membrane compartment characteristic of autophagosomes that loose previous ER markers and acquire late endosome markers associated with autophagosome maturation such as LAMP-1 (Starr *et al.,* 2012). Moreover, it was demonstrated that the formation of such autophagic BCV was determinant for the

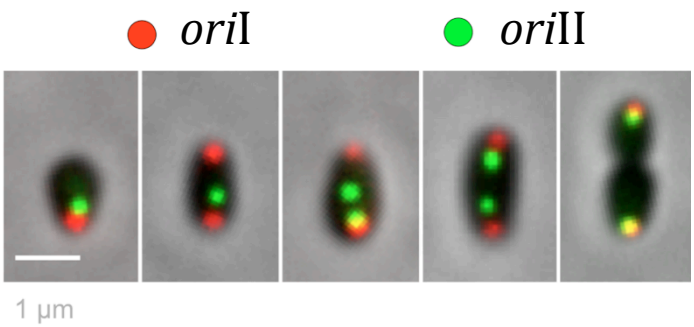


Figure 3. Characterization of *B. abortus* cell cycle. (Deghelt *et al.,* 2014) Fluorescent microscopy images after *in vitro* culture showing the consistent initiation of replication of chr I prior to chr II. *ori*I is stained in red, and *ori*II is stained in green.

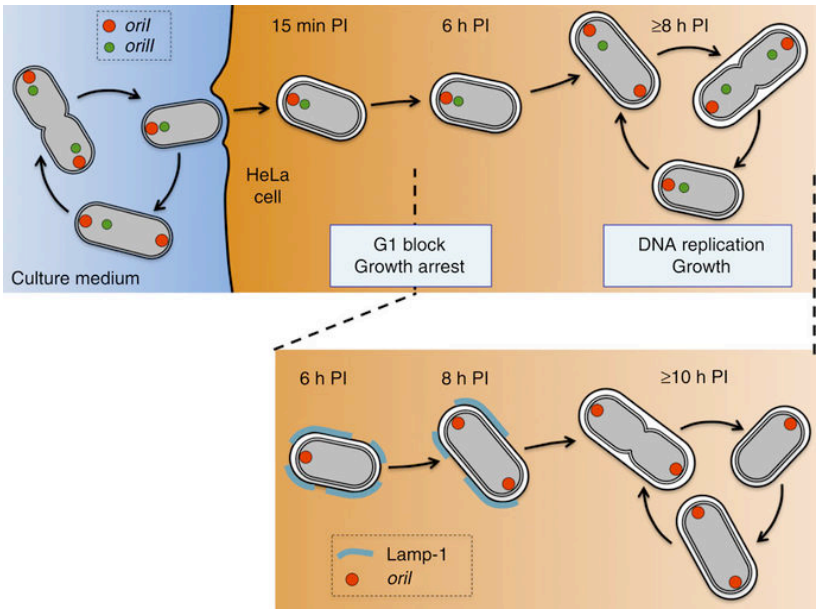


Figure 4. Schematic representation of the newborn bias during *B. abortus* cellular infection of HeLa cells. (Deghelt *et al.,* 2014).

infection of cells neighboring the primoinfected host cell, thus highlighting a plausible route for *Brucella* spreading (Starr *et al.,* 2012).

As a side note, it is worth mentioning that *Brucella* is commonly considered as a stealth pathogen (Moreno & Moriyon, 2006). In fact, it has been demonstrated in mice that *Brucella* infection typically results in the absence of leukocytes recruitment as well as low proinflammatory cytokines induction or complement activation and that the intracellular replication of *Brucella* does not seem to alter the cell cycle of infected host cells (Barquero-Calvo *et al.,* 2007). Interestingly, it has been postulated that the unusually low immunogenicity of *Brucella* could be dependent on inherent properties (Barquero-Calvo *et al.,* 2007) as exemplified by its characteristic LPS composition (Cardoso *et al.,* 2006) (Barquero-Calvo *et al.,* 2007). Indeed, while the LPS of most Gram-negative bacteria induces the activation of bactericidal mechanisms such as the production of reactive oxygen/nitrogen species and bactericidal peptides (Silverstein & Steinberg 1990 *Host defense against bacterial and fungal infections* Microbiology 4th ed 485-505), the LPS of *Brucella* has been shown to circumvent the activation of such mechanisms (Rasool *et al.,* 1992; Jiang *et al.,* 1993). In addition, reports have shown that the absence of the LPS O-chain, typically resulting in rough bacteria, leads to increased invasiveness and reduced intracellular survival (Detilleux *et al.,* 1990; Porte *et al.,* 2003). However, it should be stressed that recent evidence suggest that the role of LPS in host-*Brucella* interaction appears much more complex. In fact, it has been shown that the lipid A portion of *Brucella* LPS was able to induce cell death in neutrophils but not in lymphocytes in a dose dependent fashion (Barquero-Calvo *et al.,* 2015), thus highlighting diametrically opposed cell type specific effects of LPS exposure.

5.b) Cell cycle characterization

A recent study conducted in our laboratory characterized the cell cycle of *B. abortus* at the chromosome replication initiation level (Deghelt *et al.,* 2014). Briefly, by fluorescently labeling the origin of replication of each chromosome (*ori*I and *ori*II, respectively), it was possible to assess chromosome replication initiation at the single cell level by monitoring the duplication of replication origins (Figure 3). With these tools, it was observed that the chromosome replication initiation of *B. abortus* is a sequential and robust process in which the replication of chr I is steadily initiated prior to chr II. Accordingly, the cell cycle of *Brucella* was characterized by distinguishing three bacterial cell types based on their respective chromosome replication status and morphology: the "newborn" defined by the absence of chromosome replication initiation (*i.e.* a single copy of each chromosome per bacterium), the "intermediate" in which chromosome replication has been initiated, and the "predivisional" which displays signs of cytokinesis. Remarkably, when assessing the proportion of each cell types during *in vitro* culture and in infection, a strong selection bias in favor of the newborn cells was observed both in HeLa cells and RAW 264.7. Indeed, while the amount of newborn per population of *Brucella* roughly reached 25 % *in vitro* and during intracellular proliferation, this percentage rose up to 80 % during the non-proliferative phase of infection, showing that newborn bacteria are the preferentially infectious cell type in these infection models (Figure 4) (Deghelt *et al.,* 2014). It is worth mentioning that while the

reasons underlying such newborn selection are currently unknown, the rapid establishment of this bias (*i.e.* as early as 15 min PI) suggests the presence of structures mediating newborn-specific adhesion and/or internalization. Intriguingly, while bacterial proliferation occurs in the ER, both the initiation of chromosome replication and the resuming of bacterial growth were shown to occur in LAMP-1 positive compartments prior to accessing the ER. The same observation regarding growth has also been surprisingly observed in a Δ*virB* mutant which is inherently unable to reach the ER, suggesting that intracellular growth resuming is independent from genuine replication niche access (Deghelt *et al.,* 2014). Interestingly, it should be noted that the molecular tools described above are currently being used in order to characterize *Brucella* cell cycle during mice infection (Georges Potemberg, ongoing PhD thesis).

As a side note, it should be stressed that the kinetics of non-proliferative and proliferative phases of infection appear to be highly dependent on the host cell type. Indeed, in other cell line such as the JEG-3 trophoblasts, bacterial growth is seemingly reinitiated as early as 2 h PI (Thi Anh Phuong Ong, ongoing PhD thesis). Nonetheless, this is likely due to the fact *Brucella* intracellular trafficking strongly differs in this cell type, as it has been reported that *B. abortus* actively proliferates inside LAMP-1 positive acidic compartments in this host cell line (Salcedo *et al.,* 2013).

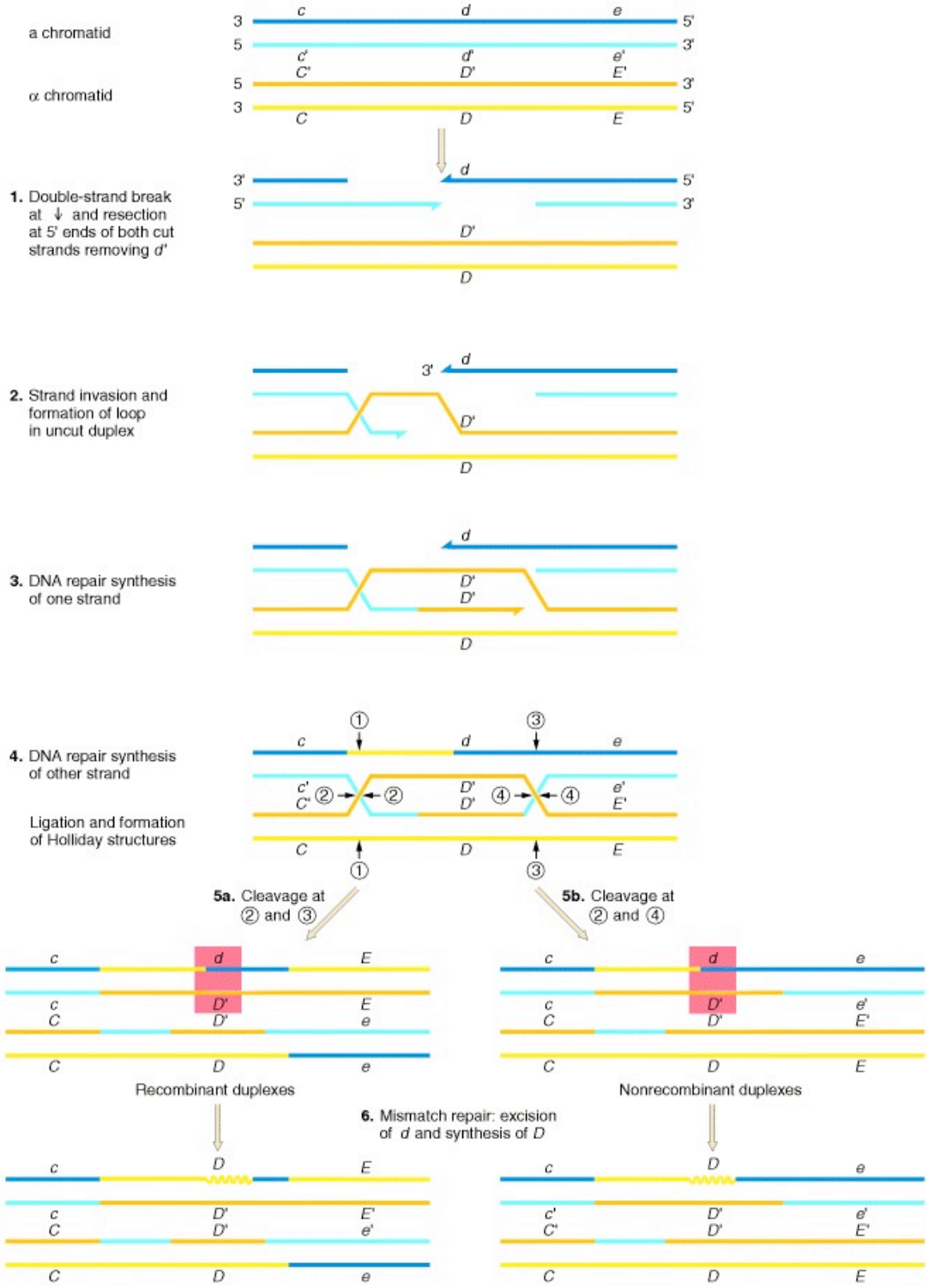


Figure 5. Double-strand-break repair model (From H. Lodish *et al.*, *Molecular Cell Biology*, 3rd ed. 1995). The two strands of each of two chromatids (in blue and yellow) are shown. The darker and lighter colors indicate complementary strands, as do the matching primes (for example, *C* and *C′* or *c* and *c′* for the allele designations). A key point of this model is shown in steps 3 and 4, which picture the generation of a Holliday structure with two crossover points. Two different ways of resolving this structure lead to the recombinant or nonrecombinant duplexes seen in step 5a and b. The red patch shows the heteroduplex mismatched regions, for the alleles *d* and *D′*. Mismatch repair results in the "conversion" of *d* into *D*.

II. Transpositional mutagenesis

1) DNA recombination

1.a) Homologous recombination

As demonstrated by Morgan in the beginning of the 20th century, large portions of genetic material can be exchanged within a single chromosome but also between different chromosomes. It was later discovered that this mechanism relies on the presence of homologous DNA sequences present at both sites of recombination. Homologous recombinations in eukaryotes and prokaryotes mainly occur as a repair mechanism for DNA damages such as double-strand breaks. Such damages are among the most severe types of damages that DNA can encounter and are defined as the concomitant cleavage of both phosphodiester backbones of a DNA double helix, thus resulting in a spatial separation of the two cleaved ends. Briefly, upon occurrence, DNA strands are locally denatured and hybridize on a homologous template for resynthesizing the damaged region. However, Such repair process involves the creation of one or several Holliday Junctions which subsequent sliding and resolution will dictate the occurrence and extent of genetic material transfer (Griffiths *et al.,* An Introduction to Genetic Analysis. 7th ed. Holliday model[1]) (Figure 5). In eukaryotes, while this type of rearrangements also serves beneficial purposes such as the generation of variability during meiosis, more problematic scenarios may also arise. Indeed, recombination can lead to genetic deregulation such as gene disruption and changes in expression levels which have been shown to be involved in the biogenesis of numerous types of cancer (Lengauer *et al.,* 1998; Bishop & Schiestl 2002).

In prokaryotes, besides its role in DNA damage repair, homologous recombination has also been shown to have key implications in pathogenesis, such as enabling phase variation, as demonstrated in *Neisseria gonorrhoeae.* Briefly, *N. gonorrhoeae* possesses phase variable pili which are surface exposed macromolecular structures promoting adhesion to host cells and modulating the phagocytosis and degranulation ability of neutrophils during infection (Punsalang & Sawyer, 1973; Johnson & Criss, 2011). To accomplish pili phase variation, *N. gonorrhoeae* uses two major pilin expression loci called *pilE1* and *pilE2* that undergo homologous recombinations with a reservoir of silent pilin encoding genes scattered in *N. gonorrhoeae* genome (Hagblom *et al.,* 1985; For review: Hills & Davies, 2009). Likewise, it has also been shown that the naturally occurring switching from pili-producing to non-piliated bacteria was dependent on the deletion of pilin producing genes originating from homologous recombination occurring within the *pilE1* and/or *pilE2* loci (Segal *et al.,* 1985).

[1] Available from: https://www.ncbi.nlm.nih.gov/books/NBK22099/

1.b) Site-specific recombination and mobile elements

Another example of DNA rearrangements able to promote mutagenesis could be the cases of site-specific recombination. Site-specific recombination can be compared to homologous recombination in the sense that it implies sequence homology for recombination to occur. However, it differs from the latter category given the relatively small size of needed homology sequences as well as the requirement of dedicated recombination machineries. For convenience, this category has been grouped with the mobile elements category as both share similar mechanics.

In prokaryotes, site-specific recombination ensures housekeeping functions such as chromosome dimer resolution. In fact, due to the circular nature of most prokaryotic chromosomes, the occurrence of an odd number of crossing overs during DNA replication results in chromosome dimers. Such events are deleterious for the bacterium as it implies an unequal chromosome repartition between the two daughter cells, and their frequent nature (*i.e.* estimated at 10 to 15 % of bacterial cell divisions in *E. coli*) requires a dedicated repair mechanism (Cornet *et al.,* 1996). Concisely, chromosome dimers are resolved thanks to the recruitment of the two site-specific DNA recombinases XerC and XerD at *dif* sites located at the chromosomes replication termination locus at the end of septation (Blakely *et al.,* 1993), while failure to resolve dimers results in cell division induced DNA damages (Blakely *et al.,* 1991).

The involvement of site-specific recombination and subsequent horizontal transfer in bacterial pathogenesis has also been documented, as exemplified by the *Shigella flexneri* "she" pathogenicity island. Briefly, this island consists in a set of laterally acquired genes involved in virulence and integrated in the 3' terminus of the phenylalanine carrying tRNA gene *pheV* (Al-Hasani *et al.,* 2001), and it has been demonstrated that the island can be precisely excised from its genome thanks to the recognition of a 22 bp sequence located both upstream and downstream of the island (Sakellaris *et al.,* 2004). Such excision is dependent on the action of a site-specific integrase *int* encoded within the pathogenicity island (Sakellaris *et al.,* 2004). Upon excision, the pathogenicity island acquires a circular conformation, and when transferred to *E. coli,* it can spontaneously integrate in *E. coli phe* tRNA gene (Sakellaris *et al.,* 2004). Similarly, it has been shown in *Yersinia pseudotuberculosis* that its high pathogenicity island can be excised from its *asn* tRNA locus thanks to the island-encoded integrase *int* and associated factor *hef* which recognize *attB*-R and *attB*-L sites located on both sides of the island (Lesic *et al.,* 2004).

Interestingly, the features developed above are highly reminiscent of latent phage infections, with numerous cases of pathogenicity islands displaying features typical of prophages (Schmidt & Hensel, 2004). It fact, besides their own genetic material, phages can transfer genetic information laterally from one bacterium to another by accidentally encapsidating a portion of host bacterium DNA in a process called transduction (Campbell & Reece *Biology* 7th ed. 379-380). Consequently, due to the enormous amount of phages present in the environment, the implications of phages as a driving force of evolution is widely accepted (Fuhrman, 1999; For reviews: Brüssow *et al.,* 2004; Stern & Sorek, 2011); Accordingly, while transduction relies on random DNA uptake and thus mostly account for poor fitness modulation, rare transduction events may result in the acquisition of genes

providing significant fitness increase, thus resulting in a strong positive selection of this event (Hacker & Carniel, 2001). Such postulates have been proven likely regarding bacterial pathogenesis, as numerous reports have shown that phages frequently carry genes that are not required for their own synthesis, some of which play an important role in bacterial virulence. One of the most famous examples of phage-mediated virulence is observed in the cholera-causing bacterium *Vibrio cholerae,* where the bacterium virulence is dependent on the presence of a phage carrying genes responsible for the synthesis of the cholera toxin (Waldor & Mekalanos, 1996)**.** Additionally**,** it should also be noted that the full virulence of *V. cholera* is also dependent on the expression of the toxin co-regulated pilus encoding genes, which are part of a horizontally acquired pathogenicity island (Faruke & Mekalanos, 2003). Another example could be the case of the plague causing bacterium *Yersinia pestis*, in which it was demonstrated that the bacterium full virulence was dependent on the presence of an integrated filamentous phage (Debrise *et al.,* 2007). Similarly, it was shown in *N. meningitidis* that the presence of an 8 kb prophage/genomic island is associated with the occurrence of hyperinvasive phenotypes that are common in cases of cerebrospinal meningitidis (Bille *et al.,* 2005). While the exact role of this prophage in *N. meningitidis* pathogenesis is still unclear (Meyer *et al.,* 2016), it is worth mentioning that it can still produce viral particles that can thus spread in the environment in a non-lytic fashion (Bille *et al.,* 2005; Meyer *et al.,* 2016). As a final example, it was also demonstrated in *Salmonella typhimurium* that several genes encoding type III secretion system effector proteins are carried by bacteriophages, as shown for the effector SopE (Mirold *et al.,* 1999).

As a side note, it should be stressed that as for all mobile elements, phages do not only import new functions but also create diversity through the disruption of (non-)coding sequences and the addition of homologous sequences that can be used for further genome remodeling (Brüssow *et al.,* 2004).

2) Transposon-mediated mutagenesis

2.a) Introduction to transposons

Transposons are mobile DNA sequences that can move from one locus to another within the same genome by a process called transposition and performed by an enzyme resembling site-specific recombinases called transposase. The discovery of transposons occurred as a result of investigations conducted in 1953 by B. McClintock aiming at understanding the causes of unusual color patches on maize leaves. Briefly, it was discovered that unusual colorations within individuals that had been self-pollinated for numerous generations were due to changes in genes positions, which were still believed to be statically anchored to chromosomes at the time (McClintock, 1953). Simplest transposons typically contain a transposase-encoding gene flanked by inverted repeats (IR) that are recognized by the transposase during transposition and thus delimitate the transposon. While transposon excision relies on site-specific recombination involving the recognition of both IR by the transposase, insertion of the excised transposon does not require sequence homology, making target sequence selection a random process (Campbell & Reece *Biology* 7th ed. 381-383).
Transposons have been demonstrated to confer evolutionary advantages through enhanced mutability (Chao *et al.,* 1983). Accordingly, as for the biogenesis of pathologically significant mobile elements, cases of transposons positively selected for conferring virulence associated fitness increase to their host have been described. For example, it has been shown that extensive local genome rearrangements in the fungal plant pathogen *Verticillium dahliae* strongly influence its virulence (de Jong *et al.,* 2013) and that such rearrangements are directly dependent on the presence of both active and passive transposons (Faino *et al.,* 2016). However, transposon-mediated virulence is not restricted to eukaryotes. In fact, it has been recently reported that shigellosis outbreaks are due at least in parts to multidrug resistance phenotype originating from the acquisition of Tn*7* transposons (Holt *et al.,* 2012; Kozyreva *et al.,* 2016). Likewise, another example could be the high proportion of transposons and transposon derivatives found within in the genome of *Neisseria spp.* as exemplified by the Correia element (Buisine *et al.,* 2002; Siddique *et al.,* 2011). Indeed, while this element is present in more than 100 copies in *Neisseria* genome and is likely used as a platform for recombinations and associated genome rearrangements, it has also been shown to encode strong promoters that are able to drive the expression of adjacent genes, thus further contributing to phenotypical diversity (Siddique *et al.,* 2011).

While numerous types of transposons have been described, the rest of this work will only address examples of non-replicative transposon (*i.e.* following a "cut and paste" transposition process) and associated mutagenesis applications.

2.b) Transposons as mutagens

A major contributing factor in the development of modern genetics screening was the use of transposons as mutagenic factors. Indeed, because of their ability to randomly insert

within a genome, transposon insertions can theoretically result in the disruption of any coding or non-coding sequences. The subsequent transposon-mediated disruption would then result in fitness alterations that vary in amplitude depending on the nature of the target sequence. One of the main advantages of transposon-mediated mutagenesis resides in the ease of localizing the mutated locus. In fact, while chemical mutagenesis requires the painstaking identification of phenotypically relevant point mutations, the use of a transposon of known sequence intrinsically provides a platform for subsequent mutated locus identification through degenerated PCR or direct sequencing. Another major advantage of transposon as mutagenic factors is their straightforward modifiability. Indeed, due to their DNA nature and simple functional architecture, genes can easily be swapped and artificial transposon (*e.g.* carrying selection markers such as antibiotic resistance cassettes) can be readily constructed by genetic engineering to fit the experimenter needs. Common examples include the use of modified Tn*5* and *mariner* transposons and will be presented hereunder.

The *mariner* transposon was discovered in 1981 by Bingham when investigating the causes of an unusual and unstable mosaic eye color phenotype observed in *Drosophila melanogaster* (Bingham *et al.,* 1981). Since then, numerous similar transposons were discovered and *mariner* grew to a large superfamily of transposable elements (Plasterk *et al.,* 1999). A common *mariner* transposon used in mutagenesis of eukaryotes is *Mos1* as initially demonstrated in the nematode *Caenorhabditis elegans* (Bessereau *et al.,* 2001; Williams *et al.,* 2005), with the example of pathogenicity related study at the host level by Yook and Hodgkin identifying genes involved in *C. elegans* response to the infection of the bacterial pathogen *Microbacterium nematophilus* (Yook & Hodgkin, 2007). As a side note, it should be noted that besides mutagenesis, *Mos1* has also been extensively used as a transformation vector for various eukaryotes, notably in the zebrafish (Fadool *et al.,* 1998), the chicken (Sherman *et al.,* 1999) and in the parasitic protist *Leishmania major* (Gueiros-Filho & Beverley, 1997). Regarding prokaryotes, a commonly used transposon in *mariner* mutagenesis is *Himar1* as non-exhaustively demonstrated in bacterial pathogens such as *H. influenzae* and *Streptococcus pneumoniae* (Arkerley *et al.,* 1999) as well as in *E. coli* and *Mycobacterium smegmatis* (Rubin *et al.,* 1999). Problematically, it should be stressed that *mariner* transposons suffers from strong transposition biases (*i.e.* hot spots and cold spots effect) (D. Mazel, *personal communication*) due at least in parts to the requirement of a TA di-nucleotide sequence for insertion (Plasterk *et al.,* 1999).

The Tn5 transposon was originally discovered by Berg and coworkers in 1974 as a naturally occurring transposon carrying aminoglycosides resistance genes such as kanamycin resistance (Berg, 2017). Tn*5* displays poor target sequence specificity and can be integrated in the genome of either eukaryotes or prokaryotes (de Bruijn & Lupski, 1984). Examples of Tn*5* and engineered Tn*5* derivatives used for mutagenesis are numerous and include investigations conducted on various bacterial pathogens, such as for example *Pseudomonas aeruginosa,* (Jacobs *et al.,* 2003), *N. meningitidis* (Geoffroy *et al.,* 2003), *Helicobacter pylori* (Salama *et al.,* 2004) and *Brucella suis* (Köhler *et al.,* 2002).

2.c) Limitations

Despite its advantages, the standard transpositional mutagenesis method is also characterized by two main technical limitations, the first one being transposition frequency. In fact, due to inherent toxicity of excessive transposition, naturally occurring transposons have been selected for low transposition frequency, thus impairing the straightforward generation of large mutant collections. Nevertheless, while scaling up the experimental setup can be a solution to some extent, the introduction of key point mutations within the transposase sequence was shown to result in increased transposase activity as shown in both *Himar1* (Lampe *et al.,* 1999) and Tn*5* (Zhou & Reznikoff, 1997). The second main problem associated with standard transpositional mutagenesis is the relationship between mutagenesis genome coverage and the time/labor required for individual mutant characterization. As a clarifying example, the generation of transpositional mutants for all genes of an organism has been demonstrated technically achievable in *Mycoplasma genitalium*, where 1354 transposon insertion sites were obtained for a ~580 kb genome composed of 517 genes (Hutchison *et al.,* 1999). However, such setup suffers from poor genome coverage. Indeed, genome coverage can be defined as the number of mutant per gene and is thus equivalent to 2.6 transposons per gene in the latter example when ignoring intergenic regions. Problematically, this low value implies increased statistical risks of experiencing genes fortuitously avoiding transpositional mutagenesis. Moreover, having only few transposon insertions per genes is also not sufficient since it has been demonstrated that even essential genes can tolerate transposon disruption when occurring close to genes ends (Christen *et al.,* 2011). Consequently, a perfect scenario would require having a highly saturating mutagenesis (*i.e.* with the highest possible genome coverage) so that the undesired statistical effects detailed above become negligible. However, the number of mutants needed to reach saturation requires extremely time and labor consuming downstream investigations in order to localize all mutated loci. Similarly, it should be stressed that comprehensive single-deletion mutant collections covering entire genomes have been constructed for several models species, including *S. cerevisiae* (Winzeler *et al.,* 1999) *Bacillus subtilis* (Kobayashi *et al.,* 2003) and *E. coli* (Baba *et al.,* 2006). However, one can easily have sense of the time and effort needed to accomplish such tasks. As a concluding remark, one can understand that a straightforward large-scale transposon mutagenesis requires technical modifications, which are presented in the Tn-seq section (see below).

3) Tn-seq

3.a) Principle

Transposon-sequencing (abbreviated Tn-seq) is a method initially developed by van Opijnen and coworkers in 2009 that enables straightforward saturating transposon mutagenesis by coupling the generation of large transpositional mutant libraries to next generation sequencing (NGS). Briefly, Tn-seq is based on the assumption that for each mutant genotype, the number of sequencing reads obtained post NGS directly dependends on the

Figure 6. Schematic representation of Tn-seq data. (Adapted from Christen *et al.,* 2011). In this genome segment where arrows symbolize genes, each vertical red line represents a transposon insertion site. Essential genes are shown in pink and are defined as regions where (almost) no transposon insertion can be found, suggesting the lethality of the corresponding mutated alleles. Consequently, Tn-seq highlights mutable genes, and the absence of insertions denotes essentialness.

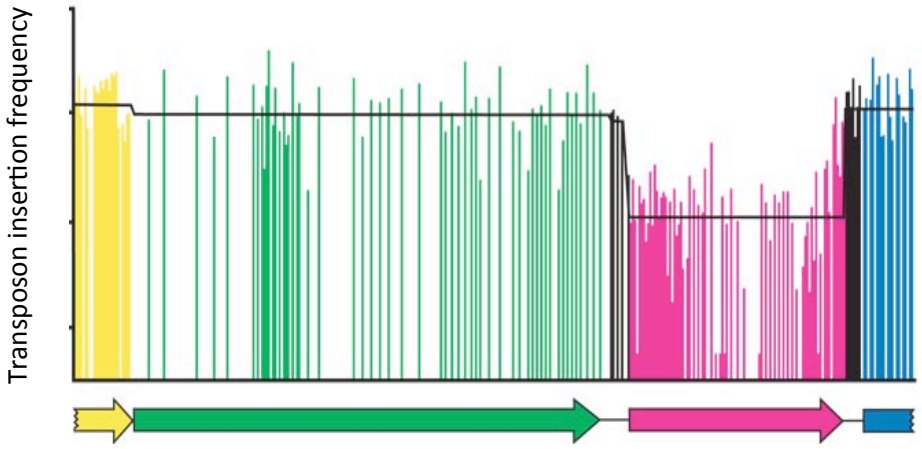


Figure 7. Schematic representation of the quantitative analyses potential of Tn-seq. (Adapted from van Opijnen *et al.,* 2009). In this genome segment, colored arrows symbolize genes and each vertical line represents a transposon insertion site. The height of each line corresponds to the number of sequencing reads found per insertion site and thus translates to the abundance/fitness of the corresponding mutants. The black line represents the global transposon insertion frequency per gene and shows that mutations in the magenta gene result in fitness reduction as opposed to the other genes found in this segment.

abundance of the corresponding mutant in the library, consequently correlating reads abundance to mutant fitness. Accordingly, due to its saturating characteristic, the absence of reads for a given locus can be translated to a fitness reduction instead of a fortuitous transposon absence. As a result, a gene for which numerous and diverse corresponding sequencing reads are found suggest that this gene could be readily mutated and therefore does not contribute to the bacterium fitness in the tested condition. Conversely, a gene for which no sequencing reads can be found suggests that this gene strongly contribute to the bacterium fitness as no mutated alleles can be found in the library (van Opijnen *et al.,* 2009) (Figure 6).

One of the main strength of Tn-seq is that it considerably reduces the amount of work needed for the determination of transposon insertion sites. Indeed, all mutants are pooled for genomic DNA (gDNA) extraction and NGS, resulting in the automated identification of all transposon insertion sites simultaneously, while previous method required the individual streaking and subsequent PCR analyses for each mutants (van Opijnen *et al.,* 2009). Moreover, the use of NGS as opposed to the classical Sanger sequencing allows for the parallel sequencing of several millions of templates in one run, making it appropriate for the straightforward sequencing of highly diverse samples such as gDNA pools obtained from Tn-seq mutant libraries. Additionally, as a result of its highly saturating nature, Tn-seq data enable quantitative analyses thanks to the assessment of the number of insertion events per insertion site (Figure 7), as opposed to the strictly qualitative data (*i.e.* presence or absence of transposon insertion) obtained from previous methods. Indeed, after NGS and downstream *in silico* analyses, all transposon insertion events can be computed to create a genome map displaying transposon insertion frequency. From this dataset, a mean transposon insertion frequency per genome can be calculated, where loci specific deviations from this value reveal mutagenesis-dependent fitness modulation (See section 8 of the Material & Methods).

One of the technical particularities of Tn-seq is that only one transposon insertion occurs per genome, thus meaning that Tn-seq only identifies non-redundant functions since no redundant gene can display fitness decrease due to functional rescue. Additionally, Tn-seq differs from anterior methodologies in the sense that mutants are lost during the gDNA extraction process and thus cannot be further used.

3.b) Overview of Tn-seq applications

Initially, Tn-seq aimed at identifying all essential genes of a bacterium in standard growth conditions as well as genes with high fitness cost when mutated. Examples include bacterial pathogens such as *S. pneumoniae* (van Opijnen *et al.,* 2009), *V. cholerae* (Kamp *et al.,* 2013), *Shigella flexneri* (Freed *et al.,* 2016), *Mycobacterium tuberculosis* (Zhang *et al.,* 2012) and *Porphyromonas gingivalis* (Klein *et al.,* 2012), and bacterial models such as *Caulobacter crescentus* (Christen *et al.,* 2011), *Agrobacterium tumefaciens* (Curtis & Brun, 2014), and *Brevundimonas subvibrioides* (Curtis & Brun, 2014). In essence, such data already yield a wealth of information regarding minimal gene requirement for survival and growth in bacterial models as well as the identification of essential genes as potential therapeutic targets in bacterial pathogens. However, Tn-seq is not restricted to the identification of either essential or low-fitness genes and can provide much additional information when

experimental conditions are compared. For example, Tn-seq allows the identification of genetic interactions. Indeed, by comparing Tn-seq data obtained from a wild type strain to data obtained from an isogenic background lacking a gene of interest, one can identify genetic partners of the deleted gene via synthetic lethality or synthetic fitness decrease (van Opijnen *et al.,* 2009; van Opijnen & Camilli, 2010). Likewise, this method can be used to discriminate between redundant genetic partners. Examples of genetic interaction identifications have been gathered from diverse bacterial pathogens such as *S. pneumoniae* (van Opijnen *et al.,* 2009), and *M. tuberculosis* (DeJesus *et al.,* 2017). Tn-seq has also proven to be a useful tool for assigning new gene functions. In fact, by comparing Tn-seq data obtained from different conditions, new gene functions can be uncovered, as demonstrated in *S. pneumoniae* where 17 different *in vitro* conditions have been compared, identifying more than 1,800 gene-condition interactions (van Opijnen & Camilli, 2012). In extension, this principle can be applied to the investigation of pathogenesis. In fact, by comparing data obtained from a standard *in vitro* growth condition to data obtain from a given infection condition, one can identify genes involved in virulence due to infection-specific fitness decrease, as initially demonstrated in *S. pneumoniae* recently (van Opijnen & Camilli, 2012). Comparable experiments have been performed using bacterial pathogens, such as *Moraxella catarrhalis* during pharyngeal and lung epithelial cells infection (de Vries *et al.,* 2013) and *Mycobacterium marinum* during THP1 monocytes infection (Weerdenburg *et al.,* 2015). In addition Tn-seq using animal infection models have been recently reported, such as mice infections using *Staphylococcus aureus* (Wilde *et al.,* 2015) and uropathogenic *E. coli* (Subashchandrabose *et al.,* 2013), as well as zebrafish infection using *Mycobacterium marinum* (Weerdenburg *et al.,* 2015). Taken together, these recent observations highlight the surge of a global interest in using Tn-seq for investigating bacterial virulence and consequently advocate its use in other bacterial pathogens.

Objectives

Recently, it has been established in *Brucella* that the newborn bacterial cell type, characterized by a 1n genome replication status, is the predominantly infectious cell type at early PI time points when infecting model cell lines such as HeLa cells and RAW 264.7 macrophages (Deghelt *et al.,* 2014). Due to the rapid occurrence of the newborn bias in infection (*i.e.* observed as early as 15 min PI) in conjunction with stable CFU counts (Deghelt *et al.,* 2014), it was postulated that newborns selection was unlikely to result from an intracellular selection process. Rather, it was hypothesized that newborn selection was dependent on bacterial factors anterior to cellular infection, such as the presence of newborn-specific adhesion and/or internalization promoting factors. Consequently, due to their recent identification in *Brucella,* four adhesins were investigated, namely *bmaC* (Posadas *et al.,* 2012)*, btaE* (Ruiz-Ranwez *et al.,* 2013a)*, btaF* (Ruiz-Ranwez *et al.,* 2013b)*,* and *bigA* (Czibener *et al.,* 2016). Unexpectedly, none of the reported attenuation phenotypes associated to the mutation of such adhesins could be reproduced in our hands (see section III.1 of the Results). As a consequence, it was decided to reorient the work for the remaining 32 months towards the identification of new virulence factors using Tn-seq as a method without *a priori*.

Tn-seq is a saturating random mutagenesis method that allows the identification of all genes essential to a given test condition in prokaryotes such as *in vitro* culture (van Opijnen *et al.,* 2009), and which can also be used to discover new virulence factors involved in bacterial pathogenesis (van Opijnen & Camilli, 2012). In our experimental setup, this method will be used, one the one hand, to identify all genes necessary for *B. abortus* to grow on rich medium, and on the other hand, to determine specific gene requirements at key PI time points (*i.e.* 2 h PI, 5 h PI, and 24 h PI) in RAW 264.7 macrophages. Moreover, as a result of its strong infection-specific fitness decrease, the pyrimidine biosynthesis pathway will be extensively investigated, generating new tools for the study of *Brucella* intracellular trafficking. In addition, a second Tn-seq will be performed, helping raise awareness regarding the impact of preexisting growth defects in virulence assessment as well as technical considerations associated to Tn-seq experiments design.

With this work, we hope to supply a comprehensive fundamental knowledge background regarding *Brucella* biology both *in vitro* and *in infectio* as well as bringing key insights into specific requirements underlying *Brucella* intracellular proliferation.

Results

I. Article under review

Remark : The article presented here is formatted as required for *PLoS Genetics* submission. For convenience, all supplementary tables and texts are only available as *in silico* files due to their large sizes. Table 1 may be found the end of section I.3 of the Results, p. 33-38.

Tn-seq on the pathogen *Brucella abortus* uncovers essential functions for culture and critical pathways for macrophages infection including pyrimidines biosynthesis.

Jean-François Sternon[*], Pierre Godessart[*], Rosa Gonçalves de Freitas[*], Mathilde Van der Henst[*], Katy Poncin[*], Nayla Francis[*], Kevin Willemart[*], Matthias Christen[†], Beat Christen[†], Jean-Jacques Letesson[*], Xavier De Bolle[*‡]

[*]Research Unit in Microorganisms Biology (URBM), Narilis, University of Namur, 61 rue de Bruxelles, 5000 Namur, Belgium
[†]Institute of Molecular Systems Biology, ETH Zürich, HPT E 71, Auguste-Piccard-Hof 1, 8093 Zürich, Switzerland
Short title : **Tn-seq on *Brucella abortus* in culture and during macrophages infection**
Classification : BIOLOGICAL SCIENCES, Microbiology
[‡]Corresponding author.
Prof. Xavier De Bolle
61 rue de Bruxelles
5000 Namur
Phone : +32 81 72 44 38
E-mail: xavier.debolle@unamur.be

1) Abstract and author summary

Brucella abortus is a class III zoonotic bacterial pathogen able to survive and replicate inside host cells, including macrophages. Here we applied a highly saturating mutagenesis (Tn-seq) to identify genes necessary for growth on rich medium as well as genes required for

the infection of RAW 264.7 macrophages, either for survival in the first step of intracellular trafficking, or for proliferation inside the endoplasmic reticulum. Our approach first allowed the functional annotation of conserved essential pathways for growth on plates. Time-resolved Tn-seq then enabled the discrimination between two categories of mutants, " aggravated mutants" which display a growth defect on plates that is just amplified in the course of infection, such as the purine biosynthesis pathway, and "attenuated mutants" which show decreased infectiosity while being able to grow as the wild type strain on plates. We found that about a third the previously reported *Brucella* attenuated mutants are actually "aggravated mutants". Nine genes generated increased invasion potential when mutated and all are predicted to be involved in lipopolysaccharide O chain synthesis. Besides virulence factors previously characterized such as the VirB secretion system, we identified 30 genes involved in early survival and 48 required for intracellular growth. In particular, we report the need of an intact pyrimidine nucleotides biosynthesis pathway in order for *B. abortus* to proliferate inside macrophages, but not for anterior infection steps or *in vitro* culture. Indeed, a *B. abortus pyrB* mutant is still able to reach the ER in HeLa cells, suggesting that its intracellular trafficking remains intact. We thus propose that hyper-saturating Tn-seq performed in a time-resolved manner is a useful strategy to identify genes required at different steps of an infection by a bacterial pathogen. In addition, we also advocate the use of our R200 analysis methodology as it strongly reduces analyses biases due to the non-requirement of anterior gene prediction and allow for deeper analyses such as essential domain mapping.

Establishing gene functions in pathogens yields a strong background knowledge that is fundamental for the development of efficient treatments and/or vaccines. Unfortunately, such investigations are typically tedious and time-consuming, especially for pathogens requiring a particular containment like the class III bacterial pathogen *Brucella abortus*. Here we applied a genomewide transposon-mediated mutagenesis strategy (Tn-seq) to readily identify all essential genes in a standard culture condition as well as genes required for survival and growth in macrophages. In addition, further investigations demonstrated that *B. abortus* needs a complete pyrimidine biosynthesis pathway to proliferate inside host cells, but not for anterior intracellular trafficking. Thanks to its straightforward setup and high throughput characteristics, this approach could be applied to various other pathogens and infection models in the future. In addition, the R200 analysis methodology presented in this work allows for performing Tn-seq even in models with poor genome annotations.

2) Introduction

Brucella abortus is a class III bacterial pathogen from the *Brucella* genus known for being the causative agent of brucellosis, a worldwide anthropo-zoonosis generating major economical losses and public health issues (Moreno & Moriyon, 2006). These bacteria are Gram-negative and belong to the Rhizobiales order within the *Alphaproteobacteria* group and share common characteristics such as the DivK-CtrA regulation network which governs cell

cycle regulation (De Bolle *et al.,* 2015) and unipolar growth as observed in *Agrobacterium tumefaciens* or *Sinorhizobium meliloti* (Brown *et al.,* 2012), respectively. Another feature of *B. abortus* common among Rhizobiales is its multipartite genome, composed of two replicons of 2.1 and 1.2 Mb named chromosome 1 (chr I) and chromosome 2 (chr II), respectively (Michaux-Charachon *et al.,* 1997). The chr I replication origin is similar to the one of *C. crescentus*, while chr II replication origin resembles those found in megaplasmids of the Rhizobiales (Deghelt *et al.,* 2014; Pinto *et al.,* 2012). A main aspect of *B. abortus* infections is the ability of the bacteria to invade, survive, and proliferate within host cells, including macrophages. Recently, the cellular infection process of *B. abortus* in both RAW 264.7 macrophages and HeLa cells has been extensively characterized at the single cell level in terms of growth and genome replication, highlighting a typical biphasic infection profile (Deghelt *et al.,* 2014). Indeed, during the course of cellular infections, *Brucella* first enters host cells through the endosomal pathway where it remains in a "*Brucella* containing vacuole" (BCV) for several hours without proliferating while preventing the maturation of its compartment into phagolysosomes (Celli, 2015). During that period, typical markers such as LAMP1 are acquired (Celli, 2015). In most cell types, surviving bacteria are able to maturate their BCV to an endoplasmic reticulum (ER) derived compartment where they actively proliferate (Detilleux *et al.,* 1990). This maturation is dependent on the type IV secretion system called VirB (Comerci *et al.,* 2001; O'Callaghan *et al.,* 1999). The chemical composition of the replicative BCV is difficult to study directly, but mutant *Brucella* strains may be used as probes to gain a better knowledge of the bacterial environment in these compartments. Screening of transpositional mutants collections for attenuated strains generated hypotheses, such as the availability of histidine that was proposed to be limited (Kohler *et al.,* 2002), which is consistent with the ability of histidinol dehydrogenase inhibitors to impair growth of *Brucella* in macrophages (Abdo *et al.,* 2011). However, a major drawback of previous screenings for attenuated mutants was the size of the library, typically limited to a few thousands of mutants, which does not allow saturation at the genome-wide level and therefore cannot yield quantitative results.

In the present study, we have performed Tn-seq on *B. abortus* both before and after the infection of RAW 264.7 macrophages using a highly saturating library of approximately 3.10^6 mutants. This approach first allowed the identification of genes involved in several essential processes for growth on rich medium. In addition, three post-infection (PI) time points were then chosen in order to gain insights into genes required for *B. abortus* at discrete steps of a cellular infection. Accordingly, crucial factors involved in survival during the endosomal stage, as well as consistent pathways required for initial growth in the ER of host cells have been identified, further deciphering the content of *Brucella* replication niche, and supporting the use of time-resolved Tn-seq for dissecting infection processes in different pathogens and various infection models.

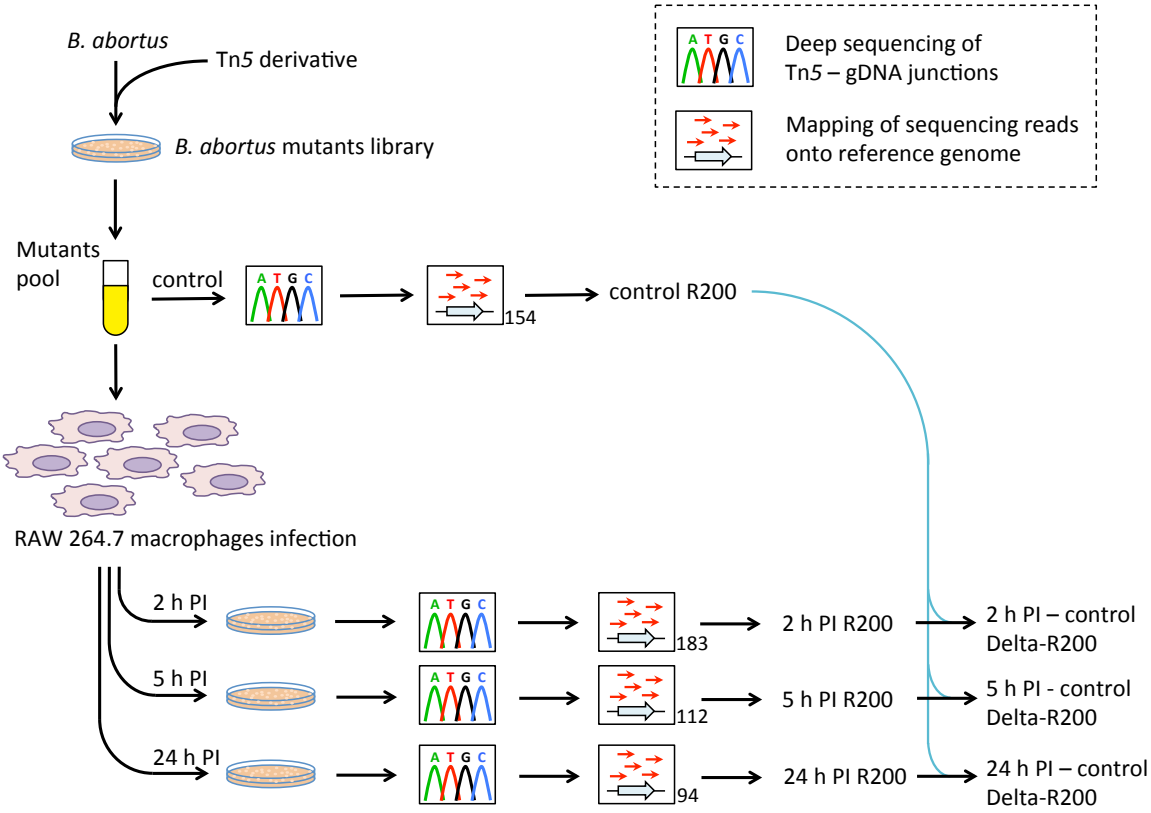


Figure 8. Summary of the Tn-seq approach. A transpositional mutant library was initially created in *B. abortus* on plates using a mini-Tn*5* derivative. Three millions mutants were then pooled and the resulting suspension was split in two. One half of the pool underwent direct sequencing of Tn*5*-gDNA junctions, allowing the identification of mini-Tn*5* insertion sites by mapping on the genomic sequence, resulting in the control dataset. The second half of the pool was used to infect RAW 264.7 macrophages in three separated infection events. At given time points post-infection (2 h, 5 h and 24 h PI), bacteria were extracted, grown on plates, colonies were collected and their gDNA was subsequently extracted to be sequenced as for the control, resulting in 2 h PI, 5 h PI and 24 h PI-specific datasets. Transposon tolerance maps are in the form of R200 values, and all post-infection lists were separately compared to the control list using the Delta-R200 method (see Material & Methods). The number of mapped read (in millions) for each dataset is displayed besides its respective mapping icon, and the number of insertion sites is 929 10^3 for the control condition, and 742 10^3, 713 10^3, and 579 10^3 for the 2 h, 5 h and 24 h PI datasets, respectively.

3) Results

3.a) Identification of essential genes in *B. abortus* 2308

A *B. abortus* 2308 library of 3.10^6 random transpositional mutants was grown on rich medium and transposon insertions sites were identified by deep sequencing (Figure 8). We identified 929,769 insertion sites from 154,630,306 mapped reads, saturating the *B. abortus* genome with an insertion site every 3.5 bp on average. To allow a genomewide scanning analysis method independent of gene annotations and predictions, we created a parameter assessing transposon insertion frequency termed R200 (see Experimental procedures), equal to the $\log_{10}$ of the number of transposon insertions + 1 found within a 200 bp sliding window (Figure 9). The average R200 for the two chromosomes was 3.47 in this condition, illustrating that multiple insertions at the same site in the sequence were taken into account in the computing of R200.

We considered as essential all genes where at least one R200 value was equal to 0 in the control condition (growth on rich medium), since the probability of such events to happen randomly was estimated to be approximately 4.10^{-15} (see Experimental procedures) (Christen *et al.,* 2011). In order to test the validity of this analysis, we checked that genes required for supposedly essential processes were indeed scored as essential, if they do not have functional paralogs. As expected, genes coding for all four RNA polymerase core subunits (α, β, β' and ω), the housekeeping σ^{70}, and 51 out of the 54 ribosomal proteins encoding genes were scored essential. Additionally, the previously established essentiality of *pdhS*, *ccrM*, *omp2b*, *divK, cckA* and *chpT* genes was also confirmed (Hallez *et al.,* 2007; Laloux *et al.,* 2010; Robertson *et al.,* 2000; Van der Henst *et al.,* 2012; Willett *et al.,* 2015).

Out of the 3419 predicted genes annotated on the *B. abortus* genome, 491 genes were found to be essential for *in vitro* culture, *i.e.* 14.4% of the predicted genes. This percentage is in agreement with those previously reported for other *Alphaproteobacteria* such as *C. crescentus* (12.4%), *Brevundimonas subvibrioides* (13.4%), and *Agrobacterium tumefaciens* (6.9%) (Christen *et al.,* 2011; Curtis & Brun, 2014). The transposon tolerance map for *B. abortus* genome is available as Supporting Information. A list of all essential genes for *in vitro* culture is also available as Supporting Information (S1 Table).

Based on the presence of a plasmid-like replication and segregation system on chr II and differences in gene content, it has been postulated that chr II could originate from an ancestrally acquired megaplasmid (Deghelt *et al.,* 2014; Paulsen *et al.,* 2002). We thus tested the distribution of essential genes between the two chromosomes of *B. abortus*. Accordingly, 429 out of the 2236 genes (19%) of chr I were essential. This is 3.7 times more than the 5 % found on chr II with 62 essential genes out of 1183. This result further supports the megaplasmid hypothesis. One can thus hypothesize that essential genes have started to be transferred from chr I to chr II but that the frequency of this transfer was not sufficient to equilibrate the proportion of essential genes on both chromosomes yet.

Tn-seq allows for the reconstruction of essential pathways. Here, we specifically focused on pathways relative to *B. abortus* cell cycle, which will be divided into four categories, the replication of DNA, the growth of the envelope, the cell division, and the cell

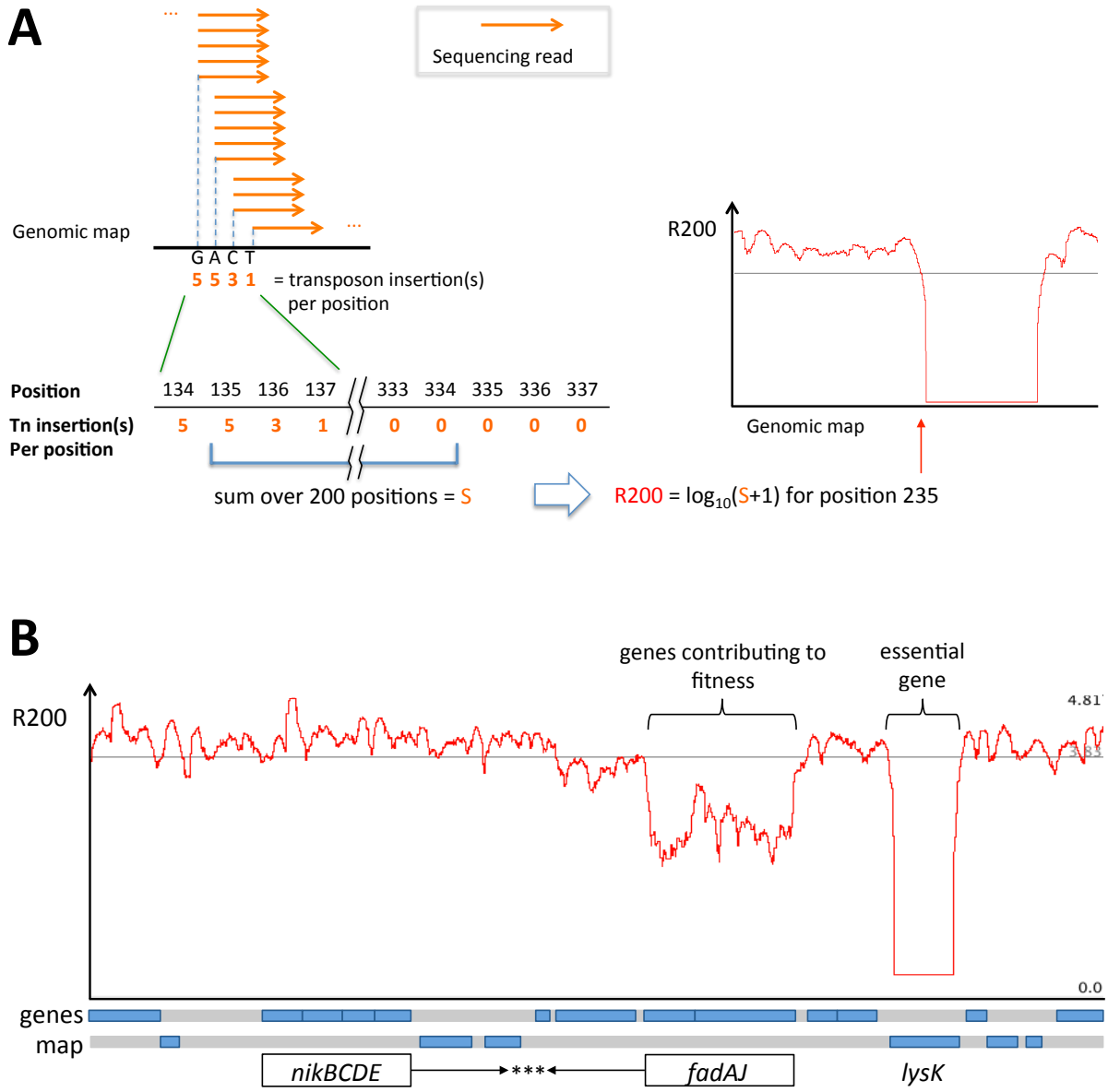


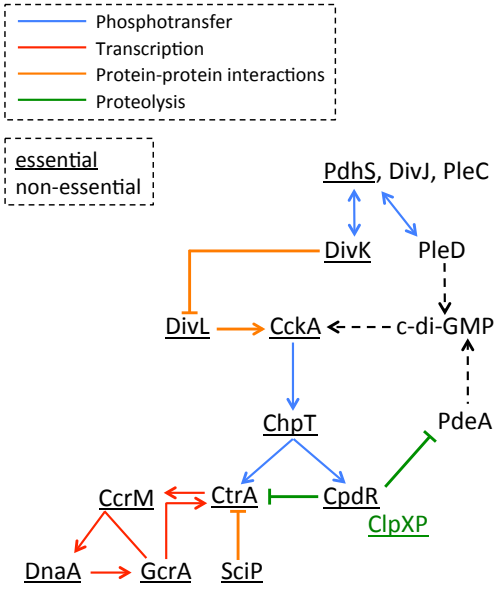
Figure 9. R200 computing and genome mapping. (**A**) Sequencing reads allow the identification of mini-Tn*5* insertion sites, and the number of reads at each insertion site is computed. A score (S) is computed by summing the number of reads aligned in a region of 200 bp. This score is used to calculate R200 and the R200 values are aligned to the map of predicted genes of each chromosome. (**B**) Example of a region in which non-essential genes such as the *nikBCDE* operon are found close to genes contributing to fitness (*fadAJ* operon). In the same region, an essential gene (*lysK*) is also identified. The R200 values of *nikBCDE* and *fadAJ* operons are statistically different ($p < 10^{-12}$ in a t test with independent samples, n = 16 non-overlapping windows for each operon).

cycle regulation network. Regarding the replication of DNA, α, β, γ, δ, χ and τ subunits of the DNA polymerase III as well as the helicase *dnaB* and the primase *dnaG* are essential for growth. None of the three ε-subunits (BAB1_2072, BAB2_0617, and BAB2_0967) were scored essential, which is likely due to functional redundancy. Genes responsible for the initiation of DNA replication for chromosome I (*dnaA*) and for chromosome II (*repC*) and for the segregation of chromosome I (*parA* and *parB*) and chromosome II (*repA* and *repB*) were also essential. Interestingly, only one homolog of the structural maintenance of chromosome gene (*smc,* BAB1_0522) could be found in the genome and was not essential in our Tn-seq, either suggesting that a functional analog is present, or that this function is not essential in *B. abortus*. Genes responsible for the synthesis of peptidoglycan precursors from fructose-6-phosphate and their export to the periplasm (namely, *glmS*, *glmM*, *glmU*, *murA*, *murB*, *murC*, *murD*, *murE*, *murF*, *murG*, *mraY*, *ftsW*) were all scored essential. Interestingly, out of the three predicted class-A penicillin-binding proteins (PBP), only one (BAB1_0932) was found to be essential for growth on rich medium as well as the only predicted class-B PBP, *ftsI*. The two other class A PBP were either only needed for optimal growth on rich medium while not being strictly essential (BAB1_0607) or seemingly not required at all (BAB1_0114).

The entire pathway responsible for the synthesis of lipopolysaccharide (LPS) lipid A from UDP-GlcNAc (namely, *lpxA*, *lpxB*, *lpxC*, *lpxD*, *lpxK*, *lpxXL* and *kdtA*), as well as the pathway responsible for its export to the outer membrane (namely, *msbA*, *lptA*, *lptB*, *lptC*, *lptD*, *lptE*, *lptF*, *lptG*) appeared essential for growth on rich medium. Conversely, no gene known for being involved in the LPS core synthesis (Fontana *et al.,* 2016) was scored essential for culture on plates. Moreover, none of the genes responsible for the LPS O-chain synthesis were scored essential, with the exception of *wbkC* (see Discussion).

Genes involved in the export of outer membrane proteins (OMPs) such as *bamADE* were essential for growth on rich medium (see Discussion). However, no predicted homolog of *bamB* nor *bamC* could be found. Interestingly, none of the three homologs of the OMPs periplasmic chaperone *degP* were scored essential, which is probably due to functional redundancy. It should be noted that no clear predicted homolog of the OMPs periplasmic chaperones *skp* and *surA* could be identified *in silico*. Regarding cell division, genes coding for the divisome proteins *ftsZ*, *ftsA*, *ftsQ*, *ftsK*, *ftsW*, *ftsY*, *ftsI*, as well as for the outer membrane invagination system *tolQRAB-pal* were all essential on plates. The cell division regulatory operon *minCDE* as well as *ftsEX* were not essential, and no predicted homologs could be found for *ftsB*, *ftsN* and *zipA*. Additionally, among genes responsible for lipoprotein export to the outer membrane, both *lgt* and *lspA* were essential, but surprisingly, *lnt* was not. Lnt is the phospholipid/apolipoprotein transacylase that is N-acylating the N-terminal cysteine during the biogenesis of lipoproteins. The *lnt* gene is also not essential in *B. subvibrioides*, suggesting that the dispensability of *lnt* could be a shared feature among *Alphaproteobacteria*. Interestingly, regarding the lipoprotein export system LolAB, only *lolA* is essential while no clear homolog of *lolB* can be identified by sequence homology. As a side note, it should be stressed that a *lolB* homolog has been suggested thanks to protein structure homology and appears essential in *Brucella* (Sean Crosson, *personal communication*).

Regarding the regulation of the bacterial cell cycle in *Brucella*, one of the key features of *Brucella* cell cycle is the DivK-CtrA pathway, conserved among many *Alphaproteobacteria* (Brilli *et al.,* 2010). Interestingly, most but not all members of the DivK-



Adapted from Kirkpatrick *et al.*, 2011

Figure S1. Outline of the DivK-CtrA regulation network.
The network is predicted from *B. abortus* genome analysis and adapted from Kirkpatrick & Viollier, *FEMS Microbiol Rev* 36 (2012) 193–205. Essential genes are underlined. It is striking that most proteins in this network are encoded by essential genes. The presence of several redundant enzymes able to synthesize or degrade cyclic di-GMP (c-di-GMP) probably explain the non-essential score of *pleD* and *pdeA* genes.

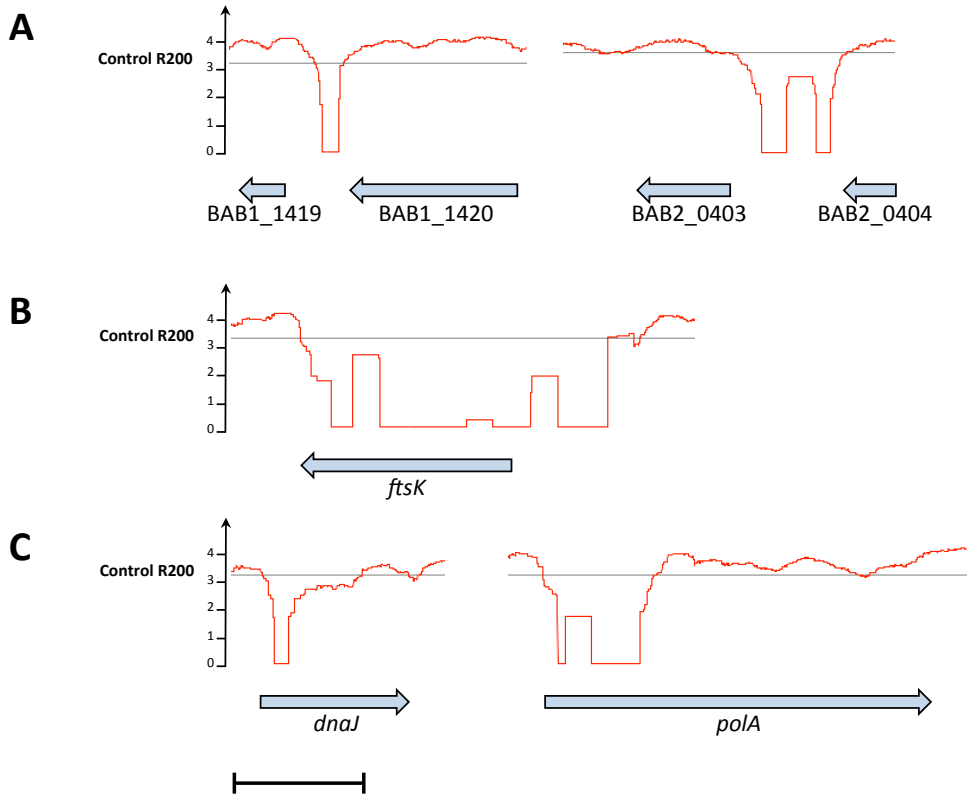


Figure 10. Tn-seq applications in genome annotations using the R200 methodology. By processing Tn-seq data with the R200 method, we were able to (**A**) identify previously unannotated essential coding sequences, (**B**), correct previously misannotated ORFs as exemplified by *ftsK* for which an essentiality region spans beyond its 5' end and matches with *C. crescentus* full length *ftsK* allele, and (**C**), identify essential domains within essential coding sequences.

CtrA pathway were found to be essential for growth on plates in Tn-seq. Indeed, *pdhS*, *divK*, *divL*, *cckA*, *chpT*, *ctrA*, *cpdR* and *clpXP* were found to be essential (Figure S1).

Taken together, these data demonstrate that Tn-seq enables the reconstitution of essential pathways, in a single experiment. In particular, it allows the identification of a crucial homolog within a family of several potential paralogs, such as BAB1_0932, encoding the only essential bifunctional (transglycosidase and transpeptidase) PBP involved in peptidoglycan (PG) biosynthesis. Moreover, Tn-seq in conjunction with our R200 methodology also allowed genome annotation and corrections such as the identification of new coding sequences, the extension of the *ftsK* ORF, and essential domain mapping as exemplified by *dnaJ* and *polA* Figure 10).

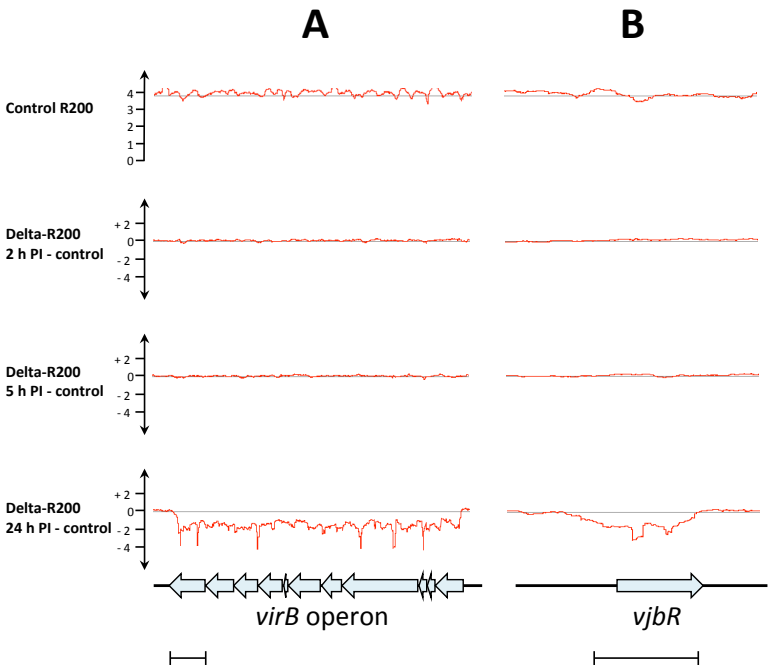


Figure S2. Tn-seq profile of *virB* and *vjbR* loci.
Tn-seq profiles reporting the attenuation of the mutants in the *virB* operon, encoding the type IV secretion system (**A**) and in the *vjbR* gene, coding for a transcriptional activator of *virB* (**B**). Attenuation is detected only at 24 h PI, but not before, which is consistent with the expected role of VirB in the intracellular trafficking of *B. abortus*. Scale bars correspond to 1 kb.

3.b) Screening for genes required for macrophage infection

One of the main objectives of this study was to identify genes specifically required for macrophage infection. For this purpose, three large-scale infections of RAW 264.7 macrophages were carried out in parallel using the transpositional mutants library described above (Figure 8). For each infection, a specific post-infection (PI) time point was selected in order to obtain a better understanding of the specific gene requirement at different stages of the cellular infection process. Firstly, the 2 h PI time point was chosen to highlight genes required for either adhesion or invasion of host cells as well as early survival. Secondly, the 5 h PI time point was selected to identify genes required for the intracellular survival and trafficking step in endosomal compartments (Celli, 2015) as it represents the end of the non-proliferative phase of the infection of RAW 264.7 macrophages (Deghelt *et al.,* 2014). Finally, the 24 h PI time point was chosen to identify genes required for the access to the replication niche and for intracellular proliferation. After each time point, bacteria were extracted from infected macrophages and grown prior to transposon insertion site identification and R200 calculation as explained above (Figure 8).

Attenuation corresponds to a decrease of fitness specific to infection. Therefore, in the context of an attenuated mutant, one expects that the R200 values for the mutated gene would be lower after infection compared to the control condition. Consequently, for analyzing post-infection data, the control R200 values were subtracted from the corresponding PI R200 for each PI dataset. This resulted in three lists of Delta-R200 values, namely "2 h PI - control", "5 h PI - control", and "24 h PI - control", corresponding to Delta-R200 2 h PI, Delta-R200 5 h PI and Delta-R200 24 h PI, respectively. A total of 164 candidates have been identified using the Delta-R200 analysis. Among these candidates, 75 were found at 2 h PI, 98 were found at 5 h PI, and 164 were found at 24 h PI (see Table 1).

In order to validate the Delta-R200 analysis, we checked for genes known for causing attenuation during a cellular infection when mutated. For this purpose, we have chosen a 2 h PI control, the response regulator *bvrR*, and two 24 h PI controls, the type IV secretion system operon *virB* (Comerci *et al.,* 2001; Delrue *et al.,* 2001), required for intracellular proliferation, and *vjbR,* an important transcriptional activator of *virB* which is not part of the *virB* operon (Arocena *et al.,* 2010; Delrue *et al.,* 2005). As expected, the *bvrR* mutants were strongly attenuated from 2 h PI, and both *virB* and *vjbR* mutants were only attenuated at 24 h PI (Figure S2)(Table 1), which is in agreement with the timing required by the bacterium to reach its replicative niche and proliferate. Taken together, these data lean toward the validation of the candidates identified by the Delta-R200 analysis.

3.c) Few genes are required for short term infection in macrophages

We decided to assess the validity of our datasets by testing for the reproducibility of the observed phenotypes by mutating candidate genes and testing resulting mutants survival by colony forming units -CFU- countings after infecting RAW 264.7 macrophages for two hours. We chose four candidates, two that displayed no attenuation in Tn-seq as negative controls (*omp2a* and *ftsK*-like, corresponding to BAB1_0659 and BAB2_0709 coding

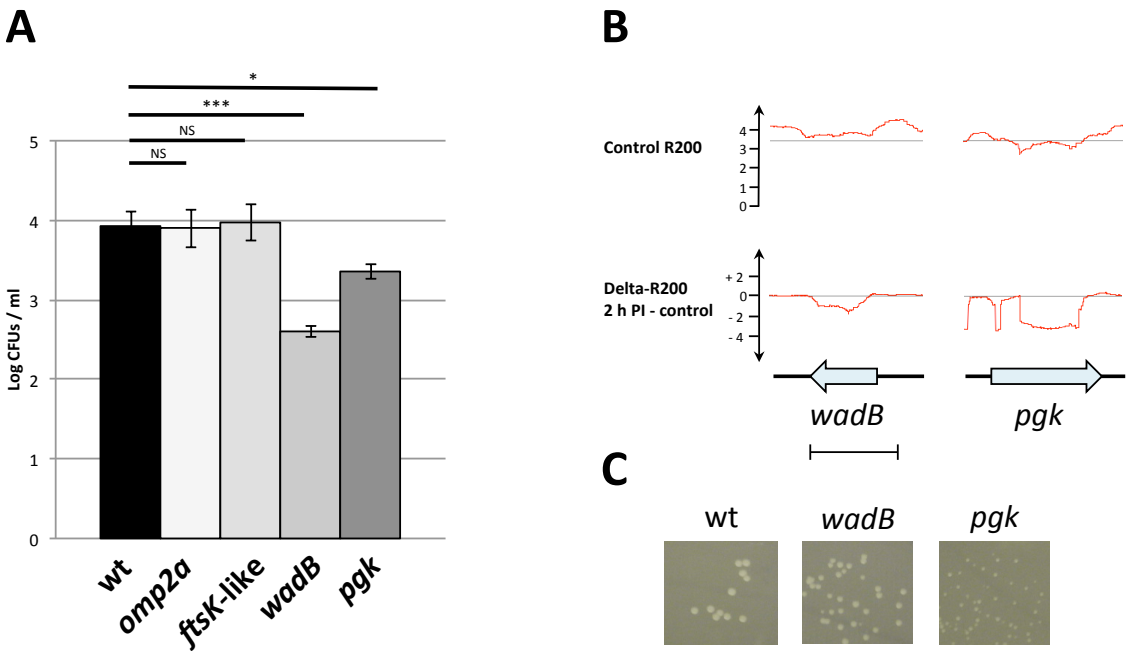


Figure 11. Validation of Tn-seq data using reconstructed mutants. (**A**) CFUs counting after a 2 h infection of RAW 264.7 macrophages with individual mutants in *omp2a*, *ftsK*-like, *wadB* and *pgk*. The Tn-seq data indicate that *omp2a* and *ftsK*-like mutants are fully virulent at 2 h PI while the *wadB* and *pgk* mutants are attenuated. The wild type (wt) strain was used as a virulent control. * = p > 0.05, *** = p > 0.001, ns = not significant. (**B**) Comparison of the Tn-seq profiles of the *wadB* and *pgk* mutants, highlighting the aggravated profile of the *pgk* mutant. (**C**) Pictures of the wild type strain, *wadB* mutant, and *pgk* mutant colonies on rich medium, supporting the hypothesis of the *pgk* growth defect results in an aggravated profile.

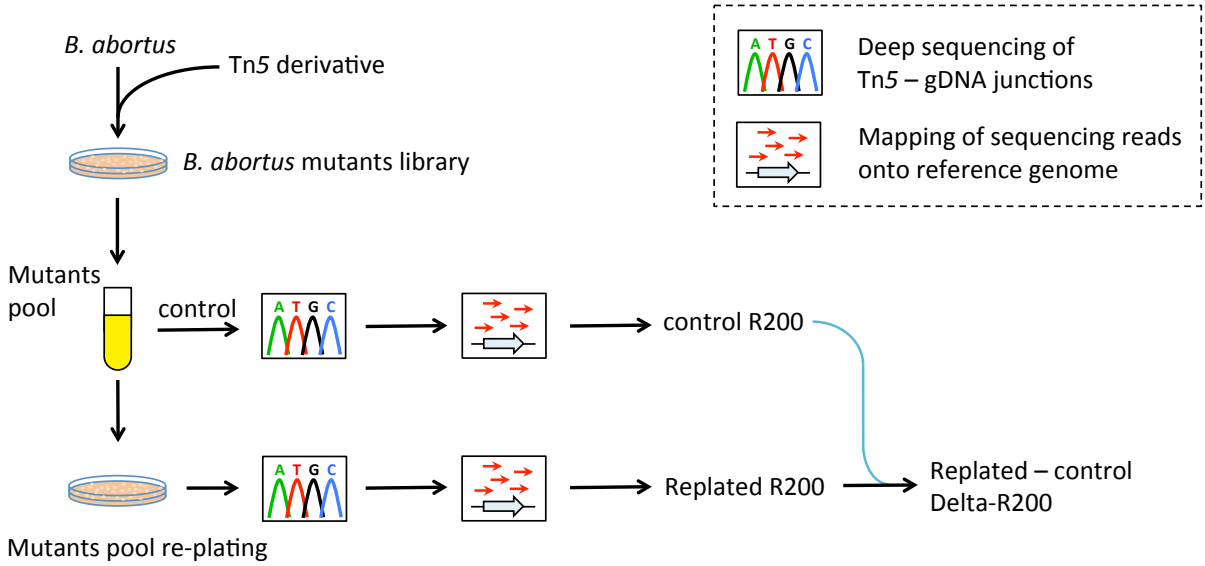


Figure 12. Schematic representation of the "re-plating" Tn-seq setup. For this experiment, a mutant library was generated and pooled as explained in Figure 8, but instead of infecting macrophages, the pool was re-plated on rich medium in order to assess the effects of an additional growth cycle on mutants fitness and to estimate to what extent growth defects may contribute to attenuation in infection.

sequences, respectively), and two that displayed attenuation in Tn-seq as positive controls (*wadB*, and *pgk*, corresponding to BAB1_0351 and BAB1_1742, respectively). As expected, the two negative control strains did not show any sign of attenuation compared to the wild type strain, while the positive control strains exhibited a statistically significant attenuation phenotype (Figure 11A).

One can distinguish two categories of attenuated mutants. On the one hand, mutants can be attenuated due to a failure to perform a successful infection, but on the other hand, mutants can be attenuated due to growth impairments that are actually already observed when grown on rich medium and amplified during infection. In fact, as opposed to candidates displaying a typical attenuation profile as exemplified by the *wadB* mutant, others displayed attenuation in conjunction with low control R200 values as exemplified by the *pgk* mutant (Figure 11B). This second type of profile suggested growth deficiencies anterior to infection. In addition, while the *pgk* mutant strain displayed a minute but yet significant decrease in CFU (Figure 11A), it was clear that colonies were smaller in size than the wild type colonies (Figure 11C). Therefore, the attenuation of candidates sharing a *pgk*-like profile in Tn-seq is likely to be due, at least in part, to growth impairments already present in the control condition. Actually, the second round of culture taking place after the infection (Figure 8) could simply amplify the disadvantage of clones that already display growth delays on plates.

In order to investigate the effect of additional rounds of culture on Tn-seq candidates attenuation, we performed a new Tn-seq experiment (refered to as "re-plating") in which the colonies of the library of transpositional mutants were collected and re-plated prior to sequencing instead of infecting host cells (the R200 lists for these two new conditions are available as Supporting Information) see Figure 12. By comparing these two new datasets, we were able to monitor the fitness loss of mutants due to growth deficiencies during culture in an infection-independent manner. Surprisingly, 54 % of the candidates harboring attenuation at 2, 5 or 24 h PI in our initial Tn-seq also displayed a similar attenuation when simply regrown on rich medium (see "re-plating" in Table 1), thus meaning that the fitness loss of those candidates was due to growth defects instead of infection. Consequently, "aggravated mutant" (*i.e.* attenuated mutants displaying aggravation/worsening of pre-existing growth impairments during infection) were not selected for further analysis (see Discussion). It is striking that consistent pathways fall into this category, like the purine biosynthesis (*purB*, *purC*, *purD*, *purF, purH*, *purL*, *purMN* and *purQ*) and the cytochrome c maturation (*ccmABC* and *ccmIEFH*) pathways (Table 1).

3.d) Identification of hyper-invasive mutants

Tn-seq can theoretically highlight hyper-invasive mutants (defined throughout this work as mutants displaying increased internalization inside host cells compared to the wild type strain) in addition to attenuated mutants. Indeed, such mutants would display higher R200 values compared to the control, meaning that proportionally more mutant bacteria would be found inside host cells when such genes are disrupted, thus resulting in positive Delta-R200 values. Using this criterion, nine genes (namely *wbkD, wbkF, per, gmd, wbkA, wbkE, wboA, wboB,* and $manB_{core}$ [BAB2_0855]) were identified and remarkably all of them

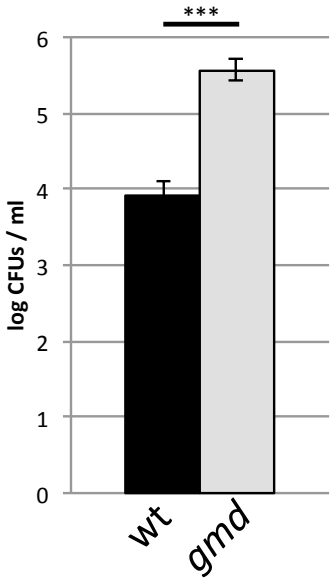


Figure S3. CFUs counting of the wild type strain and a *gmd* mutant after a 2 h macrophage infection. CFUs counting performed on RAW 264.7 macrophages 2 h PI using a *gmd* mutant and a wild type strain, confirming the hyper-invasive profile of the *gmd* mutant as suggested by its Tn-seq profile. *** = $p < 0.001$ in a Student's t test.

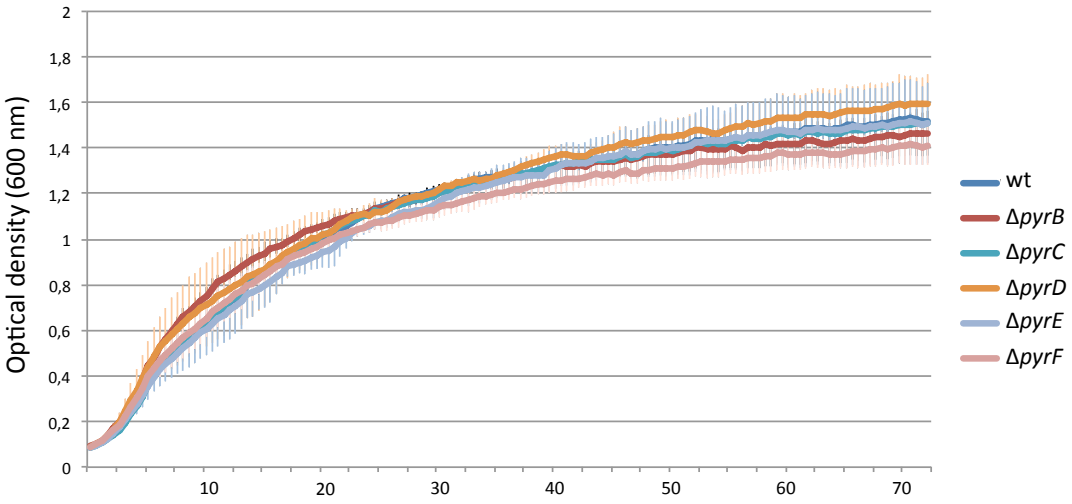


Figure S4. Growth curves of the different *pyr* mutants in rich medium
The optical density for each mutant was measured every 30 minutes for 72 h. Each curve represents the mean of values gathered from three independent experiments, and error bars display the standard deviation for these values.

are part of the lipopolysacharride O-chain synthesis pathway (Fontana *et al.,* 2016). Indeed, such mutants have a rough LPS and are known to be more invasive than the smooth parental strain (Detilleux *et al.,* 1990; Porte *et al.,* 2003). To confirm this using our settings, a mutant of the GDP-mannose dehydratase gene (*gmd*) was constructed and CFU were counted after infecting RAW 264.7 macrophages for 2 h. As expected, the resulting strain displayed increased invasiveness typical of rough strains (Figure S3).

Consequently, this confirms that Tn-seq is a suitable method for identifying hyper-invasive mutants, and more importantly that no other hyper-invasive mutants than predicted rough mutants could be identified using this method.

3.e) Identification of genes required up to 24 h PI

The 24 h PI-specific candidates mainly include genes predicted to be involved in trafficking, growth and metabolism (see Table 1). The biosynthesis of histidine seems to be crucial, as well as the synthesis of pyrimidines, suggesting that *B. abortus* cannot find or uptake enough of these compounds from the ER compartments in which it is proliferating. The 24 h PI specific candidates also comprise expected virulence genes such as the *virB* operon (O'Callaghan *et al.,* 1999), transcriptional regulators like *vjbR* (Delrue *et al.,* 2005) and *vtlR* (Sheehan *et al.,* 2015), the cytochrome bd biosynthesis operon *cydABCD* (Endley *et al.,* 2001). When comparing our data with the list of attenuated transpositional mutants from previous studies performed on various infection models and *Brucella* species (Cha *et al.,* 2012; Delrue *et al.,* 2004; Kim *et al.,* 2012; Liautard *et al.,* 2007; Wu *et al.,* 2006), it is striking that only 43 % of the attenuated mutants identified here had already been identified. Indeed, 33 of the 76 candidates attenuated in our time-resolved Tn-seq analysis but not when re-plated were part of the 259 candidates compiled from the previous studies. Therefore, the Tn-seq approach reported here has generated 43 new candidates, suggesting that the comprehensive analysis of the *Brucella* genome could yield new insights into the genes required for a macrophage infection (see Discussion).

3.f) The pyrimidines biosynthesis pathway allows intracellular proliferation

One of the major hits of the 24 h PI dataset is the pyrimidines (pyr) biosynthesis pathway. In fact, with the exception of genes already essential for culture on rich medium, all pyr biosynthesis genes became strongly attenuated at 24 h PI (namely *pyrB*, *pyrC*, *pyrD*, *pyrE*, and *pyrF*), while none of them were impacted when re-plated.

We further investigated the pyr pathway by creating a deletion mutant of its first non-essential gene, the Δ*pyrB* strain. The Δ*pyrB* mutant grew as the wild type strain in rich medium (Figure S4). We then evaluated the infectious potential of the Δ*pyrB* strain and its complementation strain by enumerating CFU in RAW 264.7 macrophages at both 2 h PI and 24 h PI. Estimation of the intracellular growth ratio (CFU at 24 h PI divided by CFU at 2 h PI) clearly showed that the Δ*pyrB* strain is strongly attenuated compared to the wild type and the complementation strains, validating the Tn-seq profile of the mutant (Figure 13A).

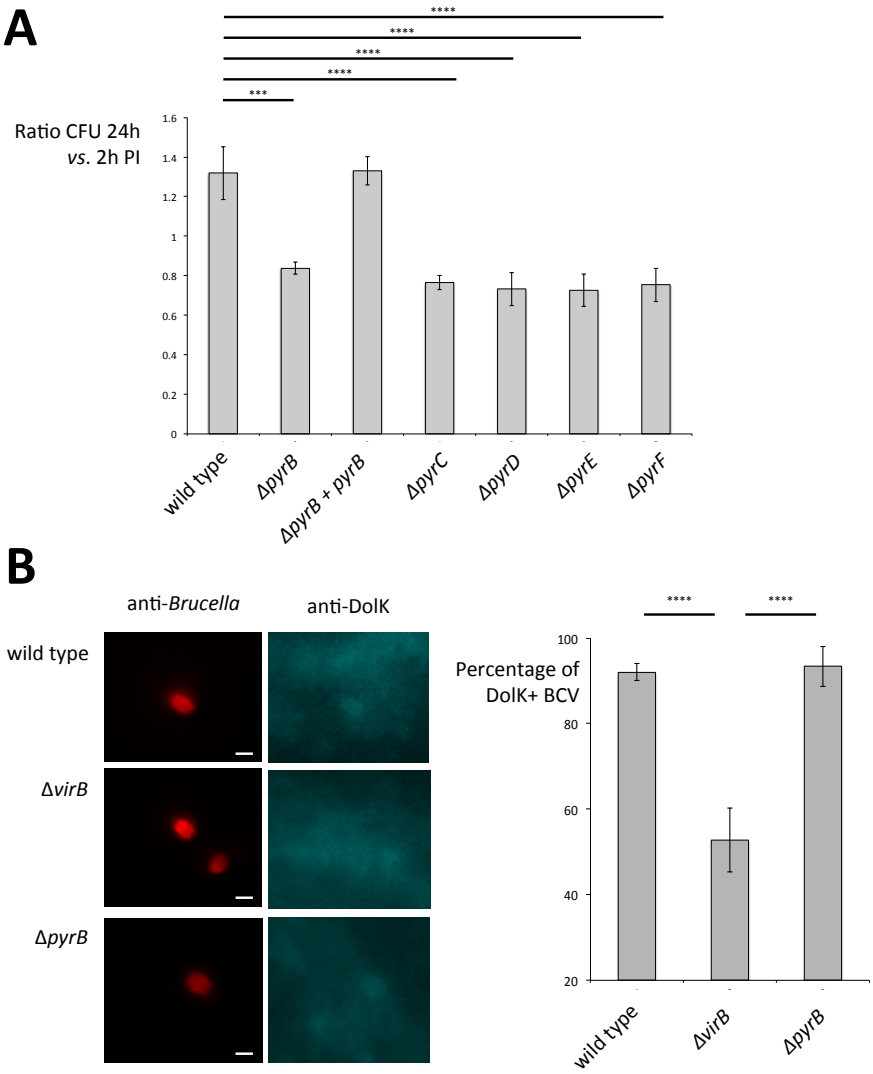


Figure 13. Intracellular growth and trafficking of *pyr* mutants. (**A**) CFUs ratio for each *pyr* mutant after infection of RAW 264.7 macrophages at 2 h PI and 24 h PI. For each strain, the $\log_{10}$CFUs counting after a 24 h infection of RAW 264.7 macrophages was divided by the corresponding $\log_{10}$CFUs counting after a 2 h infection, resulting in a ratio depicting the evolution of the bacterial load from 2 h PI to 24 h PI. Accordingly, an identical load will give a ratio of 1, while an increased load will give a ratio > 1, and a decreased load will give a ratio < 1. Unprocessed data are reported in Figure S5. *** = $p < 0.001$, **** = $p < 0.0001$. (**B**) Fluorescence microscopy images of HeLa cells that were infected for 8 h with the Δ*pyrB* strain using anti-LPS antibodies to label *Brucella* (red), and anti-DolK antibodies to label the ER (cyan). The wild type and Δ*virB* strains were used as controls able and unable to reach the ER, respectively. The scale bar corresponds to 1 µm. (**C**) Quantitative measurement of the percentage of bacteria stained with the anti-DolK staining. The total number of observations for the wild type, the Δ*virB*, and the Δ*pyrB* strains was 263, 200, and 203 bacteria, respectively; divided into three biological replicates. **** = $p < 0.0001$.

Consistently, deletion mutants for all other non-essential pyr genes (Δ*pyrC*, Δ*pyrD*, Δ*pyrE*, and Δ*pyrF*) behaved as the wild type when cultured in rich medium (Figure S4) and were impaired for intracellular growth, further validating the involvement of the pyr pathway for proliferation inside RAW 264.7 macrophages (Figure 13A). The unprocessed data of Figure 13A can be found in Figure S5).

Since the most likely hypothesis was that the pyr pathway is required for intracellular proliferation but not for intracellular trafficking, we tested the intracellular localization of the Δ*pyrB* strain. We thus infected HeLa cells for 8 h with the Δ*pyrB* mutant, and in parallel with the Δ*virB* and the wild type controls. HeLa cells were chosen due to their flat morphology which allows a clearer labeling of the *Brucella* containg vacuoles, and time 8 h PI was chosen as the majority of wild type bacteria are supposed to have reached the ER at that time. We then performed simultaneous anti-*Brucella* anti-DolK immunofluorescence stainings, using the wild type strain as a positive control (*i.e.* able to reach the replication niche) and a Δ*virB* mutant as a negative control (*i.e.* known for being blocked in endosomes and unable to reach the replication niche) (Figure 13B). It should be noted the dolichol kinase DolK, involved in host protein N-glycosylation, was chosen as an ER marker in this work as recent experiments performed in our laboratory (Francis *et al.,* 2017) suggested that it generates a clearer staining than other typical ER markers (Katy Poncin, *personal communication*). Remarkably, we observed that the Δ*pyrB* mutant is predominantly found in DolK-positive compartments as opposed to the Δ*virB* mutant, thus meaning that the Δ*pyrB* strain is able to reach the endoplasmic reticulum, but does not proliferate, likely due to a lack or shortage of pyrimidines precursors (see Discussion).

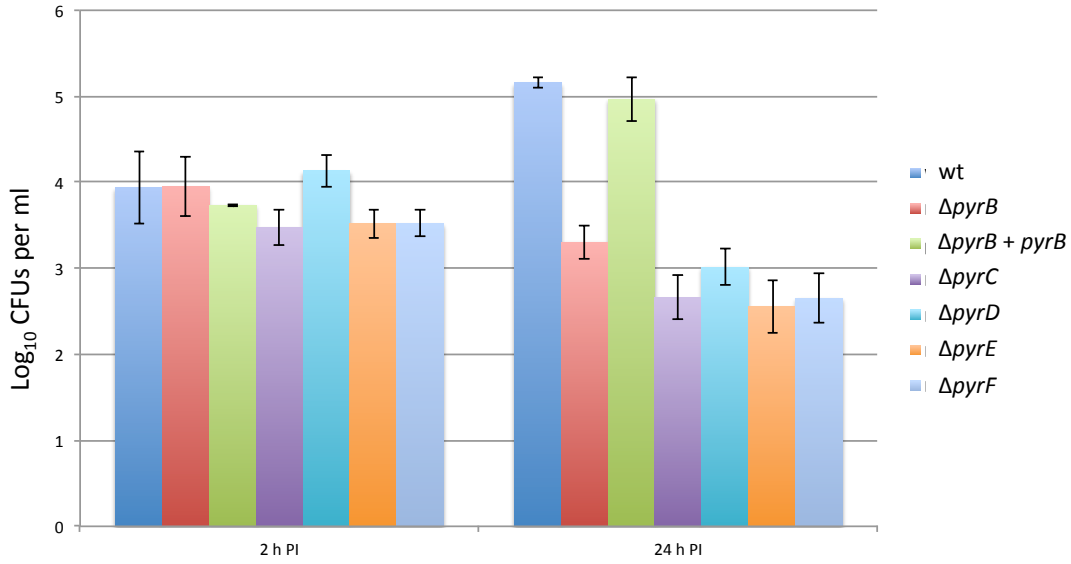


Figure S5. CFUs counting for the wild type (wt) and each *pyr* mutant after the infection of RAW 264.7 macrophages at 2 h PI and 24 h PI.
These data were used to calculate the post infection ratio as shown in Figure 13.

Table 1

ORF	Gene name	2 h PI attenuation	5 h PI attenuation	24 h PI attenuation	attenuation when replated	Predicted function
BAB1_0024	*cmk*	+	+	+	-	CDP synthesis from CMP
BAB1_0091	*ccmA*	+	+	+	+	cytochrome c maturation
BAB1_0092	*ccmB*	+	+	+	+	cytochrome c maturation
BAB1_0093	*ccmC*	+	+	+	+	cytochrome c maturation
BAB1_0143	*glnD*	+	+	+	+	glutamine pool regulation
BAB1_0160	*ptsN*	+	+	+	-	pts, metabolism, carbon catabolite repression
BAB1_0304	*cenR*	+	+	+	+	envelope
BAB1_0351	*wadB*	+	+	+	-	envelope, LPS core synthesis
BAB1_0386	*copA*	+	+	+	+	cation metal exporter, ATPase
BAB1_0387		+	+	+	+	cation exporter associated protein
BAB1_0435		+	+	+	-	metabolism, glycolate oxidase
BAB1_0442	*purD*	+	+	+	+	DNA bases, purines
BAB1_0491		+	+	+	+	invasion protein B
BAB1_0631	*ccmI*	+	+	+	+	cytochrome c maturation
BAB1_0632	*ccmE*	+	+	+	+	cytochrome c maturation
BAB1_0633	*ccmF*	+	+	+	+	cytochrome c maturation
BAB1_0634	*ccmH*	+	+	+	+	cytochrome c maturation
BAB1_0730	*purN*	+	+	+	+	purines biosynthesis
BAB1_0731	*purM*	+	+	+	+	purines biosynthesis
BAB1_0739		+	+	+	+	Electron Transport Chain (ETC)- complex I subunit
BAB1_0773	*gppA*	+	+	+	+	exopolyphosphatase
BAB1_0855		+	+	+	+	glutaredoxin
BAB1_0857	*purL*	+	+	+	+	purines biosynthesis
BAB1_0860	*purQ*	+	+	+	+	purines biosynthesis

BAB1_0862	*purC*	+	+	+	+	purines biosynthesis
BAB1_0868	*purB*	+	+	+	+	purines biosynthesis
BAB1_0962		+	+	+	+	protein-L-isoAsp O-methyltransferase
BAB1_1030	*gor*	+	+	+	+	glutathione reductase
BAB1_1038	*mlaE*	+	+	+	+	ABC transporter, permease
BAB1_1039	*mlaF*	+	+	+	+	ABC transporter, ATPase
BAB1_1040	*mlaD*	+	+	+	+	ABC transporter-associated inner membrane protein
BAB1_1041	*mlaA*	+	+	+	+	predicted lipoprotein
BAB1_1045	*fdx*	+	+	+	+	ferredoxin
BAB1_1115	*tgt*	+	+	+	-	tRNA modification
BAB1_1191	*clpA*	+	+	+	+	protease
BAB1_1437	*pepP*	+	+	+	-	peptidase, Xaa-Pro aminopeptidase
BAB1_1460	*mntH*	+	+	+	-	manganese transport, ion transport
BAB1_1485		+	+	+	-	uncharacterized conserved protein
BAB1_1553	*ychF*	+	+	+	+	translation-associated GTPase
BAB1_1665	*rpoH2*	+	+	+	-	RNA polymerase sigma factor
BAB1_1669		+	+	+	-	signal transduction, HWE-family histidine kinase
BAB1_1757	*purE*	+	+	+	-	purines biosynthesis
BAB1_1766	*hfaC*	+	+	+	-	double ATPase, ABC-like, no permease
BAB1_1824	*purH*	+	+	+	+	purines biosynthesis
BAB1_1918	*lpd*	+	+	+	-	dihydrolipoyl dehydrogenase
BAB1_2006	*cenR*	+	+	+	+	envelope (also called otpR)
BAB1_2025		+	+	+	-	DnaJ-like chaperone
BAB1_2084	*hisH*	+	+	+	+	histidine biosynthesis, first part of the pathway
BAB1_2085	*hisA*	+	+	+	+	histidine biosynthesis, first part of the pathway
BAB1_2086	*hisF*	+	+	+	+	histidine biosynthesis, first part of the pathway
BAB1_2092	*bvrR*	+	+	+	+	response regulator, regulation of envelope
BAB1_2093	*bvrS*	+	+	+	+	histidine kinase, regulation of envelope

BAB1_2094	*hprK*	+	+	+	+	kinase, regulator of phsphotranferase system (PTS)
BAB1_2103		+	+	+	+	NDP-sugar pyrophosphorylase
BAB1_2135	*gshB*	+	+	+	+	glutathione synthetase
BAB1_2158	*lnt*	+	+	+	-	lipoprotein synthesis
BAB1_2159	*hipB*	+	+	+	+	transcriptional regulator
BAB1_2175	*irr*	+	+	+	+	ion transport, transcriptional regulator, ferric uptake regulator
BAB2_0076	*omp10*	+	+	+	+	outer membrane lipoprotein
BAB2_0366	*eryI*	+	+	+	+	erythritol metabolism (also called rpiB)
BAB2_0367	*eryH*	+	+	+	+	erythritol metabolism (also called tpiA2)
BAB2_0370	*eryC*	+	+	+	+	erythritol metabolism
BAB2_0411		+	+	+	+	conserved hypothetical
BAB2_0412	*tauE*	+	+	+	+	sulfite exporter
BAB2_0442		+	+	+	+	acetyl-CoA dehydrogenase
BAB2_0476	*gshA*	+	+	+	+	gamma-glutamylcysteine synthetase
BAB2_0511		+	+	+	+	NAD-dependent dehydrogenase
BAB2_0656	*ccdA*	+	+	+	+	cytochrome c maturation (dsbD homolog)
BAB2_0668		+	+	+	+	metabolism, phosphatidylserine synthase
BAB2_0699		+	+	+	-	ABC transporter, substrate binding component, oligopeptide trans
BAB2_0701		+	+	+	-	ABC transporter, permease, oligopeptide transport
BAB2_0702		+	+	+	-	ABC transporter, permease, oligopeptide transport
BAB2_0703		+	+	+	-	ABC transporter, ATPase
BAB2_1012	*dapB*	+	+	+	+	dihydrodipicolinate reductase, diaminopimelate biosynthesis
BAB2_1013	*gpm*	+	+	+	+	phosphoglycerate mutase
BAB1_0051		-	+	+	-	copper binding prot, ion transport
BAB1_0388	*ccoG*	-	+	+	+	cytochrome c oxidase
BAB1_0389	*ccoP*	-	+	+	+	cytochrome c oxidase
BAB1_0392	*ccoN*	-	+	+	+	cytochrome c oxidase
BAB1_0427		-	+	+	-	O-methyltransferase

BAB1_0471		-	+	+	+	oxoacyl-(acyl carrier protein) reductase
BAB1_0477	*rplI*	-	+	+	-	ribosomal protein L9
BAB1_0688	*pyrC2*	-	+	+	-	pyrimidines biosynthesis
BAB1_0704	*ksgA*	-	+	+	+	16S ribosomal RNA methyltransferase
BAB1_0898	*bglX*	-	+	+	-	beta-glucosylase-related glycosidase
BAB1_1435		-	+	+	-	related to cytochrome c oxidase synthesis
BAB1_1462	*ampD-like*	-	+	+	-	N-acetyl-anhydromuramyl-L-alanine amidase
BAB1_1476		-	+	+	-	lysophospholipase
BAB1_1557		-	+	+	+	cytochrome c1 family
BAB1_1559		-	+	+	+	ubiquinol cytochrome c reductase, iron-sulfur component
BAB1_1589		-	+	+	+	major facilitator superfamily (transporter)
BAB1_1846		-	+	+	+	membrane bound metallopeptidase
BAB1_1962	*gntR13*	-	+	+	+	transcriptional regulator
BAB1_2145	*phoU*	-	+	+	-	phosphate transport control
BAB1_2167	*truB*	-	+	+	+	tRNA modification
BAB2_0643		-	+	+	-	endonuclease
BAB2_0678	*rirA*	-	+	+	-	transcriptional regulator, iron responsive
BAB2_1027	*upp*	-	+	+	+	phosphoribosyl transferase
BAB1_0003	*recF*	-	-	+	-	recombination, DNA
BAB1_0045	*tamA*	-	-	+	-	envelope OM protein
BAB1_0046	*tamB*	-	-	+	-	uncharacterized conserved
BAB1_0084	*yccA*	-	-	+	+	inner membrane prot, uncharacterized
BAB1_0096	*ilvD*	-	-	+	-	valine isoleucine metabolism
BAB1_0113		-	-	+	-	short chain dehydrogenase
BAB1_0139		-	-	+	-	thioredoxin-like
BAB1_0162	*ibpA*	-	-	+	-	chaperone
BAB1_0168		-	-	+	-	adeonise kinase
BAB1_0172	*rph*	-	-	+	-	ribonuclease

BAB1_0285	*hisD*	-	-	+	-	histidine biosynthesis, second part of the pathway
BAB1_0318		-	-	+	-	predicted phosphatase
BAB1_0341	*pyrD*	-	-	+	-	pyrimidines biosynthesis
BAB1_0478		-	-	+	-	hypothetical protein
BAB1_0638	*glnE*	-	-	+	-	glutamine pool regulation
BAB1_0640	*pleC*	-	-	+	-	histidine kinase, regulation of cell cycle
BAB1_0861	*purS*	-	-	+	+	purines biosynthesis
BAB1_1019	*rluA*	-	-	+	-	pseudouridylate synthase, 23S RNA-specific
BAB1_1092		-	-	+	+	dihydropteroate synthase
BAB1_1098	*hisI*	-	-	+	+	histidine biosynthesis, second part of the pathway
BAB1_1206		-	-	+	-	dehydrogenase, involved in queuosine biosynthesis
BAB1_1217	*murI*	-	-	+	-	glutamate racemase
BAB1_1283		-	-	+	-	uncharacterized conserved membrane prot
BAB1_1345		-	-	+	-	Kef-type potassium transporter
BAB1_1399	*ilvC*	-	-	+	-	keto-acid reductoisomerase involved in Val and Ile biosynthesis
BAB1_1517	*vtlR*	-	-	+	-	LysR transcriptional regulator cotrolling sRNA expression
BAB1_1657	*dsbB*	-	-	+	+	chaperone, dislufide bond formation in periplasm
BAB1_1670		-	-	+	+	hypothetical protein
BAB1_1695	*purA*	-	-	+	-	purines biosynthesis
BAB1_1761	*pykM*	-	-	+	+	pyruvate kinase
BAB1_1930	*omp19*	-	-	+	+	outer membrane lipoprotein
BAB1_1988	*hisC*	-	-	+	-	histidine biosynthesis
BAB1_2022	*cobT*	-	-	+	+	cobalamin biosynthesis
BAB1_2023	*cobS*	-	-	+	+	cobalamin biosynthesis
BAB1_2069	*maf-2*	-	-	+	-	nucleotide binding protein Maf
BAB1_2082	*hisB*	-	-	+	-	histidine biosynthesis
BAB1_2087	*hisE*	-	-	+	-	histidine biosynthesis
BAB1_2132	*pyrF*	-	-	+	-	pyrimidines biosynthesis

BAB2_0058	*virB11*	-	-	+	-	type IV secretion system
BAB2_0059	*virB10*	-	-	+	-	type IV secretion system
BAB2_0060	*virB9*	-	-	+	-	type IV secretion system
BAB2_0061	*virB8*	-	-	+	-	type IV secretion system
BAB2_0062	*virB7*	-	-	+	-	type IV secretion system
BAB2_0063	*virB6*	-	-	+	-	type IV secretion system
BAB2_0064	*virB5*	-	-	+	-	type IV secretion system
BAB2_0065	*virB4*	-	-	+	-	type IV secretion system
BAB2_0066	*virB3*	-	-	+	-	type IV secretion system
BAB2_0067	*virB2*	-	-	+	-	type IV secretion system
BAB2_0068	*virB1*	-	-	+	-	type IV secretion system
BAB2_0118	*vjbR*	-	-	+	-	quorum sensing transcriptional regulator
BAB2_0143	*deoR1*	-	-	+	-	transcriptional regulator
BAB2_0182	*hisZ*	-	-	+	+	histidine biosynthesis
BAB2_0183	*hisG*	-	-	+	+	histidine biosynthesis
BAB2_0513	*gcvT*	-	-	+	+	glycine cleavage system
BAB2_0514	*gcvH*	-	-	+	+	glycine cleavage system
BAB2_0515	*gcvP*	-	-	+	+	glycine cleavage system
BAB2_0640	*pyrC*	-	-	+	-	pyrimidines biosynthesis
BAB2_0641	*pyrB*	-	-	+	-	pyrimidines biosynthesis
BAB2_0727	*cydB*	-	-	+	-	cytochrome d oxidase
BAB2_0728	*cydA*	-	-	+	-	cytochrome d oxidase
BAB2_0729	*cydC*	-	-	+	-	ABC transporter, cytochrome bd
BAB2_0730	*cydD*	-	-	+	-	ABC transporter, cytochrome bd
BAB2_1010	*glk*	-	-	+	-	glucokinase
BAB2_1079	*znuA*	-	-	+	+	ABC transporter for zinc
BAB2_1080	*znuC*	-	-	+	+	ABC transporter for zinc
BAB2_1081	*znuB*	-	-	+	+	ABC transporter for zinc

4) Discussion

In this work, a Tn-seq approach has been designed with *B. abortus* in order to identify genes essential for growth on rich medium as well as genes required for different steps of the infection of RAW 264.7 macrophages.

As most of the current therapies against bacterial pathogens aim at targeting essential processes such as translation or cell wall synthesis, identifying all essential genes for growth on rich medium generates a valuable knowledge base for pinpointing novel therapeutical targets. In this study, a total of 491 candidate genes were qualified as essential for growth on rich medium plates. Interestingly, our quantitative analysis also highlights genome sections that show a reduced fitness and statistical comparison of any regions (inside or outside predicted genes) in the genome is thus feasible (Figure 9B, see section I.3.a of the Results). The quantitative analysis also allows the identification of regions in which mutagenesis generates a growth defect on the control plates, and thus "aggravated" mutants for which a growth defect after infection is probably non-specific to the intracellular condition.

While PG biosynthesis is clearly essential, only one of the three predicted class-A penicillin-binding proteins, BAB1_0932, was found to be essential for growth on rich medium, showing that Tn-seq could allow the identification of the main functional gene among a group of paralogs. Tn-seq also allowed to reshape essential processes compared to other bacteria, as exemplified by the *bam* genes, responsible for the OMPs export system. *Brucella* possesses an incomplete OMPs export system composed of only *bamADE* and lacking *bamB* and *bamC*, and here, we showed that *bamE* was scored essential in Tn-seq, whereas it is not essential in *E. coli* (Sklar *et al.,* 2007). One possibility would be that *bamE* is functionally redundant with another gene in *E. coli*, while this redundancy is absent in *B. abortus*.

Intriguingly, a few genes (*mucR*, *sodA*, *pgm* and *wbkC*) for which a viable deletion mutant was previously reported (Caswell *et al.,* 2013; Godfroid *et al.,* 2000; Martin *et al.,* 2012; Mignolet *et al.,* 2010; Mirabella *et al.,* 2013; Ugalde *et al.,* 2000) are actually scored as essential in our Tn-seq analysis. This particularity has already been observed in previous Tn-seq experiments performed on other bacterial species (Hubbard *et al.,* 2016). One cannot rule out the possibility that such contradictions could be due to strain-to-strain differences. However, another tempting hypothesis for this would be that, as Tn-seq intrinsically sets the same time period for all mutants to grow and generate colonies, the probability of obtaining suppressor strains is much reduced compared to the classical "pop in - pop out" (*i.e.* replacement of the coding sequence by an antibiotic resistance marker or a deletion allele) strategy generally used to obtain deletion strains. Gene deletion thus requires several rounds of culture on plates, which are usually maintained until colonies become detectable, thus resulting in an increased risk of selecting suppressor strains. Conversely, the absence of subcultures as performed in Tn-seq decreases the risk of isolating suppressor strains.

Remarkably, a second Tn-seq aiming at assessing the impact of growth deficiences in attenuation (termed "re-plating") revealed that 54 % of these candidates displayed fitness decreases similar to those found in infection when undergoing a second round of culture.

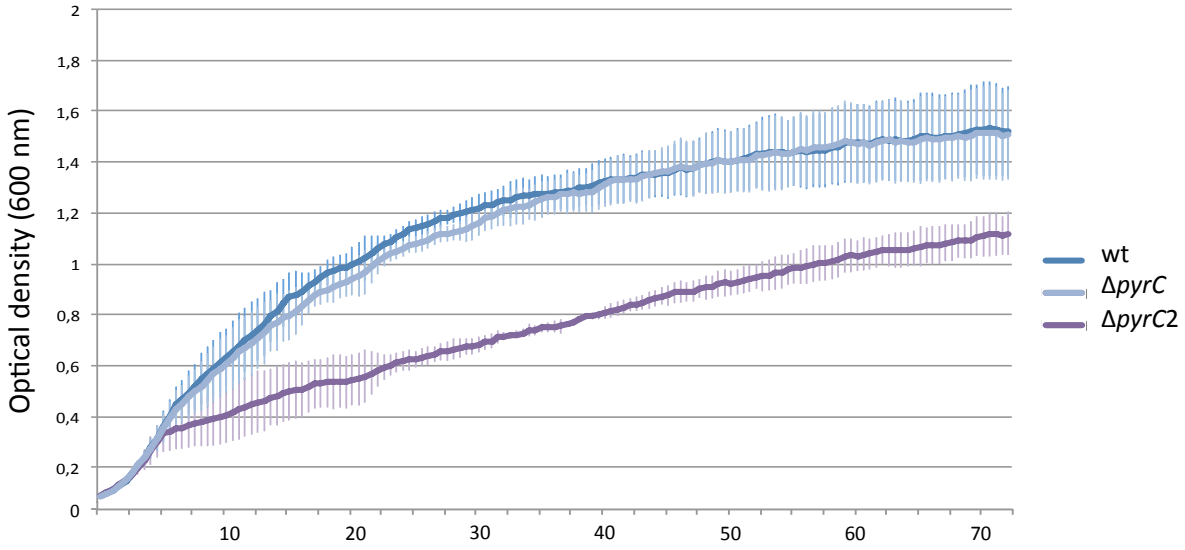


Figure S6. Growth curves of the different *pyrC* mutants in rich medium
The optical density for each mutant was measured every 30 minutes for 72 h. Each curve represents the mean of values gathered from three independent experiments, and error bars display the standard deviation for these values. The *pyrC*2 mutant is displaying a growth defect that is not observed for the other *pyr* mutants.

Such findings demonstrate the utmost importance of taking into account growth defect in culture before proposing virulence attenuation for mutant strains.

When comparing our candidates to the list of 259 attenuated mutants previously available from different infection models (Cha *et al.,* 2012; Delrue *et al.,* 2004; Kim *et al.,* 2012; Liautard *et al.,* 2007; Wu *et al.,* 2006), it was surprising that only 33 could be found in common (S3 Table). This means that 43 additional candidates were identified by Tn-seq. However, this also means that 216 attenuated mutants previously reported were not identified using Tn-seq in a RAW 264.7 macrophages infection, which is not surprising since these screening were done with different strains/species and different infection models. However, it should be noted that intriguingly, out of the remaining 216 candidates, 80 were categorized here as either strictly essential for growth on rich medium or essential when re-plated (S3 Table). It is thus likely that several attenuated mutants previously reported are actually suppressors of mutants in essential genes that display a growth defect in the infection model like the "aggravated" mutants cited above.

The 24 h PI time point revealed several pathways involved in trafficking and metabolism. The type IV secretion system VirB was needed but the effectors proposed to be translocated to the host cell (Ke *et al.,* 2015) were untouched in our Tn-seq analysis, which is probably the result of functional redundancy between effectors. Regarding metabolism, the biosynthesis of histidine was strongly impacted, as previously documented (Kohler *et al.,* 2002; Abdo *et al.,* 2011), and more specifically the second part of the pathway (namely, *hisB, hisC,* and *hisD*). This could be due to the fact that the first half of the histidine biosynthesis pathway (composed of *hisZ, hisG, hisE, hisI, hisA, hisH, hisF*) is shared with the purine biosynthesis pathway, and both are already essential when re-plated (see below). Therefore, Tn-seq demonstrates that genes responsible for the synthesis of histidine from imidazole-glycerol-3-phosphate are required for the proliferation of *B. abortus* in the endoplasmic reticulum of RAW 264.7 macrophages. Another major hit is the biosynthesis of pyrimidines. Indeed, Tn-seq showed that all non-essential pyr biosynthesis genes (*i.e.* starting from carbamoyl-phosphate to UMP, namely *pyrB, pyrC, pyrD, pyrE,* and *pyrF*) were consistently attenuated at 24 h PI in RAW 264.7 macrophages while none of the associated mutants displayed growth defects on rich medium. Interestingly, none of the *pyr* genes were impacted when re-plated as opposed to almost all genes involved in purine biosynthesis (*i.e. purB, purC, purD, purH, purL*-1*, purL*-2*, purN, purM, purS*). This is likely due to the composition of the culture medium and the respective half-life and/or heat resistance of precursors during medium sterilization. The *pyrB* and *pyrD* genes were already hit in a previous screening for attenuated mutants (Kohler *et al.,* 2002) but the pyrimidines biosynthesis pathway was never investigated. Here we show that all the mutants in genes of the pyr pathway are attenuated for growth inside macrophages, hence validating the Tn-seq data. It should be noted that a second homolog was found for *pyrC* (BAB1_0688, here called *pyrC*2), however a *B. abortus* Δ*pyrC*2 strain displays a growth defect in rich medium (Figure S6) and attenuation at 5 h PI in Tn-seq (Table 1), suggesting pleiotropic defects in this mutant compared to the *pyr* mutants characterized in this work. Altogether, these results strongly suggest that the ability of *B. abortus* to synthesize pyrimidines in the host cell is decisive for its proliferation inside macrophages. Furthermore, we show that the inability to synthesize pyrimidines is independent of the intracellular trafficking because the Δ*pyrB* mutant is still able to reach the

endoplasmic reticulum (Figure 13B). This reveals that the bacterium has no access to (enough) pyrimidines inside the ER. Interestingly, the *Brucella pyrB* mutant could also be tested in vaccine trials, since it is able to traffic in host cells like the wild type, but does not proliferate in the usual replication compartment.

Tn-seq data also generate surprising observations, such as the attenuation of the *lnt* mutants at 2 h PI, while *lnt* is not required for growth on rich medium, suggesting that a redundant function is present for growth on plates but not for short term survival in RAW 264.7 macrophages. Alternatively, it is also possible that the activity of Lnt is dispensable in *B. abortus*, at least in the control condition. Interestingly, our screening also revealed a role for a TamAB system homolog (BAB1_0045 and BAB1_0046) for intracellular proliferation, TamAB being proposed to be involved in the translocation of outer membrane proteins (Selkrig *et al.,* 2012). These data thus open new investigations pathways to better understand the molecular processes required for *B. abortus* survival and growth inside host macrophages.

In conclusion, Tn-seq is a comprehensive method that allowed the identification of attenuated *B. abortus* mutants for macrophages infection. Moreover, unlike the typical gene-by-gene analysis, the R200 analysis method presented here proved to be an efficient tool allowing the identification of previously unannoted ORFs and enabling genome annotations corrections as it is independent of gene prediction (thus being usable even in models with poorly annotated genomes). Regarding *Brucella*, the high coverage of the genome with transposons has allowed for identification of mutants with essential, attenuated, aggravated and non-essential phenotypes. It would be interesting to perform such experiments on other *Brucella* strains as well as other host cell types (including activated macrophages and trophoblasts (Salcedo *et al.,* 2013)) and more complex infection models such as animal models like a mouse intranasal infection (Hanot Mambres *et al.,* 2016), as this would generate a fundamental knowledge of the molecular arsenal required for *Brucella*.

5) Supporting Information

S1 Table. List of all essential genes for growth on rich medium.

S2 Table. List of PCR primers used in this study.

S3 Table. Comparison between Tn-seq candidates and attenuated mutants identified in previous studies.

S1 Text. TTM_ctrl_chrI.txt Transposon tolerance map (R200) for chromosome I when grown on rich medium.

S2 Text. TTM_ctrl_chrII.txt Transposon tolerance map (R200) for chromosome II when grow on rich medium.

S3 Text. Delta-R200_2hPI_chrI.txt Attenuation profile at 2 h post-infection for chromosome I.

S4 Text. Delta-R200_2hPI_chrII.txt Attenuation profile at 2 h post-infection for chromosome II

S5 Text. Delta-R200_5hPI_chrI.txt Attenuation profile at 5 h post-infection for chromosome I

S6 Text. Delta-R200_5hPI_chrII.txt Attenuation profile at 5 h post-infection for chromosome II

S7 Text. Delta-R200_24hPI_chrI.txt Attenuation profile at 24 h post-infection for chromosome I

S8 Text. Delta-R200_24hPI_chrII.txt Attenuation profile at 24 h post-infection for chromosome II

S9 Text. TTM_replated_ctrl_chrI.txt Transposon tolerance map (R200) for chromosome I when grown on rich medium (prior to replating)

S10 Text. TTM_replated_ctrl_chrII.txt Transposon tolerance map (R200) for chromosome II when grown on rich medium (prior to replating)

S11 Text. TTM_replated_chrI.txt Transposon tolerance map (R200) for chromosome I when replated on rich medium

S12 Text. TTM_replated_chrII.txt Transposon tolerance map (R200) for chromosome II when replated on rich medium.

S13 Text. ChrI.gb Annotated chromosome I of *B. abortus* 2308.

S14 Text. ChrII.gb Annotated chromosome II of *B. abortus* 2308.

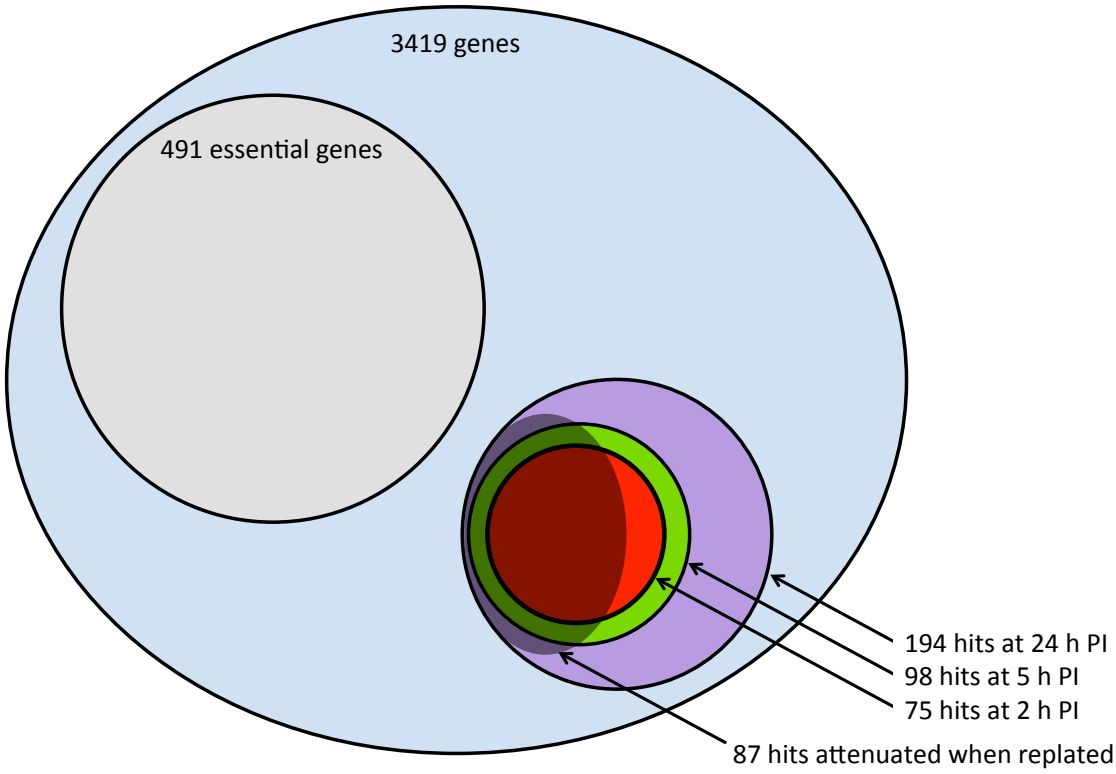
6) Funding, acknowledgments, author contributions

This research has been funded by the Interuniversity Attraction Poles Programme initiated by the Belgian Science Policy Office (https://www.belspo.be/) to J.-J.L. and by grants from Fonds de la Recherche Scientifique-Fonds National de la Recherche Scientifique [PDR T.0053.13, PDR Brucell-cycle T.0060.15, CDR J.0091.14, FRFC 2.4.541.08 F to X.D.B.]. This work was also supported by Swiss National Science Foundation [Grant 31033A_166476 to B.C.].

We would like to thank M. Waroquier for his flawless technical support, F. Tilquin for lab managing, as well as Dr. F. Renzi, J.-Y. Matroule and R. Hallez for stimulating and helpful discussions.

Conceptualization : JFS, XDB, **Formal Analysis :** JFS, XDB **Funding acquisition :** XDB **Investigation :** JFS, PG, RGDF, MVDH **Methodology :** JFS, XDB, BC, MC **Resources :** BC, MC, KP, NF, KW **Supervision :** JFS, XDB, JJL **Validation :** JFS **Visualization :** JFS, XDB **Writing– original draft :** JFS, XDB **Writing– review & editing :** JFS, XDB, JJL

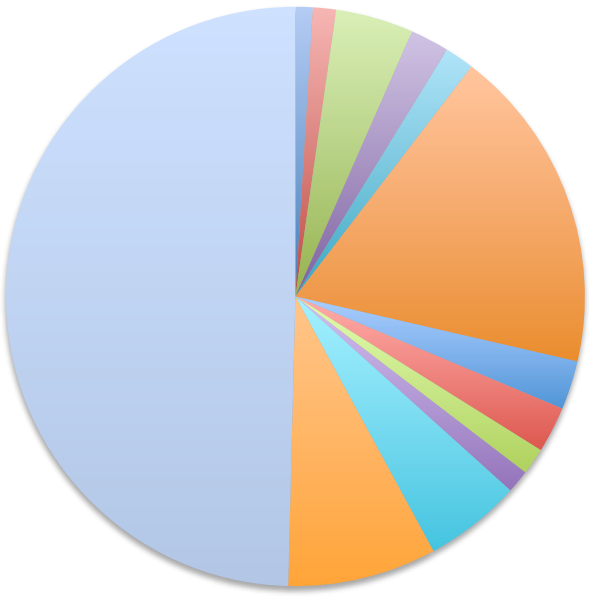
II. Visual overview of Tn-seq data



Schematic representation of Tn-seq data highlighting the number of hits per PI time points as well as the number of hits invalidated due to growth defects when regrown on rich medium ("re-plating" Tn-seq). As depicted above, all hits from 2 h PI are contained within the 5 h PI subset, and all hits from 5 h PI are contained within the 24 h PI subset. Out of the 194 initially identified candidates, the "replating" Tn-seq has allowed questioning the nature of the attenuation of 87 of these candidates (mostly in the 2 h PI and 5 h PI subset) due to various growth defects when re-grown on rich medium.

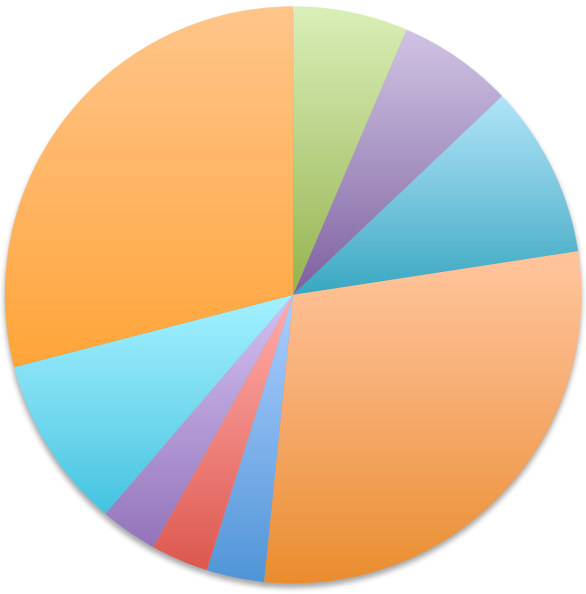
The three following pie charts display the succinct functional categorization of all *Brucella* genes (n= 3419), of the 2 h and 5 h PI candidates that do not display growth defects (n= 31), and of the 24 h PI candidates that to do not display growth defects (n= 46), respectively.

All



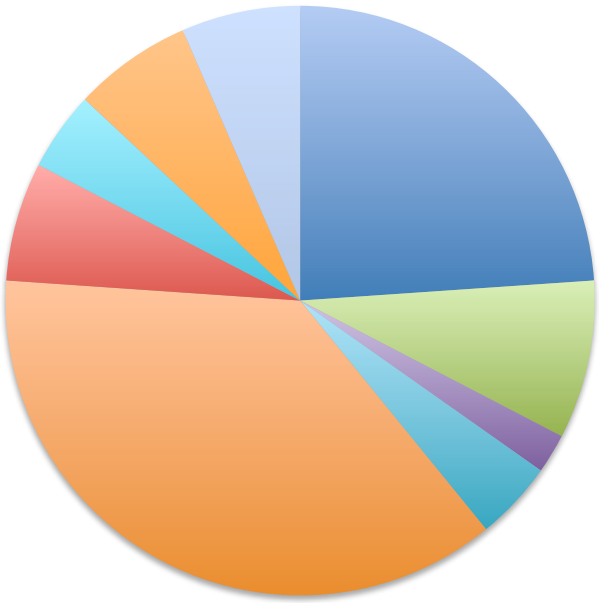
- Appendices (flagellum and secretion systems)
- Cell cycle (DNA replication, division)
- Control of transcription
- DNA processing (recombination, repair, RM systems, transcription, transposition)
- Enveloppe biogenesis and homeostasis
- Metabolism (central carbon, nitrogen, sulfur and phosphorus, aa, ...)
- Other functions (NOT hypotheticals)
- Respiration, redox processes and denitrification
- RNA processing
- Signalling (phosphorelays)
- Translation (incl. aa-tRNA synthesis, protein folding and proteolysis)
- Transport (substrates, ions)
- Uncharacterized conserved/ unknown

2h and 5h PI



- appendices (flagellum and secretion systems)
- cell cycle (DNA replication, division)
- control of transcription
- DNA processing (recombination, repair, RM systems, transcription, transposition)
- enveloppe biogenesis and homeostasis
- metabolism (central carbon, nitrogen, sulfur and phosphorus, aa, ...)
- other functions (NOT hypotheticals)
- respiration, redox processes and denitrification
- RNA processing
- signalling (phosphorelays)
- translation (incl. aa-tRNA synthesis, protein folding and proteolysis)
- transport (substrates, ions)
- uncharacterized conserved/ unknown

24h PI



- appendices (flagellum and secretion systems)
- cell cycle (DNA replication, division)
- control of transcription
- DNA processing (recombination, repair, RM systems, transcription, transposition)
- enveloppe biogenesis and homeostasis
- metabolism (central carbon, nitrogen, sulfur and phosphorus, aa, ...)
- other functions (NOT hypotheticals)
- respiration, redox processes and denitrification
- RNA processing
- signalling (phosphorelays)
- translation (incl. aa-tRNA synthesis, protein folding and proteolysis)
- transport (substrates, ions)
- uncharacterized conserved/ unknown

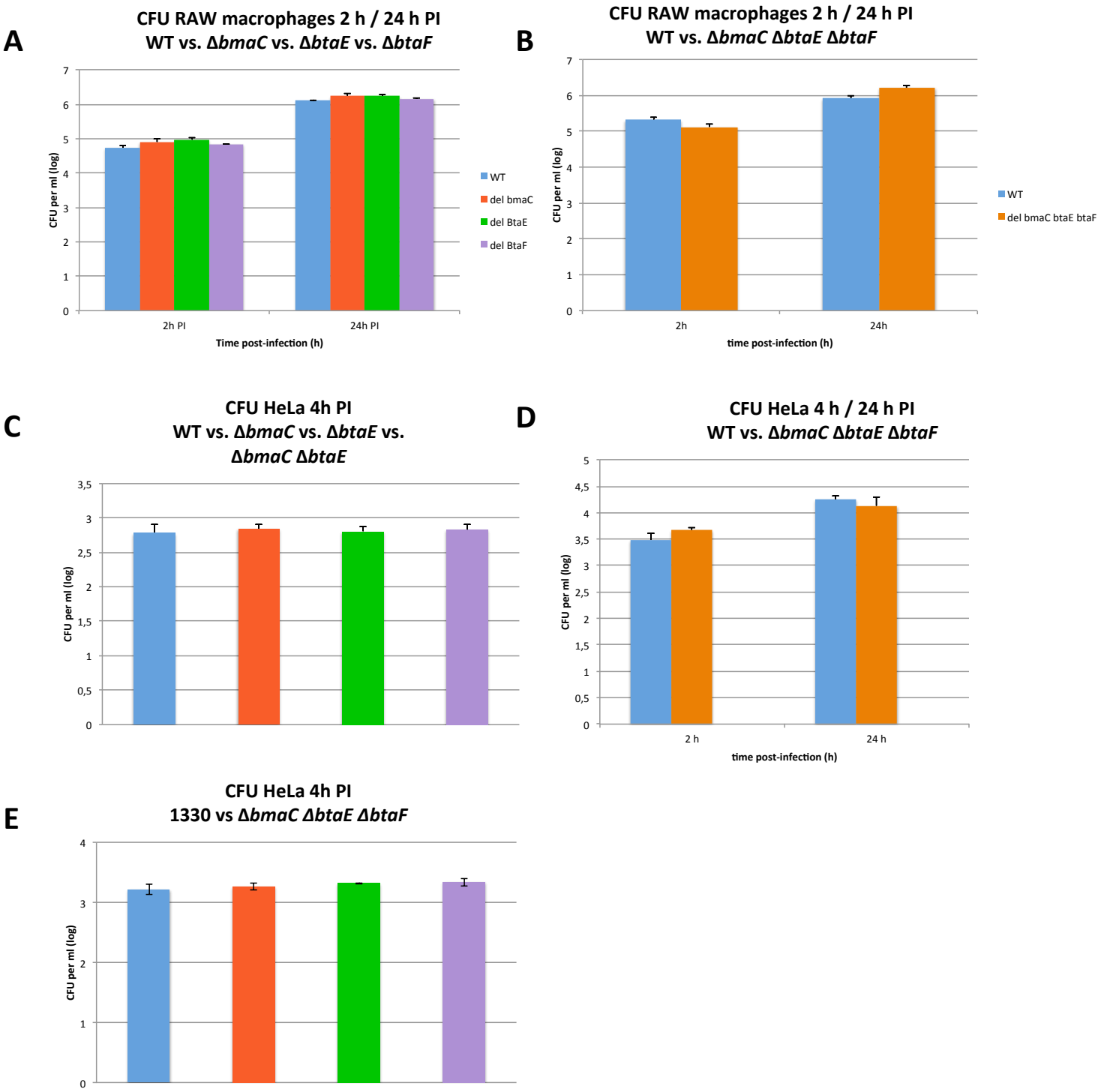


Figure 14. CFU counting of the *bmaC, btaE, btaF* and triple deletion mutants in different *Brucella* backgrounds and at different time points in HeLa cells and RAW 264.7 macrophages. WT strain in blue, Δ*bmaC* in red, Δ*btaE* in green, Δ*btaF* in purple, and triple deletion strain in orange. The error bars indicate the standard deviation calculated for a single experiment.

III. Additional investigations

Remark: Unless stated otherwise, the infectiosity (*i.e.* the ability to achieve a genuine cellular infection process) of each mutant presented in this work was assessed by performing infections of RAW 264.7 macrophages at either 2, 5, or 24 hours post infection followed by CFU counting (see sections 5 and 6 of Material & Methods).

1) *Brucella* adhesins

Given that the preferential selection of newborn bacteria was shown to happen at the onset of infection (see section I.5.b of the Introduction), we hypothesized that the underlying selection mechanism could be dependent on newborn-specific adhesion factors. In order to test this hypothesis, deletion mutants for *bmaC, btaE,* and *btaF* have been constructed in the *B. abortus* 2308 background (see Material and Methods for a description of targeted deletion strains constructions) and their infectious potential has been tested by CFU at 2 h PI, since adhesion impairment was shown to translate into invasiveness decrease observable by performing CFU at early PI time points (Posadas *et al.,* 2012; Ruiz-Ranwez *et al.,* 2013a; Ruiz-Ranwez *et al.,* 2013b). Unexpectedly, no decrease in infectiosity could be observed for any of these mutants either in HeLa cells or in RAW 264.7 (Figure 14A and C). Consequently, a triple deletion mutant (Δ*bmaC*Δ*btaE*Δ*btaF*) harboring the simultaneous deletion of *bmaC*, *btaE* and *btaF* was constructed based on the hypothesis that these three adhesins could have redundant functions, impairing the generation of a clear phenotype for the single deletion strains. The Δ*bmaC*Δ*btaE*Δ*btaF* strain was tested for its ability to infect HeLa cells or RAW 264.7 macrophages for 2 h or 4 h, respectively, but again, no attenuation phenotype could be observed in terms of CFU (Figure 14B and D). Given that our experimental settings differed from those presented in the literature by the use of *B. abortus* instead of *B. suis*, we decided to test the implication of the background strain on data reproducibility by reconstructing the Δ*bmaC,* Δ*btaE,* and Δ*btaF* single deletion mutants in *B. suis* 1330. However, no attenuation could be observed at 2 h PI in HeLa cells (Figure 14E).
Finally, as it was shown for *bigA* that adhesin overexpression could result in increased bacterial adhesion and invasiveness (Czibener *et al.,* 2016 *Cell Microbiol*), a *bmaC* overexpression strain was constructed during M. Van der Henst master thesis by replacing *bmaC* endogeneous promoter by the more active *PsodC* promoter. Nevertheless, despite a 20-fold increase in *bmaC* expression measured by RT-qPCR, no infectiosity increase could be observed.

The fact that no attenuation could be observed with neither single deletion mutants nor the triple deletion mutant in either background strains, including the *B. abortus* strain able to select newborn bacteria for infection, prompted us to reorient the work by trying to identify new adhesion factors using an approach without *a priori*.

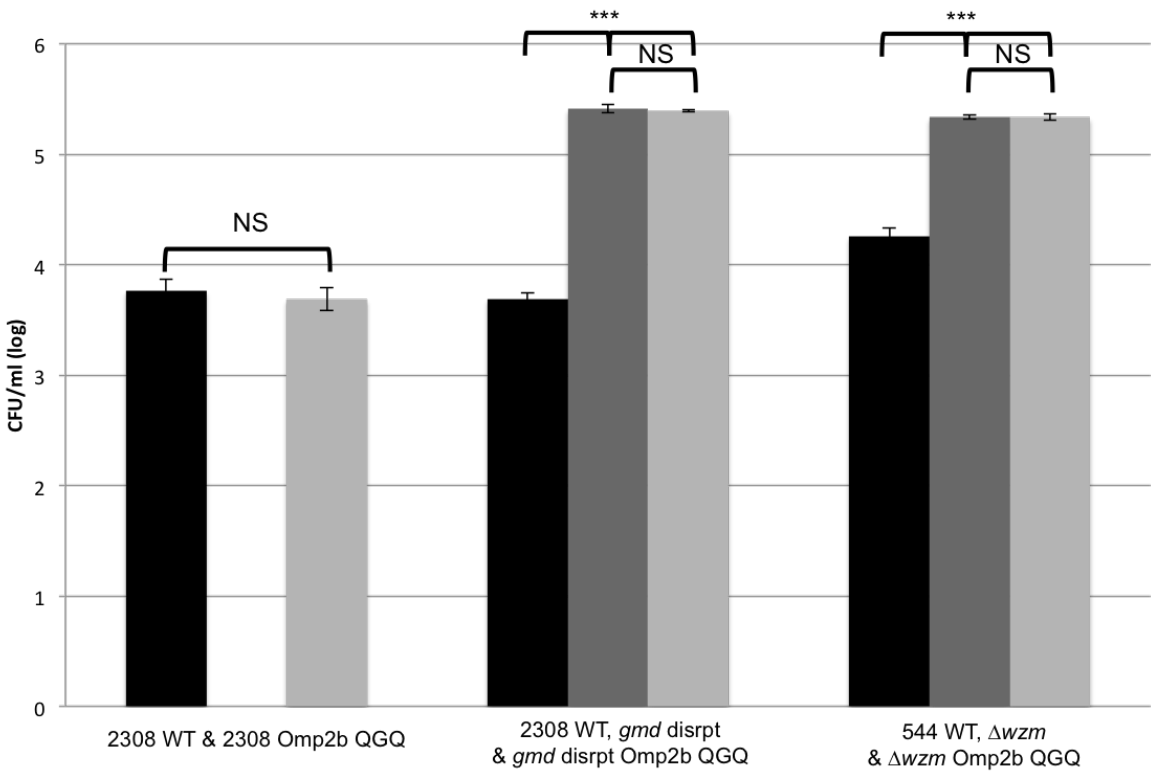


Figure 15. CFU counting at 2h PI of *B. abortus* 2308 and 544 both in the presence or absence of *omp2*B QGQ loop in conjunction with a wild type or a disrupted *gmd* allele. The error bars indicate the standard deviation calculated for a single experiment.

2) RGD motif in Omp2b extracellular loop

Adhesion of *B. abortus* to host cells was further investigated by assessing the role of an RGD loop from the outer membrane protein (OMP) Omp2b on *Brucella* infectiosity. Indeed, Omp2b is one of the most abundant OMP in *B. abortus,* and a predicted surface-exposed loop possesses a RGD motif, such motives being known for interacting with several integrins in eukaryotes (Ruoslahi, 1996), thus anchoring extracellular matrix components to the cellular membrane. Consequently, it was hypothesized that the RGD-containing loop of Omp2b could mediate adhesion between *Brucella* and its host cell. To test this hypothesis, a mutant lacking this RGD loop was constructed and tested in CFU. However, no change in internalization could be observed compared to the wild type strain (Figure 15). It was then postulated that, due to its size, the LPS O-chain could interfere with the hypothetical integrin-RGD loop recognition mechanism. Additionally, this could have explained why rough strains (lacking the LPS O-chain) are typically more adherent and invasive than smooth strains (Detilleux, *et al.,* 1990; Porte *et al.,* 2003). For this reason, a mutant lacking the Omp2b RGD loop was constructed in a rough disrupted *gmd* and a Δ*wzm* background (Figure 15), but again, no impact on internalization could be observed, showing that the RGD loop from Omp2b is not crucial for the invasion of RAW 264.7 macrophages.

3) *ialB, oppB* and other Tn-seq candidates prior to replating

Prior to perform the "re-plating" Tn-seq, several candidates from the initial Tn-seq have also been investigated. In particular, BAB1_0491 (*ialB*) and BAB2_0701 (*oppB*) were studied during the course of P. Godessart master thesis. BAB1_0491 is homologous to the *Bartonella bacilliformis* gene *ialB,* known for being involved in human erythrocyte parasitism, and BAB2_0701 (*oppB*) is part of an operon homologous to the oligopeptide ABC transporter operon *opp*. Due to their apparent attenuation in our initial Tn-seq, mutants have been constructed by plasmidic disruption (*i.e.* integration of a pNPTS138 plasmid in the middle of the *ialB* or *oppB* coding sequence by homologous recombination) for the sake of time. When testing these mutants in infection, we observed that the *ialB* mutant was attenuated at 2 h PI and that the *oppB* mutant, even though yielding CFU values similar to the wild type strain, displayed colonies of smaller size after infection (Figure 16A). Moreover, the *ialB* mutant displayed morphological changes seemingly reminiscent of rounder bacteria (*i.e.* a darker cell edge and a clearer cell center). Problematically, those results suffered from poor reproducibility, which prompted us to construct *bona fide* deletion mutants, as reproducibility issues could originate from inherent genomic instability of plasmidic disruption mutants.

Interestingly, no Δ*ialB* mutant could be obtained, which is in agreement with the "re-plating" Tn-seq showing a near-essential profile for *ialB* as almost no sequencing reads can be recovered from the replated condition for this gene. A deletion mutant for *oppB* was obtained, and the same phenotype as for the *oppB* disruption strain was observed (*i.e.* smaller colonies than the wild type after infection but similar CFU values at 2 h PI), and the "replating" Tn-seq profile of *oppB* suggests no growth defect on rich medium. Therefore, one can hypothesize

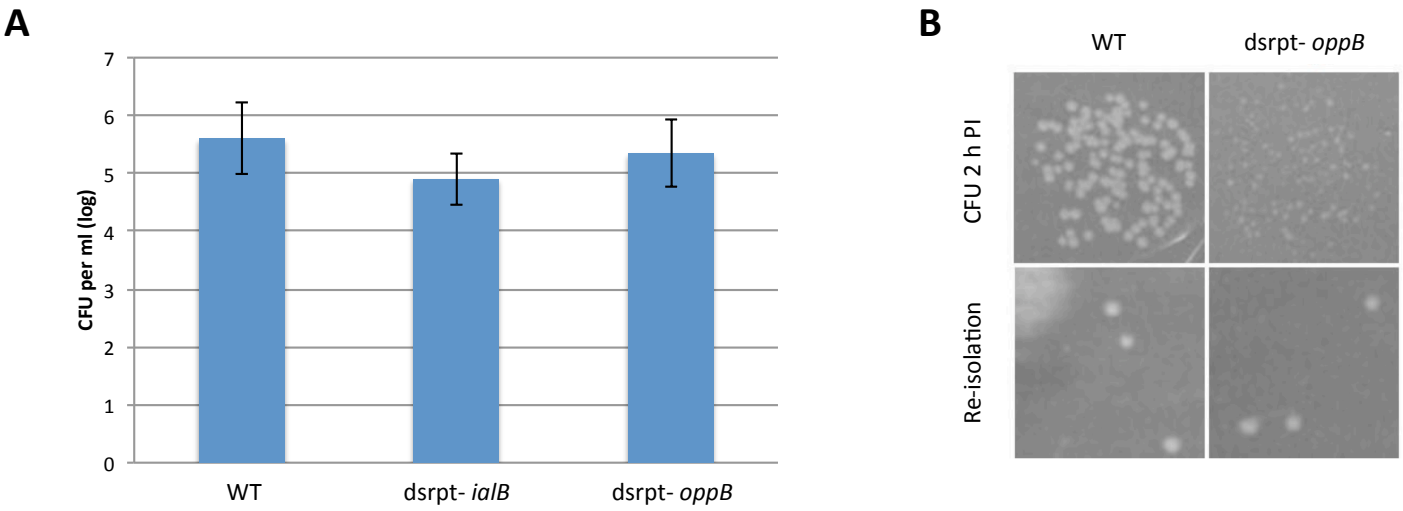


Figure 16. Investigation of *ialB* and *oppB*. (**A**) CFU counting of the *ialB* and *oppB* disruption strains showing no significant difference between these strains and the wild type strain. The error bars indicate the standard deviation calculated for three independent experiments. (**B**) Picture of the wild type strain and the *oppB* disruption strain when grown after infection (CFU) and after re-isolation from their respective CFU, highlighting the post-infection growth defect of *oppB* disruption strain.

that the observed Δ*oppB* growth defect (*i.e.* smaller colonies) observed only after infection but not when strains are regrown on rich medium (Figure 16B) could be a consequence of infection itself, such as specific damages occurring inside the host cell and which would require repairs prior to being able to resume cell cycle on rich medium. One could also hypothesize that such phenotype could be due to infection-related treatments such as Triton X100 exposure. Further investigations would be required to answer this question.

It should be noted that disruption mutants for *ccmE* and *purH* were also initiated due to their strong attenuation observed in our initial Tn-seq. However, the "replating" Tn-seq revealed that the observed attenuations were very likely to be consequences of strong growth defects. Consequently, these candidates were not further investigated.

4) Localization of the type IV secretion system machinery

Given that this thesis project initially aimed at characterizing crucial steps of *Brucella* cellular infection (the adhesion/invasion of hosts cells and the access to the replication niche), work was conducted in order to localize *B. abortus* type IV secretion system VirB by fusing fluorescent proteins to specific VirB components. Indeed *Brucella* is a polarized pathogen (Van der Henst *et al.,* 2013) and it is known for *A. tumefaciens*, *Coxiella burnetti* and *Legionella pneumophila* that the type IV secretion system can found at polar sites in the bacterial envelope (Judd *et al.,* 2005; Morgan *et al.,* 2010; Jeong *et al.,* 2017). In conjunction with the fact that the *Brucella* newborn bacteria are characterized by the presence of fresh new pole generated after bacterial cell division, it was postulated that the type IV secrection system could be located at these poles, thus justifying the newborn enhanced infectiosity. In order to assess VirB localization in *Brucella,* three fusion proteins were engineered based on structural predictions in order to have either periplasm-exposed (mCherry) or cytoplasm exposed (YFP) fluorescent proteins. Genes encoding these fusions where then integrated in *B. abortus* genome by allelic replacement, generating marker-less strains without plasmidic sequence in their chromosomes.

First, a mCherry-VirB6 fusion was constructed and checked by fluorescence microscopy, however, no usable fluorescence could be observed when grown in rich medium. The same results were obtained for a second fusion, VirB4-YFP. Finally, a YFP-VirB8 fusion protein was constructed. This mutant displayed an interesting diffuse membrane signal when grown on rich medium, however, a strong CFU decrease observed at 24 h in HeLa cells suggested a non-functional VirB system. Consequently, we decided to use the same fusion protein but to produce it from a replicative plasmid. It was postulated that a different expression system allowing the simultaneous expression of YFP-VirB8 and wild type VirB8 could result in a mixed VirB8 oligomer that could still display fluorescence while maintaining VirB function. The subsequent mutant possessed wild type-like infectiosity after a 24 h infection period in HeLa cells. Additionally, a diffuse membrane signal could be obtained for that mutant when grown in rich medium and slightly enhanced when incubated for 1 h in acidic PBS (pH = 4.5), suggesting that the fusion could be induced in conditions known for inducing VirB expression in *Brucella* (Boschiroli *et al.,* 2002). Unfortunately, exploitable

fluorescence signal could not be obtained for this fusion in infection. Moreover, the same fluorescence intensity and membrane localization as detected in *Brucella* when grown on rich medium could also be observed in *E. coli* after overnight cultures. The latter observation demonstrates that even in the absence of other VirB proteins, the YFP-VirB8 fusion itself is sufficient for fluorescence and membrane localization. Such data suggest that while being an interesting tool for membrane staining, the YFP-VirB8 fusion is impractical for VirB localization.

Discussion and perspectives

1) Lessons from the control condition

In this work, a highly saturating Tn-seq has been performed on the pathogen *B. abortus*. The first objective was to identify all genes required for the bacterium to grow on rich medium. Then, using time-resolved infections, the second objective was to characterize *Brucella* molecular arsenal allowing a cellular infection, by identifying genes causing attenuation at given times post infection when mutated.

Analyzing the control dataset allowed the identification, at the genomewide level, of essential genes for growth *in vitro* and represented the control condition for all subsequent datasets comparisons. However, this dataset also proved to generate a wealth of information by itself when extensively analyzed. Regarding general observations, a first example could be that the percentage of *Brucella* genome taken by essential genes (*i.e.* 14.4 % for 3.3 Mb) is quite low and similar to other Alphaproteobacteria, such as *C. crescentus* (12.4% for 4 Mb) and *Brevundimonas subvibrioides* (13.4% for 3.4 Mb) (Christen *et al.,* 2011)(Curtis & Brun 2014), which are non-pathogenic and free-living species. To some extents, this may challenge the genome reduction hypothesis that characterizes numerous pathogens (Moran, 2002) and that was addressed for *Brucella* (Wattam *et al.,* 2014) since non-pathogenic species with similar genome sizes also have similar percentages of essential genes. Indeed, one could have expected to find a higher percentage of essential genes, as observed for *M. genitalum,* with an estimated essential genes proportion of at least 55 % for a genome of 517 genes (Hutchison *et al.,* 1999). Nevertheless, one should keep in mind that the genome reduction hypothesis is not limited to the proportion of essential genes as it also dependent on the number of redundant homologs for a given function as well as pseudogenes proportions (Batut *et al.,* 2004) which have been suggested to play a significant role in the genetic variation and subsequent host adaptation of some *Brucella* species (Suárez-Esquivel *et al.,* 2017).

Another general observation is that the proportion of essential genes differs from one chromosome to another in respect to their different sizes. Indeed, chrI housed 3.7 times more essential genes than chrII. In conjunction with the fact that chrII possesses a plasmid-like replication and segregation machineries as opposed to chrI, this data further validates the hypothesis that *Brucella* chrII may have been acquired as an ancestral megaplasmid. Indeed, the unbalanced essential gene percentage could be the consequence of a relatively recent plasmid acquisition, where the frequency of random gene transfer was not sufficient to presently equilibrate the proportions of essential genes. *In extenso*, it would be interesting to perform Tn-seq on *B. suis* biovar 3, since variation in essential genes density throughout its single chromosome could suggest that this atypical genome structure results form a former fusion of both chromosomes, whereas a homogeneous essential gene density would indicate an ancestral single chromosome feature.

Regarding more specific observations, we were able to reshape essential pathways. For example, only one of the three Class-A PBP present in *Brucella* genome was essential for growth. This strongly suggests that this gene (BAB1_0932) codes for *Brucella* housekeeping PBP1. Interestingly, while one of the two remaining candidates (BAB1_0607) seems required for optimal growth on plates though not being strictly essential and displays aggravation in infection, the question regarding the function of the additional PBP encoding gene is raised.

One possibility could be that this gene is not expressed or that its basal expression level is extremely low, resulting in functional status quo. However, RNA-seq data tend to invalidate this postulate (data now shown). Another tempting hypothesis could be that its expression is induced upon specific conditions that necessitate particular peptidoglycan synthesis/remodeling/repair. However, one should keep in mind that this gene does not display any attenuation in the conditions tested in this work, suggesting that its inducing conditions were not touched in our work. Another example of specific observation was found by comparing our Tn-seq data to other data obtained from different species. In fact, *bamE* which codes for a component of the Bam system involved in OMPs export was surprisingly found to be essential in *Brucella* while being unessential in *E. coli*. Moreover, *Brucella* Bam system is incomplete in the sense that it lacks predicted homologs for both BamB and BamC as opposed to the archetypal *E. coli* BamABCDE system. One way to interpret these data could be that functional homology can be drawn between *Brucella bamE* and *E. coli bamB* or *bamC*, even though no sequence homology can be found by protein sequence alignment. Another possibility could be that *Brucella* has managed to get rid of accessory elements while maintaining a fully functional core system. Actually, both hypotheses can be combined knowing that in *E. coli,* both *bamA* and *bamD* are essential whereas *bamB*, *bmaC* and *bamE* are not (Malinverni *et al.,* 2006; Sklar *et al.,* 2007). Consequently, in *E. coli* the Bam core function is achieved by BamA and BamD while BamB, BamC and BamE likely perform functionally redundant duties justifying their dispensableness (Malinverni *et al.,* 2006; Sklar *et al.,* 2007). Therefore, one can easily portray that in *Brucella,* the seeming absence of *bamB* and *bmaC* results in a loss of redundancy observed in *E. coli*, thus making *bamE* essential alongside *bamA* and *bamD*. These hypotheses are in line with the evolutionary models suggesting that BamA and BamD are the ancestral subunits and biochemical data available for *C. crescentus* regarding the composition of the Bam complex (Anwari *et al.,* 2012).

Given the examples detailed above, Tn-seq demonstrates that essential pathways can be redrawn and that numerous strong hypotheses can be generated from a single experiment using only data gathered from the rich medium condition. Due to conceivable therapeutic outputs, this method is of great use when working on pathogens, especially when restrictive handling measures are required as for class III pathogens. Nonetheless, such findings necessitate substantial data mining efforts that require specific case-by-case investigations. For this reason, much additional work is needed to explore the full potential of this dataset.

2) Analysis method: the goods and the bad

Due to the very large size of Tn-seq datasets after reads mapping (*i.e.* each dataset consisting of a text file containing approximately three million lines each housing a numerical value ranging from 0 to ~10^4), the straightforward comparison of datasets imposed data processing. We decided to create a parameter to assess transposon insertion density termed R200, defined by the $\log_{10}$ of the number of mapped sequencing reads +1, on a 200 bp window that is sled every 5 bp on *Brucella* genome. Beside smaller datasets, the two major advantages of this technique are that the output data does not depend on either gene prediction

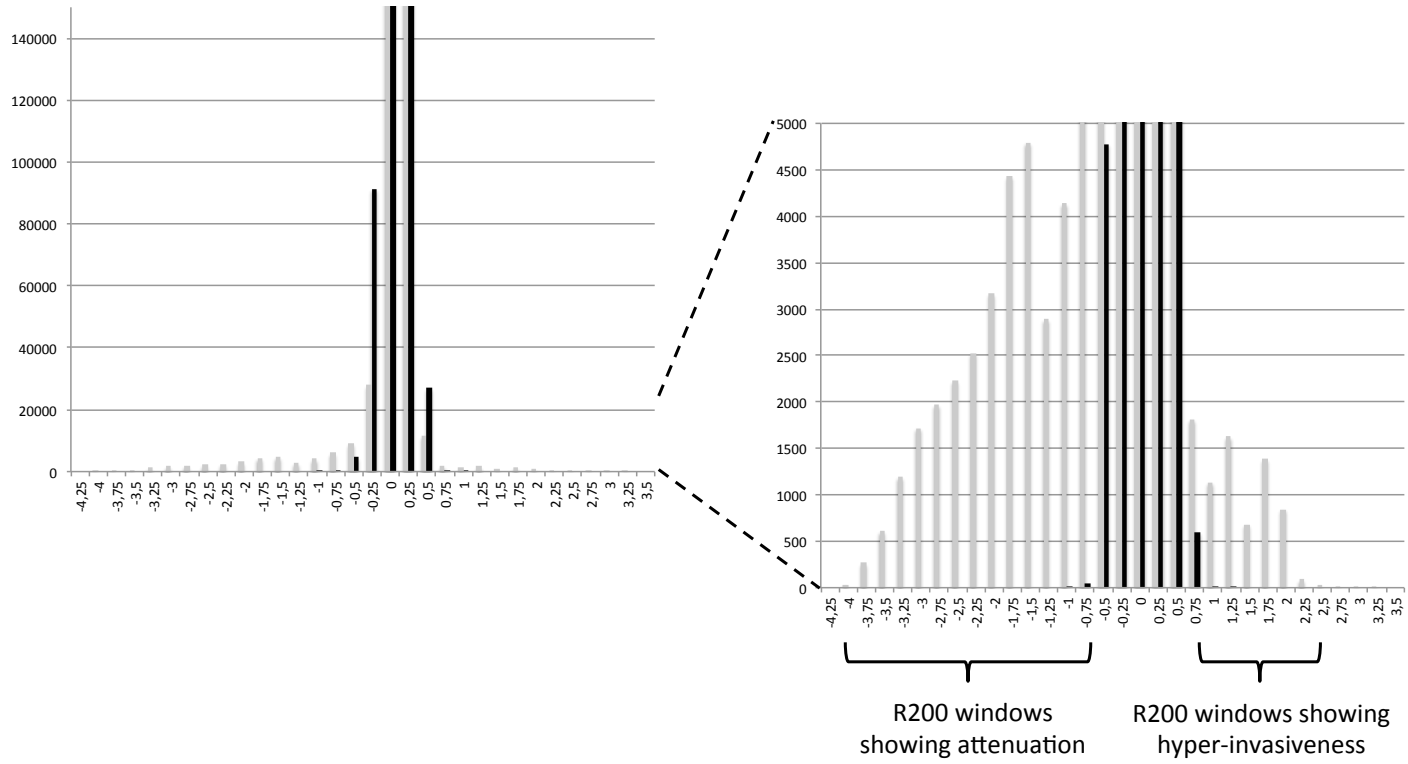


Figure 17. Principle of Delta-R200 threshold determination. This graph displays classes of Delta-R200 values on the X-axis and the frequency of such classes on the Y-axis. By fitting a normal curve (black profile) on the observed distribution of Delta-R200 values (grey profile), we show that observed values both negatively and positively exceed values excepted from the normal curve. Consequently, this suggests that the wide distribution of observed values is unlikely to be randomly obtained and rather corresponds to a biological effect (*i.e.* attenuation for lower values, and hyper-invasiveness for higher values) (see section 8 of Material & Methods).

or gene size, making it useful for analyzing previously uncharacterized genomic regions. With this method, datasets could be simply analyzed by subtracting them to one another depending on the desired comparison.

To our knowledge, a softer spot of our analysis is probably the threshold determination for R200. Indeed, when dealing with numerical value data regarding mutant fitness, one has to arbitrarily draw a line to best distinguish between background noise and actual information. Consequently, this decision is directly dependent on the subjectivity of the experimenter. Fortunately, this problem does not actually apply to the determination of essential genes, as it is a particularly strong phenotype (Figure 9B, see section I.3.a of the Results). Indeed, based on our R200 analysis, essential genes were defined in this work as having at least one 200 bp window with no transposon insertion. Given the size of *Brucella* genome, the number of insertions sites, and the size of the window to be analyzed, it was estimated that the probability of fortuitously finding an empty window was approximately $3.8\ 10^{-15}$ (see section 7 of the Material & Methods). Knowing that a genome processed by the R200 method contains roughly $6.6\ 10^{5}$ R200 values, obtaining even a single empty window by chance proved to be extremely unlikely, assuming that mini-Tn*5* insertion occurs only randomly. While not being a problem for essential genes identification, the arbitrary threshold determination proved to be more challenging for phenotypes of weaker amplitude, such as the low fitness reduction observed for some genes at short infection time points. As a reminder, the control R200 was subtracted from the dataset to be analyzed (Delta-R200), resulting in negative values for attenuated regions and values around zero for unaffected regions. Here, because of the wide distribution of Delta-R200 values (*i.e.* ranging from 0 to -4), we created a method where frequency classes were established for the Delta-R200 values. Briefly, this computing uncovered a curve showing a main peak centered on 0 (containing all non-affected regions) and a wide range of negative R200 values (containing all attenuated regions). By fitting a normal curve on the 0-centered peak, we were able to set a threshold beyond which negative Delta-R200 values were statistically more frequent than what could be expected by chance (Figure 17). Due to the distribution width, the position of the threshold impacts downstream experiments as it dictates the noise *vs.* information limit. Indeed, when assessing attenuation, moving the threshold closer to 0 increases information together with the proportion of false-positive candidates, while moving it further away decreases the proportion of false-positive candidates but also removes information. This demonstrates that when dealing with "grey color gradient" data distributions instead of "black or white" distributions as it is often the case in biology, opting for the best approximation is an imperfect but unavoidable decision for analyzing large datasets.

3) *Brucella* doesn't need many tools at first, or does it ?

A striking observation regarding our short-term infection times (2 h and 5 h PI) is the low number of candidates required for the initial steps on RAW 264.7 macrophages infection. In fact, out of the 3419 predicted genes of *B. abortus*, only 30 were scored as genuinely attenuated when mutated at early time points in Tn-seq. Moreover, most of these candidates

have predicted functions that could lead to pleiotropic phenotypes and/or functions that can only be indirectly linked to infectiosity, such as metabolism and regulation of gene expression. This leads us to several hypotheses.

First, one could postulate that the bacterial entry inside host is a passive process mainly driven by the host cell itself. This seems unlikely because G1 bacteria have a better chance to be internalized compared to other stages of the bacterial cell cycle, suggesting that the bacterium is able to control its uptake in this cell line. However, survival inside RAW 264.7 macrophages could be dependent on a dormancy-like state. This would be in agreement with the apparent absence of early virulence factors. In addition, presumed early intracellular conditions such as nutrient starvation and endosome acidification are known for triggering the type IV secretion system expression (Boschiroli *et al.,* 2002), which in turn allows the bacterium to access its replication niche and resume proliferation. Therefore, such a dormancy-like state could be maintained until the adequate conditions for *virB* expression are encountered.

A second possibility could be that many functions required for bacterial entry and early survival are carried by redundant genes, as this would explain why no such candidates could be highlighted by Tn-seq.

Finally, a third hypothesis could be that such early pathogenesis functions are carried out by genes that are already essential for growth on rich medium. Indeed, due to its professional pathogen lifestyle, one can imagine that *Brucella* could have coordinated its cellular infection cycle with its own bacterial cell cycle. Following this idea, genes performing essential tasks for *Brucella* homeostasis could have evolved to share a role in its infection process. It should be noted that this hypothesis was one of the reason why the RGD loop found in Omp2b was investigated (see section III.2 of the Results). Indeed, the *omp2b* gene is essential and thus its role in cellular infection cannot be tested with a conventional genetic strategy. Similarly, genes leading to an aggravated phenotype could also fall into the same category. Moreover, the line between aggravation and essentialness being sometimes blurry, aggravation should be addressed as an individual point (see below).

4) A few precisions on hyper-invasive mutants

An interesting outcome of our Tn-seq was the ability to identify mutants displaying increased internalization inside host cells compared to the wild type strain. In fact, due to their unusually positive Delta-R200 values at early times PI, 9 genes stood out as resulting in bacteria being more found inside host cells than the wild type strain when mutated, hence being more invasive. All 9 genes are involved in the synthesis of the LPS O-chain which absence is already known for resulting in rough strains that display increased invasiveness (Detilleux *et al.,* 1990; Porte *et al.,* 2003). However, the real strength of this observation lies in the saturating nature of our Tn-seq. Indeed, all genes present on *Brucella* genome have been tested and failed at producing a hyperinvasive phenotype beside those 9 candidates. Therefore, we can conclude that no other genetic screen relying on transposon-mediated gene disruption should be able to highlight mutants causing hyperinvasion of RAW 264.7

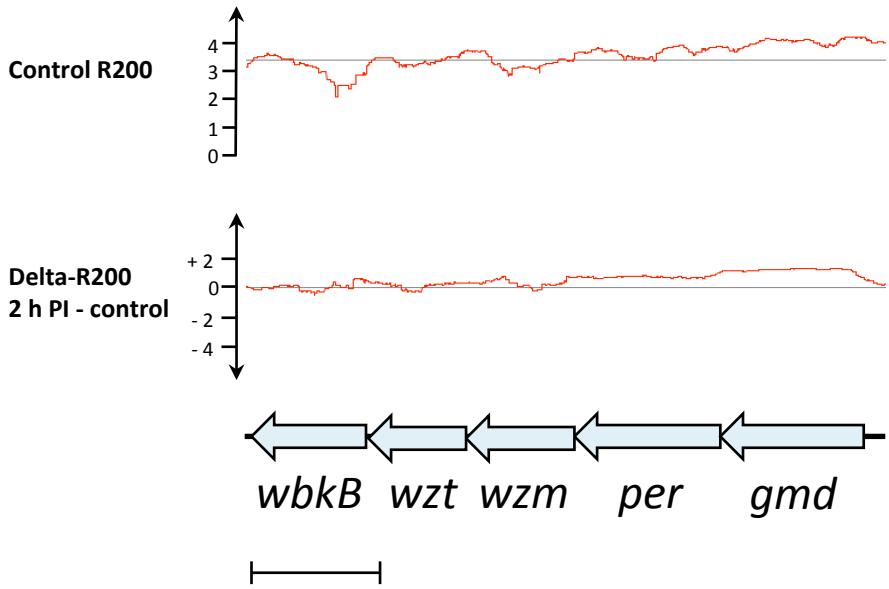


Figure 18. Balance between hyper-invasiveness and inherent fitness reduction of some rough mutants. This comparison of control R200 and Delta-R200 values obtained at 2 h PI suggest a correlation between the absence of hyper-invasiveness for *wzm* mutants (*i.e.* with values close to 0 on the lower graph), and the altered fitness on plates (*i.e.* values below the chromosomal average on the upper graph). Likewise, both curves globally share a similar shape, suggesting that the progressive diminution of fitness on plates observed from right to left in this genomic segment results in a progressive balancing of hyper-invasiveness.

macrophages besides genes involved in the LPS O-chain biosynthesis. As a side note, not all genes contributing to the formation of a *bona fide* LPS O-chain generated hyper-invasiveness in Tn-seq. This could likely be explained by the fact that disruption of some genes involved in LPS O-chain biosynthesis result in fitness reduction on plates causing various degrees of aggravation that counterbalance their increase in invasiveness. This hypothesis can be exemplified by the *wzm* (Figure 18), which is part an ABC transporter involved in O-chain export across *Brucella* inner membrane (Godfroid *et al.,* 2000).

5) Characterizing *Brucella* intracellular lunchbox

Another outcome of the Tn-seq approach was the ability to use mutants as metabolic probes to further define "the" *Brucella* replicative compartment. Classically, metabolites characterization requires the extraction of given organelles or tissues, a tedious task highly prompt to sample contamination and/or alteration (Liebeke & Bundy 2012; Schock *et al.,* 2016). Given these issues, another method could be the use of pathogens themselves as means for gathering information regarding their immediate environment in the host. In regards to this matter, the understanding of host-pathogen interactions for host characterization has already been initiated (*Microbes* as *Tools* for *Cell Biology, Methods in Cell Biology, vol. 45). Unfortunately, the latter example has been criticized for missing most of the initial goal by focusing primarily on technical information* (*Goldberg, 1995. Trends Microbiol*). Here we suggest that the use of a wide range of mutants, and more specifically metabolic mutants, could be used for a qualitative characterization of the ER composition thanks to their respective ability or inability to replicate in given conditions. The main example of niche characterization in this work is the pyrimidine biosynthesis pathway. Indeed, we have shown that the mutation of any of the non-essential genes involved in the biosynthesis of the pyrimidines precursor UMP greatly impairs *B. abortus* intracellular replication while being dispensable for growth on rich medium or for survival in the first 5 hours of *Brucella* intracellular trafficking inside RAW 264.7 macrophages. Moreover, the trafficking of the Δ*pyrB* mutant has been characterized, demonstrating that unlike a Δ*virB* mutant, the Δ*pyrB* mutant is able to reach its replicative niche while not being able to proliferate. The data summarized above thus suggest that the precursors involved in all non-essential steps leading to UMP biosynthesis are either absent from the ER or not present in a sufficient amount to permit *Brucella* duplication.

It would be interesting to further validate the pyrimidine requirement in the ER by rescuing the pyrimidine scarcity phenotype. A first risky but simple way to test this hypothesis would be to perform infections using the Δ*pyrB* mutant and add pyrimidines or pyrimidine precursors into the extracellular medium after 5 h PI and check for bacterial cell cycle resuming. Another tempting perspective would be to create a depletion allele, such as a Δ*pyrB* strain for which *pyrB* expression would be triggered by the presence of IPTG. In this case, bacteria would be cultured and used for initial infection time points without IPTG as *pyrB* is dispensable for growth and for the first hours of intracellular trafficking. Then, IPTG would be added after 5 h PI to check if bacterial duplication can be restored by inducing *pyrB*

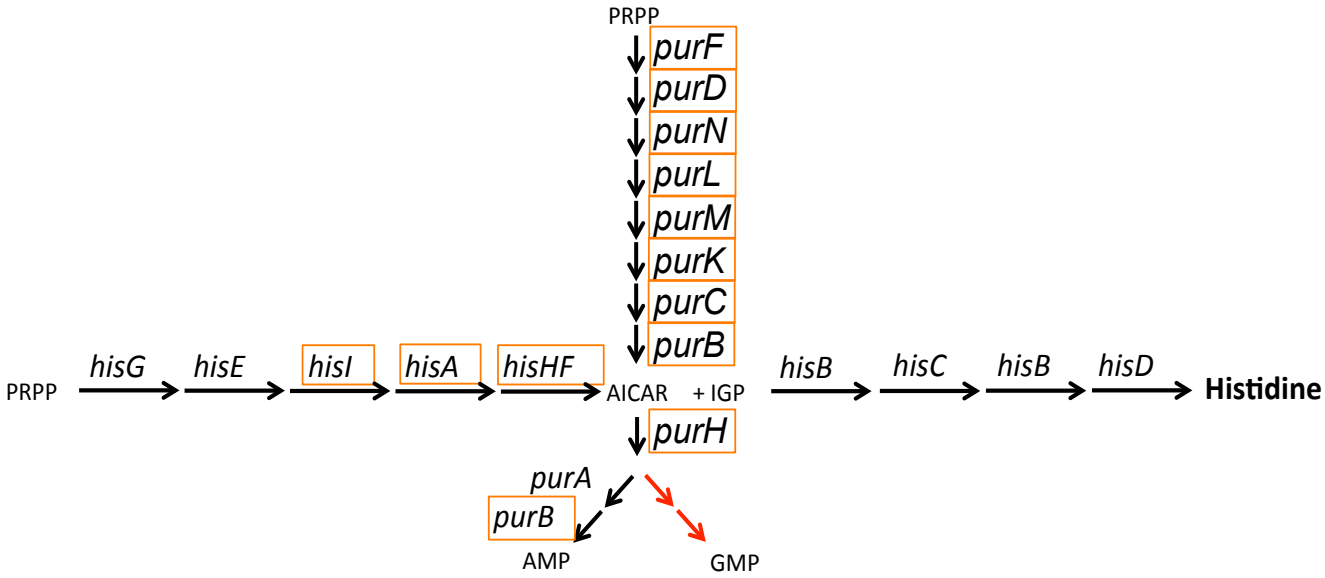


Figure 19. Links between the purine and the histidine biosynthesis pathways in *Brucella*. Reactions performed by non-essential genes are indicated by black arrows and reactions performed by essential genes are indicated by red arrows. Orange boxes show genes displaying aggravation. This metabolic map shows that besides the common phosphoribosyl pyrophosphate (PRPP) precursor, a major connection between the purine and histidine biosynthesis pathway is the 5-Aminoimidazole-4-carboxamide ribonucleotide (AICAR). The fact that none of the four enzymes involved in histidine synthesis from imidazole glycerol phosphate (IGP) display aggravation suggests that the aggravation of *his* genes upstream of AICAR is likely independent of an impaired histidine production. Likewise, the concomitant attenuation of *hisB, hisC* and *hisD* mutants at 24 h PI suggests that histidine biosynthesis is not mandatory prior to that PI time point.

expression only once bacteria have reached the replication niche. The advantage of this technique is that a depletion vector has already been designed and validated for use during *Brucella* infection by Nayla Francis, thus proving that adding extracellular IPTG is sufficient for reaching the ER lumen and trigger P*lac* activation inside *Brucella* (Francis *et al.,* 2017). Another interesting perspective would be to assess whether the intracellular proliferation inhibition is specific to a pyrimidine shortage or rather to a more general nucleobases shortage. One way to answer this question would be to use the same depletion allele strategy as described above, but using a purine biosynthesis gene instead of a *pyr* gene. Given that most *pur* genes are strongly aggravated on plates, one should initially express the chosen *pur* gene constitutively by adding IPTG to the culture medium and the infection medium. This time however, IPTG should be removed from the infection medium once bacteria have reached the ER. Thanks to this method, we should be able to test the effect of a purine shortage inside the ER while circumventing the issues associated with the assessment of aggravated/essential genes implication in infection. Also, it would be interesting to assess if a Δ*pyrB* mutant is able to reinitiate DNA replication within the ER (*e.g.* thanks to the remains of a hypothetical pyrimidines stock) by using the chromosome replication initiation monitoring tools described earlier (see section I.5.b of the Introduction).

As for pyrimidines, analysis of data from the control condition in conjunction with post infection data allows to postulate strong hypotheses regarding other components of *B. abortus* intracellular niche. An interesting example is the histidine requirement, where we have observed that transpositional mutants of either *hisB, hisC,* or *hisD* are strongly impacted at 24 h PI, while being able to grow as the wild type on rich medium. These data suggest that the ability of *Brucella* to produce histidine is mandatory for proliferating inside the host cell, thus meaning that the ER does not contain (enough) histidine. These data validate anterior studies showing strong attenuation phenotype of *his* mutants (namely, *hisD* and *hisF*) (Köhler *et al.,* 2002). However, the saturating nature of our Tn-seq also allowed us to observe that the histidine biosynthesis pathway could be phenotypically split in two categories. In fact, while the second half of the pathway is impacted at later stages of the infection as detailed above, the first half of the pathway (composed of *hisZ, hisG, hisE, hisI, hisA, hisH, hisF*) is already strongly impaired for growth on plates. A likely reason for this observation could be that this portion of the *his* pathway is connected to the purine biosynthesis pathway by the production of the purine precursor AICAR (see Figure 19). Knowing that the purine biosynthesis pathway is strongly aggravated on plates and essential when re-plated, the interconnection between the *his* and the *pur* pathway likely implies pleiotropic effects that question data relative to the requirement histidine in infection. However, the second part of the *his* pathway (*hisB, hisC,* and *hisD*) is predicted to be solely dedicated to the synthesis of histidine from imidazole-glycerol-3-phosphate, and Tn-seq showed that all three genes are required for the proliferation of *B. abortus*. Consequently, thanks to high saturation, we can strongly hypothesize that histidine biosynthesis is required for *bona fide B. abortus* infection despite most *his* genes being aggravated on plates.

Another example regarding *Brucella* replication niche composition would be lysine biosynthesis. The main pathway predicted to be involved *Brucella* lysine biosynthesis is the decarboxylation of diaminopimelate performed by the enzyme LysA. Interestingly, this enzyme is neither essential in rich medium nor attenuated at any of the tested PI time points,

and the only potent predicted homolog is an ornithine decarboxylase. Consequently, these data allow us to propose that *Brucella* has access to sufficient amounts of lysine both on rich medium plates and in the ER of RAW 264.7 macrophages to sustain growth and division. Unexpectedly, anterior transpositional screenings have scored *lysA* as being required for the proliferation of *B. suis* inside macrophages (Köhler *et al.,* 2002). However, such differences could be explained by the fact that these data were obtained using a different experimental setup, such as the use of *B. suis* as the *Brucella* strain, THP1 macrophages as the host cell line, and TSB as rich medium. Thanks to similar methods, the in-depth characterization of the different amino-acids requirements for *Brucella* inside RAW 264.7 macrophages and other host cell types is planned to be addressed in upcoming PhD theses.

To conclude this point, we would like to stress that a major parameter regarding the assessment of *Brucella* intracellular requirements as well as the likely foundation of all metabolic aggravations is the culture medium. Indeed, due to its composition, the culture medium intrinsically sets a strong selection throughout all experiments. Accordingly, though being a rich medium, the main weakness of the 2YT used in this work is its undefined nature. In fact, having few information regarding the presence and/or amount of numerous metabolites on plates makes it difficult to compare to infection, thus hampering most downstream investigations. Consequently, we strongly advocate the use of a both rich and defined medium when possible, such as RPMI in the case of *B. abortus*.

6) Behind the curtains of aggravation

Aggravation was initially defined in this work as a specific phenotype where mutated genes result in attenuation during infection in conjunction with apparent growth defects during *in vitro* culture. This is the case for the *fadAJ* operon as shown in Figure 9B (see section I.3.a of the Results). In this first case, growth defects can already be observed in the control dataset as regions being noticeably lower in R200 compared to the surrounding background while not reaching 0, which characterizes essential genes. To our surprise however, many of the candidates thought as attenuated in our initial Tn-seq turned out to have various degrees of growth deficiencies when grown on rich medium as highlighted by our "re-plating" Tn-seq. In addition, it should be noted that most of the candidates characterized as aggravated in the initial Tn-seq became essential in the replated Tn-seq dataset. Consequently, the definition of aggravation was extended to candidates displaying fitness reduction when replated.

Given the observations detailed above, the quantification of aggravated phenotypes was considered. Indeed, one can imagine that, even though having a slight to mild growth defect on plates, a given mutant could be strongly impacted in infection due to the same impairment. Following this idea, an example would be a metabolic mutant for which the impacted pathway slows its growth on plates while also being required for intracellular proliferation. This could be the case for the *znuABC* operon, responsible for zinc uptake (Kim *et al.,* 2004). Another example could be the two components system BvrRS that is involved in outer membrane homeostasis (Sola-Landa *et al.,* 1998) and for which aggravation is found

only for *bvrS* while *bvrR* is strongly attenuated at 2 h PI. One way to interpret these data could be that BvrS has other roles than simply (de)phosphorylating BvrR, thus resulting in pleitropic effects justifying aggravation on plates. Indeed, the *bvrRS* operon itself is part of bigger operon that includes members of the phosphotransferase system (PTS), thus suggesting a plausible yet undescribed functional interaction with this pathway. Another similar hypothesis could be that *bvrS* also senses signals for other two components systems. Finally, we cannot rule out that in *bvrS* mutants, a completely unphosphorylated BvrR is somehow toxic for *B. abortus*.

Problematically, the great diversity in terms of aggravation amplitude makes it a very heterogeneous category of mutants. Consequently, we advise that the investigation of aggravated candidates should be addressed case by case. Nevertheless, one can speculate that the softer end of the aggravation spectrum probably includes several mutants that would be worth investigating. As explained above, the composition of the medium sets a strong selective pressure. Therefore, having a perfect medium that would fulfill every single need for *Brucella* growth would be ideal to circumvent metabolism-mediated aggravation and identify genes involved in infection only, and as stated above, RPMI appears to be a suitable candidate.

7) Other infection models

One of the strengths of the Tn-seq approach as described in this work is the ability to easily swap infection models while retaining a globally similar setup. In fact, other cultured cells can be used as the host moiety, thus enabling the comparison of essential gene panels depending on the tested eukaryotic cell type. Such experiments have already been conducted in the course of Phuong Thi Anh Ong ongoing PhD thesis (unpublished), the JEG-3 trophoblastic cell line has been used instead of RAW 264.7 macrophages as host cells. Surprisingly, few genes were found to be specifically required for trophoblast infection compared to macrophages infections, thus suggesting that the molecular arsenal required for *Brucella* infection could be a more straightforward than expected and less specific to host cell types despite the fact that the intracellular trafficking of *Brucella* inside trophoblastic cells strongly differs from macrophages (see section I.5.b of the Introduction).

Additional models also include animal models. For example, a Tn-seq with a mice intranasal infection model has been performed using *B. melitensis* during the ongoing PhD thesis of Georges Potemberg (unpublished) and is currently being analyzed. *In extenso,* one could imagine using *B. abortus* natural host, the cattle, as a model for performing Tn-seq. Indeed, biosafety level 3 animal facilities allowed to handle and infect bovines permit such experiments to be conducted. In these settings, the use of spatiotemporally resolved infections (*i.e.* collecting given organs at given times points) could allow the in depth characterization of *Brucella* infectious cycle while escaping the majority of biases imposed by the use of a non-natural host models.

Nevertheless, it should be noted that, especially for animal infections, one of the major parameter of Tn-seq is the infection bottleneck. In fact, in order to yield usable data, Tn-seq

requires a high diversity of sequenced reads to reach sufficient genome coverage. Consequently, while the number of reads may be high after next generation sequencing, low diversity datasets (*i.e.* conditions where numerous colonies are collected but only a few different genotypes are present in terms of transposon insertions) will undermine or even prohibit further data analyses. This issue has already been encountered by Georges Potemberg (unpublished), where the reads diversity was too low to enable gene fitness estimation when collecting bacteria from mice lungs after a 3 days infection period. Therefore, preliminary assessment of read diversity per animal is mandatory to approximate a proper number of animals to be used.

8) On the origin of the pathogenicity

As a general conclusion, one can postulate that the emergence of *Brucella* as a professional pathogen could have been a quick process on an evolutionary time scale (Wattam *et al.,* 2014) and where its intracellular trafficking could be a mostly passive process dependent on preexisting structures. Indeed, based on phylogenetical studies, the closest genus to *Brucella* is the *Ochrobactrum* genus that regroups soil bacteria able to colonize a wide variety of environments (Velasco *et al.,* 2000; Scholz *et al.,* 2008). Interestingly, *Ochrobactrum* is also known as an opportunistic and nosocomial pathogen, mostly able to infect humans with preexisting medical conditions (Alnor *et al.,* 1994; Romano *et al.,* 2009). In addition, it has been shown that *Ochrobactrum* outer membrane shares numerous similarities with *Brucella* outer membrane and is impermeable to hydrophobic probes and resistant to EDTA (Velasco *et al.,* 2000). Remarkably however, *Ochrobactrum* is susceptible to outer membrane destabilizing cationic peptides mix polymixin B while *Brucella* is much less susceptible thanks to changes in the structure of its LPS core (Velasco *et al.,* 2000). Consequently, one could imagine that due to the seeming absence of specific mechanisms for internalization and early survival, *Brucella* infection predominantly relies mechanisms acquired as a soil bacterium. Indeed, the ability of *Ochrobactrum* to infect unhealthy hosts suggests the presence of imperfect virulence factors and/or high environmental stress resistance underlying its opportunistic pathogenicity. Following this idea, *Brucella* could have evolved as a soil bacterium/opportunistic pathogen that has acquired a handful of key pathogenesis factors such as the type IV secretion system *virB* and LPS modifications (Wattam *et al.,* 2014). Given that such factors may have been acquired by horizontal gene transfer, the actual transition from an opportunistic to professional pathogen lifestyle could have happened over a short period of time. While difficult to validate, this transition hypothesis would be in agreement with the postulate that the intracellular trafficking of *Brucella* could be a rather passive process mostly dependent on mechanisms originating from environmental stresses resistance. Additionally, the blend between *Brucella* infectious cycle and cell cycle could have merged related functions into single genes, thus likely producing essential or aggravated Tn-seq profiles. Taken together, these two hypotheses could explain why Tn-seq did not identifying early virulence factors. Eventually, the resulting proto-pathogen would then have evolved depending on the specific selective pressures of its host,

such as the presence or absence of given metabolites in the replication niche, to become the professional pathogen now known as *Brucella.*

Material and methods

1) Bacterial strains and media

The wild type strain *Brucella abortus* 2308 Nal^R was cultivated in 2YT (1% yeast extract, 1.6% peptone, 0.5% NaCl). The conjugative *Escherichia coli* S17-1 strain was cultivated in rich medium (Luria-Bertani broth). When required, antibiotics were used at the following concentrations: ampicillin, 100 µg ml^{-1}; kanamycin 50 µg ml^{-1} for *E. coli* and 10 µg ml^{-1} for *B. abortus*; Nalidixic acid, 25 µg ml^{-1}.

2) RAW 264.7 macrophages culture

RAW 264.7 cells are murine macrophages derived from an Abelson murine leukemia virus induced tumor. These cells were cultivated in DMEM (Invitrogen) supplemented with 10% fetal bovine serum (Gibco), 4.5 g l^{-1} glucose, 1.5 g l^{-1} $NaHCO_3$, and 4 mM glutamine at 37 °C with 5% CO_2.

3) Mini-*Tn*5 Kan^R plasmid construction

The pXMCS-2 mini-*Tn*5 $Genta^R$ plasmid (Christen *et al*., 2016) was manipulated to exchange the gentamycin resistance cassette ($Genta^R$) with a Kan^R gene, using a dual joining PCR strategy. The region upstream of the $Genta^R$ cassette was amplified by PCR from the pXMCS-2 mini-*Tn*5 $Genta^R$ plasmid using primers "Tn-Kan part 1 F" and "Tn-Kan part 1 R" and fused by overlapping PCR to the Kan^R coding sequence, amplified from the pNPTS138 plasmid using primers "Tn-Kan part 2 F" and "Tn-Kan part 2 R". Then, this DNA fragment was fused to the region downstream of the $Genta^R$ cassette amplified by PCR using primers "Tn-Kan part 3 F" and "Tn-Kan part 3 R" from the pXMCS-2 mini-*Tn*5 $Genta^R$ plasmid. In parallel, the pXMCS-2 mini-*Tn*5 $Genta^R$ plasmid was restricted using *Eco*RI and *Nde*I to excise the $Genta^R$ fragment. The DNA fragment bearing the new Kan^R cassette was digested with *Eco*RI and *Nde*I was then ligated in the previously restricted pXMCS-2 mini-*Tn*5 $Genta^R$ plasmid. Primers used for this construct are listed in S2 Table.

4) Mutant library generation

One ml of an overnight culture of *B. abortus* 2308 was mixed with 50 µl of an overnight culture of the conjugative *E. coli* S17-1 strain carrying the pXMCS-2 mini-*Tn*5 Kan^R plasmid. This plasmid possesses a hyperactive Tn*5* transposase allowing the straightforward generation of a high number of transpositional mutants. The resulting *B. abortus* transpositional mutants were selected on 2YT plates (2% Agar) supplemented with both kanamycin and nalidixic acid. Tn*5* mutagenesis generates insertion of the transposon in

only one locus per genome, as demonstrated previously for *Brucella* (Lestrate *et al.,* 2000). Each Tn*5* derivative contains a *C. crescentus xyl* promoter that is constitutively active in *B. abortus* since when it is fused to YFP coding sequence on a pBBR1-derived plasmid, it generates a fluorescent signal of uniform intensity similar to the *E. coli lac* promoter fused to YFP coding sequence.

5) RAW 264.7 macrophages infection using the transpositional mutant library

Transpositional mutants were pooled using 2YT medium, diluted in RAW 264.7 macrophages culture medium to reach a MOI (Multiplicity of Infection) of 50, and added onto the macrophages, which were previously seeded in 6-well plates to a concentration of 1.5 10^5 cells per ml. A total of 16 6-well plates were planned per time point. Macrophages were then centrifuged 10 min at 400 *g* at 4°C and subsequently incubated for 1 hour at 37°C with 5% CO_2. The culture medium was then removed and replaced with fresh medium containing gentamycin 50 µg ml^{-1} in order to kill extracellular bacteria, and macrophages were then further incubated for 1, 4, and 23 hours at 37°C with 5% CO_2. For each time post-infection (2 h, 5 h or 24 h PI), culture medium was removed, each well was washed twice with PBS, and macrophages were lysed using PBS 0.1 % Triton X-100 for 10 min at 37°C. Macrophages lysates were then plated on 100 2YT plates per time point, each supplemented with kanamycin and incubated at 37°C for four days in order to obtain colonies that were collected for gDNA preparation and sequencing of Tn*5*-gDNA junctions.

6) RAW 264.7 macrophages infection and CFU counting

The infection protocol for performing CFU is identical to the one described above with the exception of the inoculum, which originates from an overnight liquid culture. After infection, infected macrophages are lysed and the resulting extracts are cultivated on 2YT plates supplemented with kanamycin. Once grown, colonies were counted to calculate the number of colonies per ml of lysate.

7) Analysis of essential genes for growth on plates

In order to assess the overall transposon insertion across *B. abortus* genome, we have created a parameter called R200, defined by the $\log_{10}$(number of Tn*5* insertion + 1) for a 200 bp sliding window. This sliding window was shifted every 5 bp to generate a collection of R200 scanning the whole genome for the control condition, *i.e.* bacteria on plates. Given that *B. abortus* genome is 3,278,307 bp, a list of 655,662 R200 values was created, with an average value of 9481 transposon insertions mapped per window. As previously published

(Christen *et al.,* 2011), the probability of obtaining a window of a given size with **no** transposon insertion event can be estimated by the following formula: $P = (1-(w/g))^n$ where w is the window size, g is the genome size, and n is the number of independent Tn*5* insertion events. In our case, the resulting probability was $3.8\ 10^{-15}$, with g = 3,278,307; w = 200, and n = 544,094. It should be noted that this value only accounts for a single window, whereas essential genes are typically characterized by a series of overlapping empty windows rather than a single 200 bp window, thus further lowering the probability of finding such profiles fortuitously. Essential genes were defined as all genes having at least one R200 equal to 0. Defined essential genes usually have many R200 equal to 0, as indicated in the Supporting Information files TTM_ctrl_chrI.txt and TTM_ctrl_chrII.txt, which can be aligned using Artemis (Sanger institute, http://www.sanger.ac.uk/science/tools/artemis) with the Supporting Information files ChrI.gb and ChrII.gb for the two annotated chromosomes of *B. abortus* 2308. Analysis of the effect of re-plating on rich medium was tested in an independent Tn-seq in which a new mutant library was constructed with the same mini-Tn*5* derivative as described above. All colonies were collected and the resulting suspension was used on the one hand for control analysis (data in TTM_replated_ctrl_chrI and TTM_replated_ctrl_chrII) and on the other hand for re-plating on the same rich medium. Colonies generated after re-plating were collected and analyzed by Tn-seq as described above (data in TTM_replated _chrI and TTM_replated _chrII).

8) Attenuation in infection analysis

For each post infection sample (2 h PI, 5 h PI and 24 h PI), a list of R200 values was calculated as for the control condition. Then, the control R200 values list was subtracted from each post infection sample R200 list separately, generating three Delta-R200 datasets, available as Supporting Information (Delta-R200_2hPI_chrI.txt, Delta-R200_2hPI_chrII.txt, Delta-R200_5hPI_chrI.txt, Delta-R200_5hPI_chrII.txt, Delta-R200_24hPI_chrI.txt, Delta-R200_24hPI_chrII.txt). Therefore, regions with a neutral Delta-R200 value have no impact during infection when mutated, while regions with a negative Delta-R200 are attenuated during infection, and regions with a positive Delta-R200 depict hyper-invasiveness for the corresponding mutants. For each time PI, the frequency distribution of Delta-R200 value was computed to define a normal distribution with an average and a standard deviation covering the main peak of this distribution. The threshold for negative Delta-R200 values was set at -0.75 $\log_{10}$ for the "2 h - control" and "5 h - control" Delta-R200 analyses, selecting respectively 5.5% and 6.7% of windows from the total genome. The threshold for negative Delta-R200 values was set at -1 $\log_{10}$ for the "24h - control" Delta-R200 analysis, allowing selection of 10.3% of the windows. The threshold for positive Delta-R200 values for the "2 h - control" condition was set at +0.6 $\log_{10}$, selecting 1.1% of the windows. The genes covered by selected windows were considered as required for the infection, with usually most of their coding sequences covered.

9) Generation of the *B. abortus* targeted mutants

Unless stated otherwise, all *B. abortus* mutants were generated by insertion of a plasmid in the targeted gene, according to a previously published procedure (Haine *et al.,* 2005). The primers sequences used to generate PCR products cloned in the disruption plasmids are available in S2 Table.
All deletion strains were constructed using a previously described allelic exchange strategy (Deghelt *et al.,* 2014) The primers used to amplify upstream and downstream regions of the target genes required for homologous recombination are also available in S2 Table.

10) Growth curves

Growth was monitored in 2YT at 37 °C during 72 h by measuring the Optical Density at 600 nm using a permanently shaking plate reader (Epoch2 microplate spectrophotometer, Biotek).

11) Immunofluorescence

One day before infection, HeLa cells were seeded on coverslips and infected the next day using overnight cultures of different *B. abortus* strains (wild type, Δ*virB* (Comerci *et al.,* 2001) and Δ*pyrB*) as indicated above (see "RAW 264.7 macrophages infection"). At 8 h PI, cells were fixed using 2 % paraformaldehyde (PFA) diluted in PBS. Once fixed, the samples were permeabilized using saponin 0.2 % and simultaneously stained with both primary antibodies (rabbit anti-DolK antibody, Abcam®, and A76-12G12, an anti-lipopolysaccahride O-chain monoclonal antibody recognizing *B. abortus*) resuspended in saponin 0.2 % and bovine serum albumin (BSA) 3 %. Samples were incubated at RT for 1 h, then washed 3 times in PBS, and simultaneously stained with both secondary antibodies (Goat anti-Rabbit IgG (H+L) Secondary Antibody, Pacific Blue™, and Goat anti-Mouse IgG (H+L) Cross-Adsorbed Secondary Antibody, Texas Red®-X, from Thermo Fisher ®) resuspended in saponin 0.2 % and bovine serum albumin (BSA) 3 %. Microscopy was then performed using a Nikon® i80 and Hammamatsu® ORCA-ER camera. Images were gathered from three independent experiments and were subsequently treated using the NIS Elements Advanced Research software. Anti-DolK staining was considered positive when displaying homogenous signal around the bacterium.

References

Abdo MR, Joseph P, Mortier J, Turtaut F, Montero JL, Masereel B, et al. Anti-virulence strategy against *Brucella suis*: synthesis, biological evaluation and molecular modeling of selective histidinol dehydrogenase inhibitors. Org Biomol Chem. 2011;9(10):3681-90. doi: 10.1039/c1ob05149k. PubMed PMID: 21461427.

Akerley B, Rubin E, Camilli A, Lampe D, Robertson H, Mekalanos J. Systematic identification of essential genes by in vitro mariner mutagenesis. *Proc Natl Acad Sci* U S A. 1998 Jul 21; 95(15):8927-32.

Al Dahouk S, Köhler S, Occhialini A, Jiménez de Bagüés MP, Hammerl JA, Eisenberg T, Vergnaud G, Cloeckaert A, Zygmunt MS, Whatmore AM, Melzer F, Drees KP, Foster JT, Wattam AR, Scholz HC. Brucella spp. of amphibians comprise genomically diverse motile strains competent for replication in macrophages and survival in mammalian hosts. Sci Rep. 2017;7:44420.

Al Dahouk S, Scholz H, Tomaso H, Bahn P, Göllner C, Karges W, Appel B, Hensel A, Neubauer H, Nöckler K. Differential phenotyping of *Brucella* species using a newly developed semi-automated metabolic system. *BMC Microbiol*. 2010; 10: 269.

Al-Hasani K, Rajakumar K, Bulach D, Robins-Browne R., Adler B, Sakellaris H. Genetic organization of the she pathogenicity island in Shigella flexneri 2a. *Microb Pathog.* 2001. 30: 1–8.

Alnor D, Frimodt-Møller N, Espersen F, Frederiksen W. Infections with the unusual human pathogens *Agrobacterium* species and *Ochrobactrum anthropi*. *Clin Infect Dis*. 1994 Jun; 18(6):914-20.

Anderson TD, Cheville NF. Ultrastructural morphometric analysis of *Brucella abortus*-infected trophoblasts in experimental placentitis. Bacterial replication occurs in rough endoplasmic reticulum. *Am J Pathol*. 1986 Aug;124(2):226-37.

Anwari K, Webb CT, Poggio S, Perry AJ, Belousoff M, Celik N, Ramm G, Lovering A, Sockett RE, Smit J, Jacobs-Wagner C, Lithgow T. The evolution of new lipoprotein subunits of the bacterial outer membrane BAM complex. Mol Microbiol. 2012 Jun;84(5):832-44.

Archambaud C, Salcedo SP, Lelouard H, Devilard E, de Bovis B, Van Rooijen N, Gorvel JP, Malissen B. Contrasting roles of macrophages and dendritic cells in controlling initial pulmonary *Brucella* infection. *Eur J Immunol*. 2010 Dec;40(12):3458-71.

Arocena GM, Sieira R, Comerci DJ, Ugalde RA. Identification of the quorum-sensing target DNA sequence and N-Acyl homoserine lactone responsiveness of the Brucella abortus virB promoter. J Bacteriol. 2010;192(13):3434-40. doi: 10.1128/JB.00232-10. PubMed PMID: 20400542; PubMed Central PMCID: PMC2897657.

Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko K, Tomita M, Wanner B, Mori H. Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol*. 2006; 2: 2006.0008.

Barbier T, Nicolas C, Letesson J. *Brucella* adaptation and survival at the crossroad of metabolism and virulence. *FEBBS Lett*. 2011; 585(19):2929–2934.

Barquero-Calvo E, Chaves-Olarte E, Weiss DS, Guzmán-Verri C, Chacón-Díaz C, Rucavado A, Moriyón I, Moreno E. *Brucella abortus* uses a stealthy strategy to avoid activation of the innate

immune system during the onset of infection. *PLoS One*. 2007 Jul 18;2(7):e631.

Barquero-Calvo E, Chaves-Olarte E, Weiss DS, Guzmán-Verri C, Chacón-Díaz C, Rucavado A, Moriyón I, Moreno E. *Brucella abortus* uses a stealthy strategy to avoid activation of the innate immune system during the onset of infection. *PLoS One*. 2007 Jul 18;2(7):e631.

Barquero-Calvo E, Mora-Cartín R, Arce-Gorvel V, de Diego J, Chacón-Díaz C, Chaves-Olarte E, Guzmán-Verri C, Buret A, Gorvel JP, Moreno E. *Brucella abortus* Induces the Premature Death of Human Neutrophils through the Action of Its Lipopolysaccharide. *PLoS Pathog*. 2015 May; 11(5): e1004853.

Batut J, Andersson S, O'Callaghan D. The evolution of chronic infection strategies in the α-proteobacteria. *Nat. Rev. Microbiol*. 2004; 2(12), 933–945.

Berg DE. Julian Davies and the discovery of kanamycin resistance transposon Tn5. *J Antibiot (Tokyo)*. 2017 Apr; 70(4):339-346.

Bessereau J, Wright A, Williams D, Schuske K, Davis M, Jorgensen E. Mobilization of a Drosophila transposon in the *Caenorhabditis elegans* germ line. *Nature*. 2001 Sep 6; 413(6851):70-4.

Bille E, Zahar J, Perrin A, Morelle S, Kriz P, Jolley K, Maiden M, Dervin C, Nassif X, Tinsley C. A chromosomally integrated bacteriophage in invasive meningococci. *J Exp Med*. 2005; 201 1905–1913.

Bingham P, Levis R, Rubin G. Cloning of DNA sequences from the white locus of *D. melanogaster* by a novel and general method. *Cell*. 1981 Sep;25(3):693-704.

Bishop A, Schiestl R. Homologous Recombination and Its Role in Carcinogenesis. *J Biomed Biotechnol*. 2002; 2(2): 75–85.

Blakely G, Colloms S, May G, Burke M, Sherratt D. *Escherichia coli* XerC recombinase is required for chromosomal segregation at cell division. *New Biol*. 1991; 3: 789–798.

Blakely G, May G, McCulloch R, Arciszewska L, Burke M. Lovett S, Sherratt D. Two related recombinases are required for site-specific recombination at *dif* and *cer* in *E. coli* K12. *Cell*. 1993; 75: 351–361.

Boschiroli ML, Ouahrani-Bettache S, Foulongne V, Michaux-Charachon S, Bourg G, Allardet-Servent A, Cazevieille C, Liautard JP, Ramuz M, O'Callaghan D. The *Brucella suis* virB operon is induced intracellularly in macrophages. *Proc Natl Acad Sci* U S A. 2002 Feb 5;99(3):1544-9.

Barlozzari G, Franco A, Macrì G, Lorenzetti S, Maggiori F, Dottarelli S, Maurelli M, Di Giannatale E, Tittarelli M, Battisti A, Gamberale F. First report of Brucella suis biovar 2 in a semi free-range pig farm, Italy. Vet Ital. 2015 Apr-Jun;51(2):151-4.

Briles EB, Tomasz A. Radioautographic evidence for equatorial wall growth in a gram-positive bacterium. Segregation of choline-3H-labeled teichoic acid. J Cell Biol. 1970 Dec; 47(3):786-90.

Brilli M, Fondi M, Fani R, Mengoni A, Ferri L, Bazzicalupo M, et al. The diversity and evolution of cell cycle regulation in alpha-proteobacteria: a comparative genomic analysis. BMC Syst Biol. 2010;4:52. doi: 10.1186/1752-0509-4-52. PubMed PMID: 20426835; PubMed Central PMCID:

PMCPMC2877005.

Brown PJ, de Pedro MA, Kysela DT, Van der Henst C, Kim J, De Bolle X, et al. Polar growth in the Alphaproteobacterial order Rhizobiales. Proc Natl Acad Sci U S A. 2012;109(5):1697-701. doi: 10.1073/pnas.1114476109. PubMed PMID: 22307633; PubMed Central PMCID: PMCPMC3277149.

Brüssow H, Canchaya C, Hardt W. Phages and the Evolution of Bacterial Pathogens: from Genomic Rearrangements to Lysogenic Conversion. Microbiol Mol Biol Rev. 2004 Sep; 68(3): 560–602.

Buisine N, Tang C, Chalmers R. Transposon-like Correia elements: structure, distribution and genetic exchange between pathogenic *Neisseria sp*. *FEBS Lett*. 2002 Jul 3;522(1-3):52-8.

Burman L, Raichler J, Park J. Evidence for diffuse growth of the cylindrical portion of the Escherichia coli murein sacculus. J Bacteriol. 1983 Sep;155(3):983-8.

Cardoso PG, Macedo GC, Azevedo V, Oliveira SC. *Brucella* spp noncanonical LPS: structure, biosynthesis, and interaction with host immune system. *Microb Cell Fact*. 2006 Mar 23;5:13.

Castañeda-Roldán E, Ouahrani-Bettache S, Saldaña Z, Avelino F, Rendón MA, Dornand J, Girón JA. Characterization of SP41, a surface protein of Brucella associated with adherence and invasion of host epithelial cells. Cell Microbiol. 2006 Dec;8(12):1877-87.

Caswell CC, Elhassanny AE, Planchin EE, Roux CM, Weeks-Gorospe JN, Ficht TA, et al. Diverse genetic regulon of the virulence-associated transcriptional regulator MucR in *Brucella abortus* 2308. Infect Immun. 2013;81(4):1040-51. Epub 2013/01/16. doi: 10.1128/IAI.01097-12. PubMed PMID: 23319565; PubMed Central PMCID: PMC3639602.

Celli J, de Chastellier C, Franchini DM, Pizarro-Cerda J, Moreno E, Gorvel JP. *Brucella* evades macrophage killing via VirB-dependent sustained interactions with the endoplasmic reticulum. *J Exp Med*. 2003 Aug 18;198(4):545-56.

Celli J, Salcedo SP, Gorvel JP. *Brucella* coopts the small GTPase Sar1 for intracellular replication. *Proc Natl Acad Sci* U S A. 2005 Feb 1;102(5):1673-8.

Celli J. The changing nature of the *Brucella*-containing vacuole. Cell Microbiol. 2015;17(7):951-8. doi: 10.1111/cmi.12452. PubMed PMID: 25916795; PubMed Central PMCID: PMCPMC4478208.

Cha SB, Rayamajhi N, Lee WJ, Shin MK, Jung MH, Shin SW, et al. Generation and envelope protein analysis of internalization defective *Brucella abortus* mutants in professional phagocytes, RAW 264.7. FEMS immunology and medical microbiology. 2012;64(2):244-54. doi: 10.1111/j.1574-695X.2011.00896.x. PubMed PMID: 22066675.

Chao L, Vargas C, Spear B, Cox E. Transposable elements as mutator genes in evolution. *Nature*. 1983 Jun; 303: 633-635.

Christen B, Abeliuk E, Collier JM, Kalogeraki VS, Passarelli B, Coller JA, et al. The essential genome of a bacterium. Mol Syst Biol. 2011;7:528. doi: 10.1038/msb.2011.58. PubMed PMID: 21878915; PubMed Central PMCID: PMC3202797.

Christen M, Beusch C, Bosch Y, Cerletti D, Flores-Tinoco CE, Del Medico L, et al. Quantitative

Selection Analysis of Bacteriophage phiCbK Susceptibility in *Caulobacter crescentus*. J Mol Biol. 2016;428(2 Pt B):419-30. doi: 10.1016/j.jmb.2015.11.018. PubMed PMID: 26593064.

Comerci DJ, Martinez-Lorenzo MJ, Sieira R, Gorvel JP, Ugalde RA. Essential role of the VirB machinery in the maturation of the *Brucella abortus*-containing vacuole. Cell Microbiol. 2001;3(3):159-68. Epub 2001/03/22. PubMed PMID: 11260139.

Cornet F, Louarn J, Patte J, Louarn J. Restriction of the activity of the recombination site dif to a small zone of the Escherichia coli chromosome. *Genes Dev.* 1996; 10: 1152–1161.

Cross A. What is a virulence factor ? *Crit Care* 2008; 12(6): 196.

Cunningham, B. Experiences with strain 45/20 vaccines in ireland. *In*: R. P. Crawford and R. J. Hidalgo (Eds.) *Bovine Brucellosis*. 1977. Texas A&M University Press. College Station, TX. 188–200.

Curtis PD, Brun YV. Identification of essential alphaproteobacterial genes reveals operational variability in conserved developmental and cell cycle systems. Mol Microbiol. 2014;93(4):713-35. doi: 10.1111/mmi.12686. PubMed PMID: 24975755; PubMed Central PMCID: PMCPMC4132054.

Czibener C, Merwaiss F, Guaimas F, Del Giudice MG, Serantes DA, Spera JM, Ugalde JE. BigA is a novel adhesin of *Brucella* that mediates adhesion to epithelial cells. *Cell Microbiol*. 2016 Apr;18(4):500-13.

De Bolle X, Bayliss C, Field D, van de Ven T, Saunders N, Hood D, Moxon E. The length of a tetranucleotide repeat tract in Haemophilus influenzae determines the phase variation rate of a gene with homology to type III DNA methyltransferases. *Mol Microbiol*. 2000 Jan; 35(1):211-22.

de Jong MF, Starr T, Winter MG, den Hartigh AB, Child R, Knodler LA, van Dijl JM, Celli J, Tsolis RM. Sensing of bacterial type IV secretion via the unfolded protein response. *MBio*. 2013 Feb 19;4(1):e00418-12.

de Jonge R, Bolton M, Kombrink A, van den Berg GC, Yadeta KA, Thomma BP. Extensive chromosomal reshuffling drives evolution of virulence in an asexual pathogen. *Genome Res*. 2013 Aug; 23(8): 1271-82.

de Vries S, Eleveld M, Hermans P, Bootsma H. Characterization of the Molecular Interplay between *Moraxella catarrhalis* and Human Respiratory Tract Epithelial Cells. *PLoS One*. 2013; 8(8): e72193.

Deghelt M, Mullier C, Sternon JF, Francis N, Laloux G, Dotreppe D, et al. G1-arrested newborn cells are the predominant infectious form of the pathogen *Brucella abortus*. Nat Commun. 2014;5:4366. Epub 2014/07/10. doi: 10.1038/ncomms5366. PubMed PMID: 25006695.

DeJesus M, Nambi S, Smith C, Baker R, Sassetti C, Ioerger T. Statistical analysis of genetic interactions in Tn-Seq data. *Nucleic Acids Res*. 2017 Jun 20; 45(11): e93.

Delrue RM, Deschamps C, Leonard S, Nijskens C, Danese I, Schaus JM, et al. A quorum-sensing regulator controls expression of both the type IV secretion system and the flagellar apparatus of *Brucella melitensis*. Cell Microbiol. 2005;7(8):1151-61. Epub 2005/07/13. doi: CMI543 [pii]

Delrue RM, Lestrate P, Tibor A, Letesson JJ, De Bolle X. *Brucella* pathogenesis, genes identified from

random large-scale screens. FEMS Microbiol Lett. 2004;231(1):1-12. Epub 2004/02/26. PubMed PMID: 14979322.

Delrue RM, Martinez-Lorenzo M, Lestrate P, Danese I, Bielarz V, Mertens P, et al. Identification of *Brucella* spp. genes involved in intracellular trafficking. Cell Microbiol. 2001;3(7):487-97. Epub 2001/07/05. doi: cmi131 [pii]. PubMed PMID: 11437834.

Derbise A, Chenal-Francisque V, Pouillot F, Fayolle C, Prévost MC, Médigue C, Hinnebusch B, Carniel E. A horizontally acquired filamentous phage contributes to the pathogenicity of the plague bacillus. *Mol Microbiol*. 2007 Feb;63(4):1145-57.

Detilleux PG, Deyoe BL, Cheville NF. Entry and intracellular localization of *Brucella* spp. in Vero cells: fluorescence and electron microscopy. Veterinary pathology. 1990;27(5):317-28. Epub 1990/09/01. PubMed PMID: 2122572.

Detilleux PG, Deyoe BL, Cheville NF. Penetration and intracellular growth of *Brucella abortus* in nonphagocytic cells in vitro. Infect Immun. 1990;58(7):2320-8. Epub 1990/07/01. PubMed PMID: 2114362; PubMed Central PMCID: PMC258815.

Drazek E, Houng H, Crawford R, Hadfield T, Hoover D, Warren R. 1995. Deletion of *purE* attenuates *Brucella melitensis* 16M for growth in human monocyte-derived macrophages. *Infect. Immun*. 1995; 63:3297–3301.

Endley S, McMurray D, Ficht TA. Interruption of the cydB locus in Brucella abortus attenuates intracellular survival and virulence in the mouse model of infection. J Bacteriol. 2001;183(8):2454-62. doi: 10.1128/JB.183.8.2454-2462.2001. PubMed PMID: 11274104; PubMed Central PMCID: PMC95161.

Fadool J, Hartl D, Dowling J. Transposition of the mariner element from Drosophila mauritiana in zebrafish. *Proc Natl Acad Sci* U S A. 1998 Apr 28; 95(9):5182-6.

Faino L, Seidl M, Shi-Kunne X, Pauper M, van den Berg G, Wittenberg A, Thomma B. Transposons passively and actively contribute to evolution of the two-speed genome of a fungal pathogen. *Genome Res*. 2016 Aug; 26(8): 1091–1100.

Ficht T. Brucella taxonomy and evolution. Future Microbiol. 2010 Jun;5(6):859-66.

Fontana C, Conde-Alvarez R, Stahle J, Holst O, Iriarte M, Zhao Y, et al. Structural Studies of Lipopolysaccharide Defective Mutants from *Brucella melitensis* Identify a Core Oligosaccharide Critical in Virulence. J Biol Chem. 2016. doi: 10.1074/jbc.M115.701540. PubMed PMID: 26867577.

Francis N, Poncin K, Fioravanti A, Vassen V, Willemart K, Ong TA, Rappez L, Letesson JJ, Biondi EG, De Bolle X. CtrA controls cell division and outer membrane composition of the pathogen *Brucella abortus*. *Mol Microbiol*. 2017 Mar;103(5):780-797.

Freed N, Bumann D, Silander O. Combining *Shigella* Tn-seq data with gold-standard *E. coli* gene deletion data suggests rare transitions between essential and non-essential gene functionality. *BMC Microbiol*. 2016; 16(1): 203.

Fugier E, Salcedo SP, de Chastellier C, Pophillat M, Muller A, Arce-Gorvel V, Fourquet P, Gorvel JP.

The glyceraldehyde-3-phosphate dehydrogenase and the small GTPase Rab 2 are crucial for *Brucella* replication. *PLoS Pathog*. 2009 Jun;5(6):e1000487.

Fuhrman J. Marine viruses and their biogeochemical and ecological effects. *Nature*. 1999 Jun 10;399(6736):541-8.

Geoffroy M, Floquet S, Métais A, Nassif X, Pelicic V. Large-scale analysis of the meningococcus genome by gene disruption: resistance to complement-mediated lysis. *Genome Res*. 2003 Mar; 13(3):391-8.

Godfroid F, Cloeckaert A, Taminiau B, Danese I, Tibor A, de Bolle X, et al. Genetic organisation of the lipopolysaccharide O-antigen biosynthesis region of brucella melitensis 16M (wbk). Res Microbiol. 2000;151(8):655-68. Epub 2000/11/18. PubMed PMID: 11081580.

Goldberg M. Or cells to study microbes ? Trends Microbiol. 1995; 3(8):331.

Guzmán-Verri C, González-Barrientos R, Hernández-Mora G, Morales JA, Baquero-Calvo E, Chaves-Olarte E, Moreno E. Brucella ceti and brucellosis in cetaceans. Front Cell Infect Microbiol. 2012;6,2-3.

Guzman-Verri C, Manterola L, Sola-Landa A, Parra A, Cloeckaert A, Garin J, Gorvel JP, Moriyon I, Moreno E, Lopez-Goni I. The two-component system BvrR/BvrS essential for *Brucella abortus* virulence regulates the expression of outer membrane proteins with counterparts in members of the *Rhizobiaceae*. *Proc Natl Acad Sci* U S A. 2002 Sep 17; 99(19):12375-80.

Hacker J, Carniel E. Ecological fitness, genomic islands and bacterial pathogenicity A Darwinian view of the evolution of microbes. *EMBO Rep*. 2001 May 15; 2(5): 376–381.

Hagblom P, Segal E, Billyard E, So M. Intragenic recombination leads to pilus antigenic variation in Neisseria gonorrhoeae. *Nature*. 1985 May 9-15;315(6015):156-8.

Haine V, Sinon A, Van Steen F, Rousseau S, Dozot M, Lestrate P, et al. Systematic targeted mutagenesis of *Brucella melitensis* 16M reveals a major role for GntR regulators in the control of virulence. Infect Immun. 2005;73(9):5578-86. Epub 2005/08/23. doi: 10.1128/IAI.73.9.5578-5586.2005. PubMed PMID: 16113274; PubMed Central PMCID: PMC1231144.

Hallez R, Bellefontaine AF, Letesson JJ, De Bolle X. Morphological and Functional Asymmetry in Alpha-Proteobacteria. *Trends Microbiol*. 2004; 12(8): 361-365. 8 2004.

Hallez R, Mignolet J, Van Mullem V, Wery M, Vandenhaute J, Letesson JJ, et al. The asymmetric distribution of the essential histidine kinase PdhS indicates a differentiation event in *Brucella abortus*. EMBO J. 2007;26(5):1444-55. Epub 2007/02/17. PubMed PMID: 17304218; PubMed Central PMCID: PMC1817626.

Hancock J, Santibáñez-Koref M. Trinucleotide expansion diseases in the context of micro- and minisatellite evolution, Hammersmith Hospital, April 1-3, 1998. *EMBO J*. 1998 Oct 1; 17(19):5521-4.

Hanot Mambres D, Machelart A, Potemberg G, De Trez C, Ryffel B, Letesson JJ, et al. Identification of Immune Effectors Essential to the Control of Primary and Secondary Intranasal Infection with *Brucella melitensis* in Mice. J Immunol. 2016;196(9):3780-93. doi: 10.4049/jimmunol.1502265. PubMed PMID: 27036913.

Hill S, Davies J. Pilin gene variation in Neisseria gonorrhoeae: reassessing the old paradigms. *FEMS Microbiol Rev*. 2009 May;33(3):521-30.

Holt K, Baker S, Weill F, Holmes E, Kitchen A, Yu J, Sangal V, Brown D, Coia J, Kim D, Choi S, Kim S, da Silveira W, Pickard D, Farrar J, Parkhill J, Dougan G, Thomson N. Shigella sonnei genome sequencing and phylogenetic analysis indicate recent global dissemination from Europe. *Nat Genet.* 2012; 44:1056–1059.

Hubbard TP, Chao MC, Abel S, Blondel CJ, Abel Zur Wiesch P, Zhou X, et al. Genetic analysis of *Vibrio parahaemolyticus* intestinal colonization. Proc Natl Acad Sci U S A. 2016;113(22):6283-8. doi: 10.1073/pnas.1601718113. PubMed PMID: 27185914; PubMed Central PMCID: PMCPMC4896720.

Hutchison C, Peterson S, Gill S, Cline R, White O, Fraser C, Smith H, Venter J. Global transposon mutagenesis and a minimal Mycoplasma genome. *Science*. 1999 Dec 10;286(5447):2165-9.

Jacobs C, Hung D, Shapiro L. Dynamic localization of a cytoplasmic signal transduction response regulator controls morphogenesis during the *Caulobacter* cell cycle. *Proc Natl Acad Sci* U S A. 2001 Mar 27;98(7):4095-100.

Jacobs M, Alwood A, Thaipisuttikul I, Spencer D, Haugen E, Ernst S, Will O, Kaul R, Raymond C, Levy R, Chun-Rong L, Guenthner D, Bovee D, Olson M, Manoil C. Comprehensive transposon mutant library of Pseudomonas aeruginosa. *Proc Natl Acad Sci* U S A. 2003 Nov 25; 100(24):14339-44.

Jeong KC, Ghosal D, Chang YW, Jensen GJ, Vogel JP. Polar delivery of *Legionella* type IV secretion system substrates is essential for virulence. *Proc Natl Acad Sci* U S A. 2017 Jul 25;114(30):8077-8082.

Jiang X, Leonard B, Benson R, Baldwin CL. Macrophage control of *Brucella abortus*: role of reactive oxygen intermediates and nitric oxide. *Cell Immunol*. 1993 Oct 15;151(2):309-19.

Johnson M, K. Criss A. Resistance of Neisseria Gonorrhoeae to Neutrophils. *Front Microbiol*. 2011; 2: 77.

Judd PK, Kumar RB, Das A. Spatial location and requirements for the assembly of the Agrobacterium tumefaciens type IV secretion apparatus. *Proc Natl Acad Sci* U S A. 2005 Aug 9;102(32):11498-503.

Jumas-Bilak E, Michaux-Charachon S, Bourg G, O'Callaghan D, Ramuz M. Differences in chromosome number and genome rearrangements in the genus Brucella. Mol. Microbiol. 1998; 27(1), 99-106.

Kamp H, Patimalla-Dipali B, Lazinski D, Wallace-Gadsden F, Camilli A. Gene Fitness Landscapes of *Vibrio cholerae* at Important Stages of Its Life Cycle. *PLoS Pathog*. 2013 Dec; 9(12): e1003800.

Ke Y, Wang Y, Li W, Chen Z. Type IV secretion system of *Brucella* spp. and its effectors. Frontiers in cellular and infection microbiology. 2015;5:72. doi: 10.3389/fcimb.2015.00072. PubMed PMID: 26528442; PubMed Central PMCID: PMC4602199.

Kim DH, Lim JJ, Lee JJ, Kim DG, Lee HJ, Min W, et al. Identification of genes contributing to the intracellular replication of *Brucella abortus* within HeLa and RAW 264.7 cells. Vet Microbiol.

2012;158(3-4):322-8. doi: 10.1016/j.vetmic.2012.02.019. PubMed PMID: 22397928.

Kim S, Watanabe K, Shirahata T, Watarai M. Zinc uptake system (*znuA* locus) of *Brucella abortus* is essential for intracellular survival and virulence in mice. *J Vet Med Sci*. 2004 Sep; 66(9):1059-63.

Kirkpatrick C, Viollier P. Decoding *Caulobacter* development. *FEMS Microbiol Rev*. 2012; 36(1): 193–205.

Klein B, Tenorio E, Lazinski D, Camilli A, Duncan M, Hu L. Identification of essential genes of the periodontal pathogen *Porphyromonas gingivalis*. *BMC Genomics*. 2012 Oct 31; 13:578.

Kobayashi K, Ehrlich SD, Albertini A, Amati G, Andersen KK, Arnaud M, Asai K, Ashikaga S, Aymerich S, Bessieres P, Boland F, Brignell SC, Bron S, Bunai K, Chapuis J, Christiansen LC, Danchin A, Débarbouille M, Dervyn E, Deuerling E, Devine K, Devine SK, Dreesen O, Errington J, Fillinger S, Foster SJ, Fujita Y, Galizzi A, Gardan R, Eschevins C, *et al*. Essential *Bacillus subtilis* genes. *Proc Natl Acad Sci* U S A. 2003 Apr 15; 100(8): 4678-83.

Kohler S, Foulongne V, Ouahrani-Bettache S, Bourg G, Teyssier J, Ramuz M, et al. The analysis of the intramacrophagic virulome of *Brucella suis* deciphers the environment encountered by the pathogen inside the macrophage host cell. Proc Natl Acad Sci U S A. 2002;99(24):15711-6. doi: 10.1073/pnas.232454299. PubMed PMID: 12438693; PubMed Central PMCID: PMC137781.

Köhler S, Foulongne V, Ouahrani-Bettache S, Bourg G, Teyssier J, Ramuz M, Liautard J. The analysis of the intramacrophagic virulome of Brucella suis deciphers the environment encountered by the pathogen inside the macrophage host cell. *Proc Natl Acad Sci* U S A. 2002 Nov 26; 99(24): 15711–15716.

Kozyreva V, Jospin G, Greninger A,Watt J, Eisen J, Chaturvedi V. Recent Outbreaks of Shigellosis in California Caused by Two Distinct Populations of Shigella sonnei with either Increased Virulence or Fluoroquinolone Resistance. *mSphere*. 2016 Nov-Dec; 1(6): e00344-16.

Laloux G, Deghelt M, de Barsy M, Letesson JJ, De Bolle X. 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. 2010;5(10):e13274. doi: 10.1371/journal.pone.0013274. PubMed PMID: 20949000; PubMed Central PMCID: PMCPMC2952587.

Lampe D, Akerley B, Rubin E, Mekalanos J, Robertson H. Hyperactive transposase mutants of the Himar1 mariner transposon. *Proc Natl Acad Sci* U S A. 1999 Sep 28; 96(20): 11428–11433.

Latre S, Bon N, Eblag O, Unett E. Evidence that nobody ever reads references. *J. stup. Sci.* 2014 Sep; 12 1429-1433.

Lengauer C, Kinzler KW, Vogelstein B. Genetic instabilities in human cancers. *Nature*. 1998 Dec 17;396(6712):643-9.

Lesic B, Bach S, Ghigo J, Dobrindt U, Hacker J, Carniel E. Excision of the high-pathogenicity island of Yersinia pseudotuberculosi requires the combined actions of its cognate integrase and Hef, a new recombination directionality factor. *Mol Microbiol*. 2004, 52: 1337-1348.

Lestrate P, Delrue RM, Danese I, Didembourg C, Taminiau B, Mertens P, et al. Identification and

characterization of in vivo attenuated mutants of Brucella melitensis. Mol Microbiol. 2000;38(3):543-51. Epub 2000/11/09. PubMed PMID: 11069678.

Liautard J, Ouahrani-Bettache S, Jubier-Maurin V, Lafont V, Kohler S, Liautard JP. Identification and isolation of *Brucella suis* virulence genes involved in resistance to the human innate immune system. Infect Immun. 2007;75(11):5167-74. doi: 10.1128/IAI.00690-07. PubMed PMID: 17709411; PubMed Central PMCID: PMCPMC2168268.

Liebeke M, Bundy JG. Tissue disruption and extraction methods for metabolic profiling of an invertebrate sentinel species. Metabolomics. 2012; 8:819–830.

Malinverni JC, Werner J, Kim S, Sklar JG, Kahne D, Misra R, Silhavy TJ. YfiO stabilizes the YaeT complex and is essential for outer membrane protein assembly in *Escherichia coli*. *Mol Microbiol.* 2006 Jul;61(1):151-64.

Martin DW, Baumgartner JE, Gee JM, Anderson ES, Roop RM, 2nd. SodA is a major metabolic antioxidant in *Brucella abortus* 2308 that plays a significant, but limited, role in the virulence of this strain in the mouse model. Microbiology. 2012;158(Pt 7):1767-74. doi: 10.1099/mic.0.059584-0. PubMed PMID: 22556360; PubMed Central PMCID: PMCPMC3542146.

Massouh A, Schubert J, Yaneva-Roder L, Ulbricht-Jones ES, Zupok A, Johnson MT, Wright SI, Pellizzer T, Sobanski J, Bock R, Greiner S. Spontaneous Chloroplast Mutants Mostly Occur by Replication Slippage and Show a Biased Pattern in the Plastome of *Oenothera*. *Plant Cell*. 2016 Apr;28(4):911-29.

Matroule JY, Lam H, Burnette D. Jacobs-Wagner C. Cytokinesis Monitoring during Development: Rapid Pole-to-Pole Shuttling of a Signaling Protein by Localized Kinase and Phosphatase in *Caulobacter*. 2004; 118(5): 579-590.

McClintock B. Induction of Instability at Selected Loci in Maize. *Genetics*. 1953 Nov; 38(6):579-99.

Meyer J, Brissac T, Frapy E, Omer H, Euphrasie D, Bonavita A, Nassif X, Bille E. Characterization of MDAΦ, a temperate filamentous bacteriophage of *Neisseria meningitidis*. *Microbiology*. 2016 Feb;162(2):268-82.

Michaux-Charachon S, Bourg G, Jumas-Bilak E, Guigue-Talet P, Allardet-Servent A, O'Callaghan D, et al. Genome structure and phylogeny in the genus *Brucella*. J Bacteriol. 1997;179(10):3244-9. PubMed PMID: 9150220; PubMed Central PMCID: PMCPMC179103.

Mignolet J, Van der Henst C, Nicolas C, Deghelt M, Dotreppe D, Letesson JJ, et al. PdhS, an old-pole-localized histidine kinase, recruits the fumarase FumC in *Brucella abortus*. J Bacteriol. 2010;192(12):3235-9. doi: 10.1128/JB.00066-10. PubMed PMID: 20382762; PubMed Central PMCID: PMCPMC2901695.

Mirabella A, Terwagne M, Zygmunt MS, Cloeckaert A, De Bolle X, Letesson JJ. *Brucella melitensis* MucR, an orthologue of Sinorhizobium meliloti MucR, is involved in resistance to oxidative, detergent, and saline stresses and cell envelope modifications. J Bacteriol. 2013;195(3):453-65. Epub 2012/11/20. doi: 10.1128/JB.01336-12. PubMed PMID: 23161025; PubMed Central PMCID: PMC3554010.

Mirold S, Rabsch W, Rohde M, Stender S, Tschäpe H, Rüssmann H, Igwe E, Hardt W. Isolation of a

temperate bacteriophage encoding the type III effector protein SopE from an epidemic Salmonella typhimurium strain. *Proc Natl Acad Sci* U S A. 1999 Aug 17; 96(17):9845-50.

Moran NA. Microbial minimalism: genome reduction in bacterial pathogens. Cell. 2002 Mar 8;108(5):583-6.

Moreno E, Moriyon I. The Genus *Brucella*. Prokaryotes. 2006;5:315-456.

Morgan JK, Luedtke BE, Shaw EI. Polar localization of the *Coxiella burnetii* type IVB secretion system. *FEMS Microbiol Lett*. 2010 Apr;305(2):177-83.

Nicoletti P. Vaccination against Brucella. *Adv Biotechnol Processes*. 1990;13:147-68.

O'Callaghan D, Cazevieille C, Allardet-Servent A, Boschiroli ML, Bourg G, Foulongne V, et al. A homologue of the Agrobacterium tumefaciens VirB and Bordetella pertussis Ptl type IV secretion systems is essential for intracellular survival of Brucella suis. Mol Microbiol. 1999;33(6):1210-20. PubMed PMID: 10510235.

Pappas G, Papadimitriou P, Akritidis N, Christou L, Tsianos E. The new global map of human brucellosis. *Lancet Infect Dis*. 2006 Feb;6(2):91-9.

Paulsen IT, Seshadri R, Nelson KE, Eisen JA, Heidelberg JF, Read TD, et al. The Brucella suis genome reveals fundamental similarities between animal and plant pathogens and symbionts. Proc Natl Acad Sci U S A. 2002;99(20):13148-53. doi: 10.1073/pnas.192319099. PubMed PMID: 12271122; PubMed Central PMCID: PMC130601.

Pinto UM, Pappas KM, Winans SC. The ABCs of plasmid replication and segregation. Nat Rev Microbiol. 2012;10(11):755-65. doi: 10.1038/nrmicro2882. PubMed PMID: 23070556.
Pizarro-Cerdá J, Méresse S, Parton RG, van der Goot G, Sola-Landa A, Lopez-Goñi I, Moreno E, Gorvel JP. *Brucella abortus* transits through the autophagic pathway and replicates in the endoplasmic reticulum of nonprofessional phagocytes. *Infect Immun*. 1998 Dec;66(12):5711-24.

Plasterk R, Izsvák Z, Ivics Z. Resident aliens: the Tc1/mariner superfamily of transposable elements. Trends Genet. 1999 Aug; 15(8): 326–332.

Pooley H, Schlaeppi J, Karamata D. Localized insertion of new cell wall in *Bacillus subtilis*. *Nature*. 1978; 274:264-266.

Porte F, Liautard JP, Köhler S. Early acidification of phagosomes containing *Brucella suis* is essential for intracellular survival in murine macrophages. *Infect Immun*. 1999 Aug;67(8):4041-7.

Porte F, Naroeni A, Ouahrani-Bettache S, Liautard JP. Role of the Brucella suis lipopolysaccharide O antigen in phagosomal genesis and in inhibition of phagosome-lysosome fusion in murine macrophages. Infect Immun. 2003;71(3):1481-90. PubMed PMID: 12595466; PubMed Central PMCID: PMC148865.

Posadas DM, Ruiz-Ranwez V, Bonomi HR, Martín FA, Zorreguieta A. BmaC, a novel autotransporter of *Brucella suis*, is involved in bacterial adhesion to host cells. *Cell Microbiol*. 2012 Jun;14(6):965-82.

Punsalang A, Sawyer W. Role of Pili in the Virulence of Neisseria gonorrhoeae. *Infect Immun*. 1973

Aug; 8(2): 255–263.

Qin QM1, Pei J, Ancona V, Shaw BD, Ficht TA, de Figueiredo P. RNAi screen of endoplasmic reticulum-associated host factors reveals a role for IRE1alpha in supporting *Brucella* replication. *PLoS Pathog*. 2008 Jul 25;4(7):e1000110.

Rasool, Freer E, Moreno E, Jarstrand C. Effect of *Brucella abortus* lipopolysaccharide on oxidative metabolism and lysozyme release by human neutrophils. *Infect Immun*. 1992 Apr; 60(4):1699-702.

Reed P, Atilano ML, Alves R, Hoiczyk E, Sher X, Reichmann NT, Pereira PM, Roemer T, Filipe SR, Pereira-Leal JB, Ligoxygakis P, Pinho MG. Staphylococcus aureus Survives with a Minimal Peptidoglycan Synthesis Machine but Sacrifices Virulence and Antibiotic Resistance. PLoS Pathog. 2015 May 7;11(5):e1004891.

Robertson GT, Reisenauer A, Wright R, Jensen RB, Jensen A, Shapiro L, et al. The *Brucella abortus* CcrM DNA methyltransferase is essential for viability, and its overexpression attenuates intracellular replication in murine macrophages. J Bacteriol. 2000;182(12):3482-9. Epub 2000/06/15. PubMed PMID: 10852881; PubMed Central PMCID: PMC101938.

Romano S, Aujoulat F, Jumas-Bilak E, Masnou A, Jeannot JL, Falsen E, Marchandin H, Teyssier C. Multilocus sequence typing supports the hypothesis that *Ochrobactrum anthropi* displays a human-associated subpopulation. *BMC Microbiol*. 2009; 9: 267.

Rubin E, Akerley B, Novik V, Lampe D, Husson R, Mekalanos J. In vivo transposition of mariner-based elements in enteric bacteria and mycobacteria. *Proc Natl Acad Sci* U S A. 1999 Feb 16; 96(4): 1645–1650.

Ruiz-Ranwez V, Posadas DM, Estein SM, Abdian PL, Martin FA, Zorreguieta A. The BtaF trimeric autotransporter of *Brucella suis* is involved in attachment to various surfaces, resistance to serum and virulence. PLoS One. 2013 Nov 13;8(11):e79770.

Ruiz-Ranwez V, Posadas DM, Van der Henst C, Estein SM, Arocena GM, Abdian PL, Martín FA, Sieira R, De Bolle X, Zorreguieta A. BtaE, an adhesin that belongs to the trimeric autotransporter family, is required for full virulence and defines a specific adhesive pole of *Brucella suis*. Infect Immun. 2013 Mar;81(3):996-1007.

Ruoslahti E. RGD and other recognition sequences for integrins. Annu Rev Cell Dev Biol. 1996;12:697-715.

Sakellaris H, Luck S, Al-Hasani K, Rajakumar K, Turner S, Adler, B. Regulated site-specific recombination of the she pathogenicity island of *Shigella flexneri*. *Mol. Microbiol.* 2004; 52, 1329–1336.

Salama N, Shepherd B, Falkow S. Global transposon mutagenesis and essential gene analysis of Helicobacter pylori. *J Bacteriol*. 2004 Dec;186(23):7926-35.

Salcedo S, Holden D. Bacterial interactions with the eukaryotic secretory pathway. *Curr. Opin. Microbiol*. 2005 ; 8: 92–98.

Salcedo SP, Chevrier N, Lacerda TL, Ben Amara A, Gerart S, Gorvel VA, de Chastellier C, Blasco JM,

Mege JL, Gorvel JP. Pathogenic *brucellae* replicate in human trophoblasts. *J Infect Dis*. 2013 Apr;207(7):1075-83.

Salcedo SP, Chevrier N, Lacerda TL, Ben Amara A, Gerart S, Gorvel VA, et al. Pathogenic *brucella*e replicate in human trophoblasts. The Journal of infectious diseases. 2013;207(7):1075-83. Epub 2013/01/11. doi: 10.1093/infdis/jit007. PubMed PMID: 23303808.

Schmidt H, Hensel M. Pathogenicity Islands in Bacterial Pathogenesis. *Clin Microbiol Rev*. 2004 Jan; 17(1): 14–56.

Schock TB, Strickland S, Steele EJ, Bearden DW. Effects of heat-treatment on the stability and composition of metabolomic extracts from the earthworm *Eisenia fetida*. *Metabolomics*. 2016;12:47.

Scholz HC, Al Dahouk S, Tomaso H, Neubauer H, Witte A, Schloter M, Kämpfer P, Falsen E, Pfeffer M, Engel M. Genetic diversity and phylogenetic relationships of bacteria belonging to the *Ochrobactrum-Brucella* group by recA and 16S rRNA gene-based comparative sequence analysis. *Syst Appl Microbiol*. 2008 Mar;31(1):1-16.

Scholz HC, Revilla-Fernández S, Al Dahouk S, Hammerl JA, Zygmunt MS, Cloeckaert A, Koylass M, Whatmore AM, Blom J, Vergnaud G, Witte A, Aistleitner K, Hofer E. *Brucella vulpis* sp. nov., isolated from mandibular lymph nodes of red foxes (Vulpes vulpes). *Int J Syst Evol Microbiol*. 2016 May;66(5):2090-8.

Segal E, Billyard E, So M, Storzbach S, Meyer T. Role of chromosomal rearrangement in N. gonorrhoeae pilus phase variation. *Cell*. 1985 Feb;40(2):293-300.

Selkrig J, Mosbahi K, Webb CT, Belousoff MJ, Perry AJ, Wells TJ, et al. Discovery of an archetypal protein transport system in bacterial outer membranes. Nat Struct Mol Biol. 2012;19(5):506-10, S1. doi: 10.1038/nsmb.2261. PubMed PMID: 22466966.

Sheehan LM, Budnick JA, Blanchard C, Dunman PM, Caswell CC. A LysR-family transcriptional regulator required for virulence in Brucella abortus is highly conserved among the alpha-proteobacteria. Mol Microbiol. 2015;98(2):318-28. doi: 10.1111/mmi.13123. PubMed PMID: 26175079.

Sherman A, Dawson A, Mather C, Gilhooley H, Li Y, Mitchell R, Finnegan D, Sang H. Transposition of the Drosophila element mariner into the chicken germ line. *Nat Biotechnol*. 1998 Nov; 16(11):1050-3.

Siddique A, Buisine N, Chalmers R. The Transposon-Like Correia Elements Encode Numerous Strong Promoters and Provide a Potential New Mechanism for Phase Variation in the Meningococcus. *PLoS Genet*. 2011 Jan; 7(1): e1001277.

Sieira R, Comerci DJ, Sánchez DO, Ugalde RA. A homologue of an operon required for DNA transfer in Agrobacterium is required in Brucella abortus for virulence and intracellular multiplication. *J Bacteriol*. 2000 Sep;182(17):4849-55.

Sklar JG, Wu T, Gronenberg LS, Malinverni JC, Kahne D, Silhavy TJ. Lipoprotein SmpA is a component of the YaeT complex that assembles outer membrane proteins in Escherichia coli. Proc Natl Acad Sci U S A. 2007 Apr 10;104(15):6400-5.

Sklar JG, Wu T, Gronenberg LS, Malinverni JC, Kahne D, Silhavy TJ. Lipoprotein SmpA is a component of the YaeT complex that assembles outer membrane proteins in Escherichia coli. Proc Natl Acad Sci U S A. 2007;104(15):6400-5. doi: 10.1073/pnas.0701579104. PubMed PMID: 17404237; PubMed Central PMCID: PMCPMC1851043.

Slater S, Goldman B, Goodner B, Setubal J, Farrand S, Nester E, Burr T, Banta L, Dickerman A, Paulsen I, Otten L, Suen G, Welch R, Almeida N, Arnold F, Burton O, Du Z, Ewing A, Godsy E, Heisel S, Houmiel K, Jhaveri J, Lu J, Miller N, Norton S, Chen Q, Phoolcharoen W, Ohlin V, Ondrusek D, Pride N, Stricklin S, Sun J, Wheeler C, Wilson L, Zhu H, Wood D. Genome Sequences of Three *Agrobacterium* Biovars Help Elucidate the Evolution of Multichromosome Genomes in Bacteria. *J. Bacteriol*. 2009 Apr; 191(8):2501-2511.

Smith JA, Khan M, Magnani DD, Harms JS, Durward M, Radhakrishnan GK, Liu YP, Splitter GA. Brucella induces an unfolded protein response via TcpB that supports intracellular replication in macrophages. PLoS Pathog. 2013;9(12):e1003785.

Sola-Landa A, Pizarro-Cerdá J, Grilló M, Moreno E, Moriyón I, Blasco J, Gorvel J, López-Goñi I. 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. *Mol Microbiol*. 1998 Jul;29(1):125-38.

Soler-Lloréns P, Quance C, Lawhon S, Stuber T, Edwards J, Thomas A Ficht T, Robbe-Austerman S, O'Callaghan D, Keriel A. A *Brucella* Spp. Isolate From a Pac-Man Frog (Ceratophrys Ornata) Reveals Characteristics Departing From Classical *Brucellae*. *Front Cell Infect Microbiol*. 2016; 6: 116.

Starnes C, Talwani R, Horvath J, Duffus W, Bryan C. Brucellosis in two hunt club members in South Carolina. J. S. C. Med. Assoc. 2004: 100, 113–115.

Starr T, Child R, Wehrly TD, Hansen B, Hwang S, López-Otin C, Virgin HW, Celli J. Selective subversion of autophagy complexes facilitates completion of the Brucella intracellular cycle. Cell Host Microbe. 2012 Jan 19;11(1):33-45.

Starr T, Ng TW, Wehrly TD, Knodler LA, Celli J. Brucella intracellular replication requires trafficking through the late endosomal/lysosomal compartment. Traffic. 2008 May;9(5):678-94.

Stern A, Sorek R. The phage-host arms-race: Shaping the evolution of microbes. *Bioessays*. 2011 Jan; 33(1): 43–51.

Suárez-Esquivel M, Baker KS, Ruiz-Villalobos N, Hernández-Mora G, Barquero-Calvo E, González-Barrientos R, Castillo-Zeledón A, Jiménez-Rojas C, Chacón-Díaz C, Cloeckaert A, Chaves-Olarte E, Thomson NR, Moreno E, Guzmán-Verri C. Brucella Genetic Variability in Wildlife Marine Mammals Populations Relates to Host Preference and Ocean Distribution. Genome Biol Evol. 2017 Jul 1;9(7):1901-1912.

Subashchandrabose S, Smith S, Spurbeck R, Kole M, Mobley H. Genome-Wide Detection of Fitness Genes in Uropathogenic Escherichia coli during Systemic Infection. *PLoS Pathog*. 2013 Dec; 9(12): e1003788.

Szwedziak P, Löwe J. Do the divisome and elongasome share a common evolutionary past? *Curr Opin*

Microbiol. 2013 Dec;16(6):745-51.

Taguchi Y, Imaoka K, Kataoka M, Uda A, Nakatsu D, Horii-Okazaki S, Kunishige R, Kano F, Murata M. Yip1A, a novel host factor for the activation of the IRE1 pathway of the unfolded protein response during Brucella infection. *PLoS Pathog*. 2015 Mar 5;11(3):e1004747.

Tavares AC, Fernandes PB, Carballido-López R, Pinho MG. MreC and MreD Proteins Are Not Required for Growth of *Staphylococcus aureus*. *PLoS One*. 2015 Oct 15;10(10):e0140523.

Theodorides J. Historical Brief on the Fever of Malta. *Hist Sci Med.* 1996; 30(1): 87-90.

Tsokos C, Laub M. Polarity and cell fate asymmetry in Caulobacter crescentus. Curr Opin Microbiol. 2012 Dec; 15(6): 744–750.

Ugalde JE, Czibener C, Feldman MF, Ugalde RA. Identification and characterization of the *Brucella abortus* phosphoglucomutase gene: role of lipopolysaccharide in virulence and intracellular multiplication. Infect Immun. 2000;68(10):5716-23. PubMed PMID: 10992476; PubMed Central PMCID: PMCPMC101528.

Uza, F, Samartino L, Schurig G, Carrasco A, Nielsen K, Cabrera R, Taddeo H. Effect of vaccination with *Brucella abortus* strain RB51 on heifers and pregnant cattle. *Vet. Res. Commun*. 2000; 24:143–151.

Van der Henst C, Beaufay F, Mignolet J, Didembourg C, Colinet J, Hallet B, et al. The histidine kinase PdhS controls cell cycle progression of the pathogenic alphaproteobacterium Brucella abortus. J Bacteriol. 2012;194(19):5305-14. doi: 10.1128/JB.00699-12. PubMed PMID: 22843843; PubMed Central PMCID: PMCPMC3457221.

Van der Henst C, de Barsy M, Zorreguieta A, Letesson JJ, De Bolle X. The Brucella pathogens are polarized bacteria. Microbes Infect. 2013 Dec;15(14-15):998-1004.

van Opijnen T, Camilli A. A fine scale phenotype-genotype virulence map of a bacterial pathogen. *Genome Res*. 2012 Dec; 22(12):2541-51.

van Opijnen T, Camilli A. Genome-Wide Fitness and Genetic Interactions Determined by Tn-seq, a High-Throughput Massively Parallel Sequencing Method for Microorganisms. *Curr Protoc Microbiol*. 2010 Nov; 0 1: 10.1002.

Velasco J, Bengoechea JA, Brandenburg K, Lindner B, Seydel U, González D, Zähringer U, Moreno E, Moriyón I. *Brucella abortus* and its closest phylogenetic relative, *Ochrobactrum* spp., differ in outer membrane permeability and cationic peptide resistance. *Infect Immun*. 2000 Jun;68(6):3210-8.

Waldor, M. K., and J. J. Mekalanos. Lysogenic conversion by a filamentous phage encoding cholera toxin. *Science*. 1996; 272**:**1910-1914.

Watarai M, Makino S, Fujii Y, Okamoto K, Shirahata T. Modulation of Brucella-induced macropinocytosis by lipid rafts mediates intracellular replication. Cell Microbiol. 2002 Jun;4(6):341-55.

Wattam A, Foster J, Mane, Beckstrom-Sternberg S, Beckstrom-Sternberg J, Dickerman A, Keim P,

Pearson T, Shukla M, Ward D, Williams K, Sobral B, Tsolis R, Whatmore A, O'Callaghang D. Comparative Phylogenomics and Evolution of the *Brucellae* Reveal a Path to Virulence. *J Bacteriol*. 2014 Mar; 196(5): 920–930.

Weerdenburg E, Abdallah A, Rangkuti F, Abd El Ghany M, Otto T, Adroub S, Molenaar D, Ummels R, Ter Veen K, van Stempvoort G, van der Sar A, Ali S, Langridge G, Thomson N, Pain A, Bitter W. Genome-wide transposon mutagenesis indicates that *Mycobacterium marinum* customizes its virulence mechanisms for survival and replication in different hosts. *Infect Immun*. 2015 May; 83(5): 1778–1788.

Whatmore A. Current understanding of the genetic diversity of *Brucella*, an expanding genus of zoonotic pathogens. *Infect Genet Evol*. 2009 Dec;9(6):1168-84.

Wilde A, Snyder D, Putnam N, Valentino M, Hammer N, Lonergan Z, Hinger S, Aysanoa E, Blanchard C, Dunman P, Wasserman G, Chen J, Shopsin B, Gilmore M, Skaar E, Cassat J. Bacterial Hypoxic Responses Revealed as Critical Determinants of the Host-Pathogen Outcome by TnSeq Analysis of *Staphylococcus aureus* Invasive Infection. *PLoS Pathog*. 2015 Dec 18;11(12):e1005341.

Willett JW, Herrou J, Briegel A, Rotskoff G, Crosson S. Structural asymmetry in a conserved signaling system that regulates division, replication, and virulence of an intracellular pathogen. Proc Natl Acad Sci U S A. 2015;112(28):E3709-18. doi: 10.1073/pnas.1503118112. PubMed PMID: 26124143; PubMed Central PMCID: PMCPMC4507200.

Williams D, Boulin T, Ruaud A, Jorgensen E, Bessereau J. Characterization of *Mos1*-Mediated Mutagenesis in *Caenorhabditis elegans*. *Genetics*. 2005 Mar; 169(3): 1779–1785.

Winzeler E, Shoemaker D, Astromoff A, Liang H, Anderson K, Andre B, Bangham R, Benito R, Boeke J, Bussey H, Chu A, Connelly C, Davis K, Dietrich F, Dow S, El Bakkoury M, Foury F, Friend S, Gentalen E, Giaever G, Hegemann J, Jones T, Laub M, Liao H, Liebundguth N, Lockhart D, Lucau-Danila A, Lussier M, M'Rabet N, Menard P, Mittmann M, Pai C, Rebischung C, Revuelta J, Riles L, *et al.* Functional characterization of the S. cerevisiae genome by gene deletion and parallel analysis. Science. 1999 Aug 6; 285(5429):901-6.

Wood G, Hendricks J, Goodman D. Brucellosis in feral swine. J Wildl Dis. 1976; 12(4):579-82.

Wu Q, Pei J, Turse C, Ficht TA. Mariner mutagenesis of *Brucella melitensis* reveals genes with previously uncharacterized roles in virulence and survival. BMC Microbiol. 2006;6:102. doi: 10.1186/1471-2180-6-102. PubMed PMID: 17176467; PubMed Central PMCID: PMCPMC1766931.

Yook K, Hodgkin J. *Mos1* mutagenesis reveals a diversity of mechanisms affecting response of *Caenorhabditis elegans* to the bacterial pathogen *Microbacterium nematophilum*. *Genetics*. 2007 Feb; 175(2):681-97.

Zhang Y, Ioerger T, Huttenhower C, Long J, Sassetti C, Sacchettini J, Rubin E. Global assessment of genomic regions required for growth in *Mycobacterium tuberculosis*. *PLoS Pathog*. 2012 Sep;8(9):e1002946.

Zhou M, Reznikoff W. Tn5 transposase mutants that alter DNA binding specificity. *J Mol Biol.* 1997 Aug 22;271(3):362-73.