UNIVERSITÀ DEGLI STUDI DI SASSARI

SCUOLA DI DOTTORATO IN SCIENZE BIOMOLECOLARI E BIOTECNOLOGICHE

INDIRIZZO IN MICROBIOLOGIA MOLECOLARE E CLINICA

(XXIV ciclo)

Coordinatore: Prof. Bruno Masala

DIFFERENTIALLY EXPRESSED GENES IN MYCOBACTERIAL CELL WALL DEFICIENT FORMS INDUCED *IN VITRO*

***Tutor*: Prof. Leonardo A. Sechi**

Tesi di dottorato di:

Dott. ssa Valentina Rosu

This dissertation is dedicated to my son Pietro

who makes me smile everyday

and to my rock, my husband Francesco.

I thank my parents that raised me to be a free thinker.

INDEX

1. Abstract pag. 5

1.1 **Abstract (in Italian)** pag. 5

2. Introduction pag. 7

2.1 Mycobacteria pag. 7

2.2 *Mycobacterium tuberculosis* (MTB) pag. 9

2.3 *Mycobacterium smegmatis* (MSMEG) pag. 10

2.4 *Mycobacterium avium* subspecies *paratuberculosis* (MAP) pag. 11

2.5 Bacterial cell wall deficient/defective forms (CWD) pag. 14

2.6 Mycobacterial CWD forms pag. 17

3. Aim of the project pag. 21

4. Materials and methods pag. 23

4.1 Bacterial cultures and growth media pag. 23

4.2 *In vitro* generation of mycobacterial CWD forms pag. 23

4.3 Mycobacterial RNA isolation pag. 24

4.4 mRNA enrichment and linear amplification of mycobacterial RNA pag. 24

4.5 aRNA labeling and Microarray hybridization pag. 25

4.6 Microarray data analysis pag. 26

4.7 Reverse Transcritptase Real Time quantitative PCR pag. 27

5. Results pag. 28

5.1 Analysis of overall gene expression data pag. 28

5.2 Differentially expressed genes in CWD_MSMEG pag. 29

5.3 Differentially expressed genes in CWD_MTB pag. 33

5.4 Differentially expressed genes in CWD_MAP pag. 40

5.5 Real time RT-qPCR analysis of selected genes pag. 44

6. Discussion pag. 47

7. Conclusions pag. 56

8. Bibliography pag. 57

9. Appendix pag. 70

10. Acknowledgments and funding pag. 98

1. ABSTRACT

Mycobacterial cell wall deficient/defective (CWD) forms occur *in vivo* in response to various antimicrobial factors such as lysozyme released by the host or antibiotics. A wide body of research suggests that these forms are clinically significant and involved in the establishment of either subclinical or chronic infections due to their ability to resist host defense mechanisms. Although CWD mycobacteria have been identified and isolated from patients with several diseases, there is a lack of knowledge about their effective role in pathogenesis, mainly due to the difficulties in handling these pleomorphic forms. In this study, whole-genome microarrays have been used to identify for the first time differentially expressed genes of both pathogenic (*Mycobacterium tuberculosis* and *Mycobacterium avium* subspecies *paratuberculosis*) and non-pathogenic (*Mycobacterium smegmatis*) mycobacteria *in vitro* converted to the CWD state by using lysozyme and glycine. Results highlighted that *Mycobacterium smegmatis* try to respond to lysozyme activity by expressing genes encoding for peptidoglycan deacetylases probably used by the bacteria to modify its peptidoglycan layer so that it is no longer recognized by the lysozyme. On the contrary, the pathogenic species differentially regulated a *repertoire* of genes involved in mycobacterial virulence and persistence. In particular, *Mycobacterium tuberculosis* differentially regulated PE/PPE genes and many virulence genes such as *Rv2351c plcA* and *Rv2962c* that are highly specific for this species. Moreover, the down-regulation of *kasA* and *kasB* involved in mycolic acid biosynthesis, contribute to the acquisition of the altered colony morphology, the loss of the acid-fastness and the dormancy capability. Hence, the CWD state may represents an important stage in the life-cycle of pathogenic mycobacteria that potentially drives persistence.

1.1 ABSTRACT (in Italian)

La perdita totale o parziale della parete batterica è un fenomeno che si verifica *in vivo* in risposta a vari fattori antimicrobici come il lisozima, il quale rappresenta la prima arma di difesa dell'ospite nei confronti del microrganismo, oppure agli antibiotici. Numerose ricerche suggeriscono queste forme prive di parete come clinicamente significanti e coinvolte nella capacità del batterio di stabilire infezioni subcliniche o

croniche grazie alla loro aumentata capacità di resistere alle difese dell'ospite, le quali sono principalmente rivolte verso la parete batterica. Nonostante batteri in forma protoplastica o sferoplastica siano stati osservati ed isolati da pazienti con diverse malattie croniche, l'effettivo ruolo di questi batteri pleomorfi non è ancora noto, a causa della difficoltà nella loro coltivazione e manipolazione. In questo studio è stata utilizzata la tecnica del DNA-Microarray allo scopo di evidenziare, per la prima volta, i geni diversamente espressi nei micobatteri, sia patogeni (*Mycobacterium tuberculosis* e *Mycobacterium avium* subspecies *paratuberculosis*) che non patogeni (*Mycobacterium smegmatis*) dopo la conversione *in vitro* nella forma difettiva di parete utilizzando come agenti inducenti il lisozima e la glicina. In generale si è visto che le tre specie rispondono in maniera differente allo stress da superficie indicando quindi l'evoluzione di un meccanismo di difesa specie-specifico. I risultati suggeriscono che *Mycobacterium smegmatis* cerca di resistere all'azione del lisozima esprimendo dei geni che codificano per delle deacetilasi di parete, che vengono probabilmente utilizzate dal batterio per modificare la sua struttura peptidoglicanica in modo tale da renderla non più "riconoscibile" dal lisozima. Al contrario, *Mycobacterium tuberculosis* e *Mycobacterium avium* subspecies *paratuberculosis* modificano l'espressione di numerosi geni di virulenza e di persistenza. In particolare, *Mycobacterium tuberculosis* regola geni appartenenti alla famiglia delle proteine PE/PPE ma anche geni già noti in letteratura come associati alla capacità del batterio di persistere all'interno dei macrofagi e geni di virulenza altamente specifici per *Mycobacterium tuberculosis* quali *Rv2351c plcA* e *Rv2962c.* Inoltre abbiamo osservato la down-regolazione di due geni chiave nella sintesi degli acidi micolici, *kasA* e *kasB,* che probabilmente contribuiscono alla trasformazione del batterio nella forma pleomorfa, alla perdita del fattore cordale e della acido-alcol resistenza ed infine alla capacità di persistere nell'ospite. Lo stato protoplastico/sferoplastico è uno stadio importante del ciclo vitale del batterio che potenzialmente guida il fenomeno della persistenza dei micobatteri patogeni nei tessuti dell'ospite.

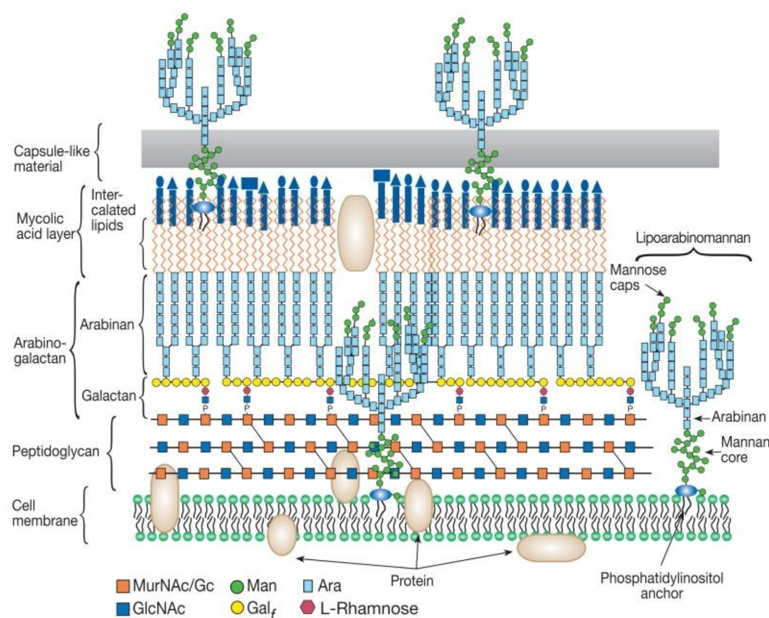
2. INTRODUCTION

2.1 Mycobacteria

Mycobacteria belong to the genus of Actinobacteria. They are aerobic, non-motile, except for the species *Mycobacterium marinum* which has been shown to be motile within macrophages, and non-sporulating bacteria, except for *Mycobacterium marinum* and perhaps *Mycobacterium bovis*. [Ghosh et al., 2009; Singh et al., 2010]. Usually they are widespread organisms, typically living in water, soil and food sources. However some of them such as *Mycobacterium tuberculosis* and *Mycobacterium leprae*, are obligate parasites and are not found as free-living microorganisms [Vaerewijck et al., 2005; Falkinham, 2010].

Mycobacteria species are characterized of a particularly thick cell wall, thicker than in many other bacteria. The structure is hydrophobic, waxy, and rich in mycolic acids/mycolates, a type of fatty acids. The mycolic acid–arabinogalactan–peptidoglycan complex is also surrounded by a polysaccharide-rich capsule of arabinomannan and mannan (Fig. 1) [Brennan and Crick, 2007].

Fig.1 Mycobacterial cell wall structure



Source: Brennan P.J. and Crick D.C. 2007. Curr. Top. Med. Chem. 7: 475–488.

The thick cell wall concurs to the slow-growing rate of many mycobacteria and most of them are pathogens. Some species, such as *Mycobacterium leprae,* can be very tedious to culture as they have a doubling time of around 2 weeks in infected tissues (for comparison *Escherichia coli* strains have a doubling time of 20 minutes), hence making laboratory culture a very slow process. Due to their growth-rate characteristics, mycobacteria can be classified in two main groups: **slowly–growing** and **rapidly–growing species**. Usually fast-growing mycobacteria form colonies within 7 days of incubation [Singleton, 2004].

Mycobacteria can also be grouped due to their capability to produce pigments (carotenoids), a process which can be either dependent or independent of light, classification is as follow:

Photochromogens are mycobacteria that produce pigments after exposure to light (e.g. *Mycobacterium kansasii, Mycobacterium marinum, Mycobacterium simiae*), **Scotochromogens** are mycobacteria that produce a yellow to orange pigment in the presence of both light and dark (e.g. *Mycobacterium scrofulaceum, Mycobacterium gordonae, Mycobacterium xenopi, Mycobacterium szulgai*) and **non-chromogens** are mycobacteria that do not produce pigments (e.g. *Mycobacterium tuberculosis, Mycobacterium avium-intra-cellulare, Mycobacterium bovis, Mycobacterium ulcerans*).

For medical purposes mycobacteria can also be classified into two main categories: the *Mycobacterium tuberculosis complex* (*Mycobacterium tuberculosis, Mycobacterium bovis, Mycobacterium africanum, and Mycobacterium canettì, Mycobacterium pinnipedi, Mycobacterium mungi and Mycobacterium microti*) and the non-tuberculous mycobacteria (NTM) [Alexander et al., 2010].

Mycobacteria are defined as acid-fast bacteria due to their characteristic to resist acids de-colorization during staining procedures, for this reason they cannot be diagnosed by standard microbiological techniques (e.g. Gram stain). Therefore, the staining process should be performed by using concentrated dyes combined with heat. The most common staining techniques are the Ziehl-Neelsen stain, in which mycobacteria are stained bright fuchsia and the fluorescence stain by using auramine-rhodamine dyes [Singleton, 2004].

2.2 *Mycobacterium tuberculosis*

Mycobacterium tuberculosis (MTB) also known as "Koch Bacillus" belongs to the *Mycobacterium tuberculosis complex* group. It is a rod shaped bacteria, aerobe and slow-grower (generation time of 15-20 hrs.), a physiological characteristic that may contribute to its virulence. MTB is not motile and for years it has been classified as deficient of pili but very recently Alteri et al. [Alteri et al., 2007] demonstrated the presence of these appendices, whose pilin subunit is encoded by the *Rv3312A* gene.

MTB is an airborne pathogen and it is the etiological agent of human tuberculosis (TB). The infection process starts with the transmission of MTB from an infected to an healthy individual. Once in the lungs, MTB enters and resides within alveolar macrophages (AMs) and dendritic cells (DCs), [Cooper, 2009abc] where the mycobacterium is able to survive and multiply. In fact, although AMs represent an active barrier against common pathogens, MTB has evolved various mechanisms to evade this barrier and survive inside these cells. This survival mechanism is not fully understood, what is known is that they are able to resist to reactive oxygen and nitrogen intermediates (ROIs and RNIs, respectively) relied by the host and to avoid the acidification of the MTB-containing phagosome. In this respect, complex lipids of the mycobacterial cell wall are thought to be important for blocking the normal biogenesis of the phagolysosome. [Flynn and Chan, 2001; Cooper, 2009abc]. Depending of the immunological status of the host, MTB bacilli can escape a process that lead to the recruitment of blood monocytes and other inflammatory cells (i.e., neutrophils) to the site of infection. Although these cells become antigen presenting cells, they are not able to effectively kill the bacteria. At this stage, MTB grows under limited tissue damage. By 6–8 weeks post-infection, antigen presenting DCs have traveled to lymph nodes where T-lymphocytes are activated and recruited. Activated T- lymphocytes that migrate to the site of infection proliferate forming an early stage granuloma, where macrophages become activated to kill intracellular MTB [Ulrichs and Kaufmann, 2006]. However, persistent T-cell activation leads to formation of granulomas and to mycobacterial latency, a stage where growth and spread of bacteria is limited. More than 90% of infected people are latently infected, they have no clinical symptoms but remain a *reservoir* of MTB. Latently infected people can be identified by a delayed hypersensitivity reaction to the purified protein derivative of MTB used in tuberculin skin test or a T-cell response to MTB-specific antigens. Decline

of the immune surveillance due to genetic (e.g. mutations in specific host cytokines or chemokines) or environmental causes (viral infection, AIDS, malnutrition, aging) can lead to reactivation of the mycobacterial infection [Geboes and Bossaert, 1977; Di Perri et al., 1993; Semba et al., 2010] and consequently to pulmonary disease [Kaplan et al., 2003; Dheda et al., 2005].

TB is a disease of poverty which mostly affects young adults in their most productive years. The vast majority of TB deaths are in the developing world. The WHO estimations reveal that 1.7 million people died from TB in 2009, including 380.000 people with HIV, which equal to 4700 deaths a day. There were also 9.4 million new TB cases (including 3.3 million women) in 2009, 440.000 new multidrug-resistant TB (MDR-TB) cases in 2008 and 150.000 deaths from MDR-TB. These numbers explain why TB research needs to be very active in order to design new tools for diagnosis, treatment and prevention of TB.

To improve the understanding of the slow-growing MTB pathogen, Cole et al., have determined in 1998 the complete genome sequence of the best-characterized strain of MTB, the H37Rv strain. The genome comprises 4,411,529 base pairs, contains around 4,000 genes, and has a very high G + C content (about 65%) that is reflected in the biased amino-acid content of the proteins.

2.3 *Mycobacterium smegmatis*

Mycobacterium smegmatis (MSMEG), is an aerobic, rapidly-growing (double time of about 3hrs), non-photochromogenic species of mycobacterium first discovered by Lustgarten in 1884. MSMEG can be found in water and soil as well as in the human genital secretions (smegma). Although few cases of human infection have been reported [Newton et al., 1993; Pennekamp et al., 1997] MSMEG is classified as a saprophytic mycobacterium and is generally considered as a non-hazardous environmental microorganism, independent on living in animals.

MSMEG is very useful for research analysis due to its non-pathogenic and fast-growing characteristic; hence it is used as a model organism and a surrogate host for genetic analysis of mycobacterial pathogens. The strain usually used in laboratory is the strain MC^2 155, which can be transformed with plasmid vectors by electroporation 10 to 100 thousand times more efficiently than the parental strain [Snapper at al., 1990].

The genome of MSMEG has been sequenced in 2006 [Fleischmann et al, 2006]. It is 6.988,209 nucleotides long and it has a 67% G+C content. 90% of the genome (6716/6938 genes) represents coding regions that encode for 6716 proteins. The 6938 genes are composed circularly with an absence of any plasmids.

2.4 *Mycobacterium avium* subspecies *paratuberculosis*

Mycobacterium avium subspecies *paratuberculosis* (MAP) belongs to the *Mycobacterium avium complex* group. It is aerobic, rod-shaped and slow-grower bacterium. As all pathogenic mycobacteria, it has a very thick cell wall which is responsible for its slow-growing rate (doubling time of about 20 hours) and which concurs to limit the flow of nutrients in and out of the cell. One major difference between MAP and other mycobacterial species is its inability to produce mycobactin *in vitro*, which is a siderophore responsible for the binding and transport of iron into the cell. A sequence comparison of the mycobactin gene cluster between MAP K-10, *M. avium* strain 104 and MTB strain H37Rv reveals that this deficiency could be related either to the presence of a short *mbtA* gene, lacking more than 200 residues important for protein function [Li et al. 2005] or to a polymorphism identified in the *mbt* operon [Semret at al., 2004].

MAP is the etiologic agent of Johne's disease, also known as animal Paratuberculosis, a chronic granulomatous enteritis which affects ruminants (such as cattle, sheep, goats, deer and bison) where the bacterium can survive latently for years (without development of any clinical signs or disease) before to cause severe disease. In fact, MAP is able to evade host defenses and to survive within phagosomal compartments for a long time; moreover it is also able to reprogram gene expression of macrophages in order to increase its chances of survival [Tessema et al., 2001; Tooker et al., 2002; Sigurethardottir et al., 2004].

Paratuberculosis develops in 4 stages. In stage 1, usually young stocks are infected silently and sub-clinically. Infected animals at this early stage cannot be detected with any current available diagnostic tests. The disease can either progresses slowly over many years or resolves completely. In stage 2, infected animals still appear healthy but start to shed MAP organisms so that can be detected by fecal culture. These animals are responsible for the contamination of the environment where MAP is able to survive,

probably in association with protozoa and insects [Rowe and Grant, 2006; Whan et al., 2006] provoking drastic problems for farmers trying to eradicate Johne's disease from their herds.

The stage 3 usually occurs in animals with advanced infection with acute or intermittent clinical features. Animals start to lose weight and drop in milk production. In the last stage, stage 4, there is an acerbation of clinical features that can lead to the death of infected animals [Biet at al., 2005]. There is no treatment for this disease and despite efforts for its eradication, it is still present. The main obstacles for complete eradication are the lack of appropriate diagnostic methods, the role played by other domestic and wild animals as *reservoirs* and the difficulty of working with MAP in the laboratory. Moreover, attempt to diagnose Paratuberulosis is very tedious, slow and expensive, especially when culture methods are performed, while serological assays still have many limitations (Bannantine et al., 2008).

MAP has also been suspected to be the causative agent of some cases of inflammatory bowel disease in humans, especially Crohn's disease, where the pathological aspects resembles Johne's disease [Brugere-Picoux, 1998; Hampson et al.,1998; Chiodini, 1989; Engstrand, 1995; Chiodini and. Rossiter, 1996; Sechi et al. 2001; Greenstein, 2003; Behr and Kapur, 2008; Rosenfeld and Bressler, 2010; Over et al., 2011]. The route of infection is believed to be the consumption of contaminated food, especially milk which has been proposed as a high risk factor for transmission of the putative zoonotic pathogen from cattle to humans (see fig. 2). In fact, MAP has been frequently found in dairy products of bovine origin, milk in particular where MAP has been demonstrated to survive standard pasteurization, but also in municipal water supplies where MAP can resist chlorination [Schoos, 2005; Collins, 1997; Stabel, 2000; Lund et al., 2002; Gill et al., 2011]

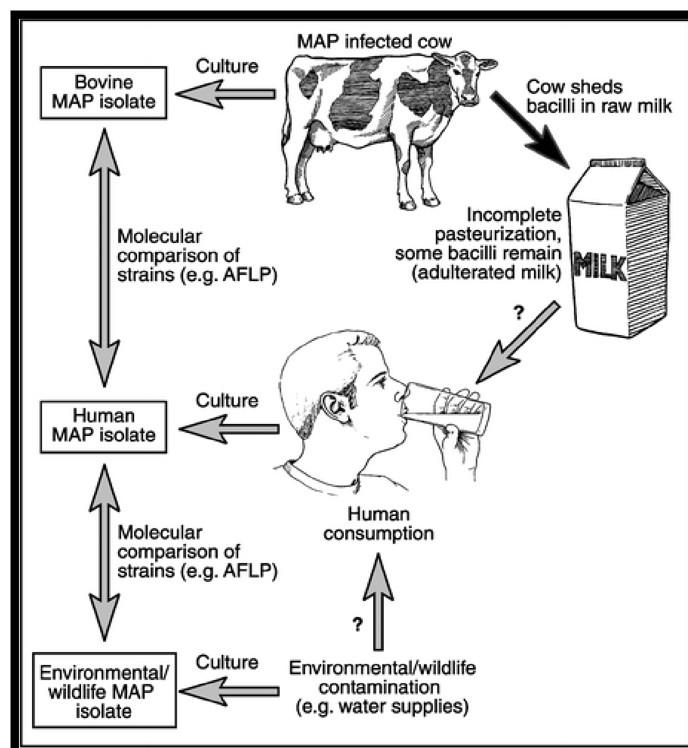
Very recently, the group of Prof. Sechi demonstrated the presence of live circulating MAP and antibodies against MAP in the blood of Sardinian patients affected with either Type 1 Diabetes or Multiple sclerosis [Sechi et al., 2008ab; Rosu et al.,2008; Rosu et al., 2009; Paccagnini et al., 2009; Cossu et al., 2011]. Moreover other researchers have found MAP in patients with Hashimoto's thyroiditis [Sisto et al., 2010].

However, MAP is still not accepted as a human pathogen in part because it has not been seen under the microscope in large numbers in the tissues of patients with Crohn's

disease (Pierce, 2009). Therefore, its effective role is this disease is still under debate although the presence of MAP in human sample is having an increasing attention.

The MAP genome (strain K-10 of bovine type II clinical isolate) has been sequenced in 2005 [Li et al., 2005]. This strain was chosen because of its high efficiency of transformation with plasmid DNA, its virulence and its amenability to transposition mutagenesis. The K-10 MAP genome has a single circular chromosome of 4,829,781 base pairs, 91.3% of which are protein coding genes.

Fig. 2. Possible transmission of MAP from infected animals to human



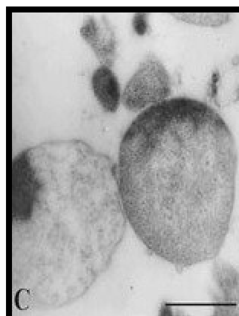
Source: Chacon,et al., 2004.

2.5 Bacterial cell wall deficient forms

The cell wall-deficient or also called cell wall-defective (CWD) state is considered one stage of the bacterial cell-cycle which could occurs in a particular group of microorganisms capable of existing in more than one form [Domingue and Schlegel, 1977ab].

CWD forms are believed to be a product of the interaction between an infectious agent and the defense system of the host [Mattman, 1961ab]. They are able to grow without their cell wall, which could be either completely (protoplast form) or partially (spheroplast form) absent. They can assume pleomorphic forms such as tiny filterable, filamentous, vesicles, giant and membranous structures (see example in Fig. 3).

Fig. 3. Examples of CWD bacteria



CWD forms of *Staphylococcus aureus*

Source: Michailova et al., Persistence of *Staphylococcus aureus* L-form during experimental lung infection in rats. FEMS Microbiol Lett 268, 88-97. 2007

The life cycle of the CWD bacteria is not well known, but they are thought to replicate by various mechanisms including budding, filamentous growth and binary fission.

A residual cell wall persists in most of the CWD bacteria but Slabyj et al., [1973] found that the length of the teichoic acid chain in CWD *Streptococcus pyogenes* was about half of that present in the parent strain. Moreover, it has been shown that the lipid composition of CWD bacteria is increased compared with the correspondent vegetative form [Slabyj and Panos, 1973; Onwuamaegbu et al., 2005].

The CWD bacteria are known from over a century and they were initially defined as L-forms or difficult-to-culture/non-cultivable microorganisms. Since then, CWD bacteria have fascinated many researchers due to their unusual nature and their possible connection with several chronic diseases. The first scientist, who observed pleomorphic

forms of the bacteria *Vibrio cholera* lacking rigid cell walls, was Richard Pfeiffer in 1895.

In 1990 and in 1997 two important scientists, respectively Lida Mattman and Gerald Domingue, gave a great contribution to the understanding of the CWD bacteria.

Lida Mattman defined the ability of MTB to convert to the CWD form as the "L-cycle" of mycobacteria and showed the presence of CWD-bacteria in the blood of diseased humans. Since then, CWD bacteria have been postulated to play a role in several diseases, including 'autoimmune' disorders [Mattman, 2001; Almenoff et al., 1996]. Also Domingue demonstrated the presence of different CWD bacteria in osmotically lysed and filtered human blood that resembled streptococcal, staphylococcal and gram-positive filamentous (cocco-bacillary) forms [Domingue and Schlegel, 1977a]. He found these forms in 71% of the blood specimens processed from diseased humans and in 7% of those from supposedly healthy humans. A finding that modify the current held view that the blood of healthy humans is sterile for bacteria.

In 2007, a group of researchers demonstrated that CWD bacteria are able to "internalize, replicate and persist" in the lungs of infected rats and suggested these forms as involved in the pathogenesis of chronic and latent lung infections [Markova et al., 2008]. In fact, CWD bacteria are believed to resist phagocytosis to a greater extent than their parent bacillary forms [Harwick et al., 1972; Harwick et al., 1977] and to replicate slowly over a long period of time maintaining infection at a subclinical or chronic level, therefore CWD forms have been suggested to be clinically significant [Rook and Stanford, 1992; Domingue and Woody, 1997; Domingue, 2010]. The hypothesized mechanism is that these cryptic bacteria could convert back to their vegetative phase thus causing relapse or the CWD forms themselves could cause disease since they may integrate with host cells organelle to create bacterial-host cell antigen complexes which could provoke immunopathology consequences. Moreover, CWD bacteria could possess peptides that mimicking host peptides could be responsible for the activation of auto-reactive T or B cells [Domingue, 2010].

The *in vivo* bacterial conversion to the CWD state is probably not a spontaneous phenomenon but a consequence to the lytic activity of the host operated inside the phagolysosomes [Gerasimov et al., 2003] and to some bacterial drugs targeting the bacterial cell walls [Wang and Chen, 2001; Toshkov et al., 1977]. Lytic activity is

achieved by enzymes such as lysozyme which, hydrolyzing the polysaccharide portion of the peptidoglycan layer, represents the first "biological weapon", with a strong bactericidal effect, used by the host against pathogens. It is interesting to underline that CWD bacteria have been found in body districts such as spinal fluids or human lungs where there are elevated levels of lysozyme. In support to this, it has also been reported a finding of around 18% of culture negative meningitis spinal fluids in untreated cases of diseased humans and around 30% of culture negative meningitis spinal fluids in diseased patients that had received therapy [Mattman, 2001].

Many researchers have reported the isolation of several strains of CWD bacteria from patients with different diseases, such as infective endocarditis, acute coronary syndromes, rheumatic fever, Sarcoidosis, urinary tract infections, meningitis etc.,. These bacteria have been found in different tissues and interestingly also in red blood cells where Bisset and Bartlett [1978] demonstrated that atypical variants of *Bacillus licheniformis* are able to persist also within the normal people.

Numerous bacterial species are known to become CWD when subject to particular stress pressures, some of them are human pathogen and are listed in Table 1.

The conversion to the CWD forms can be simulated *in vitro* by using some inducing agents such as cell wall inhibiting antibiotics, high concentration of aminoacids (especially glycine and phenylalanine), peptidases and lytic enzymes [Madoff et al., 1967; Mattman, 2001]. Removal of the inducing agent may result in the reversion to the bacillary state or in the permanent stabilization to the CWD phase.

Although these altered, pleomorphic bacteria could be responsible of several chronic disorders recently there has been an important decline in active research on these forms. Moreover, CWD bacteria are mostly ignored by clinicians due to the difficulty in their identification and handling. For example the routine heat fixing technique used for staining procedures, transforms CWD into globs due to their high lipid content, therefore they cannot be visualized by the routinely stains assay.

Table 1. List of clinical important bacteria that undergo to CWD conversion

1. *Staphylococcus aureus*	19. *Stenotrophomonas maltophilia*
2. *Streptococcus pyogenes*	20. *Pseudomonas aeruginosa*
3. *Streptococcus viridans*	21. *Corynebacterium species*
4. *Mycobacterium species*	22. *Vibrio species*
5. *Escherichia coli*	23. *Listeria monocytogenes*
6. *Shigella flexneri*	24. *Legionella pneumophilia*
7. *Enterococcus faecalis*	25. *Bacteroides funduliformis*
8. *Clostridium welchii*	26. *Rhizobium lupinus*
9. *Clostridium tetani*	27. *Proteus mirabilis*
10. *Neisseriae species*	28. *Bifidobacterium bifidum*
11. *Haemophilus influenzae*	29. *Salmonella typhimurium*
12. *Micrococcus species*	30. *Streptobacillus moniliformis*
13. *Bacillus species*	31. *Nocardia species*
14. *Lactobacillus species*	*Spirochaetes:*
15. *Brucella abortus*	32. *Treponema pallidum*
16. *Bordetella pertusis*	33. *Borrelia burgdorferi*
17. *Serratia marcescens*	34. *Leptospira interrogans*
18. *Pasteurella multocida*	

Source: Onwuamaegbu et al., 2005

2.5 Mycobacterial cell wall deficient forms

Mycobacterial CWD forms have been isolated from patients affected by several diseases such as pulmonary TB [Zhu et al., 2000; Dorozhkova et al., 1995], Sarcoidosis [Almenoff et al., 1996], Scleroderma [Cantwell et al., 1980], and Crohn's disease [Chiodini et al., 1986;] where this state might be responsible for triggering an abnormal immune or autoimmune response [Hulten et al., 2000a]. Moreover, MTB spheroplast-like agents have been found following anti-tuberculous drug-treatment and suggested to be responsible for the persistence of infection [Imaeda, 1984].

Recently Lawrence Broxmeyer and Alan Cantwell speculated acid-fast tuberculous mycobacteria as "primary agents" in AIDS and reported numerous finding of acid-fast bacteria in AIDS damaged tissues and the possible correlation between mycobacterial CWD proteins and "gag, pol" proteins, specifically ascribed to HIV [Broxmeyer and Cantwell, 2008].

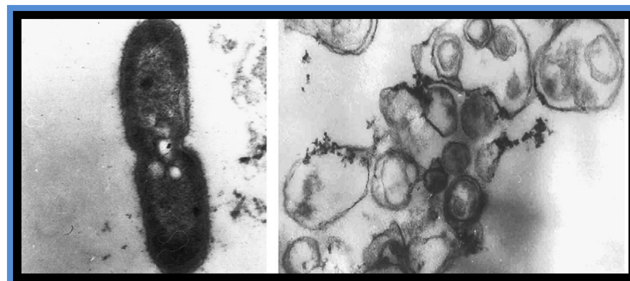
Considering that the CWD forms are believed to be the predominant forms of mycobacteria *in vivo* [El-Zaatari et al., 2001] and considering the increasing number of

potentially human pathogenic mycobacterial species [Sechi et al., 2008, Rosu et al., 2008], CWD-forms should deserve a special attention. Although they are very difficult to study, as they usually are isolated in low number and hardly cultured *in vitro* [Chiodini e al., 1986], the conversion to the CWD form can be simulated *in vitro* for many mycobacteria, MSMEG included [Udou et al., 1983; Rastogi and David, 1981; Naser et al. 1993].

In vivo isolated MTB_CWD forms are highly pleomorphic and exist as round spheroids, small granules and long slender filaments acid-fast negative in staining, as shown in Fig. 4 [Ma et al., 1989].

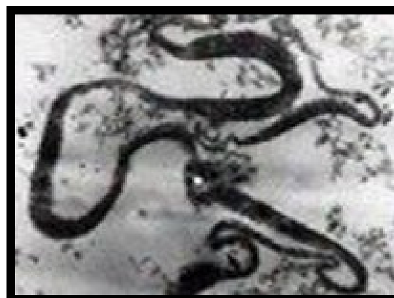
Because of their altered morphology and fully changed bacterial life-style, CWD-forms are not easy to identify in clinical specimens by ordinary methods and that is why CWD forms are considered to be both ‘‘difficult-to-identify” and ‘‘difficult-to-cultivate’’.

Fig. 4



Transmission electron microscopy of classical (left picture) and CWD-forms of MTB (right picture).

Source: Markova. Hidden face of tuberculosis, Bioscience Hypotheses (2009) 2, 441e442



Source: Michailova et.al, 2005 Int J Tuberc lung Dis 9(8):907–914

For example, MTB_CWD forms cannot grow in media used for MTB routine isolation and usually became acid-fast negative. Moreover, PCR analysis could also fail if the

DNA extraction method is not able to break-down the mycobacterial membrane which has been shown to be thicker (40.54 nm) than those of the bacillary forms (34.84 nm) [Liu and Lin, 1995]. In respect to this, Dai Yunhai [Dai et al., 1996] demonstrated that the best method to detect CWD-MTB DNA is physical grinding with glass sand compared to SDS, lysozyme and hot alkaline methods. Therefore, misdiagnosis can happen, but in order to reduce dissemination and relapse of TB, both reversible forms, also different in drug's susceptibility, needs to be identified and treated.

The MTB-CWD forms are considered a natural phenomenon ensuring the adaptive strategy of bacteria to survive and reproduce under un-favorable circumstances. It is quite likely that tubercle bacilli can use such pathway to their own advantage for long persistence *in vivo* (dormant state) and resistance to antimicrobial agents [Markova, 2009]. It has been shown in fact, that MTB can convert to the CWD form inside macrophages in guinea pigs, enhancing significantly their survival and persistence [Zhang, 2004]. Reversion to acid- fast rods might be then responsible for the disease reactivation.

CWD forms of MAP have also been observed in cattle with Johne´s disease, an important finding which explained the failure to successfully culture MAP from some infected animals and why pathology is seen in tissues (like the liver) when no acid fast bacilli are detected [Hulten et al., 2000b]. Moreover, these pleomorphic bacteria have been seen in patients with Crohn's disease [Chiodini,1986]. In respect to this, Hulten et al. in 2001 revealed CWD_MAP by *in situ* hybridization in patients with granulomatous and non-granulomatous form of Crohn's disease, Ulcerative Colitis and healthy humans [Hulten et al., 2001]. Moreover, Sechi et al. in 2001 recorded high detection rates of CWD_ MAP forms by *in situ* hybridization in more than 70% of patients with Crohn's disease but in contrast to the previous authors, they did not reveal significant differences in their detection between the granulomatous and non-granulomatous forms [Sechi et al., 2001]. Other evidences of Map spheroplasts forms were also reported in tissues and blood of patients with Crohn's disease [Schwartz et al. 2000; Sechi et al., 2005; Naser et al., 2004].

The particular characteristics of CWD-forms arise many questions about their biology and their putative mechanisms of persistence *in vivo*. Therefore, mycobacterial CWD-forms needs to be re-examined in the context of modern molecular biology. However, the

difficulty in identification and handle the *in vivo* mycobacterial CWD forms in a sufficient amount for molecular investigations makes their study very troublesome. Since today, the only way to investigate mycobacterial CWD forms at genomic level is simulating *in vitro* what happen *in vivo* through the use of inducing agents. Following mycobacterial chemical treatment is in fact possible to obtain sufficient amounts of bacteria that can be easily maintained in the spheroplast–like phase and used for further investigations.

3. AIM OF THE PROJECT

A wide body of research has shown that mycobacteria can enter into CWD forms previously referred as L-forms or non-cultivable forms. These forms have been associated to several human diseases including Sarcoidosis, Crohn's disease and extrapulmonary-Tuberculosis. Several studies underlined the clinical significance of the bacterial CWD forms as infectious agents and supposed a sort of parasitic relationship between CWD bacteria and their host, probably through the fusion with host organelle, that lead to a creation of bacterial host cell antigen complexes [Domingue, 2010]. This parasitic mechanism could be responsible for the tolerance of the microorganism by the immune system and hence for their persistence or dormancy characteristic.

Although numerous publications exist about detection of mycobacterial CWD forms in humans and animals from biological tissues (by microscopy and DNA hybridization techniques) and about their role in antibiotic resistance, the significance of CWD forms in pathogenesis is still unknown mainly due to the difficulty in handling the *in vivo* isolated CWD mycobacteria.

Hence, in the contest of the modern molecular biology, we decided to examine the gene expression changes of both pathogenic (MTB and MAP) and non-pathogenic (MSMEG) mycobacteria, *in vitro* converted to the CWD state. The *in vitro* conversion by chemically treating the mycobacteria is necessary in order to obtain sufficient amount of stable CWD bacteria that could be used for Microarray experiments. To our knowledge, this is the first work finalized to the study of the mycobacterial CWD forms at transcriptional level.

The aim of this research was to identify key genes responsible for the biochemical and physiological characteristics of these forms that we supposed to be peculiar and different from those of the bacillary parents. In particular we wanted to investigate the gene regulation of the CWD mycobacteria giving a special attention to those genes involved in persistence and pathogenesis. At the same time we aimed to highlight the stress response of mycobacteria to lysozyme burst, which is relied *in vivo* by activated phagocytic cells.

We believed that results could open new insights on mycobacterial CWD physiology and life-style also in relation to the still poorly understood host-pathogen interaction. In particular findings could have inferences in understanding several diseases such as Crohn’s disease where CWD_MAP have been found, animal Paratuberculosis where

diagnosis is sometimes difficult due to the failure to successfully culture this bacteria from some infected animals or in human Tuberculosis where coexistence of classical walled bacteria and CWD forms is considered a natural phenomenon, ensuring the adaptive strategy of the mycobacteria. Therefore, findings might also have important implications in the development of CWD specific targets useful for clinical investigations and drug design.

4. MATERIALS AND METHODS

4.1 Bacterial cultures and growth media

MAP1515 (ATCC 43015) Linda strain originally isolated from a Crohn's disease patient [Chiodini et al., 1984], MSMEG MC^2155 (ATCC 700084) and MTB H37Rv (ATCC 25618) were cultured in Middlebrook 7H9 medium (Sigma-Aldrich, MO, USA) supplemented with 0.2% glycerol (Sigma-Aldrich, MO, USA), 0.05% Tween 80 (Sigma-Aldrich, MO, USA) and 10% v/v albumine dextrose catalase (ADC, Sigma-Aldrich, MO, USA) in 25 cm^2 vented tissue culture flasks (T25) positioned up-right at 37°C. For MAP, 2 mg/L of Mycobactin J (MicJ, Allied Monitors, Fayette, MO, USA) were also added.

4.2 *In vitro* generation of mycobacterial CWD forms

MSMEG, MTB and MAP cells were converted to the CWD form as previously described [Hulten et al., 2000a; Naser et al., 1993]. Mid-log-phase cells were treated with 1% glycine solution in 7H9/ADC/Tw80 medium for 18h (MSMEG) and 72-92h (MTB and MAP) at 37°C. Pre-spheroplasts were harvested by centrifugation at 5800 x g for 15min at room temperature (RT) and washed 2 times with sucrose – MgCl2 (SM) solution (0.5M sucrose, 20mM MgCl2). Pellet was re-suspended in the original volume of 7H9/ADC/Tw80-SM (60:40) solution supplemented with 25µg/ml lysozyme (Sigma-Aldrich, MO, USA) and 1.2% glycine. Treated cultures were monitored daily by light microscopy for morphology changes, such as rounding of bacteria and un-clustered cells which are consistent with the formation of CWD forms [Rastogi and David, 1981; Udou et al., 1983; Chiodini et al., 1986]. Conversion from the bacillary form to the CWD state (spheroplast-form) was completed after 48h incubation at 37°C for MSMEG and after 5 days for the other two mycobacterial strains. Conversion was also confirmed by Ziehl-Neelsen stain that was negative.

Untreated cells, were grown in 7H9/ADC/Tw80 medium (for MAP supplemented with 2 mg/L of MicJ). They were harvested as the treated cells, washed 2 times with 7H9/ADC/Tw80 medium, then suspended in the original volume of 7H9/ADC/Tw80 medium and incubated as described for treated cells.

After conversion of treated cultures to the CWD phase, both treated and untreated cultures were harvested at 4°C, and immediately processed for RNA extraction.

4.3 Mycobacterial RNA isolation

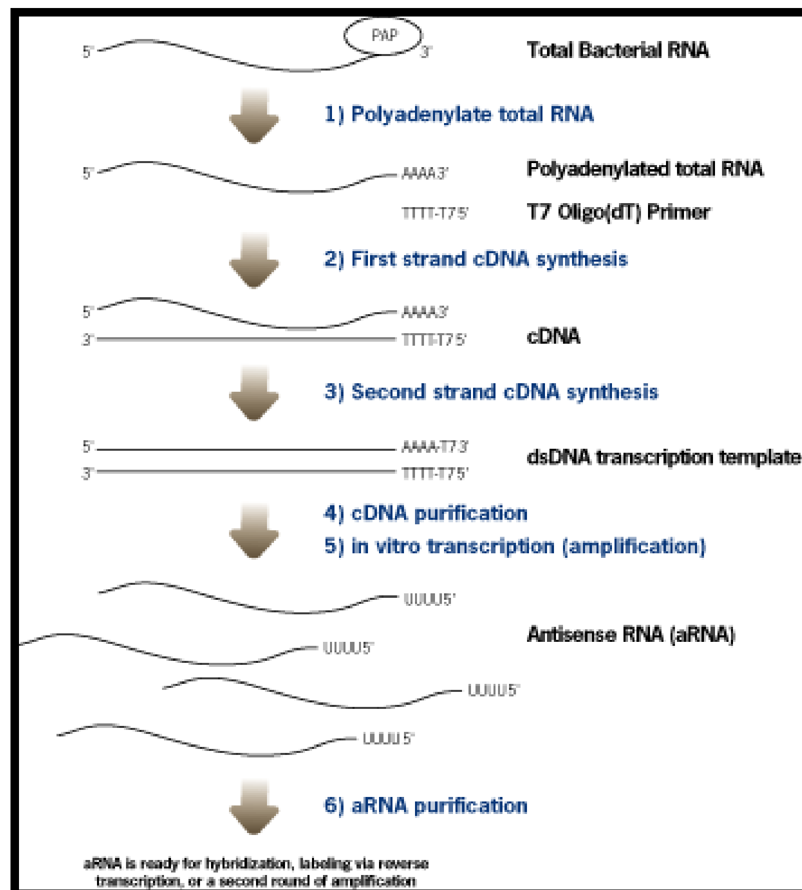
RNA was isolated by using the *RiboPure-Bacteria* Kit (Ambion, Inc) following the manufacturer's protocol with few modifications. Approximately $1x10^9$ mycobacterial cells were re-suspended in 350µl of *RNAwiz* solution and transferred to a 0.5ml RNAse-free microcentrifuge tube containing 250µl of 0.1mm ice-cold Zirconia Beads. Tubes were immediately processed by using the *RiboLyser RNA Lysing* machine (RiboLyser FP120-HY-230, Hybaid) for 2 cycles (30s at speed 6.5 with 1 min of cooling interval between each bead beating step) for MSMEG and MTB, and 3 cycles for MAP. Remaining steps were performed according to the manufacturer's instructions. All RNA samples were treated with *Dnase I* (Ambion, Inc) to remove trace amounts of genomic DNA and purified by *RNeasy* column (Qiagen). RNA yield and purity was evaluated with the *Nanodrop spectrophotometer* (NanoDrop1000, Thermo Scientific) while RNA quality was examined by Denaturing Gel Electrophoresis.
For each species we carried out three individual experiments on three different flasks. Good quality RNAs obtained from each pellet were then pulled together and used for microarray experiments to get an average of three separate experiments.

4.4 mRNA enrichment and linear amplification of mycobacterial RNA.

The *16S* and *23S* ribosomal RNAs (rRNAs) were removed from total RNA (tot-RNA) by using the *MICROBExpress Bacterial mRNA Purification* Kit (Ambion, Inc). Ten micrograms of in-put tot-RNA was used to get an average of 1-2µg out-put enriched m-RNA. rRNAs removal was confirmed by Denaturing Gel Electrophoresis. Mycobacterial mRNA was then amplified by using the *MessageAmp II Bacterial Kit* (Ambion, Inc) to generate amplified RNA (aRNA). Briefly, enriched m-RNA was first polyadenylated by *E. coli* poly(A) polymerase. Poly(A) tailed RNA was then reverse transcribed by using an oligo(dT) primer that incorporated a *T7* promoter. The resulting single-stranded cDNA was then subjected to a second-strand production reaction to obtain a double-stranded cDNA that was used as the template for *in vitro* transcription (IVT) by *T7* RNA polymerase which generated antisense aRNA copies of each transcript in the sample.

Amino-allyl UTP was incorporated into aRNA during the IVT reaction to allow subsequently fluorescence dye coupling. The amplification of RNA was required in order to increase the small amount of aRNA obtained especially for MAP (see Fig.5). This is a linear amplification method where the composition ratios of mRNA in the total mRNA are well preserved.

Fig.5. Schematic of MessageAmp™ II-Bacteria Procedure (Ambion)



Source: MessageAmp™ II-Bacteria kit (Ambion)

4.5 aRNA labeling and Microarray hybridization

Fourty µg of aRNA were labelled with either *Alexa Fluor dye 647* or *555* (Invitrogen, USA) following manufacturer's protocol. Purification of coupled aRNA was performed by *RNeasy purification system* (Qiagen) and incorporation of dyes was evaluated using the *Nanodrop spectrophotometer*. Before hybridization, coupled aRNA was fragmented using the RNA fragmentation reagents (Ambion, Inc) following manufacturer's protocol.

All microarray slides were purchased from MYcroarray® (Biodiscovery LLC, Ann Arbor, MI, USA). MSMEG MC^{2}155 3x20K microarray contained 18968 probes surveying 6644 genes or 98.9% of all genes with 3 replicates per probes. MTB H37Rv 3X20K microarray contained 11353 probes surveying 3975 genes or 99.4% of all genes with 5 replicates per probe. MAP K10 3x20K microarray contained 12536 probes surveying 4337 genes or 99.7% of all genes with 4 replicates per probe. The 3X20K array format consisted of a grid of 160 columns by 132 rows for a total of 21,120 spots with 1120 spots reserved for internal quality control, leaving 20,000 spots available for mycobacterial probe sequences. All 3 arrays were identical.

Microarray hybridization was performed following the *Recommended Sample Preparation and Hybridization Protocol* drawn-up by MYcroarray® (Biodiscovery LLC) with some modifications.

Briefly, 220µl hybridization solution (66µl of 20X SSPE (3M NaCl, 20mM EDTA, 118.2mM NaH2PO4, 81.8mM Na2HPO4), 10% formamide, 0.01mg/ml BSA, 0.01% Tween-20, 1mM DTT, 1% manufacturer control oligos, 10µg of each target coupled-aRNA, RNAse-free water to final volume) was prepared and pre-warmed at 56°C. All hybridizations were carried out in a water bath at 55°C for 18h with *OneArray Sealed Hybridization Chambers* (PhalanxBio Inc., Palo Alto, CA, USA). After incubation, slides were washed twice with SSPE 1X- 0.1mM DTT for 2 min followed by a 30 sec final wash with 0.25 X SSPE – 0.1mM DTT and dried out with air-spray and immediately scanned. All scans were carried out with the *Gene Pix AXON 4200A* instrument (Axon Instruments) at 5-µm resolution with full dynamic range of signal intensities at 1 – 65.000 in two-color red/green mode (635nm and 532 nm filters).

4.6 Microarray data analysis

The intensities of the two dyes for each spot were quantified by using the *Gene Pix 6.0 Software* and calculated as the mean of median intensity of each replicate spot for each gene. Spot signal intensities were normalized by using total array intensity and the ratio-mean based algorithm options. Subtraction of background was achieved with internal control probes as default. Subsequent expression analysis and hierarchical clustering was then performed by the free viable software *Multi Experiment Viewer*

(MeV) from TM4 Suite [Saeed et al., 2003]. A change of 2-fold in the intensities with a *p*-value < 0.05 was considered significant for a differential gene expression.

4.7 Reverse Transcriptase Real Time quantitative PCR

The gene expression ratios detected by Microarray analysis were confirmed by reverse transcription real-time quantitative PCR (RT-RealTimeqPCR). Total RNA was reverse transcribed with random primers by using the *GoTaq® 2-Step RT-qPCR System* (Promega) according to the manufacturer's protocol and *SYBR green* assay. cDNA was further amplified by incubating the amplification reaction mixture for 2 min at 95.0 °C, followed by 40 cycles of 15s at 95.0 °C, 1min of annealing/extension at 60.0 °C. PCR efficiencies were derived from standard curve slopes in the *iCycler software v. 5* (BioRad Laboratories, Inc., Hercules, CA). Melt-curve analysis was also performed to evaluate PCR specificity. Each reaction was performed in duplicate on a *Real Time PCR i-Cycler iq5* (Biorad). Primers were designed with the software, *Primer3* (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi) and analyzed using the BLAST algorithm against the corresponding mycobacterial genome to confirm their specificity.

Relative gene expression was determined by the 2^{-ddCT} method from calculated threshold cycle (CT) values using mycobacterial *16S* gene of the genus specific control strain as an internal normalization standard to determine transcript level changes.

5. RESULTS

5.1 Analysis of overall gene expression data

To investigate the transcriptional profile associated to the mycobacterial CWD state we used an *in vitro* approach (since the *in vivo* isolated forms are troublesome to culture and to keep in the CWD phase) in order to get sufficient amount of stable spheroplast-like forms for genetic investigations. The *in vitro* CWD induction was achieved as described in Materials and Methods by using glycine, an iper-osmotic solution and lysozyme. The transcriptome was then studied by using commercial arrays representing more than 98% of the mycobacterial protein coding genes and compared to those of the bacillary parents.

The transcriptome profile was evaluated for MSMEG, MTB and MAP grown to mid-log exponential phase (McF=3). The number of genes represented in each array and the percentages of up-regulated and down-regulated genes are displayed in **Table 2 A** and **B.** The list of genes with a fold change of ≥2 for each species is available in the Appendix section while all raw microarray files and processed data have been submitted to the GEO database at NCBI (**accession number GSE32237**).

Table. 2A. Schematic representation of overall data

The table shows the number of genome's genes, the number of genes represented in the array and the number of differentially regulated genes after conversion to the CWD form.

	Genes	Refseq	Genes in the array	N° differentially expressed genes
MSMEG *MC2 155*	6938	NC_008596	6644	209 up/346 down
MTB *H37RV*	4047	NC_000962	3975	239 up/231 down
MAP *K10*	4399	NC_002944	4337	49 up/61 down

Table. 2B. Number and percentages of differentially expressed genes

	% Differentially expressed genes	% Up	% Down
MSMEG *MC2 155*	8.4	3.2	5.2
MTB *H37RV*	11.8	6	5.8
MAP *K10*	2.5	1.1	1.4

Results for CWD_MSMEG showed that out of 6644 genes represented on the array, 8.4% were differentially expressed (3.2% up and 5.2% down); for CWD_MTB among 3975 genes on the array, 11.8% were differentially expressed (6% up and 5.8% down), while for CWD_MAP out of 4337 genes represented in the array only 2.5% were differentially regulated (1.1% and 1.4% down).

Genes were classified according to functional categories by using the *Microbial Genome Database for Comparative Analysis* (MBGD) [Uchiyama, 2007], the *Kyoto Encyclopedia of Genes and Genomes* (KEGG) [Ogata et al., 1999; Kanehisa and Goto, 2000] and in addition for MTB the *TubercuList database* [Camus et al., 2002]. Genes were grouped into the following major categories: hypothetical and hypothetical conserved proteins, antibiotic production and resistance, aminoacid metabolism, carbohydrate and xenobiotic metabolism, lipid and fatty acid metabolism, metabolism of cofactors, energy metabolism, cell envelope/membrane proteins, metabolism of terpenoids and poliketides, membrane transport, PE/PPE proteins, cell division, signal transduction, transcription, translation, DNA replication and repair, transposition, virulence genes and broad functions.

Within each category, we searched for genes involved both in the capacity of the bacteria to respond to lysozyme stress and in the persistence of pathogenic species.

5.2 Differentially expressed genes in CWD *Mycobacterium smegmatis*

Percentages of differentially regulated genes, grouped according to functional categories, are summarized in Table 3.

A large percentage (14.83 up and 13.87 down) of differentially expressed genes coded for hypothetical and hypothetical conserved proteins. This result underlines the need of assigning a function to those proteins in order to reveal their specific role in the physiology of the CWD_MSMEG state.

Within the category of membrane proteins encoding genes, 9.57% were up-regulated and deputized to membrane transport while 11.85% were down-expressed and most of them committed to the ABC transporter superfamily that are involved in metabolite trafficking transport in prokaryotes. A percentage of 4.78% over-expressed genes belonged to the class of cell envelope/membrane proteins and among them we observed *MSMEG_5644* implicated in host invasion and *MSMEG_3256* encoding for a mucin associated protein.

This protein is used by the bacteria to bind mucin, an important substrate implicated in bacterial cells growth, adhesion and protection in the gastrointestinal tract. Among down-expressed genes (2.60%), we found *MSMEG_6201*, *MSMEG_3022* encoding for transglycosylase associated proteins and *MSMEG_3528* expressing for an enzyme involved in murein degradation. PE/PPE genes were not differentially expressed in the CWD form except for the over-expressed *MSMEG_2737* gene.

In CWD_MSMEG respectively 5.74% of up-regulated and 3.76 of down-regulated genes were assigned to amino acid metabolism, while 13.88% of over-expressed genes were involved in carbohydrate and xenobiotic metabolism. Among them, we found the up-regulation of *MSMEG_5728* and *MSMEG_4373* that encode for peptidoglycan deacetylases probably used *in vivo* by the mycobacteria to modify its cell envelope so that the peptidoglycan is no longer recognized by the host lysozyme. Therefore the peptidoglycan modification might represent a defense mechanism that MSMEG uses against this stress.

In the functional class of lipid and fatty acid metabolism, up and down regulated genes were respectively 7.66% and 5.2% of the total differentially expressed. Over-expressed genes were involved both in the β-oxidation and the oxidation of organic substrates belonging to the cytochrome P450 superfamily and some genes involved in lipid biosynthesis, especially cuticular wax and fatty acid, such as *MSMEG_4023*, *MSMEG_0290* and *MSMEG_1204*. Down-expressed genes belonged both to lipid biosynthesis and lipid degradation pathways, such as genes encoding for bi-functional wax ester synthases, short chain dehydrogenase and acetyl-CoA synthases enzymes. Interestingly we observed the down-regulation of *MSMEG_2468* (L-carnitine dehydrates/bile acid-inducible) which encodes for a hydro-lyases involved in the metabolization of L-carnitine in the host intestine. In prokaryotes L-carnitine might be used as both a carbon and nitrogen source for aerobic growth.

CWD_MSMEG showed and high number of differentially expressed genes encoding for transcriptional regulators and most of them were repressors. We observed for example an high expression of *MSMEG_6913* (putative transcriptional repressor protein) involved in the development of antibiotic resistance and in the regulation of the virulence factor synthesis, *MSMEG_4487* (ferric up-take regulation protein) responsible for controlling the intracellular concentration of iron, *MSMEG_0285*, *MSMEG_6628*, *MSMEG_6604*,

that express for negative transcriptional regulators of the *TetR*, LuxR family proteins. Moreover, most of the genes involved in translation were down-regulated especially those encoding for ribosomal proteins and for some initiation translation factors. Interestingly, we also observed the down-regulation of 30 genes encoding for transposase, just 3 were up-regulated.

According to the KEGG database, few virulence genes were down expressed such as MSMEG_0834 a tuberculin related peptide and MSMEG_2404, an extracellular deoxyribonuclease hypothesized to play a role in dissemination and spread of infecting bacteria. To further analyses the presence of CWD activated or repressed virulence genes, we performed an additional search by using the database Virulence Factors for Bacterial Pathogens (VFDB) [Chen et al., 2005]. We searched for 109 know virulence genes, results are summarized in Fig. 6 and Table 4.

Fig. 6. Venn diagrams of overlapping genes between either up or down regulated genes and 109 known virulence genes, searched by using the Virulence Factors for Bacterial Pathogens (VFDB) database.

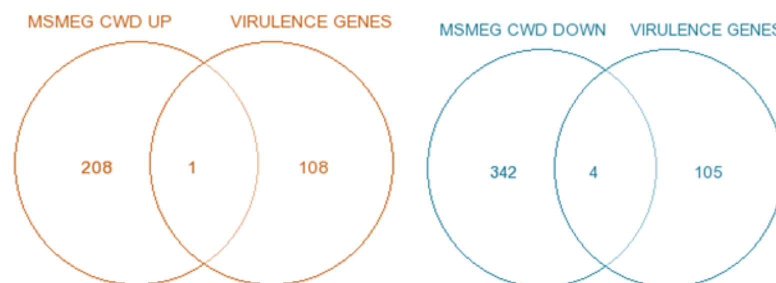


Table 3: Distribution of differentially expressed genes for CWD_SMEMG, according to functional categories. Up-regulated and down-regulated genes are expressed in percentages.

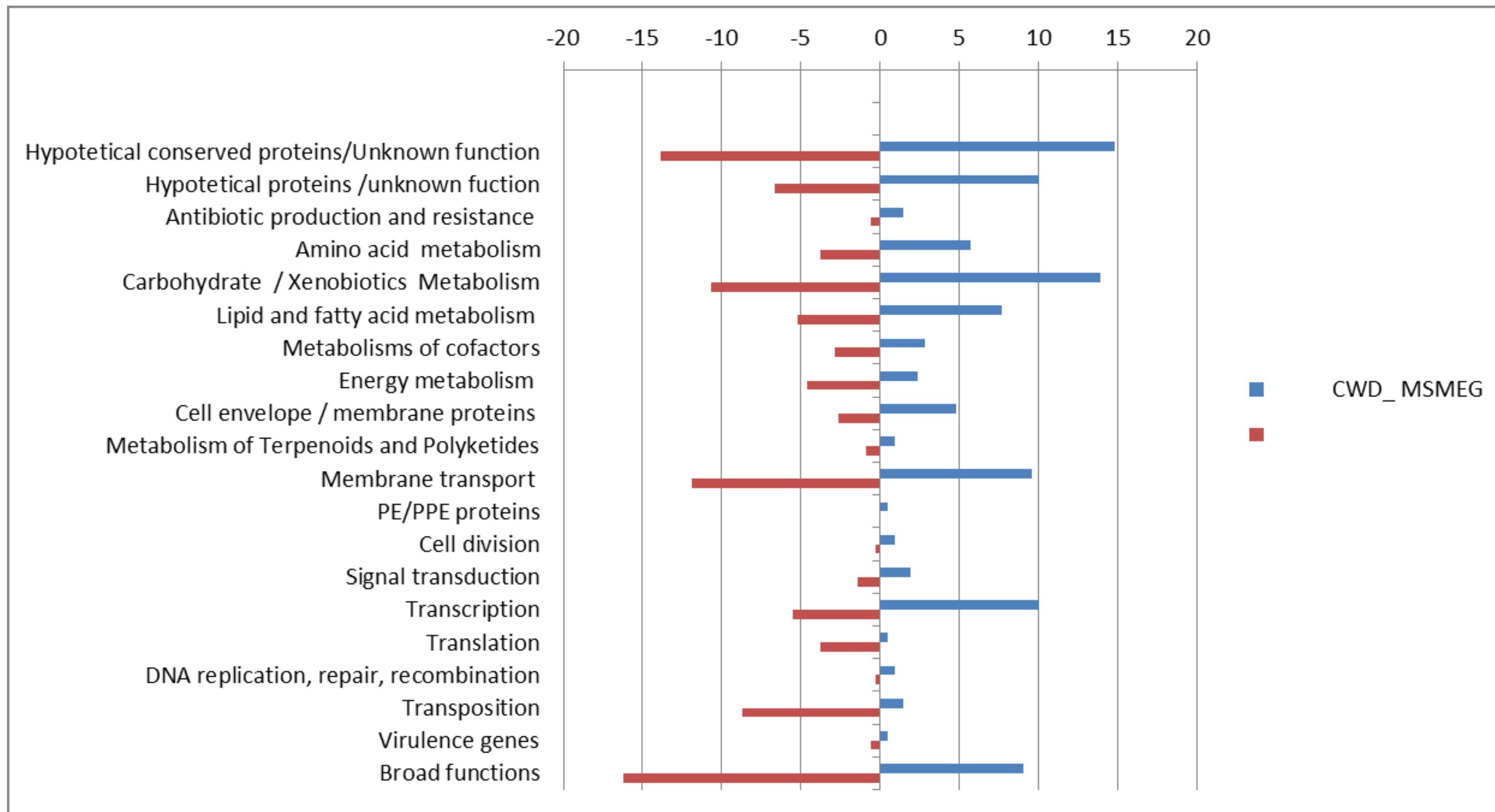


Table 4: Virulence genes differentially regulated in CWD_MSMEG, gene's homologies with MTB and MAP are also shown. These genes were found by using the Virulence Factors for Bacterial Pathogens (VFDB) database.

LOCUS	*GENE*	HOMOLOG	HOMOLOG	GENE PRODUCT/ FUNCTION
UP				
MSMEG_0082	*esx1*		*Rv3882c*	Secretion system
DOWN				
MSMEG_2965	*relA*	*MAP__1047*	*Rv2583c*	(p)ppGpp synthesis and hydrolysis
MSMEG_3932	*hspX*		*Rv2031c*	Secreted proteins
MSMEG_4258	*trpD*	*MAP_1931c*	*Rv2192c*	Amino acid and purine metabolism

In CWD_MSMEG we found one up and three down-regulated virulence genes, these are *MSMEG_0082 esx1*, *MSMEG_2965 relA* which it is important to sense nutrient deprivation and to respond to stress by synthesizing guanosine pentaphosphate (pppGpp), *MSMEG_ 4258trpD* and *MSMEG_3932 hspX*.

5.3 Differentially expressed genes in CWD *Mycobacterium tuberculosis*

Results for CWD_MTB are shown in Table 5

In CWD_MTB we identified a high proportion of differentially regulated genes encoding for cell envelope components such as PE/PPE family proteins probably implicated in the mycobacterial virulence and antigenic variation. Three PE (*Rv3746c* PE34, *Rv1169c* PE11, *Rv1430* PE16), one PE_PGRS (*Rv3507*, PE_PGRS53) and four PPE (*Rv3532* PPE61, *Rv2608* PPE42, *Rv1168c* PPE17, *Rv1918c* PPE35) were up-regulated while five PE_PGRS (*Rv0578c* PE_PGRS7, *Rv1087* PE_PGRS21, *Rv3514* PE_PGRS57, *Rv0754* PE_PGRS11, *Rv0977* PE_PGRS16) and six PPE (*Rv3135* PPE50, *Rv2356c* PPE40, *Rv3347c* PPE55, *Rv3136* PPE51, *Rv1790* PPE27, *Rv1135c* PPE16) were down-expressed. The PE_PGRS family proteins belong to a subgroup of the PE family, which is characterized to contain a highly repetitive domain rich in the amino acids glycine and alanine [Brennan and Delogu, 2002]. We also observed the concomitant down-regulation of six *esat-6* like genes (*Rv3874 esxB*, *Rv3887c*, *Rv0287 esxG*, *Rv1197 esxK*, *Rv1038c esxJ*, *Rv2347c esxP*) that encode for a large family of mycobacterial proteins (which typically consist of about 100 aminoacids and are characterized by the conservation of a central WXG motif) and two *mce* genes that are *Rv0169 mce1A* and *Rv0594 mce2F*.

In CWD_MTB only 2.5% of total genes involved in aminoacid metabolism were up-regulated and were preeminently related to leucine and cysteine synthesis, while almost 4.8% were down-expressed such as genes for the biosynthesis of phenylalanine, arginine, valine and isoleucine. This massive down-regulation was correlated to the consequent down-regulation (3.9%) of genes associated to the metabolism of cofactors and vitamins.

Within the class of carbohydrate metabolism, we observed the over-expression of genes involved in the glyoxylate cycle, carbohydrate degradation and oxidoreductase activity. In particular, it is important to underline the up-regulation of *Rv0467 icl* (isocitrate lyase) which expresses for one of the two enzymes of the glyoxylate shunt, implicated in the recycling of acetyl-CoAs produced by β-oxidation.

Within the genes involved in the lipid an fatty acid metabolism pathways, we observed the up-regulation of two genes involved in lipid biosynthesis, *Rv3174* and *Rv0439c* (putative dehydrogenases) and six genes involved in lipid degradation, four of them were *fadD34*, *fadD29*, *fadD19*, *fadD14* (putative fatty acyl-coenzyme A (CoA) synthase) needed to convert fatty acids in fatty acetyl-CoAs. Among down-expressed genes, six were involved in fatty acid and mycolic acid biosynthesis such as *Rv2244 acpM*, *Rv2245 kasA*, *Rv2246 kasB*, *Rv1483 fabG, Rv2247 accD6.* It has been suggested that KasA, functions in the initial elongation of the mycolic acid chain, while KasB elongates the chain to full length [Swanson et al., 2009]. Among down-regulated genes, we found *Rv2719c*, probably involved in the bacterial cell wall degradation and primarily expressed from a DNA damage inducible promoter independently activated from RecA [Brooks et al., 2006].

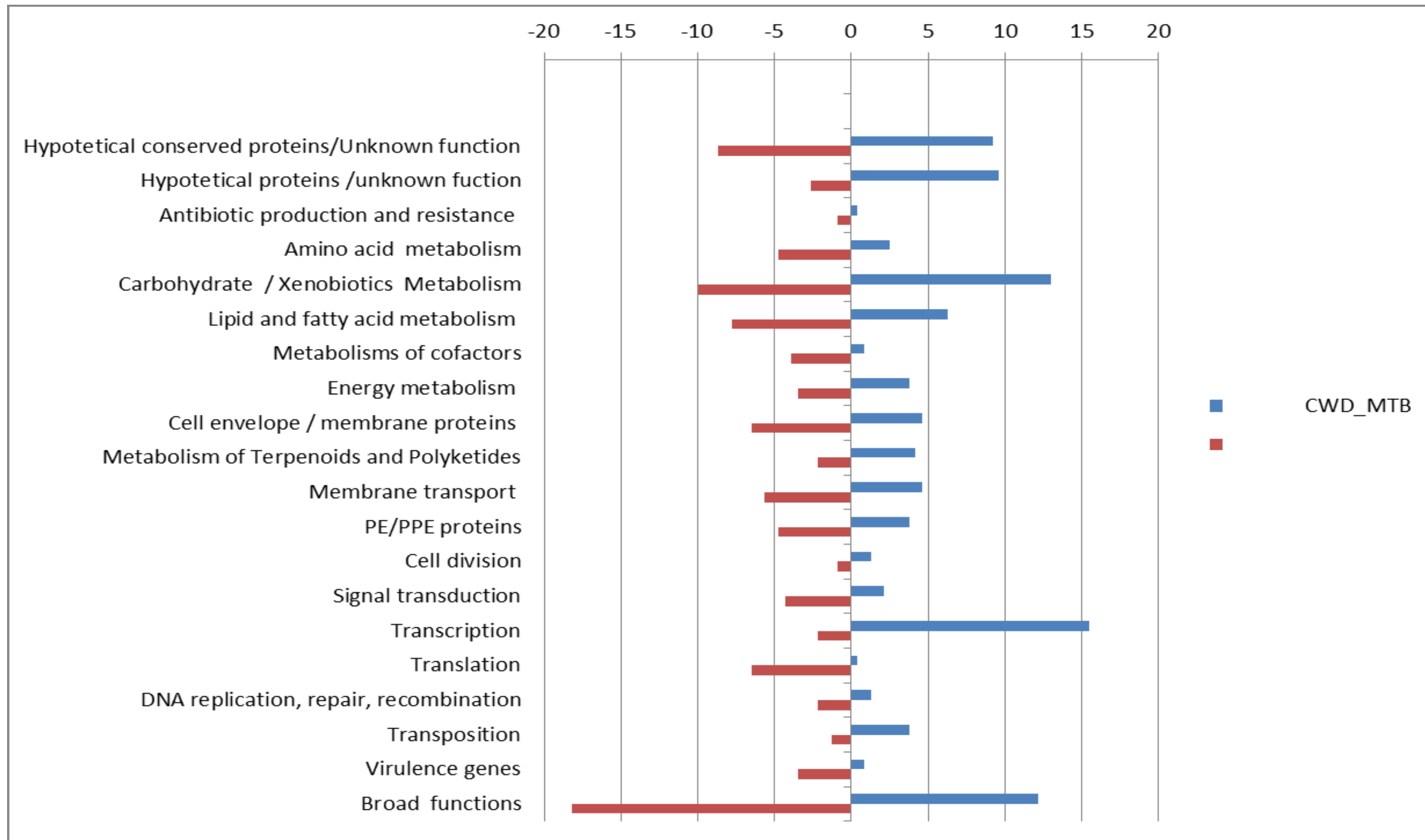
Moreover, CWD_MTB showed a high percentage (4.18%) of over-expressed genes committed to the functional class of the Terpenoids and Polyketides pathways that are important precursors of complex lipids; among them, we found the up-regulated genes *Rv1661 pks, Rv3562 fadE31*, *Rv3139 fadE24* and many *cyp* genes. However, some genes were also down-expressed such as *Rv3825c pks2*, *Rv1934c fadE17* and *Rv0972c fadE12.*

It is interestingly to note that CWD_MTB up-regulated 9 genes involved in general stress response (especially nutritional stress), some of them encodes for PIT domain proteins which are apparent cytosolic ATPases associated to type IV pilus systems but not required for its biogenesis. The majorities of PIN-domain proteins are the toxic components of toxin-antitoxin operon, and toxin-antitoxin gene cassettes and have found

to be abundant in free-living prokaryotes, including many pathogenic bacteria. These loci provide a control mechanism that helps free-living bacteria to cope with nutritional stress.

Other genes, *Rv1285cysD* and *Rv1286 cysNC,* involved in defense against stress-stimuli, were up-regulated. MTB *cysD* and *cysNC* genes form a single operon whose transcription is regulated in response to sulfur starvation. The MTB *cysD* gene, probably encodes for the adenylyl-transferase subunit of ATP sulfurylase, and it is up-regulated when MTB is inside macrophages [Pinto et al., 2004].

Table 5. Distribution of differentially expressed genes for CWD_MTB according to functional categories.



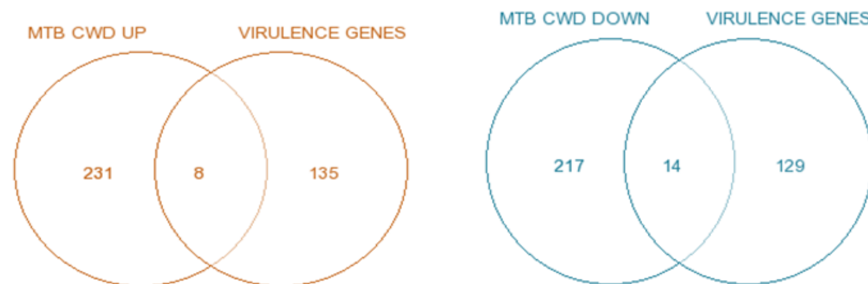
CWD_MTB showed the up-regulation of an high number of transcriptional regulators genes along with many transposase genes and, as we observed for CWD_MSMEG, most of the differentially regulated genes belonging to the translation class were down-expressed. Among up-regulated genes we found *Rv0586* (FadR-like regulator) which encodes for a putative fatty acid metabolism transcriptional regulator. This single repressor controls the transcription of the whole *fad* regulon, which regulates the expression of key enzymes required for both fatty acid β-oxidation and fatty acid biosynthesis. The *fad-R* like regulator has also been shown to have an auto-regulatory nature and to be associated to the *mce2* operon [Vindal et al., 2008]. Other differentially expressed genes were *Rv3862c whiB6* and *Rv3260c whiB2.* The last one in particular is required for differentiation and sporulation in sporulating Actinobacteria. A distinctive pattern of expression was also observed for genes encoding sigma factors such as *sigE* and *sigB* and transcriptional regulators belonging to the *TetR* (tetracycline antibiotic repressor) and *MarR* family (multiple antibiotic resistance repressor).
According to the KEGG database, CWD_MTB over-expressed some *mce* genes such as *Rv2390c* and *Rv3494c mce4F* and PE/PPE genes. Down-regulated genes belonged to the *mce* family genes such as *Rv0169 mce1A* and *Rv0594 mce2F* and ESX- family genes such as *Rv3874 esxB* which encodes for a small secreted proteins found in MTB culture filtrates, also known as an immunodominant antigen recognized by the sera of TB patients [Skjøt at al., 2000]. By using the VFDB database, we searched for 143 known virulence genes; results are shown in Table 6 and Fig. 7.

Table 6. Virulence genes differentially regulated in the CWD form of MTB

Virulence genes and their homologous were detected according to the database Virulence Factor of Pathogenic Bacteria (VFDB http://www.mgc.ac.cn/VFs/main.htm). Gene homologies with SMEG and MAP and are also shown.

LOCUS	*GENE*	HOMOLOG	HOMOLOG	GENE PRODUCT/ FUNCTION
UP				
Rv0467	*icl*	*MAP_3961*	*MSMEG_0911*	Lipid and fatty acid metabolism
Rv0981	*mprA*	*MAP_0916*	*MSMEG_5488*	Regulation
Rv1221	*sigE*	*MAP_2557c*	*MSMEG_5072*	Regulation
Rv2351c	*plcA*			Fospholipase C, lipid and fatty acid metabolism
Rv2386c	*mbtI*	*MAP_2205c*	*MSMEG_4524*	Mycobactin, metal uptake
Rv2428	*ahpC*	*MAP_1589c*	*MSMEG_4891*	Stress adaptation
Rv2962c	-			PDIM and PGL biosynthesis and transport
Rv3494c	*mce4F*	*MAP_0569*	*MSMEG_5895*	Mce operons
DOWN				
Rv0169	*mce1A*	*MAP_3604, MAP_3289c*	*MSMEG_0134*	Mce operons
Rv0594	*mce2F*	*MAP_4089*		Mce operons
Rv1736c	*narX*			Anaerobic respiration
Rv1886c	*fbpB*	*MAP_1609c*	*MSMEG_2078*	Antigen 85 complex, secreted proteins
Rv2031c	*hspX*		*MSMEG_3932*	Alpha-crystallin
Rv2192c	*trpD*	*MAP_1931c*	*MSMEG_4258*	Amino acid and purine metabolism
Rv2246	*kasB*	*MAP_1999*	*MSMEG_4328*	Lipid and fatty acid metabolism
Rv2936	*ddrA*	*MAP_1238c*		PDIM and PGL biosynthesis and transport
Rv2939	*papA5*			PDIM and PGL biosynthesis and transport
Rv2987c	*leuD*	*MAP_3025c*	*MSMEG_2388*	Amino acid and purine metabolism
Rv3132c	*devS*	*MAP_3270c*	*MSMEG_5241*	Regulation
Rv3133c	*devR/dosR*	*MAP_3271c*	*MSMEG_5244*	Regulation
Rv3825c	*pks2*	*MAP_3764c*		Sulfolipid-1 biosynthesis and transport, cell surface components
Rv3874	*esxB*		*MSMEG_0065*	ESX-1 (ESAT-6 system-1), secreted proteins

Fig.7. Venn diagram showing overlapping genes between differentially regulated genes and virulence genes detected according to the database Virulence Factor of Pathogenic Bacteria



CWD_MTB over-expressed eight virulence genes, in particular *Rv2351c plcA* and *Rv2962c* deserve a special attention, as they have no homologous not only in MSMEG and MAP but also in other mycobacterial species, *Mycobacterium bovis* included. Interestingly we observed the down-regulation of fourteen virulence genes and among them of *Rv2031c hspX* and *Rv2192c trpD* also repressed in CWD_MSMEG. Moreover, we found down-regulation of *Rv2936 ddrA* and *Rv2939 papA5* implicated in the formation and transport of the lipid phthiocerol dimycocerosate (PDIM) which is peculiar of pathogenic mycobacteria [Kolattukudy et al., 1997].

In the class of antibiotic production and resistance, CWD_MTB up-expressed genes involved in antibiotic production such as *Rv0893c* and *Rv0726c (*O-methyltransferase proteins) and *Rv1128c* (putative bacteriocin). Some genes were also involved in antibiotic resistance, such as *Rv0191* which encodes for an integral membrane protein probably involved in antibiotic transport and some genes belonging to the two-component signal transduction systems that enable bacteria to sense, respond, and adapt to a wide variety of stimuli including antibiotics. Among down-regulated genes we found *Rv1010 ksgA*, that when inactivated confers resistance to the aminoglycoside antibiotic kasugamycin [Duffin and Seifert, 2009; Ochi et al., 2009], *Rv3205c* and *Rv2004c* that encode for two conserved hypothetical proteins probably implicated in the bacterial antibiotic resistance mechanism through the inactivation of aminoglycoside antibiotics via phosphorylation.

5.4 Differentially regulated genes in CWD *Mycobacterium* subspecies *paratuberculosis*

Results for CWD_MAP are summarized in Table 7.

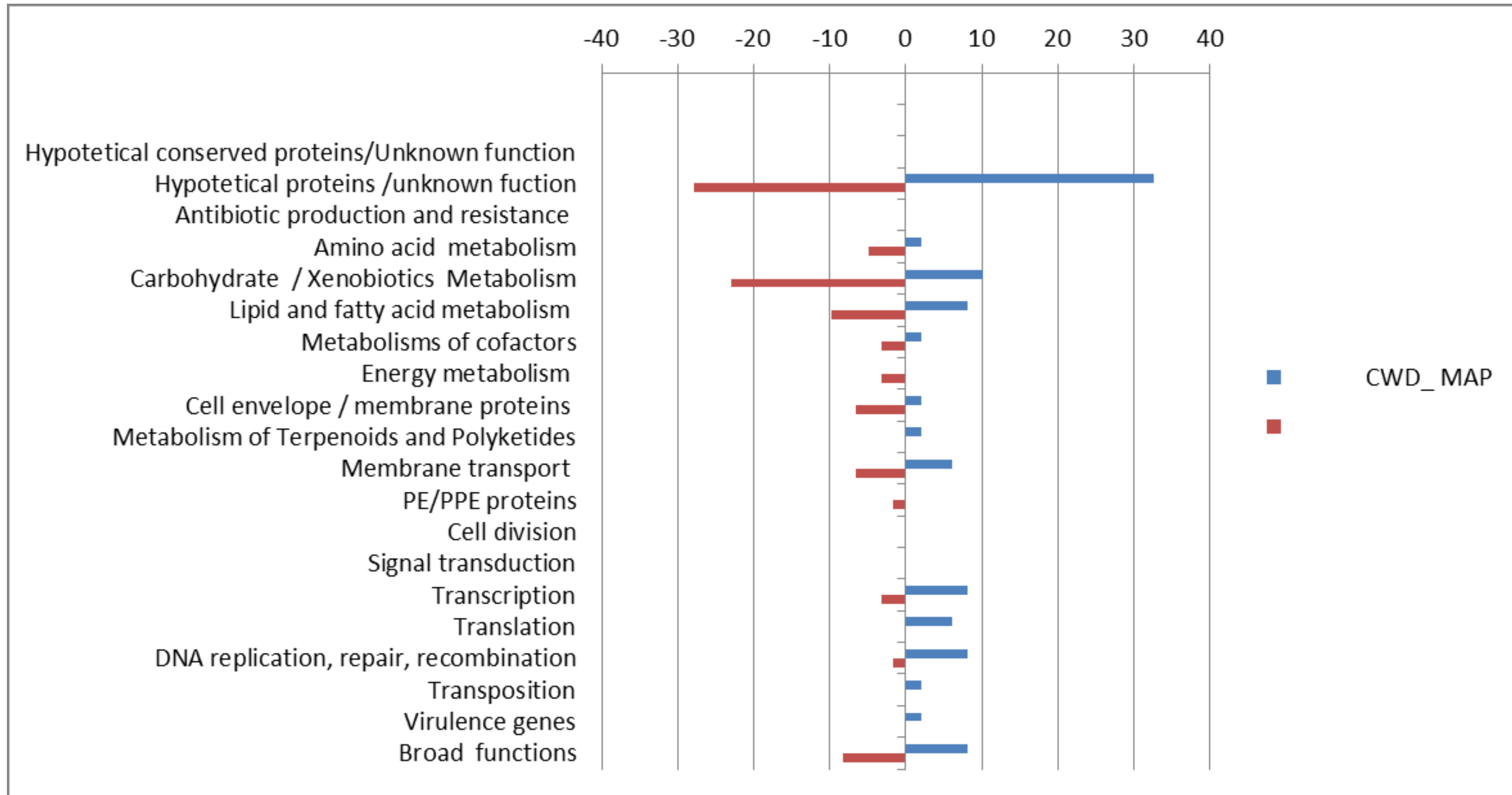
CWD_MAP differentially regulated a low number of genes. It also showed the highest percentage of up-regulated (32.6%) and down-regulated (27.8%) genes encoding for hypothetical proteins comparing to the other two species. However, no hypothetical conserved proteins were differentially regulated in CWD_MAP.

The CWD state was characterized by the over-expression of some ABC transporters genes such as *MAP_2098c*, *MAP_1236c drrC* and *MAP_1107* involved in cobalt/nickel, antibiotic and sulfonate/nitrate transport. No PE/PPE genes were up-regulated. Among down-expressed genes we found *MAP_3163* probably involved in peptidoglycan catabolic processes system, *MAP_1522* belonging to the PPE family, and few genes implicated in aminoacid and iron transport (*MAP_2807c, MAP_0044c, MAP_3727*).

Within the aminoacid metabolism category, 2% of up-regulated genes were associated to leucine biosynthesis while down-expressed genes belonged to lysine and serine biosynthesis pathways. A large percentage of genes were functionally correlated to the carbohydrate and xenobiotics metabolism class, 22.9% of them was down-regulated.

In the functional category of lipid metabolism, we observed the over-expression of *MAP_4294 fadD1* (long-chain-fatty-acid-CoA ligase) and of *MAP_1380* and *MAP_1605c* (hypothetical proteins) probably involved in lipid biosynthesis.

Table 7. Percentage of differentially regulated genes for MAP after *in vitro* conversion to the CWD form.



Down-expressed genes were *MAP_4095c mmaA2* (distal cyclopropane synthase of the alpha-mycolic acid), some genes encoding hypothetical proteins probably involved in lipid degradation and *MAP_3962 fabG2* (putative 3-hydroxybutyryl-CoA dehydrogenase repressor) and *MarR* family (multiple antibiotic resistance repressor).

CWD_MAP showed over-expression of genes involved in both transcription and translation and among them we found *MAP_0061c* which belongs to the PadR-like family of transcriptional regulators and *MAP_2895* which encodes for an iron dependent repressor of the DtxR family that orchestrates the virulence of several important human pathogens. The regulator represses genes for iron acquisition while activates iron storage genes, and it is a positive regulator of oxidative stress responses. Moreover, we observed down-regulation of *MAP_0098c* and *MAP_4151c* which belongs to the TetR family proteins that control the level of bacterial susceptibility to hydrophobic antibiotics and detergents. No transposase genes were differentially expressed in this specie.

According to the KEGG database, CWD_MAP up-regulated only one known virulence gene which is the *mce*-gene *MAP_3606 mce1c*. An additional search of 126 know virulence genes was performed by using the VFBD database, results are summarized in Fig. 8 and Table 8.

Fig.8. Venn diagram showing overlapping between differentially regulated genes and 126 known virulence genes according to the database Virulence Factor of Pathogenic Bacteria

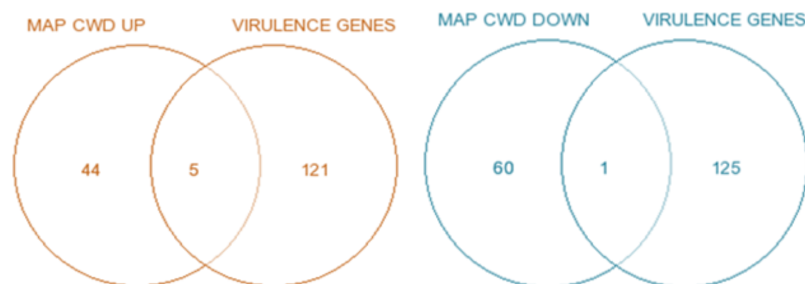


Table 8. Virulence genes differentially regulated in the CWD form of MAP

Virulence genes and their homologous were detected according to the database Virulence Factor of Pathogenic Bacteria (VFDB http://www.mgc.ac.cn/VFs/main.htm). Gene homologies with SMEG and MTB and are also shown.

LOCUS	*GENE*	HOMOLOG	HOMOLOG	GENE PRODUCT/ FUNCTION
UP				
MAP_1236c	*drrC*		*Rv2938*	PDIM and PGL biosynthesis and transport
MAP_1855	*mce5E*	*MSMEG_1147*		Mce operons
MAP_2177c	*mbtB*	*MSMEG_4515*	*Rv2383c*	Mycobactin, metal up-take
MAP_2192	*mce8D*	*MSMEG_2858*		Mce operons
MAP_3606	*mce1C*	*MSMEG_0136*	*Rv0171*	Mce operons
DOWN				
MAP_1509			*Rv1794*	ESX-5 secretion system

We found five over-expressed genes belongings respectively to the *mce* family (*MAP_1855, MAP_2192, MAP_3606*), iron up-take (*MAP_2177c*) and glycolipid biosynthesis and transport (*MAP_1236c*). The gene *MAP_1236c drrC* is also probably involved, as mentioned above, in the formation of a drug efflux proteins used for antibiotic extrusions. Just one gene, *MAP_1509,* was down-regulated.

5.5 Real Time Reverse Transcriptase quantitative PCR analysis of selected genes

To evaluate the performance of microarray experiments some microarray data were compared to those obtained from RT-Real Time qPCR. Selected genes and primer pairs are listed in Table 9. We examined the expression level of ten genes; four of them belonged to CWD_MSMEG, three to CWD_MTB and other three to CWD_MAP. The specie specific housekeeping gene *16S* was used as endogenous control. Real Time PCR data confirmed microarray results for the examined genes as shown in Table 10.

Table 9. Genes and primers for RT-Real Time qPCR used in this study.

Primer pairs for specific genes were selected using the online software Primer 3. The table also shows amplicon size and the function of selected genes.

PRIMERS	SEQUENCE	AMPLICON BP	GENE NAME	FUNCTION
MSMEG 16S FW	CAGCTCGTGTCGTGAGATGT	234	*16S*	16S ribosomal RNA
MSMEG 16S RV	AGACCGGCTTTGAAAGGATT			
MSMEG_1074 FW	TGTACGACTCCAGCCTGATG	186		polysaccharide deacetylase
MSMEG_1074 RV	ACTCCAGTTGCCACAGTTCC			
MSMEG_2404 FW	CTACTGGCCGCAACAGATG	153		extracellular deoxyribonuclease
MSMEG_2404 RV	AACTGCATCGCGTACTGACA			
MSMEG_2468 FW	ACGTACGGCCAGAACTTCAC	208		L-carnitine dehydratase
MSMEG_2468 RV	TGAACCACAGGGTGAACAGA			
MSMEG_5644 FW	GATGTACCTGTGGGCGTTCT	210		membrane protein, invasion protein
MSMEG_5644 RV	TAGTACCTGCCGCTGGACA			
MTB 16S FW	ATGCATGTCTTGTGGTGGAA	184	*16S*	16S ribosomal RNA
MTB 16S RV	GTGCAATATTCCCCACTGCT			
Rv1168c FW	TGAACCTCTACGCCACAATG	170	*PPE17*	cell wall protein, probable involved in virulence
Rv1168c RV	CTTCGAAGGCAGATTCTTGG			
Rv2351c FW	GTGGGAAAAGACCGCACTTA	200	*plcA*	probable membrane associated phospholipase C1
RV2351c RV	TGTACGGCGAAATGACAATG			virulence factor
Rv3545c FW	ATTTAACATCCTGCGCAACC	211	*cyp125*	esogenous lipid degradation
Rv3545c RV	GGTGTAGTCGACCTGCCAGT			
MAP 16S FW	GCCGTAAACGGTGGGTACTA	194	*16S*	16S ribosomal RNA
MAP 16S RV	TGCATGTCAAACCCAGGTAA			
MAP 1855 FW	AATTTCACCAACATCATCAACG	153	*mce5E*	virulence factor, invasion
MAP 1855 RV	GAGAGTGGACAATTGCGTCA			
MAP 2895 FW	GATGCTGCGCTATTTCACC	172		probable iron repressor factor
MAP 2895 RV	TTATGCCGTCACCCAGAT			
MAP_3032c FW	GCTGGACCATCACATCAACTT	190	*LeuB*	amino Acid biosynthesis
MAP_3032c RV	AGCGGTGTTCTGACTGACCT			

Table 10. Comparison between Microarray data and RT-Real Time qPCR results for a selected group of genes

Gene	Gene product	Expression	Folding Microarray	P-Value Array	Folding Real-time	SD
MSMEG_1074	Polysaccharide deacetylase	down	-3,7399	0,0431	-3,4105	1,2869
MSMEG_2404	Extracellular deoxyribonuclease	down	-4,0840	0,0026	1,36 E-19	0,1414
MSMEG_2468	L-carnitine dehydratase/bile acid-inducible	down	-2,6317	0,0142	-10,6295	2,1425
MSMEG_5644	YhgE/Pip C-terminal domain protein membrane protein	up	3,9086	0,0265	414,9433	1,3647
Rv1168c	PPE17/ PPE family protein	up	25,7054	0,0358	39,8084	0,4455
Rv2351c	Membrane-associated phospholipase C	up	10,6295	0,0072	40,6540	0,4950
Rv3545c	Cytochrome P450 125	up	10,9587	0,0133	9,5777	0,9687
MAP_1855	Mce5E	up	2,3174	0,0052	10,5561	0,3960
MAP_2895	Iron dependent repressor, DtxR family	up	2,2268	0,0015	129,7868	1,5556
MAP_3032c	3-isopropylmalate dehydrogenase	up	2,6344	0,0005	9,1261	0,2758

6. DISCUSSION

Although the mycobacterial cell wall is an important protective barrier also implicated in pathogenesis, mycobacteria such as MTB and MAP are found to survive inside the host in a CWD state also known as L-form [Almenoff et al, 1996; Chiodini et al., 1986, Wall et al., 1993; Hulten et al., 2001]. CWD forms have been suspected to play an important role in mycobacterial diseases, including TB, in which they could be responsible for the persistence of infection [Imaeda, 1984]. As a matter of fact, it has been suggested that the CWD forms could be able to avoid host phagocytosis, resist antibiotic pressure and to accumulate in the body over time [Onwuamaegbu et al., 2005]. The mycobacterial CWD forms are very difficult to identify and to culture hence the molecular basis of their pathogenesis is still unknown despite they already have been implicated in numerous human diseases such as Sarcoidosis [Almenoff et al.,1996], Amylotrophic Lateral Sclerosis [Koch et al., 2003], Kaposi's sarcoma [Cantwell, 1981] Crohn's disease [Chiodini et al., 1986, Hulten et al., 2001] and recently AIDS [Broxmeyer and Cantwell, 2008]. It has been postulated that CWD bacteria establish a sort of parasitic relationship within its host, probably through the integration with host organelles, which lead to toleration by the immune system and hence to a "silent survival" of the bacteria [Domingue, 2010; Domingue and Woody, 1997]. We believed that this mechanism was determined by modifications of the bacterial gene expression machinery and by the control of virulent determinants. Therefore, we attempted for the first time to analyze the gene expression changes of three different mycobacterial species after their *in vitro* conversion to the CWD form, a transformation which is known to occur *in vivo* inside phagocytic cells probably in response to the intracellular host lysozyme burst activity.

Despite the *in vitro* study do not completely display what happens *in vivo* inside the host, where mycobacteria simultaneously face several antimicrobial stressors, this is the first work that attempt to clarify the transcriptome of the mycobacterial CWD forms and the identification of key genes involved in their pathogenesis.

Experiments were performed with MSMEG, MTB and MAP in order to understand the gene regulation relative to the CWD state of both non-pathogenic and pathogenic

mycobacteria. Therefore, we mimed the *in vivo* CWD state by chemically treating mycobacteria.

Overall Microarray data showed that CWD_MTB differentially controls a higher number of genes (11.8%) than CWD_MSMEG (8.4%) and CWD_MAP (2.5%) suggesting either the evolution for MSMEG and MTB of a more sophisticated mechanism to respond to lysozyme activity then MAP, or a different stage of MAP in the CWD conversion. Moreover, overall data for the three mycobacteria species identified the regulation of a large percentage of genes encoding for hypothetical and conserved hypothetical proteins, a result that underline the need of assigning a function to these proteins as they might be responsible for determining the mycobacterial physiology associated to the CWD form and the distinctive capacity of pathogenic mycobacteria to survive within the host.

After the *in vitro* conversion to the CWD state, the mycobacteria started to regulates a high number of transcriptional regulators, stress responders and genes involved in lipid metabolism, membrane transport and translation. However, this regulation was quite different among the three species.

The specific response of MSMEG to surface stress, achieved by the chemical treatment, is characterized by up-regulation of two important genes, *MSMEG_5728* and *MSMEG_4373* encoding for peptidoglycan deacetylases whose function is to deacetylase peptidoglycan GlcNAc residues so that peptidoglycan cannot longer be recognized by the host lysozyme [Rae et al., 2011; Vollmer, 2008]. Modification of the peptidoglycan layer is a well-known mechanism that bacteria have evolved to resist lysozyme activity. The resistance mechanism consists in the modification of the peptidoglycan and other cell wall-linked components such as techoic acids to render them unrecognizable by the host's enzymes [Bera et al., 2007]. *In vivo* this could also be important in limiting the release of bacterial cell wall fragments at the site of infection that can be recognized by the immune cells. Peptidoglycan modifications, such as N-deacetylation and O-acetylation, have been shown to occur in many pathogenic bacteria but little is known about mycobacteria [Boneca et al., 2007; Davis and Weiser, 2011]. Here we show that MSMEG presumably utilizes this mimetic mechanism, while MTB and MAP seem to behave differently as we did not observe over-expression of peptidoglycan deacetylases-like genes in those species.

We can speculate that in MTB the response to the host lysozyme might be linked to the presence of some PE/ PPE proteins that could create a net of proteins, probably so tightly associated within the cell wall to hide proteolytic cleavage sites. Another explanation could be that MTB and MAP let the “cell wall striping” happens, leading *in vivo* to a massive immune cells recruitment that these mycobacteria use as host niche. Hence, the CWD phase could be the result of a specific strategy evolved not only for stress tolerance such as nutrient deprivation and low iron availability (the thick cell wall is in fact responsible for the restricted nutrient up-take capability) but also for survival in infected hosts. This hypothesis is also supported by the fact that MTB has evolved a peculiar mechanism to stimulate a massive immune response by the glycosylation of the peptidoglycan layer [Davis and Weiser 2011] which on the contrary of what seems to happen for MSMEG, makes the mycobacteria “more visible”. This initial step could be then followed by the conversion to an altered state which either could mimics host cells organelle or could integrate within them to create bacterial host cell antigen complexes that lead to immunopatologic consequences [Domingue, 2010].

In the CWD state MSMEG over-expresses some transcriptional regulators and among them some were involved in the control of the intracellular iron concentration such as *MSMEG_4487* (ferric up-take regulation protein).

Within the virulence class, we found the regulation of few genes compared to CWD_MTB or CWD_MAP, a result that underline the different behavior between the pathogens MTB and MAP and the saprophytic species MSMEG. The up-regulated gene was *MSMEG_0082* belonging to the ESX1 system whereas down-regulated genes were *MSMEG_2965,* involved in (p)ppGpp synthesis and hydrolysis, *MSMEG_3932 hspX* (Alpha-crystallin,) and *MSMEG_4258* (*trpD*). The last two were also down-expressed in MTB. In general, the specific response of CWD_MSMEG is mainly characterized by the induction of self-defense mechanisms and by the differential regulation, compared to the parent strain, of few viral determinants.

Data obtained for CWD_MTB revealed that the conversion to the CWD form is a finely controlled process. First of all, we observed a distinctive pattern of expression for some important transcriptional regulators such as *Rv1221 sigE* and *Rv2710 sigB*. SigE regulates expression of several determinants involved in translation, transcriptional control, and oxidative stress response. It has been demonstrated that *Rv1221 sigE* is

usually induced after exposure to different stresses such as heat-shock and detergent surface. Moreover, it is required for the growth and survival of MTB inside macrophages as well as for the full expression of *sigB* [Manganelli et al., 2001] which was also up-regulated in our experiments. In respect to this, *sigE* has been shown to be over-expressed after invasion of human macrophages [Jensen-Cain and Quinn, 2001], where the mycobacterial cell wall could be attacked by host lysozyme.
Among the transcriptional regulator class, interestingly we observed the over-expression of *Rv3862c whib6* and *Rv3260c whiB2,* the last one is the closest MTB homologue of the Streptomyces *whiB* gene [Soliveri et al., 2000; Mulder at al., 1999]. This gene is required for differentiation and sporulation in sporulating Actinobacteria [Kormanec and Homerova, 1993; Davis and Chater, 1992]. The up-regulation of this gene in non-sporulating Actinomycetes such as MTB, after its conversion to the CWD form, is very intriguing and deserves more study especially considering that MTB has seven *whiB*-like genes which are supposed to be involved in the dormancy state [Alam et al., 2009]. This gene has been also shown to be induced after exposure to surface stress, oxidative stress and heat shock [Geiman et al., 2006].
CWD_MTB regulates a high proportion of cell envelope genes belonging to the PE/PPE families which are supposed to play an important role in the virulence of mycobacteria. Approximately 10% of the MTB genome encodes for these proteins. [Cole at al.,1998]. Although their role is still under scrutiny, it is widely accepted that they are responsible for generating antigenic variation [Banu et al., 2002] therefore they play an important role in pathogenesis. Moreover, they also have been implicated in the transport of antimicrobials compounds across the outer membrane of MTB [Danilchanka et al., 2008]. In particular, in the MTB_CWD state, we observed the up-regulation of *Rv1168c* PPE17 which has been successfully used to diagnose extra-pulmonary and smear negative pulmonary TB cases, with higher sensitivity than Esat6 or Hsp60 immunodominant antigens. The protein PPE17 has also been demonstrated to act as a potent T cell antigen; stimulating high levels of IFN-γ in PBMCs obtained from TB patients but not from BCG immunized controls. Khan and colleagues [2008] postulated that PPE17 could play an important role in a certain stages of mycobacterial infection and intracellular survival. Furthermore, this protein is predominantly expressed under conditions that mimic the *in vivo* phagosomal environment and in a non-replicative state

[Schnappinger et al., 2003; Muttucumaru et al., 2004]. Hence, the over-expression of this gene in the MTB_CWD form reveals that the intra-phagosomal non-replicative state could be the CWD state and suggest PPE17 as a potential target for the identification of the CWD forms in TB patients. Among PE/PPE over-expressed genes, we found *Rv3746c* PE34 and *Rv1169c* PE11 that are demonstrated to be also up-regulated in human lung granulomas where the mycobacteria survive in a dormant state [Rachman et al. 2006; Saunders and Britton, 2007].

It is also important to underline that the regulation of some PE/PPE genes came along with the regulation of ESX genes, this result support the finding that the PE/PPE proteins are intimately associated with the ESX systems as demonstrated for example for the ESX-5 apparatus [Abdallah et al., 2009].

CWD_MTB seems also to modulate the expression of some PE_PGRS family genes such as the cell wall-associated/secretory PE_PGRS 11 (*Rv0754*) protein which has been demonstrated to induce *in vivo* maturation and activation of dendritic cells (DCs) and therefore to potentially contribute to the initiation of innate immune responses during TB infection and to regulate the clinical course of the disease [Bansal et al., 2010].

CWD_MTB also showed the regulation of many genes involved in lipid metabolism. Lipids are part of both the cell wall and cytoplasmic membranes and they represent an important source of energy for the bacteria. Recent demonstration of the close interrelationship between lipids and mycobacterial pathogenicity open new insight into the lipogenesis and lipolysis pathways of these organisms [Schorey and Sweet, 2008]. Moreover, the MTB genome confirms the importance of lipid metabolism in this human pathogen [Cole et al., 1988] as a very large proportion of the coding capacity is devoted to the production of enzymes involved in lipid metabolism.

In CWD_MTB we observed the over-expression of six genes involved in lipid degradation, four of them were *fadD34*, *fadD29*, *fadD19* and *fadD14* (putative fatty acyl-coenzyme A (CoA) synthases) needed to convert fatty acids in fatty acetyl-CoA. MTB *H37Rv* has 36 *fadD* annotated genes that encodes for a vast *repertoire* of enzymes used for a diverse range of fatty acids that might be degraded by mycobacteria as carbon and energy sources. Moreover, Fad proteins are of particular interest because some of them have been associated to virulence in mice such as FadD19 (up-regulated in our

experiments), FadD28, FadD33 and FadD5 [Schnappinger et al., 2003; Dunphy et al., 2010].

Microarray data also showed the over-expression of the fatty acid metabolism transcriptional regulator *Rv0586.* This auto-regulatory repressor controls the transcription of the whole *fad* regulon which regulates the expression of key enzymes required for both fatty acid β-oxidation and fatty acid biosynthesis. This control is very important for the manipulation of saturated and unsaturated fatty acids relative levels hence the modifications of membrane fluidity.

Data suggested that in the CWD state, MTB uses fatty acids as the main carbon source. This is also supported by the up-regulation of the *icl* gene which encodes for a protein that allows bacteria to grow on acetate or fatty acids broken down to acetyl CoA. The *icl* gene encodes for a key gene of the glyoxylate shunt which provides a source of carbon that can enter the Krebs cycle. Interestingly, the IcL enzyme and the activation of the glyoxylate shunt have been shown to be essential for the pathogenesis of MTB and for its intracellular survival in mice [Munoz-Elias and McKinney, 2005; McKinney et al., 2000].

The response to lysozyme stress is characterized in MTB by the down-regulation of genes involved in mycolic acid biosynthesis such as *Rv2244 acpM*, *Rv1483 fabG1, Rv2245 kasA* and *Rv2246 kasB*. It has been suggested that KasA, functions in the initial elongation of the mycolic acid chain, whereas KasB elongates the chain to full length [Swanson et al., 2009]. Moreover *kasB* mutant strains synthesize mycolates with shorter chain lengths, they assume an altered colony morphology and the loss of the classic serpentine growth (cording), they lose acid-fastness but the most important effect of *kasB* deletion was the ability of the mutant strain to persist in infected immunocompetent mice for up to 600 days without causing disease or mortality [Bhatt et al., 2007]. Hence, these genes have been considered as key factors in subclinical latent tuberculosis and we can speculate that they also are key factors in the CWD_MTB state.

The-down regulation of mycolic acids biosynthesis genes is in some measure balanced by the up-regulation of other genes involved in lipid biosynthesis and in the polykedites and terpenoids pathways [Mahrous et al., 2008]. This regulation is probably aimed to the reinforcement of the bacterial cell structure partially deprived of the wall.

Interestingly, we also found that CWD-MTB up-regulate a high number of genes encoding for transposases. These proteins could either mediate genomic rearrangements that *in vivo* may drive the bacterial adaptation to the hostile environment or regulate gene expression, hence concurring to the CWD strategy.

In this work special attention has been paid on virulence determinants. We searched for 109 known virulence genes in CWD_MSMEG, 143 in CWD_MTB and 126 in CWD_MAP. CWD_MTB differentially regulates many important virulence genes and interestingly some of them have already been implicated in mycobacterial persistence. Among up-regulated genes we found *Rv0467 icl* which as described before, encodes for a protein that allows bacteria to grow on acetate or fatty acids as sole carbon sources since the glyoxalate shunt provides a source of carbon that can enter the Krebs cycle. It has been shown that *Rv0467* mRNA level increases when MTB infects human macrophages and that facilitate bacterial persistence in mice [McKinney et al., 2000], moreover it has been implicated in mycobacterial dormancy [Honer Zu Bentrup et al., 1999] where indeed the mycobacteria might assume the CWD form. Another up-regulated gene was *Rv0981 mprA*, which encodes for a two-system component response regulator whose expression is required for entrance into cells and maintenance of persistent infection. Inactivation of this regulator affects MTB *H37Rv* growth *in vivo* [Zahrt and Deretic, 2001].

It is relevant to note the over-expression of a virulence gene named *Rv2351c plcA* that has no homologous in MSMEG, MAP and in other mycobacterial species. This gene encodes for a phospholipase involved in phthiocerol dimycocerosate (PDIM) and phenolic glycolipid (PGL) biosynthesis and transport. It is one of the four genes (*plcA*, *plcB*, *plcC* and *plcD*) present in MTB encoding for a putative phospholipases C that MTB uses to release fatty acids from host phospholipids through the β-oxidation cycle and the glyoxalate shunt. Hence, this gene is useful in the CWD state to exploit host lipids for energy purposes. In this regards, it is important to underline that the up-regulation of these genes is associated with the up-regulation of the *icl* gene which is essential for the production of a glycosylate shunt enzyme that follow the phospholipases C activity. Both genes are up-regulated during macrophage infection and disruption of these genes impaired the ability of MTB to persist inside the host, therefore has been suggested that both genes are involved in the same pathway [Raynaud et al., 2002].

Other virulence genes up-regulated in the CWD state were *Rv2428 ahpC* an alkyl hydroperoxide reductase useful to detoxify organic hydroxyperoxides and important for the mycobacterial antioxidant defense [Dosanjh et al., 2005]; *Rv2462c* encoding for a protein involved in phenolic glycolipid biosynthesis and transport, which is an important virulence factor. [Pérez et al., 2004].

The acquisition of iron is strongly correlated to pathogenesis; hence CWD_MTB over-expresses the gene *Rv2386 mbtI* which encodes the first enzyme in the mycobactin biosynthesis, a siderophore that mycobacteria use to acquire iron from the extracellular environment. Among down-expressed genes we found the regulation of *Rv2031c hspX* and *Rv2192c trpD* which are also down-regulated in CWD_MSMEG (*MSMEG_3932* and *MSMEG_4258*). The protein HspX, also known as Acr, is a major MTB antigen recognized by the sera of a high proportion of TB patients and it is induced under anoxic conditions and during stationary phase of growth [Yuan et al., 1998], while *Rv2192c trpD* is involved in the tryptophan biosynthesis and although this pathway is usually non-essential for many pathogens, *trpD* mutants in MTB fails to cause disease in mice [Lee et al., 2006]. Although HspX has been pointed as a target antigen for latent TB infection it seems to not be strictly required in this phase of the CWD state. However considering that its induction is prevalent under anoxic conditions and during stationary-phase of growth it could be expressed *in vivo* under additional host's pressures or in a different phase of MTB growth.

Other down-regulated genes were, *Rv2962c* involved in the PDIM and PGL biosynthesis and transport, *Rv1736c narX* necessary during anaerobic respiration and *Rv2939 papA5* also involved in PDIM and PGL biosynthesis and transport which have no homologous in MSMEG and MAP.

In our experiments we also found over-expression of genes implicated in antibiotic resistance such as *Rv0191* which encodes for an integral membrane protein probably involved in antibiotic transport and some genes belonging to the two-component signal transduction systems that enable bacteria to sense, respond, and adapt to a wide variety of stimuli including antibiotics. Few genes were down-expressed such as *Rv1010 ksgA*, that when inactivated confers resistance to the aminoglycoside antibiotic kasugamycin [Duffin and Seifert, 2009; Ochi et al., 2009].

In the pathogenic species CWD_MAP we observed the differential regulation of a small number of genes (49 up/ 61 down) compared to CWD_MTB and CWD_MSMEG and most of them encode for hypothetical proteins. This result could be explained by hypothesizing either a lack of a sophisticated mechanism to respond surface stress or the establishment of a different phase in the CWD conversion compared to those of MTB and MSMEG. In this respect, a temporal microarray analysis on a range of time points aimed to display earlier or later stages of CWD_ MAP would be informative.

CWD_MAP regulates the intracellular iron concentration through the over-expression of *MAP_2895* which encodes for an iron dependent repressor of the DtxR family that represses genes for iron acquisition while activates iron storage genes. However, we also observed the up-regulation of *MAP_2177c mbtB* that belongs to the *mbtA-J* operon, which in MTB encodes enzymes involved in the synthesis of the siderophores mycobactin and carboxymycobactin. The function of siderophores is to chelate insoluble ferric salts (the most common form present in the environment) and render them more soluble and ready for the transport into the cell. MbtB in particular catalyzes the formation of an amide bond between salicylate and serine in mycobactin synthesis. It is well known that MAP is mycobactin J dependent in laboratory-based culture systems probably due to a polymorphism identified in the *mbt* operon that potentially explains the *in vitro* mycobactin dependence of MAP [Semret at al., 2004], however MAP might have evolved a specific but still unknown way to control iron acquisition *in vivo*. The *mbtB* gene is up-regulated inside bovine macrophages where the level of iron is low [Zhu et al., 2008] and where MAP cells might assume the CWD form.

As in CWD_MTB and CWD_MSMEG, CWD_MAP regulates some virulence genes and in particular we observed the up-regulation of *mce* genes such as *MAP_1855 mce5E*, *MAP_2192 mce8D* and *MAP_3606 mce1C* that have been shown to be responsible for the entry and persistence of mycobacteria inside the host [Flesselles et al., 1999; Haile et al., 2002]. Another important up-regulated gene is *MAP_1236c drrC*, homologous to *Rv2938* which is involved in glycolipid biosynthesis and transport and it has been found to be highly up-regulated in MAP after 48h and 120h post-infection of cows monocytes derived macrophages [Zhu et al., 2008]. *MAP_1236c drrC* could also functions as a drug efflux protein responsible for the extrusion of antibiotics.

Glycan biosynthesis is repressed as shown for example by the down-expression of *MAP_3163* which encodes for a hypothetical protein probably involved in glycan and metabolism.
Lastly, we can underline that in this species, more than in the other two, hypothetical proteins needs to be studied in order to reveal their function and their role in the CWD state.

7. CONCLUSIONS

Expression profiles of pathogenic and non-pathogenic mycobacteria after *in vitro* conversion to the CWD state, which is widely known to occur *in vivo* inside the host, yield *a portrait* of genes implicated in the ability of pathogenic species to adapt and silently persist within host environments. Results provide new insights in the molecular strategy beside the CWD state. Partially characterized or already known identified genes, can now be associated to the CWD state and might be useful in the development of CWD specific targets for clinical investigations and drug design.
Findings could have inferences in understanding several inflammatory and autoimmune diseases such as Crohn's disease where CWD_MAP have been found or human TB where coexistence of classical walled bacteria and CWD forms is considered a natural phenomenon ensuring the adaptive strategy of the mycobacteria.
One goal of this thesis is also to renew the interest of researcher on mycobacterial CWD forms, an important stage in the mycobacterial life-cycle that we should not neglect any more. Future studies should be finalized to the development of innovative and easy to use diagnostics methods in order to increase the CWD forms detection from clinical specimens and potentially allowing in the future deep studies of *in vivo* isolated forms.

8. BIBLIOGRAPHY

1. Abdallah, A.M., Verboom, T., Weerdenburg, E.M., Gey van Pittius, N.C., Mahasha, P.W., Jimenez, C., Parra, M., Cadieux, N., Brennan, M.J., Appelmelk, B.J., *et al.* (2009). PPE and PE_PGRS proteins of Mycobacterium marinum are transported via the type VII secretion system ESX-5. Mol Microbiol *73*, 329-340.

2. Alam, M.S., Garg, S.K., and Agrawal, P. (2009). Studies on structural and functional divergence among seven WhiB proteins of Mycobacterium tuberculosis H37Rv. FEBS J *276*, 76-93.

3. Alexander KA, Laver PN, Michel AL, Williams M, van Helden PD, Warren RM, Gey van Pittius NC. (2010). Novel Mycobacterium tuberculosis complex pathogen, M. mungi.Emerg Infect Dis. Aug;16(8):1296-9.

4. Almenoff, P.L., Johnson, A., Lesser, M., and Mattman, L.H. (1996). Growth of acid fast L forms from the blood of patients with sarcoidosis. Thorax *51*, 530-533.

5. Alteri CJ, Xicohtencatl-Cortes J, Hess S, Caballero-Olin G, Giron JA, et al. (2007) Mycobacterium tuberculosis produces pili during human infection. Proc Natl Acad Sci U S A 104: 5145-5150.

6. Bannantine JP, Rosu V, Zanetti S, Rocca S, Ahmed N, Sechi LA. Antigenic profiles of recombinant proteins from Mycobacterium avium subsp. paratuberculosis in sheep with Johne's disease. Vet Immunol Immunopathol. 2008 Mar 15;122(1-2):116-25.

7. Bansal, K., Elluru, S.R., Narayana, Y., Chaturvedi, R., Patil, S.A., Kaveri, S.V., Bayry, J., and Balaji, K.N. (2010). PE_PGRS antigens of Mycobacterium tuberculosis induce maturation and activation of human dendritic cells. J Immunol *184*, 3495-3504.

8. Banu, S., Honore, N., Saint-Joanis, B., Philpott, D., Prevost, M.C., and Cole, S.T. (2002). Are the PE-PGRS proteins of Mycobacterium tuberculosis variable surface antigens? Mol Microbiol *44*, 9-19.

9. Behr MA, Kapur V (2008) The evidence for Mycobacterium paratuberculosis in Crohn's disease. Curr Opin Gastroenterol 24: 17-21.

10. Bera, A., Biswas, R., Herbert, S., Kulauzovic, E., Weidenmaier, C., Peschel, A., and Gotz, F. (2007). Influence of wall teichoic acid on lysozyme resistance in Staphylococcus aureus. J Bacteriol *189*, 280-283.

11. Bhatt, A., Fujiwara, N., Bhatt, K., Gurcha, S.S., Kremer, L., Chen, B., Chan, J., Porcelli, S.A., Kobayashi, K., Besra, G.S., *et al.* (2007). Deletion of kasB in

Mycobacterium tuberculosis causes loss of acid-fastness and subclinical latent tuberculosis in immunocompetent mice. Proc Natl Acad Sci U S A *104*, 5157-5162.

12. Biet F, Boschiroli ML, Thorel MF, Guilloteau LA (2005) Zoonotic aspects of Mycobacterium bovis and Mycobacterium avium-intracellulare complex (MAC). Vet Res 36: 411-436.

13. Bisset KA, Bartlett R. (1978). The isolation and characters of L-forms and reversions of Bacillus licheniformis var endoparasiticus (Benedek) associated with the erythrocytes of clinically normal persons. J Med Microbiol 11: 335 – 349

14. Boneca, I.G., Dussurget, O., Cabanes, D., Nahori, M.A., Sousa, S., Lecuit, M., Psylinakis, E., Bouriotis, V., Hugot, J.P., Giovannini, M.*, et al.* (2007). A critical role for peptidoglycan N-deacetylation in Listeria evasion from the host innate immune system. Proc Natl Acad Sci U S A *104*, 997-1002.

15. Brennan PJ, Crick DC (2007) The cell-wall core of Mycobacterium tuberculosis in the context of drug discovery. Curr Top Med Chem 7: 475-488.

16. Brennan, M.J., and Delogu, G. (2002). The PE multigene family: a 'molecular mantra' for mycobacteria. Trends Microbiol *10*, 246-249.

17. Brooks, P.C., Dawson, L.F., Rand, L., and Davis, E.O. (2006). The mycobacterium-specific gene Rv2719c is DNA damage inducible independently of RecA. J Bacteriol *188*, 6034-6038.

18. Broxmeyer, L., and Cantwell, A. (2008). AIDS: "it's the bacteria, stupid!". Med Hypotheses *71*, 741-748.

19. Brugere-Picoux J (1998) Crohn disease: uncertainty about the etiologic role of Mycobacterium paratuberculosis. Bull Acad Natl Med 182: 823-829; discussion 830-822.

20. Camus, J.C., Moszer, I., Jones, L.M., Pryor, M.J., Cole, S.T. (2002) TubercuList: a regularly updated database dedicated to the Mycobacterium tuberculosis genome. In: Danchin A editors. Genomics of GC-rich Gram-positive bacteria. Wymondham: Caister Academic Press, 159–176

21. Cantwell, A.R., Jr., Rowe, L., and Kelso, D.W. (1980). Nodular scleroderma and pleomorphic acid-fast bacteria. Arch Dermatol *116*, 1283-1290.

22. Cantwell, A.R., Jr., and Lawson, J.W. (1981). Necroscopic findings of pleomorphic, variably acid-fast bacteria in a fatal case of Kaposi's sarcoma. J Dermatol Surg Oncol *7*, 923-930.

23. Chacon, O., Bermudez, L.E., and Barletta, R.G. (2004). Johne's disease, inflammatory bowel disease, and Mycobacterium paratuberculosis. Annu Rev Microbiol *58*, 329-363.

24. Chen, L., Yang, J., Yu, J., Yao, Z., Sun, L., Shen, Y., and Jin, Q. (2005). VFDB: a reference database for bacterial virulence factors. Nucleic Acids Res *33*, D325-328.

25. Chiodini RJ (1989) Crohn's disease and the mycobacterioses: a review and comparison of two disease entities. Clin Microbiol Rev 2: 90-117

26. Chiodini RJ, Rossiter CA (1996) Paratuberculosis: a potential zoonosis? Vet Clin North Am Food Anim Pract 12: 457-467

27. Chiodini, R.J., Van Kruiningen, H.J., Merkal, R.S., Thayer, W.R., Jr., and Coutu, J.A. (1984). Characteristics of an unclassified Mycobacterium species isolated from patients with Crohn's disease. J Clin Microbiol *20*, 966-971.

28. Chiodini, R.J., Van Kruiningen, H.J., Thayer, W.R., and Coutu, J.A. (1986). Spheroplastic phase of mycobacteria isolated from patients with Crohn's disease. J Clin Microbiol *24*, 357-363.

29. Cole, S.T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S.V., Eiglmeier, K., Gas, S., Barry, C.E., 3rd*, et al.* (1998). Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. Nature *393*, 537-544.

30. Collins MT (1997) Mycobacterium paratuberculosis: a potential food-borne pathogen? J Dairy Sci 80: 3445-3448.

31. Cooper AM (2009) T cells in mycobacterial infection and disease. Curr Opin Immunol 21: 378-384 (a)

32. Cooper AM (2009) Cell-mediated immune responses in tuberculosis. Annu Rev Immunol 27: 393-422 (b)

33. Cooper AM (2009) Host versus pathogen: two sides of the same challenge in the TB world. Eur J Immunol 39: 632-633 (c)

34. Cossu, D., Cocco, E., Paccagnini, D., Masala, S., Ahmed, N., Frau, J., Marrosu, M.G., and Sechi, L.A. (2011). Association of Mycobacterium avium subsp. paratuberculosis with multiple sclerosis in Sardinian patients. PLoS One *6*, e18482.

35. Dai YH, Lin TF, Huang GL. (1996). A serial study on mycobacteria L-forms. Zhongguo Fanglao Zazhi; (Suppl): 45–14

36. Danilchanka, O., Mailaender, C., and Niederweis, M. (2008). Identification of a novel multidrug efflux pump of Mycobacterium tuberculosis. Antimicrob Agents Chemother *52*, 2503-2511.

37. Davis, N.K., and Chater, K.F. (1992). The Streptomyces coelicolor whiB gene encodes a small transcription factor-like protein dispensable for growth but essential for sporulation. Mol Gen Genet *232*, 351-358.

38. Davis, K.M., and Weiser, J.N. (2011). Modifications to the peptidoglycan backbone help bacteria to establish infection. Infect Immun *79*, 562-570.

39. Dheda K, Booth H, Huggett JF, Johnson MA, Zumla A, et al. (2005) Lung remodeling in pulmonary tuberculosis. J Infect Dis 192: 1201-1209.

40. Di Perri G, Micciolo R, Vento S, Cruciani M, Marocco S, et al. (1993) Risk of reactivation of tuberculosis in the course of human immunodeficiency virus infection. Eur J Med 2: 264-268.

41. Domingue, G.J., and Schlegel, J.U. (1977). Novel bacterial structures in human blood. II. Bacterial variants as etiologic agents in idiopathic hematuria. Trans Am Assoc Genitourin Surg *69*, 61-64 (a)

42. Domingue, G.J., and Schlegel, J.U. (1977). Novel bacterial structures in human blood: cultural isolation. Infect Immun *15*, 621-627 (b)

43. Domingue, G.J., Sr., and Woody, H.B. (1997). Bacterial persistence and expression of disease. Clin Microbiol Rev *10*, 320-344.

44. Domingue, G.J. (2010). Demystifying pleomorphic forms in persistence and expression of disease: Are they bacteria, and is peptidoglycan the solution? Discov Med *10*, 234-246.

45. Dorozhkova, I.R., Badleeva, M.V., Skotnikova, O.I., Nosova, E., Kupavtseva, E.A., Borisov, S.E., and Moroz, A.M. (2005). The composition and drug sensitivity of a mycobacterial population in patients with suspected tuberculosis. Probl Tuberk Bolezn Legk, 36-38.

46. Dosanjh, N.S., Rawat, M., Chung, J.H., and Av-Gay, Y. (2005). Thiol specific oxidative stress response in Mycobacteria. FEMS Microbiol Lett *249*, 87-94.

47. Duffin, P.M., and Seifert, H.S. (2009). ksgA mutations confer resistance to kasugamycin in Neisseria gonorrhoeae. Int J Antimicrob Agents *33*, 321-327.

48. Dunphy, K.Y., Senaratne, R.H., Masuzawa, M., Kendall, L.V., and Riley, L.W. (2010). Attenuation of Mycobacterium tuberculosis functionally disrupted in a fatty acyl-coenzyme A synthetase gene fadD5. J Infect Dis *201*, 1232-1239.

49. El-Zaatari, F.A., Osato, M.S., and Graham, D.Y. (2001). Etiology of Crohn's disease: the role of Mycobacterium avium paratuberculosis. Trends Mol Med *7*, 247-252.

50. Engstrand L (1995) Mycobacterium paratuberculosis and Crohn's disease. Scand J Infect Dis Suppl 98: 27-29.

51. Falkinham JO, 3rd (2010) Hospital water filters as a source of Mycobacterium avium complex. J Med Microbiol 59: 1198-1202.

52. Fleischmann,R.D., Dodson,R.J., Haft,D.H., Merkel,J.S., Nelson,W.C and Fraser,C.M (2006) NCBI Reference Sequence: NC_008596.1

53. Flesselles, B., Anand, N.N., Remani, J., Loosmore, S.M., and Klein, M.H. (1999). Disruption of the mycobacterial cell entry gene of Mycobacterium bovis BCG results in a mutant that exhibits a reduced invasiveness for epithelial cells. FEMS Microbiol Lett *177*, 237-242.

54. Flynn JL, Chan J (2001) Immunology of tuberculosis. Annu Rev Immunol 19: 93-129.

55. Geboes K, Bossaert H (1977) Reactivation of tuberculosis in old age. J Am Geriatr Soc 25: 318-319.

56. Geiman, D. E., Raghunand, T. R., Agarwal, N. & Bishai, W. R. (2006). Differential gene expression in response to exposure to antimycobacterial agents and other stress conditions among seven Mycobacterium tuberculosis whiB-like genes. Antimicrob Agents Chemother 50, 2836–2841.

57. Gerasimov, V.N., Biketov, S.F., Golov, E.A., Potapov, V.D., Bakhteeva, I.V., and Nikolaeva, O.G. (2003). [Specific features of minute structure of Mycobacteria cultured in murine macrophages]. Probl Tuberk Bolezn Legk, 52-58.

58. Ghosh J, Larsson P, Singh B, Pettersson BM, Islam NM, et al. (2009) Sporulation in mycobacteria. Proc Natl Acad Sci U S A 106: 10781-10786.

59. Gill, C.O., Saucier, L., and Meadus, W.J. (2011). Mycobacterium avium subsp. paratuberculosis in dairy products, meat, and drinking water. J Food Prot *74*, 480-499.

60. Greenstein RJ (2003) Is Crohn's disease caused by a mycobacterium? Comparisons with leprosy, tuberculosis, and Johne's disease. Lancet Infect Dis 3: 507-514.

61. Haile, Y., Caugant, D.A., Bjune, G., and Wiker, H.G. (2002). Mycobacterium tuberculosis mammalian cell entry operon (mce) homologs in Mycobacterium other than tuberculosis (MOTT). FEMS Immunol Med Microbiol *33*, 125-132.

62. Hampson SJ, McFadden JJ, Hermon-Taylor J (1988) Mycobacteria and Crohn's disease. Gut 29: 1017-1019.

63. Harwick, H.J., Barajas, L., Montgomerie, J.Z., Kalmanson, G.M., and Guze, L.B. (1972). Phagocytosis of microbial L-forms. Infect Immun *5*, 976-981.

64. Harwick, H.J., Kalmanson, G.M., and Guze, L.B. (1977). Chemotactic activity of L-forms and mycoplasma. Aust J Exp Biol Med Sci *55*, 431-434.

65. Honer Zu Bentrup, K., Miczak, A., Swenson, D.L., and Russell, D.G. (1999). Characterization of activity and expression of isocitrate lyase in Mycobacterium avium and Mycobacterium tuberculosis. J Bacteriol *181*, 7161-7167.

66. Hulten, K., El-Zimaity, H.M., Karttunen, T.J., Almashhrawi, A., Schwartz, M.R., Graham, D.Y., and El-Zaatari, F.A. (2001). Detection of Mycobacterium avium subspecies paratuberculosis in Crohn's diseased tissues by in situ hybridization. Am J Gastroenterol *96*, 1529-1535.

67. Hulten, K., Karttunen, T.J., El-Zimaity, H.M., Naser, S.A., Almashhrawi, A., Graham, D.Y., and El-Zaatari, F.A. (2000). In situ hybridization method for studies of cell wall deficient M. paratuberculosis in tissue samples. Vet Microbiol *77*, 513-518 (a).

68. Hulten, K., Karttunen, T.J., El-Zimaity, H.M., Naser, S.A., Collins, M.T., Graham, D.Y., and El-Zaatari, F.A. (2000). Identification of cell wall deficient forms of M. avium subsp. paratuberculosis in paraffin embedded tissues from animals with Johne's disease by in situ hybridization. J Microbiol Methods *42*, 185-195 (b).

69. Imaeda, T. 1984, Cell wal; deficient forms p. 671-680. In G. P. Kubica and L. G. Wayne (ed.), The mycobacteria. MarcelDekker, Inc., New York.

70. Jensen-Cain, D.M., and Quinn, F.D. (2001). Differential expression of sigE by Mycobacterium tuberculosis during intracellular growth. Microb Pathog *30*, 271-278.

71. Kanehisa, M., and Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res *28*, 27-30.

72. Kaplan G, Post FA, Moreira AL, Wainwright H, Kreiswirth BN, et al. (2003) Mycobacterium tuberculosis growth at the cavity surface: a microenvironment with failed immunity. Infect Immun 71: 7099-7108.

73. Khan, N., Alam, K., Nair, S., Valluri, V.L., Murthy, K.J., and Mukhopadhyay, S. (2008). Association of strong immune responses to PPE protein Rv1168c with active tuberculosis. Clin Vaccine Immunol *15*, 974-980.

74. Koch, A.L. (2003). Cell wall-deficient (CWD) bacterial pathogens: could amylotrophic lateral sclerosis (ALS) be due to one? Crit Rev Microbiol *29*, 215-221.

75. Kolattukudy, P.E., Fernandes, N.D., Azad, A.K., Fitzmaurice, A.M., and Sirakova, T.D. (1997). Biochemistry and molecular genetics of cell-wall lipid biosynthesis in mycobacteria. Mol Microbiol *24*, 263-270.

76. Kormanec, J., and Homerova, D. (1993). Streptomyces aureofaciens whiB gene encoding putative transcription factor essential for differentiation. Nucleic Acids Res *21*, 2512.

77. Lee, C.E., Goodfellow, C., Javid-Majd, F., Baker, E.N., and Shaun Lott, J. (2006). The crystal structure of TrpD, a metabolic enzyme essential for lung colonization by Mycobacterium tuberculosis, in complex with its substrate phosphoribosylpyrophosphate. J Mol Biol *355*, 784-797.

78. Li L, Bannantine JP, Zhang Q, Amonsin A, May BJ, et al. (2005) The complete genome sequence of Mycobacterium avium subspecies paratuberculosis. Proc Natl Acad Sci U S A 102: 12344-12349.

79. Liu Y, Lin TF. (1996). Studies on morphology and electronmicrographic analysis of *M. tuberculosis* filamentous L-forms. Chinese J Microbiol Immunobiol; 16 (Suppl 2): 49.

80. Lund, B.M., Gould, G.W., and Rampling, A.M. (2002). Pasteurization of milk and the heat resistance of Mycobacterium avium subsp. paratuberculosis: a critical review of the data. Int J Food Microbiol *77*, 135-145.

81. Ma XL, Huang GL, Lin TF. (1989). Isolation of M. tuberculosis L-forms from clinical specimens. Acta Acad Med Bengbu;14 (4): 229–31

82. Madoff, S., Burke, M.E., and Dienes, L. (1967). Induction and identification of L-forms of bacteria. Ann N Y Acad Sci *143*, 755-759.

83. Mahrous, E.A., Lee, R.B., Lee, R.E. (2008). A rapid approach to lipid profiling of mycobacteria using 2D HSQC NMR maps. J Lipid Res. Feb;49(2):455-63. Epub 2007 Nov 2.

84. Manganelli, R., Voskuil, M.I., Schoolnik, G.K., and Smith, I. (2001) The Mycobacterium tuberculosis ECF sigma factor SigE: role in global gene expression and survival in macrophages. Mol Microbiol 41: 423–437.

85. Markova. (2009). Hidden face of tuberculosis. Bioscience Hypotheses 2, 441e442

86. Markova N., Michailova L., Jourdanova M., Kussovski V., Valcheva V., Mokrousov I., Radoucheva T. (2008). Exhibition of persistent and drug-tolerant L-form habit of *Mycobacterium tuberculosis* during infection in rats. Cent. Eur. J. Biol. • 3(4) • 2008 • 407–416

87. Mattman, L.H., Tunstall, L.H., and Rossmoore, H.W. (1961). Induction and characteristics of staphylococcal L forms. Can J Microbiol *7*, 705-713 (a).

88. Mattman L.H., Mycobacterium tuberculosis and atypicals, In: Mattman L.H., (2001). Cell wall deficient forms. Stealth Pathogens, 3rd ed., CRC Press, Inc., Boca Raton, FL, USA, (b).

89. McKinney, J.D., Honer zu Bentrup, K., Munoz-Elias, E.J., Miczak, A., Chen, B., Chan, W.T., Swenson, D., Sacchettini, J.C., Jacobs, W.R., Jr., and Russell, D.G. (2000). Persistence of Mycobacterium tuberculosis in macrophages and mice requires the glyoxylate shunt enzyme isocitrate lyase. Nature *406*, 735-738.

90. Michailova, L., Kussovsky, V., Radoucheva, T., Jordanova, M., and Markova, N. (2007). Persistence of Staphylococcus aureus L-form during experimental lung infection in rats. FEMS Microbiol Lett *268*, 88-97.

91. Michailova, L., Kussovski, V., Radoucheva, T., Jordanova, M., Berger, W., Rinder, H., and Markova, N. (2005). Morphological variability and cell-wall deficiency in Mycobacterium tuberculosis 'heteroresistant' strains. Int J Tuberc Lung Dis *9*, 907-914.

92. Mulder, N.J., Zappe, H., and Steyn, L.M. (1999). Characterization of a Mycobacterium tuberculosis homologue of the Streptomyces coelicolor whiB gene. Tuber Lung Dis *79*, 299-308.

93. Munoz-Elias, E.J., and McKinney, J.D. (2005). Mycobacterium tuberculosis isocitrate lyases 1 and 2 are jointly required for in vivo growth and virulence. Nat Med *11*, 638-644.

94. Muttucumaru, D.G., Roberts, G., Hinds, J., Stabler, R.A., and Parish, T. (2004). Gene expression profile of Mycobacterium tuberculosis in a non-replicating state. Tuberculosis (Edinb) *84*, 239-246.

95. Naser S.A., Ghobrial G., Romero C., Valentine J.F. (2004) Culture of Mycobacterium avium subspecies paratuberculosis from the blood of patients with Crohn's disease. Lancet 364 (9439), 1039-44

96. Naser, S.A., McCarthy, C.M., Smith, G.B., and Tupponce, A.K. (1993). Low temperature protocol for efficient transformation of Mycobacterium smegmatis spheroplasts. Curr Microbiol *27*, 153-156.

97. Newton JA, Jr., Weiss PJ, Bowler WA, Oldfield EC, 3rd (1993) Soft-tissue infection due to Mycobacterium smegmatis: report of two cases. Clin Infect Dis 16: 531-533.

98. Ochi, K., Kim, J.Y., Tanaka, Y., Wang, G., Masuda, K., Nanamiya, H., Okamoto, S., Tokuyama, S., Adachi, Y., and Kawamura, F. (2009). Inactivation of KsgA, a 16S rRNA methyltransferase, causes vigorous emergence of mutants with high-level kasugamycin resistance. Antimicrob Agents Chemother *53*, 193-201.

99. Ogata, H., Goto, S., Sato, K., Fujibuchi, W., Bono, H., and Kanehisa, M. (1999). KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res *27*, 29-34.

100. Onwuamaegbu, M.E., Belcher, R.A., and Soare, C. (2005). Cell wall-deficient bacteria as a cause of infections: a review of the clinical significance. J Int Med Res *33*, 1-20.

101. Over K, Crandall PG, O'Bryan CA, Ricke SC (2011) Current perspectives on Mycobacterium avium subsp. paratuberculosis, Johne's disease, and Crohn's disease: a review. Crit Rev Microbiol 37: 141-156.

102. Paccagnini, D., Sieswerda, L., Rosu, V., Masala, S., Pacifico, A., Gazouli, M., Ikonomopoulos, J., Ahmed, N., Zanetti, S., and Sechi, L.A. (2009). Linking chronic infection and autoimmune diseases: Mycobacterium avium subspecies paratuberculosis, SLC11A1 polymorphisms and type-1 diabetes mellitus. PLoS One *4*, e7109.

103. Pennekamp A, Pfyffer GE, Wuest J, George CA, Ruef C (1997) Mycobacterium smegmatis infection in a healthy woman following a facelift: case report and review of the literature. Ann Plast Surg 39: 80-83.

104. Perez, E., Constant, P., Laval, F., Lemassu, A., Laneelle, M.A., Daffe, M., and Guilhot, C. (2004). Molecular dissection of the role of two methyltransferases in the biosynthesis of phenolglycolipids and phthiocerol dimycoserosate in the Mycobacterium tuberculosis complex. J Biol Chem *279*, 42584-42592.

105. Pierce, E. S. 2009. Where are all the Mycobacterium avium subspecies paratuberculosis in patients with Crohn’s disease? PLoS Pathog. 5:e1000234

106. Pinto, R., Tang, Q. X., Britton, W. J., Leyh, T. S. & Triccas, J. A. (2004). The Mycobacterium tuberculosis cysD and cysNC genes form a stress-induced operon that encodes a tri-functional sulfate-activating complex. Microbiology 150, 1681–1686.

107. Rachman, H., Strong, M., Ulrichs, T., Grode, L., Schuchhardt, J., Mollenkopf, H., Kosmiadi, G.A., Eisenberg, D., and Kaufmann, S.H. (2006). Unique transcriptome signature of Mycobacterium tuberculosis in pulmonary tuberculosis. Infect Immun *74*, 1233-1242.

108. Rae, C.S., Geissler, A., Adamson, P.C., and Portnoy, D.A. (2011). Mutations of the Listeria monocytogenes Peptidoglycan N-Deacetylase and O-Acetylase Result in Enhanced Lysozyme Sensitivity, Bacteriolysis, and Hyperinduction of Innate Immune Pathways. Infect Immun *79*, 3596-3606.

109. Raynaud, C., Guilhot, C., Rauzier, J., Bordat, Y., Pelicic, V., Manganelli, R., Smith, I., Gicquel, B., and Jackson, M. (2002). Phospholipases C are involved in the virulence of Mycobacterium tuberculosis. Mol Microbiol *45*, 203-217.

110. Rastogi, N., and David, H.L. (1981). Growth and cell division of Mycobacterium avium. J Gen Microbiol *126*, 77-84.

111. Rastogi, N., David, H.L. (1988). Mechanisms of pathogenicity in mycobacteria. Biochimica 70, 1101–1120.

112. Rook, G.A., and Stanford, J.L. (1992). Slow bacterial infections or autoimmunity? Immunol Today *13*, 160-164.

113. Rosenfeld G, Bressler B (2010) Mycobacterium avium paratuberculosis and the etiology of Crohn's disease: a review of the controversy from the clinician's perspective. Can J Gastroenterol 24: 619-624.

114. Rosu, V., Ahmed, N., Paccagnini, D., Pacifico, A., Zanetti, S., and Sechi, L.A. (2008). Mycobacterium avium subspecies paratuberculosis is not associated with type-2 diabetes mellitus. Ann Clin Microbiol Antimicrob *7*, 9 (a)

115. Rosu, V., Ahmed, N., Paccagnini, D., Gerlach, G., Fadda, G., Hasnain, S.E., Zanetti, S., and Sechi, L.A. (2009). Specific immunoassays confirm association of Mycobacterium avium Subsp. paratuberculosis with type-1 but not type-2 diabetes mellitus. PLoS One *4*, e4386 (b)

116. Rowe MT, Grant IR (2006) Mycobacterium avium ssp. paratuberculosis and its potential survival tactics. Lett Appl Microbiol 42: 305-311.

117. Saeed, A.I., Sharov, V., White, J., Li, J., Liang, W., Bhagabati, N., Braisted, J., Klapa, M., Currier, T., Thiagarajan, M.*, et al.* (2003). TM4: a free, open-source system for microarray data management and analysis. Biotechniques *34*, 374-378.

118. Saunders, B.M., and Britton, W.J. (2007). Life and death in the granuloma: immunopathology of tuberculosis. Immunol Cell Biol *85*, 103-111.

119. Schnappinger, D., Ehrt, S., Voskuil, M.I., Liu, Y., Mangan, J.A., Monahan, I.M., Dolganov, G., Efron, B., Butcher, P.D., Nathan, C.*, et al.* (2003). Transcriptional Adaptation of Mycobacterium tuberculosis within Macrophages: Insights into the Phagosomal Environment. J Exp Med *198*, 693-704.

120. Schoos J (2005) [Paratuberculosis in cattle. Reflections on the interactions with Crohn disease in humans. A literature review]. Bull Soc Sci Med Grand Duche Luxemb: 327-339.

121. Schorey, J.S., and Sweet, L. (2008). The mycobacterial glycopeptidolipids: structure, function, and their role in pathogenesis. Glycobiology *18*, 832-841.

122. Schwartz, D., Shafran, I., Romero, C., Piromalli, C., Biggerstaff, J., Naser, N., Chamberlin, W., and Naser, S.A. (2000). Use of short-term culture for identification of Mycobacterium avium subsp. paratuberculosis in tissue from Crohn's disease patients. Clin Microbiol Infect *6*, 303-307

123. Sechi, L.A., Mura, M., Tanda, F., Lissia, A., Solinas, A., Fadda, G., and Zanetti, S. (2001). Identification of Mycobacterium avium subsp. paratuberculosis in biopsy specimens from patients with Crohn's disease identified by in situ hybridization. J Clin Microbiol *39*, 4514-4517.

124. Sechi LA, Scanu AM, Molicotti P, Cannas S, Mura M, Dettori G, Fadda G, Zanetti S. (2005). Detection and Isolation of Mycobacterium avium subspecies paratuberculosis from intestinal mucosal biopsies of patients with and without Crohn's disease in Sardinia. Am J Gastroenterol Jul;100 (7):1529-36.

125. Sechi, L.A., Paccagnini, D., Salza, S., Pacifico, A., Ahmed, N., and Zanetti, S. (2008). Mycobacterium avium subspecies paratuberculosis bacteremia in type 1 diabetes mellitus: an infectious trigger? Clin Infect Dis *46*, 148-149 (a)

126. Sechi, L.A., Rosu, V., Pacifico, A., Fadda, G., Ahmed, N., and Zanetti, S. (2008). Humoral immune responses of type 1 diabetes patients to Mycobacterium avium subsp. paratuberculosis lend support to the infectious trigger hypothesis. Clin Vaccine Immunol *15*, 320-326 (b)

127. Semba RD, Darnton-Hill I, de Pee S (2010) Addressing tuberculosis in the context of malnutrition and HIV coinfection. Food Nutr Bull 31: S345-364.

128. Semret M, Zhai G, Mostowy S, Cleto C, Alexander D, et al. (2004) Extensive genomic polymorphism within Mycobacterium avium. J Bacteriol 186: 6332-6334.

129. Sigurethardottir OG, Valheim M, Press CM (2004) Establishment of Mycobacterium avium subsp. paratuberculosis infection in the intestine of ruminants. Adv Drug Deliv Rev 56: 819-834.

130. Singh B, Ghosh J, Islam NM, Dasgupta S, Kirsebom LA (2010) Growth, cell division and sporulation in mycobacteria. Antonie Van Leeuwenhoek 98: 165-177.

131. Singleton Paul (2004) Bacteria in Biology, Biotechnology and Medicine. Book

132. Sisto, M., Cucci, L., D'Amore, M., Dow, T.C., Mitolo, V., and Lisi, S. (2010). Proposing a relationship between Mycobacterium avium subspecies

paratuberculosis infection and Hashimoto's thyroiditis. Scand J Infect Dis *42*, 787-790.

133. Skjot, R.L., Oettinger, T., Rosenkrands, I., Ravn, P., Brock, I., Jacobsen, S., and Andersen, P. (2000). Comparative evaluation of low-molecular-mass proteins from Mycobacterium tuberculosis identifies members of the ESAT-6 family as immunodominant T-cell antigens. Infect Immun *68*, 214-220.

134. Slabyj BM, Panos C. (1973). Teichoic acid of a stabilized L-form of Streptococcus pyogenes. J Bacteriol; 114: 934 – 942.

135. Snapper SB, Melton RE, Mustafa S, Kieser T, Jacobs WR, Jr. (1990) Isolation and characterization of efficient plasmid transformation mutants of Mycobacterium smegmatis. Mol Microbiol 4: 1911-1919.

136. Soliveri, J.A., Gomez, J., Bishai, W.R., and Chater, K.F. (2000). Multiple paralogous genes related to the Streptomyces coelicolor developmental regulatory gene whiB are present in Streptomyces and other actinomycetes. Microbiology *146 (Pt 2)*, 333-343.

137. Stabel JR (2000) Johne's disease and milk: do consumers need to worry? J Dairy Sci 83: 1659-1663.

138. Swanson, S., Gokulan, K., and Sacchettini, J.C. (2009). KasA, another brick in the mycobacterial cell wall. Structure *17*, 914-915.

139. Tessema MZ, Koets AP, Rutten VP, Gruys E (2001) How does Mycobacterium avium subsp. paratuberculosis resist intracellular degradation? Vet Q 23: 153-162.

140. Tooker BC, Burton JL, Coussens PM (2002) Survival tactics of M. paratuberculosis in bovine macrophage cells. Vet Immunol Immunopathol 87: 429-437.

141. Toshkov, A., Kuiumdzhiev, A., Gumpert, I., and Sotirova, D. (1977). Persistence and reversion of stable L forms of Proteus mirabilis in white mice. Acta Microbiol Virol Immunol (Sofiia) *6*, 35-41.

142. Uchiyama, I. (2007). MBGD: a platform for microbial comparative genomics based on the automated construction of orthologous groups. Nucleic Acids Res *35*, D343-346.

143. Udou, T., Ogawa, M., and Mizuguchi, Y. (1983). An improved method for the preparation of mycobacterial spheroplasts and the mechanism involved in the reversion to bacillary form: electron microscopic and physiological study. Can J Microbiol *29*, 60-68.

144. Ulrichs T, Kaufmann SH (2006) New insights into the function of granulomas in human tuberculosis. J Pathol 208: 261-269.

145. Vaerewijck MJ, Huys G, Palomino JC, Swings J, Portaels F (2005) Mycobacteria in drinking water distribution systems: ecology and significance for human health. FEMS Microbiol Rev 29: 911-934.

146. Vindal, V., Ashwantha Kumar, E., and Ranjan, A. (2008). Identification of operator sites within the upstream region of the putative mce2R gene from mycobacteria. FEBS Lett *582*, 1117-1122.

147. Vollmer, W. (2008). Structural variation in the glycan strands of bacterial peptidoglycan. FEMS Microbiol Rev *32*, 287-306

148. Wall, S., Kunze, Z.M., Saboor, S., Soufleri, I., Seechurn, P., Chiodini, R., and McFadden, J.J. (1993). Identification of spheroplast-like agents isolated from tissues of patients with Crohn's disease and control tissues by polymerase chain reaction. J Clin Microbiol *31*, 1241-1245.

149. Wang, H., and Chen, Z. (2001). [Observations of properties of the L-form of M. tuberculosis induced by the antituberculosis drugs]. Zhonghua Jie He He Hu Xi Za Zhi *24*, 52-55.

150. Whan L, Grant IR, Rowe MT (2006) Interaction between Mycobacterium avium subsp. paratuberculosis and environmental protozoa. BMC Microbiol 6: 63.

151. Yuan, Y., Crane, D.D., Simpson, R.M., Zhu, Y.Q., Hickey, M.J., Sherman, D.R., and Barry, C.E., 3rd (1998). The 16-kDa alpha-crystallin (Acr) protein of Mycobacterium tuberculosis is required for growth in macrophages. Proc Natl Acad Sci U S A *95*, 9578-9583.

152. Zahrt, T.C., and Deretic, V. (2001). Mycobacterium tuberculosis signal transduction system required for persistent infections. Proc Natl Acad Sci U S A *98*, 12706-12711.

153. Zhang Y. Persistent and dormant tubercle bacilli and latent tuberculosis. Front Biosci 2004;9:1113e56

154. Zhu, M., Xia, P., and Zhang, Y. (2000). [Detecting mycobacteria and their L-forms in peripheral blood from pulmonary tuberculosis patients by cultivation with hemolyzed-centrifugated blood in liquid medium]. Zhonghua Jie He He Hu Xi Za Zhi *23*, 556-558.

155. Zhu, X., Tu, Z.J., Coussens, P.M., Kapur, V., Janagama, H., Naser, S., and Sreevatsan, S. (2008). Transcriptional analysis of diverse strains Mycobacterium avium subspecies paratuberculosis in primary bovine monocyte derived macrophages. Microbes Infect 10, 1274-1282.

9. APPENDIX

A) UP-REGULATED GENES FOR CWD_MSMEG

Gene ID	GeneName	GeneProduct	log2 (exp/ctrl)	P-value
MSMEG_1535	null	conserved hypothetical protein	8,26	0,0309012870
MSMEG_1859	null	hypothetical protein	7,2766666	0,0019685768
MSMEG_5356	null	hypothetical protein	7,146666	0,0466788450
MSMEG_4783	null	conserved hypothetical protein	6,8000007	0,0030418580
MSMEG_1230	null	hypothetical protein	6,4233336	0,0426939500
MSMEG_3352	null	hypothetical protein	5,82	0,0029577771
MSMEG_2066	null	conserved hypothetical protein	3,4300003	0,0464628640
MSMEG_4781	null	conserved hypothetical protein	3,3499997	0,0455840500
MSMEG_5208	null	conserved hypothetical protein	3,1866667	0,0167137700
MSMEG_1754	null	conserved hypothetical protein	2,7033331	0,0394709500
MSMEG_3377	null	hypothetical protein	2,61	0,0336624600
MSMEG_4536	null	conserved hypothetical protein	2,5600002	0,0241273460
MSMEG_3080	null	conserved hypothetical protein	2,533333	0,0229586740
MSMEG_4413	null	hypothetical protein	2,4866667	0,0342441280
MSMEG_0805	null	conserved hypothetical protein	2,2533333	0,0027480738
MSMEG_6430	null	hypothetical protein	2,1766667	0,0334005400
MSMEG_5716	null	hypothetical protein	2,1633332	0,0241285620
MSMEG_1815	null	hypothetical protein	2,0133333	0,0223388100
MSMEG_0082	null	conserved hypothetical protein	1,8766667	0,0115457040
MSMEG_0798	null	hypothetical protein	1,8166667	0,0364872100
MSMEG_2458	null	hypothetical protein	1,7366667	0,0053973147
MSMEG_2591	null	hypothetical protein	1,6566668	0,0245052050
MSMEG_1079	null	hypothetical protein	1,5866667	0,0416111760
MSMEG_5558	null	hypothetical protein	1,5600001	0,0224125890
MSMEG_0738	null	conserved hypothetical protein	1,54	0,0010524763
MSMEG_6252	null	conserved hypothetical protein	1,5099999	0,0433644900
MSMEG_3282	null	hypothetical protein	1,4333334	0,0252514980
MSMEG_2452	null	hypothetical protein	1,4033333	0,0420802760
MSMEG_1136	null	conserved hypothetical protein	1,3866667	0,0088999610
MSMEG_4747	null	hypothetical protein	1,3866667	0,0176595780
MSMEG_4841	null	conserved hypothetical protein	1,3333334	0,0465319900
MSMEG_2862	null	conserved hypothetical protein	1,2533333	0,0134393380
MSMEG_6681	null	hypothetical protein	1,2400001	0,0236119530
MSMEG_0041	null	conserved hypothetical protein	1,2333332	0,0311214290

MSMEG_6010	null	hypothetical protein	1,2166667	0,0001425851
MSMEG_3327	null	hypothetical protein	1,1433333	0,0047178220
MSMEG_4445	null	hypothetical protein	1,13	0,0040734424
MSMEG_5692	null	conserved hypothetical protein	1,0066667	0,0065574460
MSMEG_2566	null	3-alpha-(or 20-beta)-hydroxysteroid	4,27	0,0372173100
MSMEG_1302	null	alkylphosphonate uptake protein	3,1233332	0,0246264520
MSMEG_4420	null	cupin 2. conserved barrel	5,5266666	0,0231026620
MSMEG_6488	null	conserved hypothetical protein	1,3433334	0,0113725670
MSMEG_1213	null	cytochrome P450 monooxygenase	1,4233333	0,0242058110
MSMEG_6112	null	conserved hypothetical protein	1,1433333	0,0275400780
MSMEG_3797	null	putative esterase family protein	1,4633335	0,0024247197
MSMEG_6928	null	conserved hypothetical protein	7,8566666	0,0430430960
MSMEG_3163	null	gp55 protein	1,2233334	0,0292737560
MSMEG_2096	null	integral membrane protein	5,056667	0,0172580180
MSMEG_6273	null	integral membrane protein	2,7266667	0,0253405900
MSMEG_2542	null	integral membrane protein DUF6	1,32	0,0382284970
MSMEG_6755	null	isoflavone reductase	1,7733334	0,0139724750
MSMEG_3551	null	linalool 8-monooxygenase	1,3866667	0,0112877910
MSMEG_5007	null	LprB protein	1,0033334	0,0281165730
MSMEG_4689	null	conserved hypothetical protein	2,2966666	0,0279902130
MSMEG_2764	null	probable conserved alanine rich transmembrane	1,0200001	0,0097554070
MSMEG_3937	null	conserved hypothetical protein	1,88	0,0091254740
MSMEG_3256	null	mucin-associated surface protein	7,2433333	0,0050751157
MSMEG_6763	null	oxidoreductase	2,25	0,0386166200
MSMEG_0522	null	pp24 protein	5,3799996	0,0227191260
MSMEG_2392	null	conserved hypothetical protein	1,5066667	0,0343599540
MSMEG_0145	null	conserved hypothetical protein	1,2866668	0,0420213980
MSMEG_3613	null	conserved hypothetical protein	1,7666668	0,0022600254
MSMEG_0441	null	conserved hypothetical protein. putative	2,3366668	0,0145558150
MSMEG_2643	null	conserved hypothetical protein	1,29	0,0139063240
MSMEG_0231	null	conserved hypothetical protein	1,0166668	0,0111724700
MSMEG_1748	null	conserved hypothetical protein	2,1666667	0,0373634030
MSMEG_2481	null	conserved hypothetical protein	6,7000003	0,0123521430
MSMEG_5689	null	conserved hypothetical protein	1,1966667	0,0404738800
MSMEG_6400	null	probable conserved transmembrane protein	2,4999998	0,0300942060
MSMEG_5149	null	conserved transmembrane protein	1,7566667	0,0404548240
MSMEG_6173	null	morphological differentiation-associated	1,9666666	0,0368883950
MSMEG_2007	null	putative HpcE protein	2,2733333	0,0334609150
MSMEG_2170	null	3-hydroxyisobutyrate dehydrogenase family	1,3299999	0,0354634230
MSMEG_5305	null	choline dehydrogenase	2,4833333	0,0064043240
MSMEG_1181	hutF	formiminoglutamate deiminase	1,96	0,0442776120
MSMEG_4350	dapB	dihydrodipicolinate reductase	1,6366667	0,0363708920
MSMEG_5789	null	putative thiosulfate sulfurtransferase	1,8933333	0,0003376835

MSMEG_0302	null	peptidase. S9A/B/C families	1,9166666	0,0151149210
MSMEG_0482	null	dihydroxy-acid dehydratase	1,0366668	0,0104219330
MSMEG_2521	null	amidase	6,61	0,0218871480
MSMEG_4150	null	4-hydroxy-2-oxovalerate aldolase	1,3933334	0,0223445820
MSMEG_2189	atzF	allophanate hydrolase	1,2366667	0,0021797018
MSMEG_3914	null	beta-lactamase	1,5433334	0,0227434060
MSMEG_3404	null	HNH endonuclease domain protein	1,5	0,0201727430
MSMEG_3100	opcA	OpcA protein	1,1566668	0,0079354940
MSMEG_5728	null	polysaccharide deacetylase family protein	6,263334	0,0439211350
MSMEG_4373	null	polysaccharide deacetylase. putative	1,3633333	0,0308192970
MSMEG_2550	null	glyoxylate reductase	5,2066665	0,0442701650
MSMEG_0922	deoC	deoxyribose-phosphate aldolase	1,27	0,0231250920
MSMEG_4023	null	oxidoreductase	1,23	0,0111786170
MSMEG_5920	null	FMN-dependent monooxygenase	4,3433332	0,0237152860
MSMEG_2516	null	putative oxidoreductase	2,1599998	0,0437046660
MSMEG_4348	null	acetolactate synthase large subunit	1,26	0,0402071660
MSMEG_6904	null	myo-inositol-1-phosphate synthase	1,02	0,0042341375
MSMEG_4647	null	carbohydrate kinase. PfkB. putative	6,27	0,0266310130
MSMEG_1608	null	glycosyl transferase	1,9233333	0,0386577400
MSMEG_6791	null	3-hydroxybutyryl-CoA dehydrogenase	1,82	0,0304415840
MSMEG_6795	null	enoyl-CoA hydratase/isomerase family protein	1,17	0,0302942430
MSMEG_5923	null	acetyl-CoA acetyltransferase	1,5500001	0,0119502050
MSMEG_2882	null	5-exo-alcohol dehydrogenase	1,5066667	0,0257715640
MSMEG_6242	null	alcohol dehydrogenase. iron-containing	1,4499999	0,0073846905
MSMEG_5939	null	acetaldehyde dehydrogenase	1,0666666	0,0204856750
MSMEG_2181	null	cell filamentation protein	1,17	0,0359725500
MSMEG_0370	null	conserved hypothetical protein	1,1233333	0,0446431820
MSMEG_4583	null	ATPase associated with various cellular	2,11	0,0296534230
MSMEG_0367	null	O-demethylpuromycin-O-methyltransferase	2,5133333	0,0085965070
MSMEG_3700	null	peroxidase	1,2933333	0,0382122700
MSMEG_2242	null	coniferyl aldehyde dehydrogenase	3,12	0,0264092890
MSMEG_3366	null	isonitrile hydratase. putative	6,2166667	0,0119417110
MSMEG_6505	null	NfnB protein	3,3333333	0,0114763700
MSMEG_5319	null	alkylhydroperoxidase AhpD core	8,04	0,0220324140
MSMEG_5881	null	putative carbon monoxide dehydrogenase subunit	2,2966669	0,0098642930
MSMEG_5341	null	dipeptidyl aminopeptidase/acylaminoacyl	1,9666667	0,0461537170
MSMEG_4527	null	ferredoxin sulfite reductase	1,5666666	0,0001267304
MSMEG_6108	null	zinc-binding dehydrogenase	1,6733333	0,0208116320
MSMEG_6003	null	coenzyme A transferase. subunit B	6,06	0,0352642100
MSMEG_4418	msrB	methionine-R-sulfoxide reductase	7,0266666	0,0087233890
MSMEG_0972	null	cytochrome C biogenesis protein transmembrane	8,05	0,0153961560
MSMEG_3242	null	starvation-inducible DNA-binding protein or fine	8,176666	0,0130720650
MSMEG_2934	null	lipid A biosynthesis lauroyl acyltransferase	2,0266666	0,0000432797

MSMEG_6408	null	acyltransferase family protein	1,3933333	0,0214770230
MSMEG_5618	null	acyl-CoA dehydrogenase	6,2566667	0,0226348200
MSMEG_4874	null	putative acyl-CoA dehydrogenase	5,59	0,0040533366
MSMEG_4113	null	acyl-CoA dehydrogenase	2,7033336	0,0024373853
MSMEG_3131	null	AMP-binding protein	1,1666666	0,0129050100
MSMEG_2131	null	acyl-CoA synthase	5,8433337	0,0485958600
MSMEG_6320	null	diol dehydrase gamma subunit	1,38	0,0457231740
MSMEG_2923	null	dehydrogenase/reductase SDR family protein	3,29	0,0153480540
MSMEG_0258	null	oxidoreductase. short chain	2,8000002	0,0164514280
MSMEG_0655	null	glucose 1-dehydrogenase. putative	1,4233333	0,0378416070
MSMEG_0771	null	putative oxidoreductase YqjQ	1,3866667	0,0295038040
MSMEG_0930	null	serine 3-dehydrogenase	2,3133335	0,0012268821
MSMEG_0290	null	acyltransferase. ws/dgat/mgat subfamily protein	1,6	0,0036131367
MSMEG_1204	null	3-oxoacyl-[acyl-carrier-protein] synthase 2	2,9066668	0,0020309254
MSMEG_2067	null	methyltransferase type 12	1,3966666	0,0101696340
MSMEG_4148	null	lipase	1,8966666	0,0185263290
MSMEG_3665	null	TrkA domain protein	5,9966664	0,0045455084
MSMEG_5167	null	major facilitator superfamily protein MFS_1	8,643333	0,0286789130
MSMEG_2028	null	chloramphenicol resistance protein	8,46	0,0031377247
MSMEG_2847	null	bacterial extracellular solute-binding protein.	2,8100002	0,0395561600
MSMEG_4386	null	ABC transporter permease protein	2,93	0,0072168720
MSMEG_3215	null	ABC-type molybdenum transport system. ATPase	1,3533334	0,0382116850
MSMEG_3852	null	aliphatic sulfonate binding protein	1,7966667	0,0153982330
MSMEG_1137	null	amino acid permease-associated region	1,2266667	0,0143147860
MSMEG_6196	null	gaba permease	1,9699999	0,0344258600
MSMEG_4743	null	conserved hypothetical protein	1,71	0,0187580720
MSMEG_5060	null	ABC transporter. permease protein SugA	1,4766666	0,0009210237
MSMEG_2982	null	putative periplasmic binding protein	1,6833334	0,0154598530
MSMEG_5076	null	ABC transporter membrane-spanning protein	3,4666665	0,0035461383
MSMEG_6510	null	ABC-type drug export system. membrane protein	2,2933333	0,0083634260
MSMEG_2089	ftsE	cell division ATP-binding protein FtsE	1,2733334	0,0443965800
MSMEG_1177	null	cytosine/purines/uracil/thiamine/allantoin	6,5	0,0354575480
MSMEG_6307	null	glutamine-binding periplasmic protein/glutamine	2,34	0,0342499300
MSMEG_3270	null	SN-glycerol-3-phosphate ABC transporter.	1,5066667	0,0083569400
MSMEG_3725	pqqA	coenzyme PQQ biosynthesis protein A	2,3966668	0,0178063050
MSMEG_4911	pncB	putative nicotinate phosphoribosyltransferase	4,8733335	0,0147098090
MSMEG_5085	folP	dihydropteroate synthase	1,6466666	0,0035014942
MSMEG_3871	cobG	precorrin-3B synthase	1,37	0,0045682860
MSMEG_0151	null	PntAB protein	1,2733333	0,0098109470
MSMEG_2653	ribF	riboflavin biosynthesis protein RibF	1,1666666	0,0295060500
MSMEG_2395	null	D-alanine--D-alanine ligase	1,2233334	0,0300470460
MSMEG_6075	ispF	2C-methyl-D-erythritol 2.4-cyclodiphosphate	1,1	0,0449095300
MSMEG_4619	null	putative cytochrome P450 126	2,9633331	0,0026666890

MSMEG_5849	null	conserved hypothetical protein	1,6833334	0,0230829050
MSMEG_3679	null	phosphohydrolase	2,5533333	0,0351971500
MSMEG_5549	null	conserved hypothetical protein	1,7533334	0,0102635390
MSMEG_4279	null	adenylate and Guanylate cyclase catalytic domain	1,6533333	0,0165348440
MSMEG_4803	null	cytochrome P450 superfamily protein	7,1666665	0,0224972130
MSMEG_4823	null	cytochrome p450	2,86	0,0014679002
MSMEG_2563	null	cytochrome P450 superfamily protein	1,0433334	0,0249505080
MSMEG_1327	recB	exodeoxyribonuclease V. beta subunit	2,6266668	0,0417793730
MSMEG_5935	null	ATP-dependent DNA helicase	1,0100001	0,0078470590
MSMEG_3638	null	CBS domain protein	6,593333	0,0156828220
MSMEG_5629	null	conserved hypothetical protein	1,1066667	0,0142854920
MSMEG_3448	null	two-component system sensor kinase	2,0633333	0,0285470990
MSMEG_4191	null	conserved hypothetical protein	1,8966665	0,0135655190
MSMEG_4487	null	ferric uptake regulation protein	2,8466666	0,0421614000
MSMEG_1690	null	putative ECF sigma factor RpoE1	1,5033334	0,0318923300
MSMEG_2428	null	DNA-binding protein	1,1899999	0,0262135360
MSMEG_0609	null	helix-turn-helix domain protein	1,4666667	0,0178767140
MSMEG_1970	null	sigma factor	2,7	0,0083041850
MSMEG_6913	null	putative transcriptional regulatory protein	1,5499998	0,0474650860
MSMEG_6864	null	putative transcription regulator	2,2066667	0,0190105460
MSMEG_3013	null	putative transcriptional regulator family	1,33	0,0245256800
MSMEG_2009	null	regulatory protein GntR. HTH:GntR. C-	3,1966667	0,0172410800
MSMEG_6568	null	regulatory protein. LuxR. putative	1,06	0,0033061851
MSMEG_1596	null	transcriptional regulator	1,9533334	0,0195101400
MSMEG_3319	null	repressor protein	1,09	0,0209010340
MSMEG_6344	null	transcriptional regulator	1,93	0,0344083940
MSMEG_2636	null	regulatory protein. LacI	5,853334	0,0350252730
MSMEG_6849	null	LysR-family protein transcriptional regulator	1,8166666	0,0138698040
MSMEG_3874	null	transcriptional regulator. TetR family protein	5,9633336	0,0351254340
MSMEG_1197	null	transcriptional regulator. LuxR family protein	8,596667	0,0081133840
MSMEG_4144	null	transcriptional regulator. LysR family protein	2,47	0,0039755750
MSMEG_2071	null	transcriptional regulator. TetR family protein	4,096667	0,0108660680
MSMEG_5567	null	transcriptional regulator. TetR family protein	1,5833334	0,0007614562
MSMEG_2267	null	tetratricopeptide repeat domain protein	1,4266667	0,0115921510
MSMEG_4959	argS	arginyl-tRNA synthetase	1,4833332	0,0023600070
MSMEG_6889	null	Transposase	2,3033333	0,0050898776
MSMEG_4127	null	TnpC protein	1,4533333	0,0285459680
MSMEG_1717	null	ISMsm8. transposase	1,2266666	0,0411788670
MSMEG_1095	ureG	urease accessory protein UreG	6,4599996	0,0132477570
MSMEG_5644	null	membrane protein	1,9666668	0,0264583080
MSMEG_2485	null	monooxygenase. FAD-binding	2,2166665	0,0087327030
MSMEG_4344	null	putative monooxygenase	1,2166667	0,0394304470
MSMEG_4839	null	carboxymuconolactone decarboxylase	1,4566666	0,0330414800

MSMEG_6036	null	biphenyl-2.3-diol 1.2-dioxygenase	1,6966667	0,0190930090
MSMEG_0719	null	Flavohemoprotein	1,5233334	0,0499247870
MSMEG_0329	null	FMN oxidoreductase	1,3433332	0,0289213020

B) DOWN-REGULATED GENES FOR CWD-MSMEG

Gene ID	GeneName	GeneProduct	log2 (exp/ctrl)	P-value
MSMEG_5614	null	conserved hypothetical protein	-6,466667	0,03793627
MSMEG_4438	null	hypothetical protein	-5,3566666	0,004468318
MSMEG_4124	null	conserved hypothetical protein	-4,9966664	0,018499497
MSMEG_1768	null	conserved hypothetical protein	-3,2466667	0,003880629
MSMEG_2376	null	conserved hypothetical protein	-3,0366666	0,007704751
MSMEG_1766	null	conserved hypothetical protein	-2,9166667	0,000447699
MSMEG_6335	null	hypothetical protein	-2,8933334	0,010199839
MSMEG_1758	null	hypothetical protein	-2,5733335	0,005710922
MSMEG_5722	null	conserved hypothetical protein	-2,5	0,010493814
MSMEG_4489	null	conserved hypothetical protein	-2,5	0,014668169
MSMEG_2115	null	conserved hypothetical protein	-2,49	0,005713995
MSMEG_1422	null	conserved hypothetical protein	-2,47	0,041985385
MSMEG_2983	null	conserved hypothetical protein	-2,3700001	0,000231365
MSMEG_5543	null	hypothetical protein	-2,3633335	0,017150119
MSMEG_0945	null	conserved domain protein	-2,3333333	0,003225148
MSMEG_5342	null	conserved hypothetical protein	-2,2633333	0,01925195
MSMEG_5497	null	conserved hypothetical protein	-2,16	0,03704398
MSMEG_2958	null	conserved hypothetical protein	-2,1499999	0,003750206
MSMEG_5449	null	conserved hypothetical protein	-2,01	0,029822892
MSMEG_5293	null	hypothetical protein	-1,9833332	0,034234695
MSMEG_4195	null	conserved hypothetical protein	-1,9633335	0,047093958
MSMEG_0537	null	conserved hypothetical protein	-1,9566666	0,007100731
MSMEG_3329	null	hypothetical protein	-1,88	0,003219258
MSMEG_5034	null	hypothetical protein	-1,8566667	0,005697923
MSMEG_1789	null	conserved hypothetical protein	-1,8533331	0,036576
MSMEG_1772	null	conserved hypothetical protein	-1,7566667	0,001566189
MSMEG_3544	null	conserved hypothetical protein	-1,7333335	0,012726245
MSMEG_6305	null	conserved hypothetical protein	-1,6933333	0,014214596
MSMEG_5355	null	hypothetical protein	-1,6899999	0,02553862

MSMEG_0278	null	hypothetical protein	-1,68	0,007046416
MSMEG_3971	null	hypothetical protein	-1,6366668	0,014372126
MSMEG_6090	null	conserved hypothetical protein	-1,6033334	0,04503133
MSMEG_3439	null	hypothetical protein	-1,5300001	0,013573617
MSMEG_1812	null	conserved hypothetical protein	-1,5099999	0,042287577
MSMEG_0383	null	conserved hypothetical protein	-1,4966666	0,009654932
MSMEG_5111	null	conserved hypothetical protein	-1,4333334	0,002623496
MSMEG_5298	null	hypothetical protein	-1,4233333	0,006199782
MSMEG_4575	null	hypothetical protein	-1,42	0,036810253
MSMEG_5791	null	conserved hypothetical protein	-1,41	0,008393704
MSMEG_5741	null	conserved hypothetical protein	-1,3733333	0,000817876
MSMEG_1558	null	conserved hypothetical protein	-1,3633333	0,03972353
MSMEG_2592	null	conserved hypothetical protein	-1,3299999	0,045451812
MSMEG_3766	null	hypothetical protein	-1,3233333	0,000405822
MSMEG_4991	null	hypothetical protein	-1,31	0,001182756
MSMEG_0057	null	conserved hypothetical protein	-1,3066667	0,015006973
MSMEG_0726	null	hypothetical protein	-1,2766666	0,008287729
MSMEG_4021	null	hypothetical protein	-1,2666668	0,023578148
MSMEG_1251	null	conserved hypothetical protein	-1,1966666	0,033530947
MSMEG_4336	null	conserved hypothetical protein	-1,19	0,042427305
MSMEG_3736	null	hypothetical protein	-1,1866666	0,013914728
MSMEG_6566	null	hypothetical protein	-1,18	0,034795098
MSMEG_1632	null	hypothetical protein	-1,1533333	0,04669087
MSMEG_2110	null	conserved hypothetical protein	-1,1466666	0,015485873
MSMEG_2008	null	hypothetical protein	-1,1366667	0,015038956
MSMEG_3764	null	conserved hypothetical protein	-1,1	0,002059734
MSMEG_3441	null	hypothetical protein	-1,0966667	0,002384256
MSMEG_5770	null	hypothetical protein	-1,0866667	0,045877714
MSMEG_3932	null	14 kDa antigen	-1,7566667	0,023978675
MSMEG_2695	null	35 kDa protein	-1,0066667	0,017625114
MSMEG_0696	null	alanine-rich protein	-2,17	0,000176923
MSMEG_1934	null	ATP-binding protein	-1,1366667	0,041086335
MSMEG_4017	null	BmyD protein	-1,8433334	0,006338641
MSMEG_6855	null	carboxyvinyl-carboxyphosphonate	-4,7	0,030164095
MSMEG_4213	null	cytochrome P450 hydroxylase	-1,96	0,009772598
MSMEG_0944	null	DNA binding domain. excisionase family protein	-2,04	0,018595403
MSMEG_2351	etfB	electron transfer flavoprotein. beta subunit	-1,13	0,001744454
MSMEG_2484	null	conserved hypothetical protein	-2,23	0,006024732
MSMEG_4919	null	conserved hypothetical protein	-1,5933334	0,007174201
MSMEG_1045	null	integral membrane protein	-1,7000002	0,024833422
MSMEG_6665	null	integral membrane protein	-2,4466667	0,039564792
MSMEG_0896	null	integral membrane transport protein	-1,3066667	0,03501534
MSMEG_6078	null	LpqE protein	-1,8533331	0,012908191

MSMEG_0814	null	putative thiosulfate sulfurtransferase	-7,283333	0,001576209
MSMEG_4752	null	conserved hypothetical protein	-1,0466667	0,027355542
MSMEG_1932	null	MmpS3 protein	-2,1200001	0,005326781
MSMEG_6664	null	methylenetetrahydrofolate reductase family	-1,5566667	0,027912034
MSMEG_0844	null	multiple resistance and pH regulation protein F	-1,6866666	0,000612632
MSMEG_6767	null	mycocerosic acid synthase	-1,1133333	0,020599445
MSMEG_0848	null	NADH ubiquinone oxidoreductase subunit 5	-1,5433334	0,003540392
MSMEG_0846	null	NADH-ubiquinone oxidoreductase/multisubunit	-1,0033333	0,015282404
MSMEG_5339	null	nitrile hydratase regulator1	-1,6633334	0,003467808
MSMEG_4723	null	conserved hypothetical protein	-2,2066667	0,00650984
MSMEG_0451	null	oxidoreductase. FAD-linked	-2,2	0,046364248
MSMEG_2192	null	conserved hypothetical protein	-1,27	0,015996516
MSMEG_6044	null	periplasmic binding proteins and sugar binding	-1,5466666	0,021070903
MSMEG_6407	null	Cof-like hydrolase	-2,0266669	0,029788326
MSMEG_6329	null	conserved hypothetical protein	-2,2166667	0,008203658
MSMEG_3408	null	conserved hypothetical protein	-1,6666666	0,002300042
MSMEG_3645	null	conserved hypothetical protein	-1,5	0,001345429
MSMEG_4221	null	conserved hypothetical protein	-1,3633333	0,017028546
MSMEG_2446	null	conserved hypothetical protein	-5,7933335	0,003049251
MSMEG_4933	null	conserved hypothetical protein	-1,1133333	0,03425444
MSMEG_2940	null	conserved hypothetical protein	-1,0600001	0,017113071
MSMEG_3923	null	conserved hypothetical protein	-2,6000001	0,007546835
MSMEG_2404	null	extracellular deoxyribonuclease	-2,0333333	0,002572678
MSMEG_0478	null	secreted protein. putative	-6,0633335	0,023045965
MSMEG_1878	null	S30AE family protein	-1,5633334	0,021673864
MSMEG_2575	null	conserved hypothetical protein	-1,3333334	0,014138161
MSMEG_0315	null	probable conserved transmembrane protein	-1,14	0,03509872
MSMEG_6359	null	trypsin domain protein	-1,1633333	0,016112579
MSMEG_4258	trpD	anthranilate phosphoribosyltransferase	-3,1133335	0,0171176
MSMEG_4454	ltaE	L-threonine aldolase. low-specificity	-2,2266667	0,005205316
MSMEG_4584	proA	gamma-glutamyl phosphate reductase	-1,5866667	0,012490639
MSMEG_3030	aroC	chorismate synthase	-1,4699999	0,036288284
MSMEG_4181	null	phosphoribosyl-ATP pyrophosphohydrolase	-1,0066667	0,000832256
MSMEG_5211	null	aminotransferase class-III	-2,0366666	0,017561717
MSMEG_1414	null	Amidinotransferase	-2,6766665	0,003854714
MSMEG_3206	hisC	histidinol-phosphate aminotransferase	-1,4200001	0,002978771
MSMEG_1809	null	putative thiosulfate sulfurtransferase	-1,7933334	0,01167376
MSMEG_6399	null	antigen 85-C	-1,7766666	0,02638779
MSMEG_4185	metH	methionine synthase	-1,5733334	0,017957307
MSMEG_5270	null	cystathionine beta-synthase	-1,5366668	0,013151088
MSMEG_2659	ald	alanine dehydrogenase	-1,11	0,007463688
MSMEG_4680	lipP	carboxylesterase. putative	-1,0100001	0,002376886
MSMEG_3810	null	Hydrolase	-1,9	0,022253864

MSMEG_0834	null	tuberculin related peptide	-2,5900002	0,006587965
MSMEG_6914	null	conserved hypothetical protein	-2,12	0,02658712
MSMEG_5345	null	glycosyl hydrolases family protein 16	-1,18	0,003311068
MSMEG_2335	null	hexapeptide transferase family protein	-2,3300002	0,028432647
MSMEG_1074	null	polysaccharide deacetylase	-1,9033333	0,04313218
MSMEG_5078	glgC	glucose-1-phosphate adenylyltransferase	-2,1966667	0,006256968
MSMEG_4646	null	pyruvate synthase	-1,3766667	0,03089284
MSMEG_3640	glcB	malate synthase G	-1,2400001	0,037631314
MSMEG_5189	null	Oxidoreductase	-3,6299999	0,041577823
MSMEG_5492	null	acetyl-CoA carboxylase carboxyltransferase	-3,2966669	0,014643516
MSMEG_4283	sucB	2-oxoglutarate dehydrogenase. E2 component.	-1,18	0,02062603
MSMEG_5538	null	[NADP+] succinate-semialdehyde dehydrogenase	-1,0366668	0,003220487
MSMEG_4141	null	conserved hypothetical protein	-2,42	0,017061139
MSMEG_5732	null	Monooxygenase	-1,93	0,008298458
MSMEG_6703	null	N5.N10- methylenetetrahydromethanopterin	-1,84	0,010075263
MSMEG_1794	null	Dehydrogenase	-4,3333335	0,006032545
MSMEG_4921	mce	methylmalonyl-CoA epimerase	-1,7333332	0,014143167
MSMEG_6616	null	S-(hydroxymethyl)glutathione dehydrogenase	-2,54	0,000108482
MSMEG_5704	null	conserved hypothetical protein	-1,1633333	0,005891643
MSMEG_1671	null	succinate dehydrogenase hydrophobic membrane	-1,82	0,007059341
MSMEG_1097	null	glycosyl transferase. group 2 family protein	-3,0766668	0,011919439
MSMEG_3859	null	glycosyl transferase. group 2 family protein	-1,1999999	0,006260165
MSMEG_3139	null	enoyl-CoA hydratase/isomerase	-1,0033333	0,011242288
MSMEG_4077	null	enoyl-CoA hydratase	-1,8966666	0,017561244
MSMEG_0671	null	S-(hydroxymethyl)glutathione dehydrogenase	-1,8766667	0,0496601
MSMEG_5721	null	acetyl-CoA acetyltransferase	-1,2600001	0,009683657
MSMEG_4920	null	acetyl-CoA acetyltransferase	-1,2466666	0,013889178
MSMEG_2200	purU	formyltetrahydrofolate deformylase	-2,3566668	0,003494657
MSMEG_3103	tkt	Transketolase	-1,7566667	0,041892603
MSMEG_5198	null	carnitinyl-CoA dehydratase	-1,1866667	0,027169885
MSMEG_5427	null	ribose-phosphate pyrophosphokinase	-1,0766667	0,011149839
MSMEG_4620	null	NAD-dependent deacetylase 1	-2,22	0,006934207
MSMEG_1274	null	Gluconolactonase	-1,7233334	0,029211963
MSMEG_4222	ftsZ	cell division protein FtsZ	-1,9200001	0,033118717
MSMEG_6241	null	ATPase associated with various cellular	-2,18	0,024514925
MSMEG_6636	null	[Mn] superoxide dismutase	-1,5933332	0,03793528
MSMEG_2988	null	carbon-nitrogen hydrolase family protein	-1,2366667	0,040208697
MSMEG_3407	null	epoxide hydrolase	-1,9066666	0,012888911
MSMEG_0182	null	epoxide hydrolase 1	-1,8766667	0,009574601
MSMEG_0835	sodC	copper/zinc superoxide dismutase	-1,3833333	0,043441027
MSMEG_6157	topA	DNA topoisomerase I	-1,8266667	0,001856266
MSMEG_0690	null	iron-sulfur cluster-binding protein	-1,6199999	0,017164463
MSMEG_4940	null	bacteriophage lysis protein	-1,71	0,021546328

MSMEG_3661	null	conserved hypothetical protein	-1,0133333	0,04701726
MSMEG_3255	null	DoxX subfamily protein. putative	-2,2166665	0,000774721
MSMEG_0273	null	ethanolamine utilization protein EutN	-1,4466667	0,02752651
MSMEG_4267	null	probable cytochrome c oxidase polypeptide 4	-1,4233333	0,03776832
MSMEG_2202	null	dimethylaniline monooxygenase [N-oxide-forming]	-8,036667	6,20976E-05
MSMEG_4260	null	cytochrome c oxidase subunit 3	-1,8633333	0,002156336
MSMEG_6114	ppa	inorganic pyrophosphatase	-1,5333334	0,019508425
MSMEG_4262	null	ubiquinol-cytochrome c reductase iron-sulfur	-1,5266666	0,020226149
MSMEG_2594	asnB	asparagine synthase (glutamine-hydrolyzing)	-3,5	0,024000008
MSMEG_5249	glyA	serine hydroxymethyltransferase	-1,0466667	0,01325167
MSMEG_6747	null	Oxidoreductase	-2,0200002	0,001426541
MSMEG_2391	ppk	polyphosphate kinase	-1,1133333	0,020574184
MSMEG_3105	cyoE	protoheme IX farnesyltransferase	-1,68	0,005493421
MSMEG_2838	null	conserved hypothetical protein	-2,6866665	0,014901042
MSMEG_4755	null	peptidase M20	-1,8699999	0,010258704
MSMEG_5774	null	tRNA-dihydrouridine synthase. putative	-1,93	0,001455469
MSMEG_1485	map	methionine aminopeptidase. type I	-1,1633333	0,045474943
MSMEG_5701	moaE	molybdopterin converting factor. subunit 2	-6,3733335	0,023141159
MSMEG_0024	null	peptidyl-prolyl cis-trans isomerase B	-1,0699999	0,033128142
MSMEG_1344	null	Translocase	-2,0466666	0,022548245
MSMEG_6467	null	starvation-induced DNA protecting protein	-1,7599999	0,003031528
MSMEG_0768	null	conserved hypothetical protein	-2,49	0,020486856
MSMEG_0206	null	acyltransferase 3	-2,033333	0,007354242
MSMEG_5537	null	conserved integral membrane protein	-1,2366667	0,011151741
MSMEG_5696	null	Cold-shock' DNA-binding domain protein	-1,9499999	0,002262733
MSMEG_6201	null	Transglycosylase	-1,2966666	0,021720247
MSMEG_3022	null	transglycosylase associated protein	-1,6833334	0,011120411
MSMEG_3528	null	ErfK/YbiS/YcfS/YnhG family protein	-1,3433332	0,030035779
MSMEG_6422	null	ferritin family protein	-1,33	0,006216632
MSMEG_3564	bfr	Bacterioferritin	-1,8133334	0,011069271
MSMEG_6322	null	bifunctional wax ester synthase/acyl-CoA	-1,7199999	0,005375935
MSMEG_5183	null	3-Hydroxyacyl-CoA dehydrogenase	-1,0733334	0,046242997
MSMEG_4327	null	3-oxoacyl-[acyl-carrier-protein] synthase 1	-2,36	0,017984232
MSMEG_2672	null	acyl-CoA synthase	-1,9	0,010872704
MSMEG_5435	null	acyl-CoA synthase	-1,7366667	0,040740114
MSMEG_6393	null	acyl-CoA synthase	-1,0500001	0,027012808
MSMEG_5773	null	fatty acid desaturase	-1,5033334	0,01766696
MSMEG_6385	null	putative oxidoreductase in MprA 5'region	-1,0633334	0,018693771
MSMEG_5869	null	putative short chain dehydrogenase	-1,07	0,003849899
MSMEG_1563	null	short-chain dehydrogenase/reductase SDR	-2,1299999	0,00690037
MSMEG_0831	null	short chain dehydrogenase	-5,956667	0,008327594
MSMEG_1782	null	oxidoreductase. short chain	-1,7533334	0,021130607
MSMEG_6572	null	Methyltransferase	-1,7800001	0,001919719

MSMEG_2468	null	L-carnitine dehydratase/bile acid-inducible	-1,3966666	0,014204963
MSMEG_0194	null	serine esterase. cutinase family protein	-1,6233333	0,027246276
MSMEG_5246	null	conserved hypothetical protein	-1,7733334	0,005438937
MSMEG_2984	null	putative hydrolase	-1,0633334	0,025366776
MSMEG_3289	null	gp61 protein	-2,1033332	0,005568952
MSMEG_5015	null	secreted protein	-1,2266667	0,03724941
MSMEG_5016	null	conserved domain protein	-1,0433334	0,005597296
MSMEG_1064	null	phosphate/sulphate permease	-1,5966667	0,017276166
MSMEG_6544	null	transport-associated. putative	-1,4266666	0,038083233
MSMEG_2737	null	ppe family protein	-1,2733333	0,016860466
MSMEG_0564	null	xanthine/uracil permease	-1,12	0,015293443
MSMEG_0965	null	Porin	-2,29	0,002791417
MSMEG_2416	null	conserved hypothetical protein	-1,8866668	0,007411569
MSMEG_5056	corA	magnesium and cobalt transport protein CorA	-2,4399998	0,010237687
MSMEG_3689	null	sodium:solute symporter	-1,5566667	0,013244962
MSMEG_3536	null	sugar transport protein	-1,9000001	0,009906231
MSMEG_4318	null	hypothetical protein	-1,89	0,002296961
MSMEG_3563	null	drug transporter	-1,7599999	0,038304023
MSMEG_5428	arsC	arsenate reductase	-1,2800001	0,01874647
MSMEG_1802	null	ChaB protein	-3,9633338	0,01824981
MSMEG_1790	null	conserved hypothetical protein	-2,6266668	0,005981864
MSMEG_2796	null	integral membrane protein	-1,5633334	0,007409075
MSMEG_5046	null	drug transporter	-1,4733334	0,03905322
MSMEG_3705	null	major facilitator superfamily protein MFS_1	-1,7633333	0,016319018
MSMEG_0555	null	ABC transporter permease protein	-1,0533333	0,043083053
MSMEG_0345	null	conserved hypothetical protein	-1,2533334	0,021549705
MSMEG_5901	null	TrnB2 protein	-1,7533334	0,007306829
MSMEG_1219	null	ABC-type transport system permease protein II	-1,4633332	0,026066318
MSMEG_6727	null	amino acid permease-associated region	-3,0666668	0,018467715
MSMEG_2845	null	ABC transporter permease protein	-2,61	0,001528078
MSMEG_1838	null	cationic amino acid transporter	-2,3999999	0,01696018
MSMEG_2522	null	efflux ABC transporter. permease protein	-2,1200001	0,010533408
MSMEG_6792	null	inner membrane permease YgbN	-1,4633335	0,024913812
MSMEG_5781	pstC	phosphate ABC transporter. permease protein	-2,2566667	0,000113436
MSMEG_1216	null	ABC-type transport system periplasmic	-1,4699999	0,030787962
MSMEG_0662	null	putrescine transport ATP-binding protein PotG	-1,1700001	0,007035385
MSMEG_6509	null	ABC-type drug export system. ATP-binding	-1,0766667	0,016664168
MSMEG_3402	null	cytosine permease. putative	-2,28	0,004515501
MSMEG_2925	null	permease membrane component	-1,82	0,002376479
MSMEG_2926	null	glycine betaine/carnitine/choline transport	-1,59	0,015506609
MSMEG_2924	null	permease binding-protein component	-2,1733334	0,007820645
MSMEG_1770	null	conserved hypothetical protein	-2,9433334	0,031132301
MSMEG_4104	null	transporter. major facilitator family protein	-6,0266666	0,0337545

MSMEG_1806	null	conserved hypothetical protein	-2,5533333	4,26043E-05
MSMEG_5470	null	molybdopterin biosynthesis protein MoeA 1	-1,2666668	0,003417306
MSMEG_3721	pqqE	coenzyme PQQ biosynthesis protein E	-2,6533332	0,008194975
MSMEG_0964	null	pyridoxamine 5'-phosphate oxidase family	-1,0333333	0,029323801
MSMEG_0793	thiG	thiazole biosynthesis protein ThiG	-6,683333	0,01121608
MSMEG_1822	null	biotin-[acetyl-CoA-carboxylase] ligase	-2,6	0,02978321
MSMEG_2939	null	glutamine amidotransferase subunit PdxT	-1,2766666	0,017532002
MSMEG_2616	cobO	cob(I)alamin adenosyltransferase	-1,0300001	0,007487253
MSMEG_2383	gltX	glutamyl-tRNA synthetase	-6,78	0,006635791
MSMEG_0836	null	carboxylate-amine ligase	-1,3533334	0,030343898
MSMEG_1850	selA	L-seryl-tRNA selenium transferase	-5,5499997	0,01784371
MSMEG_1028	null	geranylgeranyl reductase	-1,3366667	0,008297267
MSMEG_3418	null	conserved hypothetical protein	-3,24	0,000612276
MSMEG_6076	ispD	2-C-methyl-D-erythritol 4-phosphate	-1,23	0,017654182
MSMEG_0205	null	tetracenomycin polyketide synthesis hydroxylase	-1,77	0,006020499
MSMEG_2744	null	thymidylate synthase	-1,9433333	0,007585894
MSMEG_2965	relA	GTP pyrophosphokinase	-2,4266667	0,010169473
MSMEG_3780	null	adenylate and Guanylate cyclase catalytic domain	-1,7566667	0,003164277
MSMEG_1484	null	adenylate kinase	-1,43	0,026595062
MSMEG_1033	null	ribonucleoside-diphosphate reductase. beta	-1,0366668	0,028660791
MSMEG_1019	null	ribonucleoside-diphosphate reductase. alpha	-1,0033333	0,006600658
MSMEG_0759	purA	adenylosuccinate synthetase	-2,1566668	0,049311396
MSMEG_1315	null	transporter. small conductance mechanosensitive	-1,5799999	0,001452261
MSMEG_5995	null	P450 heme-thiolate protein	-1,4699999	0,007458868
MSMEG_4829	null	cytochrome P450	-2,8933334	0,037322402
MSMEG_6934	trx	Thioredoxin	-1,7200001	0,00270442
MSMEG_2415	null	hemerythrin HHE cation binding region	-2,4733334	9,4435E-05
MSMEG_6212	null	hemerythrin HHE cation binding domain subfamily	-2,0566666	0,042007457
MSMEG_6806	uvrA	excinuclease ABC. A subunit	-1,6999999	0,02390155
MSMEG_3471	null	GTP cyclohydrolase	-1,3833333	0,03315398
MSMEG_3522	null	dopamine receptor D4	-1,9033333	0,03493448
MSMEG_1874	mtrA	DNA-binding response regulator MtrA	-1,6233333	0,00959167
MSMEG_2427	glnD	protein-P-II uridylyltransferase	-1,13	0,0257776
MSMEG_1188	null	conserved hypothetical protein	-1,2733334	0,00429632
MSMEG_2248	null	two-component system sensor kinase	-1,43	0,01721501
MSMEG_1001	null	acetyltransferase. gnat family protein.	-2,66	0,015251485
MSMEG_6127	null	anti-anti-sigma factor	-1,1999999	0,02928633
MSMEG_6129	null	conserved hypothetical protein	-2,6833332	0,041395597
MSMEG_2039	null	putative transcriptional regulator	-2,31	0,03335434
MSMEG_3008	null	putative sigma 54 type regulator	-1,0966667	0,024880158
MSMEG_1029	null	probable transcriptional regulatory protein	-2,3966668	0,028420063
MSMEG_1178	null	transcriptional regulator	-1,5933332	0,016523363
MSMEG_3141	null	conserved domain protein	-1,8033333	0,011309213

MSMEG_2743	nrdR	transcriptional regulator. NrdR family protein	-1,3466667	0,011423854
MSMEG_0285	null	transcriptional regulator TetR family protein	-1,0666666	0,02925953
MSMEG_2954	null	LacI-family protein transcriptional regulator	-1,0300001	0,040922884
MSMEG_1696	null	regulatory protein. MarR	-1,5333333	0,001729906
MSMEG_6628	null	transcriptional regulator. TetR family protein	-2,1333332	0,011632225
MSMEG_1492	null	transcriptional regulator. MarR family protein	-6,1299996	0,012934494
MSMEG_3284	null	transcriptional regulator. MarR family protein	-1,7033333	0,010375435
MSMEG_6604	null	transcriptional regulator. TetR family protein	-6,466667	0,000606314
MSMEG_1769	null	UsfY protein	-2,46	0,012300393
MSMEG_0219	null	RNA polymerase sigma-70 factor. family protein	-1,77	0,000393438
MSMEG_2389	hup	DNA-binding protein HU	-2,2866666	0,014703365
MSMEG_1951	null	conserved domain protein	-2,6999998	0,003606369
MSMEG_1556	rplM	ribosomal protein L13	-1,5300001	0,029680701
MSMEG_2400	rpmB	ribosomal protein L28	-1,2766666	0,003768917
MSMEG_4625	rplU	ribosomal protein L21	-1,26	0,022305526
MSMEG_1365	rplL	ribosomal protein L7/L12	-1,2366667	0,002006428
MSMEG_2440	rplS	ribosomal protein L19	-1,0733334	0,030920219
MSMEG_3793	infC	translation initiation factor IF-3	-1,7366667	0,0244318
MSMEG_5438	ksgA	dimethyladenosine transferase	-1,2266666	0,019032836
MSMEG_4950	prfA	peptide chain release factor 1	-1,4399999	0,003851733
MSMEG_1364	null	50S ribosomal protein L10	-2,4366665	0,010228991
MSMEG_2519	rpsB	ribosomal protein S2	-1,8066667	0,003578822
MSMEG_2435	null	30S ribosomal protein S16	-1,6966667	0,01935519
MSMEG_3796	null	lysyl-tRNA synthetase	-1,59	0,0387264
MSMEG_4926	null	IS1096. tnpA protein	-1,83	0,024186183
MSMEG_3166	null	IS1096. tnpA protein	-1,6133333	0,0462383
MSMEG_0202	null	IS1096. tnpA protein	-1,5666666	0,017874831
MSMEG_6161	null	IS1096. tnpA protein	-1,38	0,018002925
MSMEG_4403	null	IS1096. tnpA protein	-1,3666667	0,04574234
MSMEG_6148	null	IS1096. tnpA protein	-1,3400002	0,017851477
MSMEG_6462	null	IS1096. tnpA protein	-1,1833333	0,001053679
MSMEG_1258	null	IS1096. tnpA protein	-1,1633333	0,03471387
MSMEG_5378	null	IS1096. tnpA protein	-1,0333333	0,002002296
MSMEG_2003	null	IS1096. tnpA protein	-1,0333333	0,016155561
MSMEG_6149	null	IS1096. tnpR protein	-2,1000001	0,045362484
MSMEG_2282	null	IS1096. tnpR protein	-1,7333335	0,017022014
MSMEG_4402	null	IS1096. tnpR protein	-1,4399999	0,004455007
MSMEG_1002	null	IS1096. tnpR protein	-1,41	0,002921274
MSMEG_3696	null	IS1096. tnpR protein	-1,4033333	0,000293257
MSMEG_0801	null	IS1096. tnpR protein	-1,4033333	0,004577915
MSMEG_4791	null	IS1096. tnpR protein	-1,3366667	0,000639929
MSMEG_4946	null	IS1096. tnpR protein	-1,3333334	0,00826514
MSMEG_5379	null	IS1096. tnpR protein	-1,31	0,004859171

MSMEG_2002	null	IS1096. tnpR protein	-1,2866665	0,00963113
MSMEG_0398	null	IS1096. tnpR protein	-1,24	0,049019348
MSMEG_6698	null	IS1096. tnpR protein	-1,22	0,012676957
MSMEG_1257	null	IS1096. tnpR protein	-1,16	0,019551033
MSMEG_2338	null	IS1096. tnpR protein	-1,1366667	0,027352514
MSMEG_0203	null	IS1096. tnpR protein	-1,0733334	0,02370437
MSMEG_3165	null	IS1096. tnpR protein	-1,0133333	0,000140638
MSMEG_6589	null	ISMsm5. transposase	-2,3233335	0,004625914
MSMEG_5891	null	ISMsm5. transposase	-1,88	0,01984907
MSMEG_4072	null	ISMsm5. transposase	-1,66	0,015871776
MSMEG_1591	null	TnpC protein	-1,1166667	0,020831933
MSMEG_3623	ureG	urease accessory protein UreG	-2,17	0,0311631
MSMEG_4256	null	NLP/P60 family protein	-1,79	0,017308904
MSMEG_2310	null	Monooxygenase	-1,16	0,046936546
MSMEG_1030	null	Monooxygenase	-1,15	0,03233705
MSMEG_2444	null	dienelactone hydrolase family protein	-1,3233333	0,001222304

UP-REGULATED GENES FOR CWD-MTB

Gene ID	GeneName	GeneProduct	log2 (exp/ctrl)	P-value
Rv1572c	null	CONSERVED HYPOTHETICAL PROTEIN	3,3920002000	0,0451966520
Rv3531c	null	HYPOTHETICAL PROTEIN	2,8660000000	0,0136293940
Rv2517c	null	HYPOTHETICAL PROTEIN	2,8160000000	0,0028347278
Rv2016	null	HYPOTHETICAL PROTEIN	2,7780000000	0,0060629756
Rv0471c	null	HYPOTHETICAL PROTEIN	1,7220001000	0,0001789515
Rv0698	null	CONSERVED HYPOTHETICAL PROTEIN	1,7040001000	0,0462103300
Rv2516c	null	HYPOTHETICAL PROTEIN	1,6700001000	0,0052040857
Rv0810c	null	CONSERVED HYPOTHETICAL PROTEIN	1,6540000000	0,0163350810
Rv0313	null	CONSERVED HYPOTHETICAL PROTEIN	1,5560000000	0,0272294320
Rv1632c	null	HYPOTHETICAL PROTEIN	1,5560000000	0,0283333940
Rv1993c	null	CONSERVED HYPOTHETICAL PROTEIN	1,5420000000	0,0049140030
Rv0756c	null	HYPOTHETICAL PROTEIN	1,5120001000	0,0213644970
Rv3098c	null	HYPOTHETICAL PROTEIN	1,3959999000	0,0209950190
Rv2009	null	CONSERVED HYPOTHETICAL PROTEIN	1,3780000000	0,0000905628
Rv2663	null	HYPOTHETICAL PROTEIN	1,3700000000	0,0027144623
Rv1989c	null	HYPOTHETICAL PROTEIN	1,3499999000	0,0012881712
Rv0157A	null	HYPOTHETICAL PROTEIN (FRAGMENT)	1,3260000000	0,0008662927
Rv0559c	null	POSSIBLE CONSERVED SECRETED PROTEIN	1,3180000000	0,0074845330
Rv3491	null	HYPOTHETICAL PROTEIN	1,2960000000	0,0091172820
Rv3033	null	HYPOTHETICAL PROTEIN	1,2820000000	0,0138808860
Rv0660c	null	CONSERVED HYPOTHETICAL PROTEIN	1,2740000000	0,0092489505
Rv0492A	null	HYPOTHETICAL PROTEIN	1,2720001000	0,0475370660
Rv3651	null	CONSERVED HYPOTHETICAL PROTEIN	1,1560000000	0,0437578900
Rv2295	null	CONSERVED HYPOTHETICAL PROTEIN	1,1540000000	0,0022910296
Rv2827c	null	HYPOTHETICAL PROTEIN	1,1420000000	0,0052927514
Rv2307c	null	CONSERVED HYPOTHETICAL PROTEIN	1,1240000000	0,0055699780
Rv3172c	null	HYPOTHETICAL PROTEIN	1,1100000000	0,0006534284
Rv3705c	null	CONSERVED HYPOTHETICAL PROTEIN	1,0779998000	0,0129254665
Rv0692	null	CONSERVED HYPOTHETICAL PROTEIN	1,0639999000	0,0024029075
Rv1351	null	HYPOTHETICAL PROTEIN	1,0419999000	0,0491757020
Rv1137c	null	HYPOTHETICAL PROTEIN	1,0379999000	0,0281234070
Rv0921	null	POSSIBLE RESOLVASE	1,0420000000	0,0262041930
Rv1398c	null	CONSERVED HYPOTHETICAL PROTEIN	1,3540001000	0,0073955134
Rv2632c	null	CONSERVED HYPOTHETICAL PROTEIN	1,0380000000	0,0088302360
Rv3230c	null	HYPOTHETICAL OXIDOREDUCTASE	1,9740000000	0,0000527216
Rv2709	null	PROBABLE CONSERVED TRANSMEMBRANE PROTEIN	1,0239999000	0,0163064380
Rv2784c	lppU	PROBABLE LIPOPROTEIN LPPU	1,7540001000	0,0014788650
Rv0347	null	PROBABLE CONSERVED MEMBRANE PROTEIN	4,0820000000	0,0467082750
Rv2203	null	POSSIBLE CONSERVED MEMBRANE PROTEIN	1,6559999000	0,0077802250
Rv0361	null	PROBABLE CONSERVED MEMBRANE PROTEIN	1,3060000000	0,0391392260
Rv0451c	mmpS4	PROBABLE CONSERVED MEMBRANE PROTEIN MMPS4	1,1960000000	0,0263788500
Rv2473	null	POSSIBLE ALANINE AND PROLINE RICH MEMBRANE	4,3720000000	0,0255281200
Rv1227c	null	PROBABLE TRANSMEMBRANE PROTEIN	4,8480005000	0,0262166970
Rv0846c	null	PROBABLE OXIDASE	1,4000000000	0,0491627600
Rv2142c	null	HYPOTHETICAL PROTEIN	1,6440000000	0,0239039940
Rv2865	null	CONSERVED HYPOTHETICAL PROTEIN	1,6719999000	0,0407064780

Rv3385c	null	CONSERVED HYPOTHETICAL PROTEIN	1,4900000000	0,0352761200
Rv0268c	null	HYPOTHETICAL PROTEIN	1,3040001000	0,0036905417
Rv0190	null	CONSERVED HYPOTHETICAL PROTEIN	1,2880001000	0,0415129660
Rv2825c	null	CONSERVED HYPOTHETICAL PROTEIN	1,5120000000	0,0035664714
Rv1375	null	CONSERVED HYPOTHETICAL PROTEIN	1,9240001000	0,0026567099
Rv1045	null	HYPOTHETICAL PROTEIN	1,7500000000	0,0435337050
Rv0094c	null	CONSERVED HYPOTHETICAL PROTEIN	1,4880000000	0,0035185246
Rv0140	null	CONSERVED HYPOTHETICAL PROTEIN	1,0419999000	0,0456067770
Rv1321	null	CONSERVED HYPOTHETICAL PROTEIN	1,1660000000	0,0192311220
Rv3526	null	POSSIBLE OXIDOREDUCTASE	1,7219999000	0,0042123970
Rv1506c	null	HYPOTHETICAL PROTEIN	1,3860000000	0,0141670050
Rv0011c	null	PROBABLE CONSERVED TRANSMEMBRANE PROTEIN	1,7000000000	0,0009175481
Rv3792	null	PROBABLE CONSERVED TRANSMEMBRANE PROTEIN	1,1720000000	0,0012328835
Rv3848	null	PROBABLE CONSERVED TRANSMEMBRANE PROTEIN	1,1100000000	0,0004727999
Rv0837c	null	HYPOTHETICAL PROTEIN	1,8579998000	0,0183960740
Rv2122c	hisE	Probable phosphoribosyl-AMP pyrophosphatase	4,7040000000	0,0134894820
Rv3038c	null	CONSERVED HYPOTHETICAL PROTEIN	1,3360001000	0,0049461230
Rv2955c	null	CONSERVED HYPOTHETICAL PROTEIN	1,1080000000	0,0062424247
Rv0848	cysK2	POSSIBLE CYSTEINE SYNTHASE A CYSK2	1,8880001000	0,0001361560
Rv1405c	null	PUTATIVE METHYLTRANSFERASE	3,7760003000	0,0004543714
Rv1536	ileS	isoleucyl-tRNA synthetase ileS	1,1640000000	0,0316622150
Rv1128c	null	CONSERVED HYPOTHETICAL PROTEIN	2,1339998000	0,0043877275
Rv1987	null	POSSIBLE CHITINASE	1,0840000000	0,0004506594
Rv1682	null	Probable coiled-coil structural protein	6,9920006000	0,0102792310
Rv2135c	null	CONSERVED HYPOTHETICAL PROTEIN	1,1020000000	0,0043446470
Rv2465c	null	PROBABLE ISOMERASE	1,9180000000	0,0144815210
Rv0467	icl	ISOCITRATE LYASE ICL (ISOCITRASE)	1,4880000000	0,0002559969
Rv1475c	acn	PROBABLE IRON-REGULATED ACONITATE HYDRATASE ACN	1,3280000000	0,0318664570
Rv0211	pckA	PROBABLE IRON-REGULATED PHOSPHOENOLPYRUVATE	1,2980000000	0,0011121067
Rv1373	null	GLYCOLIPID SULFOTRANSFERASE	1,3460000000	0,0036895233
Rv2988c	leuC	PROBABLE 3-ISOPROPYLMALATE DEHYDRATASE (LARGE	1,9740000000	0,0008293600
Rv3710	leuA	2-ISOPROPYLMALATE SYNTHASE LEUA	1,3040000000	0,0005187828
Rv3054c	null	CONSERVED HYPOTHETICAL PROTEIN	2,2860000000	0,0001299827
Rv3463	null	CONSERVED HYPOTHETICAL PROTEIN	5,0520005000	0,0275548620
Rv0791c	null	CONSERVED HYPOTHETICAL PROTEIN	2,8900000000	0,0047651730
Rv3093c	null	HYPOTHETICAL OXIDOREDUCTASE	1,4999999000	0,0291957400
Rv1360	null	PROBABLE OXIDOREDUCTASE	1,2620001000	0,0180920960
Rv2351c	plcA	PROBABLE MEMBRANE-ASSOCIATED PHOSPHOLIPASE C 1	1,0540001000	0,0229912270
Rv2962c	null	POSSIBLE GLYCOSYL TRANSFERASE	1,6140000000	0,0053382292
Rv0768	aldA	PROBABLE ALDEHYDE DEHYDROGENASE NAD DEPENDENT	2,7260000000	0,0001014203
Rv1141c	echA11	PROBABLE ENOYL-CoA HYDRATASE ECHA11 (ENOYL	1,2900001000	0,0008145196
Rv1017c	prsA	PROBABLE RIBOSE-PHOSPHATE PYROPHOSPHOKINASE PRSA	1,1159999000	0,0057678510
Rv3068c	pgmA	PROBABLE PHOSPHOGLUCOMUTASE PGMA (GLUCOSE	1,3340001000	0,0350745000
Rv3860	null	CONSERVED HYPOTHETICAL PROTEIN	1,1880000000	0,0362700340
Rv3641c	fic	POSSIBLE CELL FILAMENTATION PROTEIN FIC	1,3860000000	0,0029454948
Rv3660c	null	CONSERVED HYPOTHETICAL PROTEIN	1,6360000000	0,0028838299
Rv2643	arsC	PROBABLE ARSENIC-TRANSPORT INTEGRAL MEMBRANE	1,0060000000	0,0050495320
Rv2020c	null	CONSERVED HYPOTHETICAL PROTEIN	1,1520001000	0,0282496870
Rv0893c	null	CONSERVED HYPOTHETICAL PROTEIN	3,5440000000	0,0162738500
Rv0726c	null	CONSERVED HYPOTHETICAL PROTEIN	1,8219999000	0,0029807878
Rv1812c	null	PROBABLE DEHYDROGENASE	1,0620000000	0,0147938170
Rv3537	null	PROBABLE DEHYDROGENASE	1,6600001000	0,0473278160

Rv1533	null	CONSERVED HYPOTHETICAL PROTEIN	1,4160000000	0,0106012730
Rv0252	nirB	PROBABLE NITRITE REDUCTASE [NAD(P)H] LARGE	1,0220001000	0,0451102600
Rv0505c	serB1	POSSIBLE PHOSPHOSERINE PHOSPHATASE SERB1 (PSP)	1,2660000000	0,0075454256
Rv0075	null	PROBABLE AMINOTRANSFERASE	1,2040001000	0,0022174623
Rv0070c	glyA2	PROBABLE SERINE HYDROXYMETHYLTRANSFERASE GLYA2	1,3060000000	0,0384317600
Rv1285	cysD	PROBABLE SULFATE ADENYLYLTRANSFERASE SUBUNIT 2	3,5680000000	0,0052913300
Rv1286	cysN	PROBABLE BIFUNCTIONAL ENZYME CYSN/CYSC: SULFATE	1,5960000000	0,0213633380
Rv3269	null	CONSERVED HYPOTHETICAL PROTEIN	2,0460000000	0,0004323217
Rv0457c	null	PROBABLE PEPTIDASE	1,7800001000	0,0402233100
Rv0790c	null	HYPOTHETICAL PROTEIN	1,2360000000	0,0136375940
Rv2444c	rne	POSSIBLE RIBONUCLEASE E RNE	2,1020000000	0,0021523144
Rv3323c	moaX	PROBABLE MOAD-MOAE FUSION PROTEIN MOAX	1,0500001000	0,0432462500
Rv0624	null	CONSERVED HYPOTHETICAL PROTEIN	1,9519999000	0,0102677010
Rv0277c	null	CONSERVED HYPOTHETICAL PROTEIN	1,5560000000	0,0422337650
Rv2546	null	CONSERVED HYPOTHETICAL PROTEIN	4,7060000000	0,0172714480
Rv0656c	null	CONSERVED HYPOTHETICAL PROTEIN	1,3679999000	0,0128547335
Rv0229c	null	POSSIBLE CONSERVED MEMBRANE PROTEIN	1,5040000000	0,0358062900
Rv0960	null	CONSERVED HYPOTHETICAL PROTEIN	2,7500000000	0,0025507542
Rv0990c	null	HYPOTHETICAL PROTEIN	1,5580000000	0,0037668840
Rv2035	null	CONSERVED HYPOTHETICAL PROTEIN	1,1800000000	0,0149157260
Rv2428	ahpC	ALKYL HYDROPEROXIDE REDUCTASE C PROTEIN AHPC	1,0820000000	0,0037078145
Rv2339	mmpL9	PROBABLE CONSERVED TRANSMEMBRANE TRANSPORT	1,0600000000	0,0319401850
Rv3174	null	PROBABLE SHORT-CHAIN DEHYDROGENASE/REDUCTASE	3,8579998000	0,0415759570
Rv0035	fadD34	PROBABLE FATTY-ACID-CoA LIGASE FADD34	1,6620000000	0,0193592830
Rv2950c	fadD29	PROBABLE FATTY-ACID-CoA LIGASE FADD29	1,2060001000	0,0122336210
Rv3515c	fadD19	PROBABLE FATTY-ACID-CoA LIGASE FADD19	3,3779998000	0,0071988064
Rv1058	fadD14	PROBABLE MEDIUM CHAIN FATTY-ACID-CoA LIGASE	1,4520000000	0,0076484290
Rv3229c	null	POSSIBLE LINOLEOYL-CoA DESATURASE	3,0040002000	0,0002698200
Rv3252c	alkB	PROBABLE TRANSMEMBRANE ALKANE 1-MONOOXYGENASE	2,5580000000	0,0011407420
Rv0711	atsA	POSSIBLE ARYLSULFATASE ATSA (ARYL-SULFATE	2,0200000000	0,0056424570
Rv3696c	glpK	PROBABLE GLYCEROL KINASE GLPK (ATP:GLYCEROL	1,5200000000	0,0338677100
Rv2182c	null	1-acylglycerol-3-phosphate O-acyltransferase	1,4939998000	0,0114480570
Rv0242c	fabG4	PROBABLE 3-OXOACYL-[ACYL-CARRIER PROTEIN]	1,0200000000	0,0306828620
Rv0439c	null	PROBABLE DEHYDROGENASE/REDUCTASE	3,4320000000	0,0335095340
Rv0839	null	CONSERVED HYPOTHETICAL PROTEIN	1,8680000000	0,0171489440
Rv3177	null	POSSIBLE PEROXIDASE (NON-HAEM PEROXIDASE)	1,4260000000	0,0433292650
Rv0261c	narK3	PROBABLE INTEGRAL MEMBRANE NITRITE EXTRUSION	5,2480000000	0,0042360346
Rv0267	narU	PROBABLE INTEGRAL MEMBRANE NITRITE EXTRUSION	1,3959999000	0,0422164600
Rv2044c	null	CONSERVED HYPOTHETICAL PROTEIN	1,2560000000	0,0476049520
Rv0191	null	PROBABLE CONSERVED INTEGRAL MEMBRANE PROTEIN	1,6920000000	0,0041088373
Rv3836	null	CONSERVED HYPOTHETICAL PROTEIN	4,9460000000	0,0087738650
Rv3863	null	HYPOTHETICAL ALANINE RICH PROTEIN	1,3960000000	0,0166113660
Rv1999c	null	PROBABLE CONSERVED INTEGRAL MEMBRANE PROTEIN	1,7619998000	0,0318294540
Rv1704c	cycA	PROBABLE D-SERINE/ALANINE/GLYCINE TRANSPORTER	1,0059999000	0,0169221800
Rv2846c	efpA	POSSIBLE INTEGRAL MEMBRANE EFFLUX PROTEIN EFPA	1,0440000000	0,0429648420
Rv0655	mkl	POSSIBLE RIBONUCLEOTIDE-TRANSPORT ATP-BINDING	1,0519999000	0,0306357820
Rv0587	yrbE2A	CONSERVED HYPOTHETICAL INTEGRAL MEMBRANE PROTEIN	1,8900000000	0,0020341037
Rv2421c	nadD	PROBABLE NICOTINATE-NUCLEOTIDE	1,2080001000	0,0305196400
Rv2438c	nadE	GLUTAMINE-DEPENDENT NAD(+) SYNTHETASE NADE	1,1020000000	0,0146229840
Rv1007c	metS	PROBABLE METHIONYL-TRNA SYNTHETASE METS (MetRS)	4,1220000000	0,0357718240
Rv0913c	null	POSSIBLE DIOXYGENASE	4,1360000000	0,0369776500
Rv0654	null	PROBABLE DIOXYGENASE	2,1320000000	0,0004543661

Rv2386c	mbtI	PUTATIVE ISOCHORISMATE SYNTHASE MBTI	1,0540001000	0,0110052380
Rv3562	fadE31	PROBABLE ACYL-CoA DEHYDROGENASE FADE31	3,8140000000	0,0470123070
Rv3545c	cyp125	PROBABLE CYTOCHROME P450 125 CYP125	3,4539998000	0,0133372970
Rv2268c	cyp128	PROBABLE CYTOCHROME P450 128 CYP128	3,2200000000	0,0381670670
Rv0766c	cyp123	PROBABLE CYTOCHROME P450 123 CYP123	3,0780000000	0,0024517651
Rv1256c	cyp130	PROBA BLE CYTOCHROME P450 130 CYP130	1,7660000000	0,0245918820
Rv3139	fadE24	PROBABLE ACYL-CoA DEHYDROGENASE FADE24	1,6340001000	0,0029463875
Rv1880c	cyp140	Probable cytochrome p450 140 CYP140	1,3900001000	0,0046221830
Rv1847	null	CONSERVED HYPOTHETICAL PROTEIN	1,2680000000	0,0242139830
Rv1661	pks7	Probable polyketide synthase pks7	4,1180000000	0,0359942400
Rv2913c	null	POSSIBLE D-AMINO ACID AMINOHYDROLASE (D-AMINO	3,9700000000	0,0008458415
Rv0809	purM	PROBABLE PHOSPHORIBOSYLFORMYLGLYCINAMIDINE	1,1299999000	0,0007738266
Rv1471	trxB1	PROBABLE THIOREDOXIN TRXB1	2,0119998000	0,0162176580
Rv2633c	null	HYPOTHETICAL PROTEIN	1,4979999000	0,0097819360
Rv1261c	null	CONSERVED HYPOTHETICAL PROTEIN	1,2700001000	0,0305854590
Rv2470	glbO	POSSIBLE GLOBIN (OXYGEN-BINDING PROTEIN) GLBO	1,1280000000	0,0194265000
Rv1628c	null	CONSERVED HYPOTHETICAL PROTEIN	1,3840000000	0,0276411850
Rv3589	mutY	PROBABLE ADENINE GLYCOSYLASE MUTY	1,7180000000	0,0193217000
Rv3062	ligB	PROBABLE ATP-DEPENDENT DNA LIGASE LIGB	1,2520000000	0,0084574100
Rv0805	null	CONSERVED HYPOTHETICAL PROTEIN	1,4479998000	0,0310385320
Rv2407	null	CONSERVED HYPOTHETCAL PROTEIN	1,7440001000	0,0152358060
Rv1057	null	CONSERVED HYPOTHETICAL PROTEIN	2,1880002000	0,0124954180
Rv0981	mprA	MYCOBACTERIAL PERSISTENCE REGULATOR MRPA (TWO	1,0760000000	0,0001977385
Rv0983	pepD	PROBABLE SERINE PROTEASE PEPD (SERINE	1,0120000000	0,0158520490
Rv1461	null	CONSERVED HYPOTHETICAL PROTEIN	1,6760000000	0,0068304640
Rv0991c	null	CONSERVED HYPOTHETICAL SERINE RICH PROTEIN	1,1660000000	0,0204680380
Rv1358	null	PROBABLE TRANSCRIPTIONAL REGULATORY PROTEIN	4,4440002000	0,0355434530
Rv0586	null	PROBABLE TRANSCRIPTIONAL REGULATORY PROTEIN	2,3540000000	0,0094402790
Rv0748	null	CONSERVED HYPOTHETICAL PROTEIN	1,3680000000	0,0387452500
Rv1942c	null	CONSERVED HYPOTHETICAL PROTEIN	1,3020000000	0,0000412495
Rv1740	null	CONSERVED HYPOTHETICAL PROTEIN	1,0699999000	0,0116480890
Rv2657c	null	PROBABLE phiRv2 PROPHAGE PROTEIN	1,6779999000	0,0035889316
Rv1909c	furA	FERRIC UPTAKE REGULATION PROTEIN FURA (FUR)	1,0540000000	0,0114630650
Rv2669	null	CONSERVED HYPOTHETICAL PROTEIN	6,1020000000	0,0008530257
Rv3288c	usfY	PUTATIVE PROTEIN USFY	1,0960000000	0,0338616500
Rv3833	null	TRANSCRIPTIONAL REGULATORY PROTEIN (PROBABLY	1,2560000000	0,0039453455
Rv0329c	null	CONSERVED HYPOTHETICAL PROTEIN	1,2960000000	0,0358222280
Rv2034	null	Probable ArsR-type repressor protein	2,3720000000	0,0051365984
Rv1049	null	PROBABLE TRANSCRIPTIONAL REPRESSOR PROTEIN	2,7600000000	0,0328106880
Rv1960c	null	CONSERVED HYPOTHETICAL PROTEIN	1,5000001000	0,0009737214
Rv2912c	null	PROBABLE TRANSCRIPTIONAL REGULATORY PROTEIN	1,7720001000	0,0094100470
Rv1167c	null	PROBABLE TRANSCRIPTIONAL REGULATORY PROTEIN	1,6220000000	0,0047691045
Rv0767c	null	CONSERVED HYPOTHETICAL PROTEIN	1,5059999000	0,0121496070
Rv3173c	null	PROBABLE TRANSCRIPTIONAL REGULATORY PROTEIN	1,0979999000	0,0202513730
Rv1080c	greA	PROBABLE TRANSCRIPTION ELONGATION FACTOR GREA	1,1260000000	0,0177319700
Rv3260c	whiB2	PROBABLE TRANSCRIPTIONAL REGULATORY PROTEIN	1,3320000000	0,0139760600
Rv1994c	null	PROBABLE TRANSCRIPTIONAL REGULATORY PROTEIN	2,8640000000	0,0224456600
Rv2640c	null	POSSIBLE TRANSCRIPTIONAL REGULATORY PROTEIN	2,1080000000	0,0014744513
Rv2642	null	POSSIBLE TRANSCRIPTIONAL REGULATORY PROTEIN	2,0040002000	0,0006777422
Rv0792c	null	PROBABLE TRANSCRIPTIONAL REGULATORY PROTEIN	1,5899999000	0,0000859397
Rv1404	null	PROBABLE TRANSCRIPTIONAL REGULATORY PROTEIN	1,6580002000	0,0207536650
Rv3249c	null	POSSIBLE TRANSCRIPTIONAL REGULATORY PROTEIN	1,6860001000	0,0023542160

Rv0302	null	PROBABLE TRANSCRIPTIONAL REGULATORY PROTEIN	1,5600000000	0,0309955270
Rv3574	null	TRANSCRIPTIONAL REGULATORY PROTEIN (PROBABLY	1,3420000000	0,0110492700
Rv1990c	null	PROBABLE TRANSCRIPTIONAL REGULATORY PROTEIN	1,0520000000	0,0003500811
Rv3862c	whiB6	POSSIBLE TRANSCRIPTIONAL REGULATORY PROTEIN	2,5460000000	0,0010592861
Rv0195	null	POSSIBLE TWO COMPONENT TRANSCRIPTIONAL	1,5280001000	0,0048440653
Rv3183	null	POSSIBLE TRANSCRIPTIONAL REGULATORY PROTEIN	4,7640000000	0,0295182730
Rv0474	null	PROBABLE TRANSCRIPTIONAL REGULATORY PROTEIN	1,2280000000	0,0145311030
Rv1221	sigE	ALTERNATIVE RNA POLYMERASE SIGMA FACTOR SIGE	1,8340000000	0,0014525834
Rv2710	sigB	RNA POLYMERASE SIGMA FACTOR SIGB	1,3520000000	0,0218609070
Rv3911	sigM	POSSIBLE ALTERNATIVE RNA POLYMERASE SIGMA FACTOR	1,2320000000	0,0049574570
Rv2057c	rpmG1	Probable ribosomal protein L33	1,0760000000	0,0306494310
Rv2014	null	POSSIBLE TRANSPOSASE	2,1079998000	0,0027781986
Rv3186	null	PROBABLE TRANSPOSASE	1,4499999000	0,0030179622
Rv2168c	null	PROBABLE TRANSPOSASE	1,3240001000	0,0129644130
Rv2278	null	PROBABLE TRANSPOSASE	1,3180000000	0,0210548020
Rv1757c	null	PUTATIVE TRANSPOSASE	1,2700001000	0,0100642050
Rv2354	null	PROBABLE TRANSPOSASE	1,1700000000	0,0175929270
Rv1763	null	PUTATIVE TRANSPOSASE	1,1620001000	0,0056986916
Rv1047	null	PROBABLE TRANSPOSASE	4,6619997000	0,0186232380
Rv2167c	null	PROBABLE TRANSPOSASE	1,0059999000	0,0173367630
Rv2390c	null	CONSERVED HYPOTHETICAL PROTEIN	1,2200000000	0,0196680400
Rv3494c	mce4F	MCE-FAMILY PROTEIN MCE4F	1,0960000000	0,0411387870
Rv3746c	PE34	PROBABLE PE FAMILY PROTEIN (PE FAMILY-RELATED	5,4360000000	0,0474946050
Rv1169c	PE11	PE FAMILY PROTEIN	6,4059997000	0,0174431950
Rv1430	PE16	PE FAMILY PROTEIN	3,7600002000	0,0326953800
Rv3507	PE_PGRS53	PE-PGRS FAMILY PROTEIN	1,7920002000	0,0083099550
Rv1184c	null	POSSIBLE EXPORTED PROTEIN	1,4760001000	0,0058014924
Rv3532	PPE61	PPE FAMILY PROTEIN	7,0140000000	0,0011338698
Rv2608	PPE42	PPE FAMILY PROTEIN	5,8060007000	0,0057337177
Rv1168c	PPE17	PPE FAMILY PROTEIN	4,6840005000	0,0358359600
Rv1918c	PPE35	PPE FAMILY PROTEIN	3,3180000000	0,0485590850
Rv3251c	rubA	PROBABLE RUBREDOXIN RUBA	2,3040001000	0,0046900620
Rv0575c	null	POSSIBLE OXIDOREDUCTASE	2,5580000000	0,0169993150
Rv3254	null	CONSERVED HYPOTHETICAL PROTEIN	1,4080000000	0,0106770890
Rv2641	cadI	CADMIUM INDUCIBLE PROTEIN CADI	4,3620005000	0,0333595870
Rv0911	null	CONSERVED HYPOTHETICAL PROTEIN	1,2260000000	0,0109156445
Rv3854c	ethA	MONOOXYGENASE ETHA	1,1560000000	0,0003284019

C) DOWN-REGULATED GENES FOR CWD-MTB

Gene ID	GeneName	GeneProduct	log2 (exp/ctrl)	P-value
Rv2183c	null	CONSERVED HYPOTHETICAL PROTEIN	-1,016	0,0119879130
Rv3224A	null	CONSERVED HYPOTHETICAL PROTEIN	-1,026	0,0385778430
Rv0185	null	CONSERVED HYPOTHETICAL PROTEIN	-1,0280001	0,0006456433
Rv1945	null	CONSERVED HYPOTHETICAL PROTEIN	-1,12	0,0369352000
Rv3647c	null	CONSERVED HYPOTHETICAL PROTEIN	-1,166	0,0038619398
Rv3716c	null	CONSERVED HYPOTHETICAL PROTEIN	-1,214	0,0006704014
Rv2628	null	HYPOTHETICAL PROTEIN	-1,232	0,0006173889
Rv0999	null	HYPOTHETICAL PROTEIN	-1,71	0,0323680080
Rv0762c	null	CONSERVED HYPOTHETICAL PROTEIN	-2,344	0,0228717920
Rv2019	null	HYPOTHETICAL PROTEIN	-4,296	0,0207640680
Rv1929c	null	CONSERVED HYPOTHETICAL PROTEIN	-4,396	0,0068157273
Rv3212	null	CONSERVED HYPOTHETICAL ALANINE VALINE RICH	-1,48	0,0030496656
Rv3013	null	CONSERVED HYPOTHETICAL PROTEIN	-1,858	0,0321416850
Rv2426c	null	CONSERVED HYPOTHETICAL PROTEIN	-1,458	0,0116023540
Rv0203	null	POSSIBLE EXPORTED PROTEIN	-1,204	0,0006651075
Rv3352c	null	POSSIBLE OXIDOREDUCTASE	-3,4239998	0,0396058600
Rv2307D	null	HYPOTHETICAL PROTEIN	-1,3360001	0,0383456600
Rv3337	null	CONSERVED HYPOTHETICAL PROTEIN	-3,942	0,0263318140
Rv0204c	null	PROBABLE CONSERVED TRANSMEMBRANE PROTEIN	-1,02	0,0412125070
Rv1678	null	PROBABLE INTEGRAL MEMBRANE PROTEIN	-1,562	0,0016092135
Rv2199c	null	Possible conserved integral membrane protein	-1,0839999	0,0018240004
Rv3629c	null	PROBABLE CONSERVED INTEGRAL MEMBRANE PROTEIN	-1,6700001	0,0411003270
Rv0227c	null	PROBABLE CONSERVED MEMBRANE PROTEIN	-1,0680001	0,0190496780
Rv1841c	null	CONSERVED HYPOTHETICAL MEMBRANE PROTEIN	-1,2019999	0,0459065770
Rv3885c	null	POSSIBLE CONSERVED MEMBRANE PROTEIN	-1,3019999	0,0011369840
Rv2597	null	PROBABLE MEMBRANE PROTEIN	-1,316	0,0017260397
Rv1576c	null	Probable phiRV1 phage protein	-1,7120001	0,0024719466
Rv1581c	null	Probable phiRv1 phage protein	-3,7540002	0,0488344470
Rv2655c	null	POSSIBLE phiRv2 PROPHAGE PROTEIN	-1,042	0,0366992760
Rv2653c	null	POSSIBLE phiRv2 PROPHAGE PROTEIN	-4,328	0,0220333600
Rv2840c	null	CONSERVED HYPOTHETICAL PROTEIN	-1,016	0,0000846340
Rv0446c	null	POSSIBLE CONSERVED TRANSMEMBRANE PROTEIN	-1,188	0,0259605370
Rv0976c	null	CONSERVED HYPOTHETICAL PROTEIN	-2,96	0,0003090526
Rv2230c	null	CONSERVED HYPOTHETICAL PROTEIN	-1,2679999	0,0378105380
Rv3163c	null	POSSIBLE CONSERVED SECRETED PROTEIN	-1,138	0,0137427320
Rv2631	null	CONSERVED HYPOTHETICAL PROTEIN	-1,122	0,0248685130
Rv3422c	null	CONSERVED HYPOTHETICAL PROTEIN	-1,28	0,0026738907
Rv0342	iniA	ISONIAZID INDUCTIBLE GENE PROTEIN INIA	-3,1100001	0,0001955722
Rv2738c	null	CONSERVED HYPOTHETICAL PROTEIN	-1,534	0,0279773120
Rv3364c	null	CONSERVED HYPOTHETICAL PROTEIN	-1,426	0,0300517730

Rv0988	null	POSSIBLE CONSERVED EXPORTED PROTEIN	-1,53	0,0444183050
Rv1100	null	CONSERVED HYPOTHETICAL PROTEIN	-4,854	0,0293099640
Rv2620c	null	PROBABLE CONSERVED TRANSMEMBRANE PROTEIN	-1,094	0,0007845012
Rv0200	null	POSSIBLE CONSERVED TRANSMEMBRANE PROTEIN	-1,1400001	0,0352676770
Rv0996	null	PROBABLE CONSERVED TRANSMEMBRANE PROTEIN	-1,3300002	0,0009204758
Rv0882	null	PROBABLE TRANSMEMBRANE PROTEIN	-1,7379999	0,0110466490
Rv3843c	null	PROBABLE CONSERVED TRANSMEMBRANE PROTEIN	-1,784	0,0022234565
Rv2726c	dapF	PROBABLE DIAMINOPIMELATE EPIMERASE DAPF (DAP	-1,1600001	0,0363024600
Rv2439c	proB	PROBABLE GLUTAMATE 5-KINASE PROTEIN PROB	-1,194	0,0052199470
Rv3838c	pheA	POSSIBLE PREPHENATE DEHYDRATASE PHEA	-1,2039999	0,0162679870
Rv1652	argC	PROBABLE N-ACETL-GAMMA-GLUTAMYL-PHOSHATE	-1,394	0,0211498460
Rv2192c	trpD	Probable anthranilate phosphoribosyltransferase	-1,454	0,0026199743
Rv1606	hisI	Probable phosphoribosyl-AMP 1.6 cyclohydrolase	-1,46	0,0152305210
Rv0926c	null	CONSERVED HYPOTHETICAL PROTEIN	-1,4539999	0,0010283624
Rv3120	null	CONSERVED HYPOTHETICAL PROTEIN	-1,104	0,0012177445
Rv1886c	fbpB	SECRETED ANTIGEN 85-B FBPB (85B) (ANTIGEN 85	-1,03	0,0407914970
Rv1595	nadB	Probable L-aspartate oxidase nadB	-1,344	0,0009381376
Rv0520	null	POSSIBLE METHYLTRANSFERASE/METHYLASE (FRAGMENT)	-3,764	0,0303004890
Rv3205c	null	CONSERVED HYPOTHETICAL PROTEIN	-1,2639999	0,0001519050
Rv2004c	null	CONSERVED HYPOTHETICAL PROTEIN	-1,414	0,0130238410
Rv0584	null	POSSIBLE CONSERVED EXPORTED PROTEIN	-1,1780001	0,0232121370
Rv0315	null	POSSIBLE BETA-1.3-GLUCANASE PRECURSOR	-1,0020001	0,0005570248
Rv0525	null	CONSERVED HYPOTHETICAL PROTEIN	-2,47	0,0489820130
Rv0952	sucD	PROBABLE SUCCINYL-CoA SYNTHETASE (ALPHA CHAIN)	-1,0200001	0,0032712552
Rv1448c	tal	PROBABLE TRANSALDOLASE TAL	-1,058	0,0279467830
Rv0462	lpd	DIHYDROLIPOAMIDE DEHYDROGENASE LPD (LIPOAMIDE	-1,212	0,0037492227
Rv3002c	ilvN	PROBABLE ACETOLACTATE SYNTHASE (SMALL SUBUNIT)	-1,266	0,0338031100
Rv2987c	leuD	PROBABLE 3-ISOPROPYLMALATE DEHYDRATASE (SMALL	-1,4300001	0,0083522630
Rv2247	accD6	ACETYL/PROPIONYL-CoA CARBOXYLASE (BETA SUBUNIT)	-1,6559999	0,0000173950
Rv3079c	null	CONSERVED HYPOTHETICAL PROTEIN	-3,654	0,0052219713
Rv1604	impA	PROBABLE INOSITOL-MONOPHOSPHATASE IMPA (IMP)	-1,4560001	0,0015640813
Rv2029c	pfkB	Probable phosphofructokinase PfkB	-1,7520001	0,0003735946
Rv0975c	fadE13	PROBABLE ACYL-CoA DEHYDROGENASE FADE13	-2,086	0,0067775426
Rv0971c	echA7	PROBABLE ENOYL-CoA HYDRATASE ECHA7 (ENOYL	-1,056	0,0019506216
Rv3373	echA18	PROBABLE ENOYL-CoA HYDRATASE ECHA18 (ENOYL	-1,214	0,0229177230
Rv0860	fadB	PROBABLE FATTY OXIDATION PROTEIN FADB	-1,8279998	0,0094312270
Rv1867	null	CONSERVED HYPOTHETICAL PROTEIN	-1,228	0,0305698660
Rv1151c	null	PROBABLE TRANSCRIPTIONAL REGULATORY PROTEIN	-1,12	0,0142850205
Rv2922A	acyP	PROBABLE ACYLPHOSPHATASE ACYP (ACYLPHOSPHATE	-4,3340006	0,0177604200
Rv1481	null	PROBABLE MEMBRANE PROTEIN	-1,232	0,0006047193
Rv1209	null	CONSERVED HYPOTHETICAL PROTEIN	-1,23	0,0071871430
Rv3917c	parB	PROBABLE CHROMOSOME PARTITIONING PROTEIN PARB	-1,028	0,0005939335
Rv2719c	null	POSSIBLE CONSERVED MEMBRANE PROTEIN	-1,458	0,0288024360
Rv2313c	null	HYPOTHETICAL PROTEIN	-1,9	0,0018717193
Rv0480c	null	POSSIBLE AMIDOHYDROLASE	-1,792	0,0110009650
Rv2756c	hsdM	POSSIBLE TYPE I RESTRICTION/MODIFICATION SYSTEM	-2,638	0,0498517860
Rv3646c	topA	DNA TOPOISOMERASE I TOPA (OMEGA-PROTEIN)	-1,128	0,0008967522

Rv1894c	null	CONSERVED HYPOTHETICAL PROTEIN	-1,2679999	0,0001933939
Rv1309	atpG	PROBABLE ATP SYNTHASE GAMMA CHAIN ATPG	-1,232	0,0083388590
Rv2196	qcrB	Probable Ubiquinol-cytochrome C reductase QcrB	-1,234	0,0007268034
Rv3148	nuoD	PROBABLE NADH DEHYDROGENASE I (CHAIN D) NUOD	-1,278	0,0007294570
Rv1307	atpH	PROBABLE ATP SYNTHASE DELTA CHAIN ATPH	-1,428	0,0260137360
Rv1736c	narX	PROBABLE NITRATE REDUCTASE NARX	-1,93	0,0006569044
Rv3341	metA	PROBABLE HOMOSERINE O-ACETYLTRANSFERASE META	-1,572	0,0312760200
Rv2259	adhE2	Probable zinc-dependent alcohol dehydrogenase	-1,334	0,0011071657
Rv2228c	null	CONSERVED HYPOTHETICAL PROTEIN	-1,364	0,0282670280
Rv2625c	null	PROBABLE CONSERVED TRANSMEMBRANE ALANINE AND	-1,348	0,0005355369
Rv2782c	pepR	PROBABLE ZINC PROTEASE PEPR	-1,5459999	0,0047957290
Rv0782	ptrBb	PROBABLE PROTEASE II PTRBB [SECOND PART]	-1,044	0,0187017400
Rv0527	ccdA	POSSIBLE CYTOCHROME C-TYPE BIOGENESIS PROTEIN	-1,152	0,0294819850
Rv3866	null	CONSERVED HYPOTHETICAL PROTEIN	-1,08	0,0059190723
Rv0291	mycP3	PROBABLE MEMBRANE-ANCHORED MYCOSIN MYCP3 (SERINE	-1,6500002	0,0015720142
Rv1540	null	CONSERVED HYPOTHETICAL PROTEIN MEMBER OF	-1,14	0,0003318316
Rv3921c	null	PROBABLE CONSERVED TRANSMEMBRANE PROTEIN	-1,15	0,0000261450
Rv2031c	hspX	HEAT SHOCK PROTEIN HSPX (ALPHA-CRSTALLIN	-1,188	0,0119154380
Rv2829c	null	CONSERVED HYPOTHETICAL PROTEIN	-1,156	0,0260663260
Rv2759c	null	CONSERVED HYPOTHETICAL PROTEIN	-1,8620001	0,0257739310
Rv3408	null	CONSERVED HYPOTHETICAL PROTEIN	-1,004	0,0000158257
Rv3134c	null	CONSERVED HYPOTHETICAL PROTEIN	-1,074	0,0000935766
Rv2623	TB31.7	CONSERVED HYPOTHETICAL PROTEIN TB31.7	-1,212	0,0008039448
Rv2005c	null	CONSERVED HYPOTHETICAL PROTEIN	-1,218	0,0000322098
Rv2624c	null	CONSERVED HYPOTHETICAL PROTEIN	-1,46	0,0287703560
Rv2028c	null	CONSERVED HYPOTHETICAL PROTEIN	-1,5400001	0,0020589954
Rv0074	null	CONSERVED HYPOTHETICAL PROTEIN	-1,4120001	0,0399384600
Rv2262c	null	CONSERVED HYPOTHETICAL PROTEIN	-2,2800002	0,0278526190
Rv0308	null	PROBABLE CONSERVED INTEGRAL MEMBRANE PROTEIN	-1,5839999	0,0037447180
Rv1912c	fadB5	POSSIBLE OXIDOREDUCTASE FADB5	-1,036	0,0219821400
Rv1062	null	CONSERVED HYPOTHETICAL PROTEIN	-1,458	0,0206803440
Rv3130c	null	CONSERVED HYPOTHETICAL PROTEIN	-1,4719999	0,0007571120
Rv0637	null	CONSERVED HYPOTHETICAL PROTEIN	-1,1159999	0,0044169470
Rv2244	acpM	MEROMYCOLATE EXTENSION ACYL CARRIER PROTEIN	-1,314	0,0002308701
Rv2245	kasA	3-OXOACYL-[ACYL-CARRIER PROTEIN] SYNTHASE 1 KASA	-1,418	0,0007828503
Rv3559c	null	PROBABLE OXIDOREDUCTASE	-1,288	0,0464100240
Rv2612c	pgsA1	PROBABLE PI SYNTHASE PGSA1 (PHOSPHATIDYLINOSITOL	-1,37	0,0015585546
Rv2746c	pgsA3	PROBABLE PGP SYNTHASE PGSA3	-1,5119998	0,0013652737
Rv2246	kasB	3-OXOACYL-[ACYL-CARRIER PROTEIN] SYNTHASE 2 KASB	-1,22	0,0020923228
Rv1483	fabG1	3-OXOACYL-[ACYL-CARRIER PROTEIN] REDUCTASE FABG1	-1,294	0,0310051570
Rv2003c	null	CONSERVED HYPOTHETICAL PROTEIN	-1,4020001	0,0181519390
Rv1984c	cfp21	PROBABLE CUTINASE PRECURSOR CFP21	-1,6520001	0,0279018430
Rv0671	lpqP	POSSIBLE CONSERVED LIPOPROTEIN LPQP	-1,174	0,0027314906
Rv3131	null	CONSERVED HYPOTHETICAL PROTEIN	-1,158	0,0012834957
Rv0045c	null	POSSIBLE HYDROLASE	-1,3019999	0,0007302748
Rv1488	null	POSSIBLE EXPORTED CONSERVED PROTEIN	-1,336	0,0007046591
Rv3331	sugI	PROBABLE SUGAR-TRANSPORT INTEGRAL MEMBRANE	-1,0220001	0,0167974620

Rv0202c	mmpL11	PROBABLE CONSERVED TRANSMEMBRANE TRANSPORT	-1,38	0,0004210238
Rv2127	ansP1	Probable L-asparagine permease ansP1	-1,0379999	0,0004201238
Rv1634	null	Possible drug efflux membrane protein	-1,268	0,0022768562
Rv0849	null	PROBABLE CONSERVED INTEGRAL MEMBRANE TRANSPORT	-1,0979999	0,0052509964
Rv2994	null	PROBABLE CONSERVED INTEGRAL MEMBRANE PROTEIN	-1,14	0,0205939540
Rv1457c	null	PROBABLE UNIDENTIFIED ANTIBIOTIC-TRANSPORT	-1,2120001	0,0282740560
Rv2936	drrA	PROBABLE DAUNORUBICIN-DIM-TRANSPORT ATP-BINDING	-1,312	0,0023181264
Rv2326c	null	POSSIBLE TRANSMEMBRANE ATP-BINDING PROTEIN ABC	-1,43	0,0465797150
Rv1858	modB	PROBABLE MOLBDENUM-TRANSPORT INTEGRAL MEMBRANE	-1,9319999	0,0096019100
Rv2692	ceoC	TRK SYSTEM POTASSIUM UPTAKE PROTEIN CEOC	-1,898	0,0453415330
Rv3236c	null	PROBABLE CONSERVED INTEGRAL MEMBRANE TRANSPORT	-1,454	0,0359774570
Rv0868c	moaD2	PROBABLE MOLYBDENUM COFACTOR BIOSYNTHESIS	-4,71	0,0335977080
Rv2897c	null	CONSERVED HYPOTHETICAL PROTEIN	-4,406	0,0271953080
Rv3607c	folB	PROBABLE DIHYDRONEOPTERIN ALDOLASE FOLB (DHNA)	-1,04	0,0054211076
Rv1207	folP2	PROBABLE DIHYDROPTEROATE SYNTHASE 2 FOLP2 (DHPS	-1,108	0,0184041620
Rv2225	panB	Probable 3-methyl-2-oxobutanoate	-1,132	0,0029209147
Rv2977c	thiL	PROBABLE THIAMINE-MONOPHOSPHATE KINASE THIL	-1,184	0,0356253800
Rv1695	ppnK	Inorganic polyphosphate/ATP-NAD kinase ppnK	-1,188	0,0004534024
Rv1631	coaE	Probable dephospho-CoA kinase coaE	-1,188	0,0202554970
Rv0414c	thiE	PROBABLE THIAMINE-PHOSPHATE PYROPHOSPHORYLASE	-2,6880002	0,0417139160
Rv2981c	ddlA	PROBABLE D-ALANINE--D-ALANINE LIGASE DDLA	-1,01	0,0144020645
Rv2155c	murD	Probable UDP-N-acetylmuramoylalanine-D-glutamate	-1,382	0,0137089910
Rv0897c	null	PROBABLE OXIDOREDUCTASE	-1,742	0,0028616975
Rv3383c	idsB	POSSIBLE POLYPRENYL SYNTHETASE IDSB (POLYPRENYL	-1,072	0,0375692000
Rv1785c	cyp143	PROBABLE CYTOCHROME P450 143 CYP143	-1,2720001	0,0382382900
Rv1934c	fadE17	PROBABLE ACYL-CoA DEHYDROGENASE FADE17	-1,332	0,0190719720
Rv0972c	fadE12	PROBABLE ACYL-CoA DEHYDROGENASE FADE12	-5,298	0,0080210810
Rv3825c	pks2	PROBABLE POLYKETIDE SYNTHASE PKS2	-1,0979999	0,0000031381
Rv0356c	null	CONSERVED HYPOTHETICAL PROTEIN	-2,32	0,0144425540
Rv2939	papA5	POSSIBLE CONSERVED POLYKETIDE SYNTHASE	-1,55	0,0013869920
Rv2030c	null	CONSERVED HYPOTHETICAL PROTEIN	-1,53	0,0000511208
Rv2584c	apt	ADENINE PHOSPHORIBOSYLTRANSFERASE APT (APRT)	-1,066	0,0000265221
Rv3645	null	PROBABLE CONSERVED TRANSMEMBRANE PROTEIN	-1,392	0,0482884870
Rv3315c	cdd	PROBABLE CYTIDINE DEAMINASE CDD (CYTIDINE	-2,52	0,0157078900
Rv3711c	dnaQ	PROBABLE DNA POLYMERASE III (EPSILON SUBUNIT)	-1,314	0,0169729300
Rv1629	polA	PROBABLE DNA POLYMERASE I POLA	-2,3700001	0,0182174520
Rv0668	rpoC	DNA-DIRECTED RNA POLYMERASE (BETA' CHAIN) RPOC	-1,3659999	0,0000624637
Rv0804	null	CONSERVED HYPOTHETICAL PROTEIN	-1,6099999	0,0125399940
Rv1278	null	HYPOTHETICAL PROTEIN	-1,08	0,0262027640
Rv0937c	null	CONSERVED HYPOTHETICAL PROTEIN	-1,218	0,0460483770
Rv1020	mfd	PROBABLE TRANSCRIPTION-REPAIR COUPLING FACTOR	-1,414	0,0008459422
Rv0003	recF	DNA REPLICATION AND REPAIR PROTEIN RECF	-1,592	0,0143736070
Rv0629c	recD	PROBABLE EXONUCLEASE V (ALPHA CHAIN) RECD	-1,7060001	0,0159112980
Rv2440c	obg	PROBABLE GTP1/OBG-FAMILY GTP-BINDING PROTEIN	-1,0239999	0,0028271358
Rv2626c	null	CONSERVED HYPOTHETICAL PROTEIN	-1,5400001	0,0012364870
Rv2234	ptpA	PHOSPHOTYROSINE PROTEIN PHOSPHATASE PTPA	-1,274	0,0258564670
Rv2226	null	CONSERVED HYPOTHETICAL PROTEIN	-1,392	0,0106451080

Rv3246c	mtrA	TWO COMPONENT SENSORY TRANSDUCTION	-1,034	0,0087337320
Rv3132c	devS	TWO COMPONENT SENSOR HISTIDINE KINASE DEVS	-1,154	0,0104231110
Rv3245c	mtrB	TWO COMPONENT SENSORY TRANSDUCTION HISTIDINE	-1,298	0,0005845502
Rv3133c	devR	TWO COMPONENT TRANSCRIPTIONAL REGULATORY PROTEIN	-1,338	0,0000106344
Rv0491	regX3	TWO COMPONENT SENSORY TRANSDUCTION PROTEIN REGX3	-1,52	0,0013809270
Rv1030	kdpB	Probable Potassium-transporting P-type ATPase B	-2,486	0,0054834410
Rv1343c	lprD	PROBABLE CONSERVED LIPOPROTEIN LPRD	-1,7260001	0,0101333040
Rv3224B	null	CONSERVED HYPOTHETICAL PROTEIN	-1,1439999	0,0142969730
Rv2621c	null	POSSIBLE TRANSCRIPTIONAL REGULATORY PROTEIN	-1,0780001	0,0001375405
Rv0353	hspR	PROBABLE HEAT SHOCK PROTEIN TRANSCRIPTIONAL	-1,6459999	0,0017308932
Rv1388	mihF	PUTATIVE INTEGRATION HOST FACTOR MIHF	-1,278	0,0117648640
Rv2681	null	CONSERVED HYPOTHETICAL ALANINE RICH PROTEIN	-1,5899999	0,0299157300
Rv3420c	rimI	PROBABLE RIBOSOMAL-PROTEIN-ALANINE	-1,474	0,0042640260
Rv2838c	rbfA	PROBABLE RIBOSOME-BINDING FACTOR A RBFA (P15B	-1,408	0,0055914405
Rv1301	null	CONSERVED HYPOTHETICAL PROTEIN	-1,066	0,0159466940
Rv2372c	null	CONSERVED HYPOTHETICAL PROTEIN	-1,074	0,0068456577
Rv1010	ksgA	PROBABLE DIMETHYLADENOSINE TRANSFERASE KSGA	-1,132	0,0002773537
Rv0706	rplV	PROBABLE 50S RIBOSOMAL PROTEIN L22 RPLV	-1,018	0,0000527692
Rv0705	rpsS	PROBABLE 30S RIBOSOMAL PROTEIN S19 RPSS	-1,038	0,0056271580
Rv0719	rplF	PROBABLE 50S RIBOSOMAL PROTEIN L6 RPLF	-1,074	0,0051838257
Rv0701	rplC	PROBABLE 50S RIBOSOMAL PROTEIN L3 RPLC	-1,076	0,0019131484
Rv1643	rplT	Probable 50S ribosomal protein L20 RplT	-1,16	0,0004834301
Rv2909c	rpsP	PROBABLE 30S RIBOSOMAL PROTEIN S16 RPSP	-1,242	0,0090088770
Rv0716	rplE	PROBABLE 50S RIBOSOMAL PROTEIN L5 RPLE	-1,2440001	0,0000289776
Rv3442c	rpsI	PROBABLE 30S RIBOSOMAL PROTEIN S9 RPSI	-1,7779999	0,0037946764
Rv2889c	tsf	PROBABLE ELONGATION FACTOR TSF (EF-TS)	-1,4380001	0,0007459435
Rv2807	null	CONSERVED HYPOTHETICAL PROTEIN	-4,526	0,0171505720
Rv1149	null	POSSIBLE TRANSPOSASE	-1,246	0,0021626574
Rv1042c	null	PROBABLE IS LIKE-2 TRANSPOSASE	-1,276	0,0002479742
Rv3874	esxB	10 KDA CULTURE FILTRATE ANTIGEN ESXB (LHP)	-1,1960001	0,0000218318
Rv3887c	null	PROBABLE CONSERVED TRANSMEMBRANE PROTEIN	-1,176	0,0052459040
Rv0287	esxG	ESAT-6 LIKE PROTEIN ESXG (CONSERVED HYPOTHETICAL	-1,542	0,0005549943
Rv1197	esxK	ESAT-6 LIKE PROTEIN ESXK (ESAT-6 LIKE PROTEIN	-1,05	0,0021503156
Rv1038c	esxJ	ESAT-6 LIKE PROTEIN ESXJ (ESAT-6 LIKE PROTEIN	-1,0979999	0,0000269651
Rv2347c	esxP	PUTATIVE ESAT-6 LIKE PROTEIN ESXP (ESAT-6 LIKE	-1,136	0,0000234108
Rv0169	mce1A	MCE-FAMILY PROTEIN MCE1A	-1,082	0,0112284210
Rv0594	mce2F	MCE-FAMILY PROTEIN MCE2F	-1,356	0,0001792358
Rv0578c	PE_PGRS7	PE-PGRS FAMILY PROTEIN	-1,012	0,0045535453
Rv1087	PE_PGRS21	PE-PGRS FAMILY PROTEIN	-1,0439999	0,0221514300
Rv3514	PE_PGRS57	PE-PGRS FAMILY PROTEIN	-1,392	0,0182082840
Rv0754	PE_PGRS11	PE-PGRS FAMILY PROTEIN	-1,456	0,0072587720
Rv0977	PE_PGRS16	PE-PGRS FAMILY PROTEIN	-2,264	0,0368512800
Rv3135	PPE50	PPE FAMILY PROTEIN	-1,132	0,0003214829
Rv2356c	PPE40	PPE FAMILY PROTEIN	-1,2019999	0,0169424660
Rv3347c	PPE55	PPE FAMILY PROTEIN	-1,238	0,0129043260
Rv3136	PPE51	PPE FAMILY PROTEIN	-1,2520001	0,0260758600
Rv1790	PPE27	PPE FAMILY PROTEIN	-1,85	0,0241605660

Rv1135c	PPE16	PPE FAMILY PROTEIN	-3,6220002	0,0301976840
Rv1466	null	CONSERVED HYPOTHETICAL PROTEIN	-1,0059999	0,0015539294
Rv2054	null	CONSERVED HYPOTHETICAL PROTEIN	-2,6440003	0,0386119700

D) UP-REGULATED GENES FOR CWD-MAP

Gene ID	GeneName	GeneProduct	log2 (exp/ctrl)	P-value
MAP_4018c	null	hypothetical protein	1,3250000000	0,0003097756
MAP_3293	null	hypothetical protein	1,2300000000	0,0389095020
MAP_1855	null	hypothetical protein	1,2125000000	0,0052072555
MAP_1475	null	hypothetical protein	1,2049999000	0,0004901399
MAP_2192	null	hypothetical protein	1,1650000000	0,0257812830
MAP_0821	null	hypothetical protein	1,1600000000	0,0013624376
MAP_3336c	null	hypothetical protein	1,1450000000	0,0169299080
MAP_2168c	null	hypothetical protein	1,1275000000	0,0256943500
MAP_0147c	null	hypothetical protein	1,0875000000	0,0280036140
MAP_2087c	null	hypothetical protein	1,0625000000	0,0084692760
MAP_3836c	null	hypothetical protein	1,0225000000	0,0022187915
MAP_3329c	null	hypothetical protein	1,0024999000	0,0427571650
MAP_3412	null	hypothetical protein	1,1350000000	0,0235254970
MAP_1104c	null	hypothetical protein	1,9399999000	0,0000395607
MAP_4069c	null	hypothetical protein	1,7200000000	0,0008803263
MAP_4006	null	hypothetical protein	1,0274999000	0,0064385720
MAP_3032c	leuB	LeuB	1,3975000000	0,0005200015
MAP_0737	null	hypothetical protein	1,0225000000	0,0263212320
MAP_4256	null	hypothetical protein	1,3800000000	0,0144225160
MAP_2089c	null	hypothetical protein	1,0025000000	0,0045997790
MAP_1467c	null	hypothetical protein	1,0525000000	0,0101021410
MAP_2823	null	hypothetical protein	1,0649999000	0,0061139646
MAP_2281c	clpP	ClpP	1,5500001000	0,0005647753
MAP_4294	fadD1	FadD1_2	1,0425000000	0,0003715328
MAP_0508	null	hypothetical protein	1,1975000000	0,0000029189
MAP_1380	null	hypothetical protein	1,1150000000	0,0050414490
MAP_1605c	null	hypothetical protein	1,0675000000	0,0375219730
MAP_2098c	null	hypothetical protein	1,1150000000	0,0454260220
MAP_1236c	drrC	DrrC	1,2850000000	0,0011098846
MAP_1107	null	hypothetical protein	1,0050000000	0,0210607960
MAP_1141	ribH	RibH	1,2749999000	0,0081589960
MAP_2177c	mbtB	MbtB	1,4550000000	0,0209647860
MAP_0426c	null	hypothetical protein	1,1975000000	0,0236776120
MAP_2493c	null	hypothetical protein	1,1200000000	0,0008664495
MAP_3297c	uvrD2	UvrD2	1,1350000000	0,0011394792
MAP_1395	null	hypothetical protein	1,2100000000	0,0008069853
MAP_2296c	null	hypothetical protein	1,0024999000	0,0031827816
MAP_3561	null	hypothetical protein	1,0200000000	0,0093254310

MAP_2895	null	hypothetical protein	1,1550000000	0,0015378081
MAP_1875c	null	hypothetical protein	1,1075001000	0,0124319130
MAP_0061c	null	hypothetical protein	1,1350000000	0,0015196612
MAP_2265c	rpmA	RpmA	1,7225001000	0,0105173840
MAP_2266c	rplU	RplU	1,0600000000	0,0007521830
MAP_4169	rpmC	RpmC	1,0450000000	0,0048484313
MAP_2155	null	IS6110	1,1050000000	0,0181166660
MAP_3606	null	hypothetical protein	1,1425000000	0,0410099500
MAP_0793c	null	hypothetical protein	1,3575001000	0,0016311525
MAP_4236c	null	hypothetical protein	2,0975000000	0,0017537677
MAP_1579c	null	hypothetical protein	1,1075001000	0,0023496633

E) DOWN-REGULATED GENES FOR CWD-MAP

Gene ID	GeneName	GeneProduct	log2 (exp/ctrl)	P-value
MAP_0776c	null	hypothetical protein	-1,005	0,0163949950
MAP_2879c	null	hypothetical protein	-1,05	0,0132340220
MAP_1509	null	hypothetical protein	-1,06	0,0000104236
MAP_1943	null	hypothetical protein	-1,085	0,0024309915
MAP_2663c	null	hypothetical protein	-1,09	0,0087806685
MAP_0213	null	hypothetical protein	-1,1724999	0,0360559670
MAP_1208	null	hypothetical protein	-1,205	0,0003856488
MAP_4015	null	hypothetical protein	-1,2249999	0,0003052869
MAP_0025	null	hypothetical protein	-1,2325	0,0006041405
MAP_3481	lpqD	LpqD	-1,2350001	0,0031403801
MAP_0885c	null	hypothetical protein	-1,2549999	0,0009914439
MAP_1447c	null	hypothetical protein	-1,255	0,0476046130
MAP_2106c	null	hypothetical protein	-1,3775	0,0135337080
MAP_2675c	null	hypothetical protein	-1,5325	0,0028660337
MAP_0042	null	hypothetical protein	-1,965	0,0001442083
MAP_2062	null	hypothetical protein	-1,4150001	0,0001122075
MAP_3418	null	hypothetical protein	-1,2199999	0,0002837415
MAP_1988	null	hypothetical protein	-1,2225001	0,0018247743
MAP_3512	null	hypothetical protein	-1,7775	0,0000317610
MAP_2013c	null	hypothetical protein	-1,0025	0,0073368020
MAP_0303c	null	hypothetical protein	-1,8275001	0,0074178395
MAP_0211	glf	Glf	-2,025	0,0033515560
MAP_0847	null	hypothetical protein	-2,92	0,0002319428

MAP_3327c	null	hypothetical protein	-1,8675	0,0001514758
MAP_2126	null	hypothetical protein	-1,335	0,0003124127
MAP_0801	null	hypothetical protein	-1,875	0,0008244950
MAP_2583c	null	hypothetical protein	-2,0825	0,0003881148
MAP_3962	fadB2	FadB2	-1,1025001	0,0292443130
MAP_3778	null	hypothetical protein	-1,1850001	0,0198294050
MAP_1989c	null	hypothetical protein	-1,8649999	0,0004510158
MAP_1660	null	hypothetical protein	-1,405	0,0003453147
MAP_2458c	null	hypothetical protein	-1,2774999	0,0005472506
MAP_3090c	serB2	SerB2	-1,5975001	0,0002517507
MAP_3787	null	hypothetical protein	-2,32	0,0028831600
MAP_1835c	prcB	PrcB	-1,2225	0,0098402025
MAP_2027c	null	hypothetical protein	-1,1275	0,0003644569
MAP_3163	null	hypothetical protein	-1,2475	0,0004176187
MAP_1638c	null	hypothetical protein	-2,715	0,0009914223
MAP_3751	mmpL4	MmpL4_5	-1,035	0,0054370820
MAP_4297c	null	hypothetical protein	-1,155	0,0017870513
MAP_4095c	mmaA2	MmaA2	-1,1624999	0,0377263580
MAP_3577	fabG3	FabG3_2	-1,1975	0,0019142635
MAP_4288	lpqP	LpqP	-1,1999999	0,0005158042
MAP_2807c	null	hypothetical protein	-1,2375	0,0062905294
MAP_0044c	null	hypothetical protein	-1,2375	0,0004277426
MAP_3727	null	hypothetical protein	-1,3	0,0014558394
MAP_1730c	null	hypothetical protein	-1,2275	0,0010332863
MAP_0449	folE	FolE	-1,625	0,0013297669
MAP_0300	null	hypothetical protein	-1,78	0,0069258763
MAP_4199	adk	Adk	-1,17	0,0041625330
MAP_2284c	null	hypothetical protein	-1,025	0,0051767290
MAP_0016c	pknB	PknB	-1,0725	0,0000059455
MAP_0098c	null	hypothetical protein	-1,2175	0,0054932730
MAP_4151c	null	hypothetical protein	-1,3725001	0,0009004368
MAP_3277c	null	hypothetical protein	-1,8675001	0,0004273292
MAP_1522	null	hypothetical protein	-1,025	0,0168877360
MAP_1727	null	hypothetical protein	-1,3199999	0,0050884370
MAP_3366	null	hypothetical protein	-1,4825	0,0006289893
MAP_1382c	null	hypothetical protein	-1,0375	0,0004758845
MAP_4062c	null	hypothetical protein	-1,2650001	0,0022112501
MAP_3879c	null	hypothetical protein	-2,7549999	0,0002753518

10. ACKNOWLEDGMENTS AND FUNDING

I am grateful to the Istituto Zooprofilattico della Sardegna and in particular to Dr. Antonino Firinu and Dr. Ennio Bandino for giving me the opportunity to continue my research project in the Depatment of Nuoro. I am also grateful to Edmondo Manca for his precious support in ordering materials, to Emiliana Satta for technical advices and to Porto Conte Ricerche for Microarray scanner facility.

A special thank to the colleague Dr. Andrea Cossu and to Dr. Nadya Markova for sharing material on bacterial CWD forms.

THIS PROJECT HAS BEEN ENTIRELY FUNDED BY THE FELLOSHIP AND GRANT "YOUNG RESEARCHER" REGIONE SARDEGNA, POR SARDEGNA, FSE 2007-2013, L.R.7/2007 PROMOZIONE DELLA RICERCA SCIENTIFICA E DELL'INNOVAZIONE TECNOLOGICA IN SARDEGNA". PROJECT: CRP1_9 AND IT HAS BEEN CARRIED ON THANKS TO THE HOSPITALITY OF BOTH THE DEPARTMENT OF BIOMEDICAL SCIENCES OF SASSARI AND THE ISTITUTO ZOPPROFILATTICO DELLA SARDEGNA, DEPARTMENT OF NUORO.