VIEW POINT
Journal of Cellular Biochemistry 113:3051–3055 (2012)

Journal of Cellular Biochemistry

Mycobacterium smegmatis Genomic Characteristics Associated With its Saprophyte Lifestyle

Quanxin Long,[1] Qi Zhou,[2,3] Lei Ji,[1] Jun Wu,[1] Wen Wang,[2] and Jianping Xie[1*]

[1] *Institute of Modern Biopharmaceuticals, State Key Laboratory Breeding Base of Eco-Environment and Bio-Resource of the Three Gorges Area, School of Life Sciences, Southwest University, Chongqing 400715, P. R. China*

[2] *Chinese Academy Science, Kunming Inst Zool, State Key Lab Genet Resources & Evolution, CAS Max Planck Jr Res Grp, Kunming 650223, Yunnan, P. R. China*

[3] *Department of Integrative Biology, Center for Theoretical Evolutionary Genomics, 3060 VLSB, University of California, Berkeley, California 94720*

ABSTRACT

Tuberculosis (TB) remains a great threat to global public health. The high biosafety level III required to tackle its causative agent *Mycobacterium tuberculosis* seriously hinders the exploration of its biology and new countermeasures. *M. smegmatis* is a widely recognized good surrogate of *M. tuberculosis*, largely due to their conserved transcriptional machinery, sigma factors, and two-component systems. However, their distinct lifestyles often confound the explanation of the results. *M. tuberculosis* leads both parasitic and free life, while *M. smegmatis* is largely saprophyte. To make full advantage of this model, it is helpful to discover the genome features associated with *M. smegmatis* unique niches, such as its saprophytic life, high salt tolerance, and relative short generation time. We employed the gene ontology enrichment analysis to characterize the unique lifestyle of *M. smegmatis*. Gene ontology enrichment analysis provided 12 terms; most are relevant to the special lifestyle of *M. smegmatis*, especially the saprophytic niche, high salt tolerance adaptation, and short generation time. In-depth functional characterization of these genes will shed new lights on the genetic basis of *M. smegmatis* saprophytic life and hasten the understanding of the unique biology of *M. tuberculosis*. J. Cell. Biochem. 113: 3051–3055, 2012.

KEY WORDS: *MYCOBACTERIUM SMEGMATIS*; *MYCOBACTERIUM TUBERCULOSIS*; COMPARATIVE GENOMICS; SAPROPHYTE

Mycobacterium is a very important microbe genus consisting of both saprophytic and pathogenic species. Pathogens such as *Mycobacterium tuberculosis* [Cole et al., 1998], *M. leprae* [Cole et al., 2001], and *M. ulcerans* [Stinear et al., 2007] are the causative agents of tuberculosis (TB), leprosy, and Buruli-ulcer disease, respectively. TB remains a serious global public health threat. Slow growth and high bio-safety level requirements have greatly retarded the investigation of the biology of pathogenic *M. tuberculosis*. Fast growing nonpathogenic *Mycobacterium* is an ideal surrogate [Tyagi and Sharma, 2002]. *M. smegmatis* has been used to address the basic genetic characteristics [Pelicic et al., 1998], virulence [Bange et al., 1999], and regulatory network [Lagier et al., 1998] of pathogenic *Mycobacterium*. In silico comparisons found that both *M. tuberculosis* and *M. smegmatis* share conserved transcriptional machinery, sigma factors, and two-component systems [Tyagi and Sharma, 2002], though some disputes remain

Quanxin Long, Qi Zhou, and Lei Ji contributed equally to this work.

Biology questions should contact Prof. Jianping Xie; bioinformatics questions should contact Prof. Wen Wang.

Conflicts of interest: None.

Additional supporting information may be found in the online version of this article.

Grant sponsor: The National key infectious disease project; Grant numbers: 2008ZX10003-006, 2012ZX10003-003; Grant sponsor: Excellent PhD thesis fellowship of southwest university; Grant number: ky2011003; Grant sponsor: New century excellent talents in Universities; Grant number: NCET-11-; Grant sponsor: The National Natural Science Foundation; Grant number: 81071316; Grant sponsor: The Fundamental Research Funds for the Central Universities; Grant number: XDJK2009A003; Grant sponsor: Natural Science Foundation Project of CQ CSTC; Grant number: CSTC, 2010BB5002.

*Correspondence to: Jianping Xie, Institute of Modern Biopharmaceuticals, School of Life Sciences, Southwest University, 1 Rd Tiansheng, Beibei, Chongqing 400715, P. R. China. E-mail: georgex@swu.edu.cn

Manuscript Received: 6 March 2012; Manuscript Accepted: 14 May 2012

Accepted manuscript online in Wiley Online Library (wileyonlinelibrary.com): 24 May 2012

DOI 10.1002/jcb.24199 • © 2012 Wiley Periodicals, Inc.

[Reyrat and Kahn, 2001]. *M. smegmatis* was originally isolated from human smegma in 1885 and saprophyte, with a generation time about 3–4 h instead of 20–24 h of *M. tuberculosis* [Bercovier and Vincent, 2001] and approximately 1.7-fold larger genome than that of *M. tuberculosis* (October 2004, http://www.tigr.org). *M. smegmatis* lacks properties of pathogen, and can not enter epithelial cells and persist in professional phagocytes. Underlying genetic basis for such remarkable difference remains to be determined. We seek to find the genome hallmarks of *M. smegmatis* unique lifestyle via gene ontology enrichment analysis. All public available *Mycobacterium* genome sequences include pathogens (*M. tuberculosis*, *M. leprae*, *M. bovis*, *M. marinum*, *M. ulcerans*, and *M. abscessus*) and non-pathogens (*M. gilvum*, *M. sp JLS*, *M. sp KMS*, *M. sp MCS*, *M. vanbaalenii*, and *M. smegmatis*) were manually curated. Multi-paranoid was employed to get a *Mycobacterium* ortholog subset which represents genes common to all Mycobacterium. New genes of particular species are defined as those *M. smegmatis* genes that differ from the ortholog subset. Species-specifically enriched functions were based on "Ontologizer" defined new genes. These data are used to explain the *M. smegmatis* unique lifestyle.

MATERIALS AND METHODS

DATA COLLECTION

All genomes (*M. tuberculosis* H37Rv, *M. leprae*, *M. bovis*, *M. marinum*, *M. ulcerans*, and *M. abscessus*, *M. gilvum*, *M. sp* JLS, *M. sp* KMS, *M. sp* MCS, *M. vanbaalenii*, and *M. smegmatis*) are freely downloaded from ftp.ncbi.nih.gov.

GENE ONTOLOGY ENRICHMENT ANALYSIS

"InParanoid" was downloaded from http://inparanoid.sbc.su.se/cgi-bin/index.cgi [Remm et al., 2001; O'Brien et al., 2005], to identify genes in two Mycobacterium species that directly evolved from a single gene in the last common ancestor are most likely to share the function, which called orthologs. "Multiparanoid" was downloaded from http://multiparanoid.sbc.su.se/ [Alexeyenko et al., 2006], which read the output from InParanoid and built multi-species clusters from these. Orthologs shared by multiple proteomes are derived from multi-paranoid. New genes of particular species are defined as those differ from the ortholog subset. "Ontologizer" was downloaded from http://compbio.charite.de/contao/index.php/ontologizer2.html [Bauer et al., 2008], analysis generated new genes of each species, the color coding indicated the enrichment of a certain term ($P > 0.05$), and the intensity of the color correlated with the significance of the enrichment. The workflow of the gene ontology enrichment analysis is shown as Figure 1. All software were ran with default settings.

RESULTS AND DISCUSSION

Twelve color coding terms indicate that they are significantly enriched in contrast with their cognates (Fig. S1). The detailed information of the enriched terms is listed in Table I.

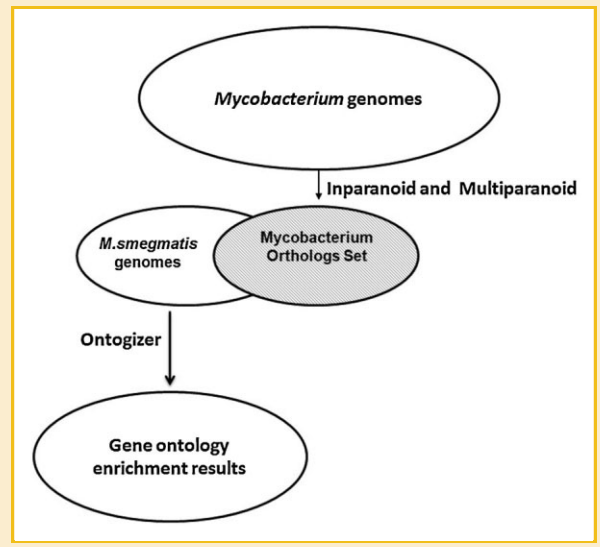

Fig. 1. The workflow of gene ontology enrichment analysis. InParanoid and Multiparanoid were employed to get a *Mycobacterium* ortholog subset which represents genes common to all Mycobacterium. New genes of particular species are defined as those *M. smegmatis* genes that differ from the ortholog subset. Species-specifically enriched functions were based on "Ontologizer" classified new genes.

M. smegmatis is a saprophyte, originally isolated from human smegma in 1885. The adaptation of *M. smegmatis* to saprophytic niche is the most distinct feature. Our data present firstly a rational explanation for this phenomenon upon the gene ontology enrichment analysis.

TOXIC SUBSTRATE METABOLISM

To thrive on saprophytic niche, microbes must metabolize whatever available, even the toxic materials, such as organic nitrogen compounds (amides, amidines, and nitrates). The enrichment of *M. smegmatis* GO:0016810 might serve this end. GO:0016810 belongs to hydrolase cut a carbon-nitrogen bonds instead of peptide bonds. The nitrogen compounds that contain carbon–nitrogen bonds instead of peptide bonds are linear or cyclic amides, amidines, and nitriles. There are 36 genes belonging to this GO in

TABLE I. The Enrichment GO Number of *Mycobacterium smegmatis*

GO number	Annotation
GO: 0018189	Pyrroloquinoline quinone biosynthetic process
GO: 0031975	Envelope
GO: 0044462	External encapsulating structure part
GO: 0016810	Hydrolase activity, acting an carbon-nitrogen bonds, other than peptide bonds
GO: 0044262	Cell carbohydrate metabolic process
GO: 0010468	Regulation of gene expression
GO: 0010556	Regulation of macromolecule biosynthesis process
GO: 0003700	Transcription factor activity
GO: 0006350	Transcription
GO: 0019219	Regulation of nucleobase nucleoside, nucleotide and nucleic acid metabolic process
GO: 0016070	RNA metabolic process
GO: 0034961	Cellular biopolymer biosynthetic process

M. smegmatis, with the total gene number of this GO in Mycobacterium orthologs set is 110. This implicates that *M. smegmatis* has reinforced its C–N bonds hydrolyzing capability beyond peptide bonds to fit surrounding environments. This number is amazing while other mycobacterium barely harbors genes within this GO (Table II). Enrichment genes of *M. smegmatis* belong to GO:0016810 were listed in Table S1.

SHORTER GENERATION TIME

Generation time variation is another salient feature need to be addressed to fully appreciate the usefulness of *M. smegmatis* as a surrogate for *M. tuberculosis*, which is about 3–4 h instead of the 20–24 h of *M. tuberculosis*. Many factors contribute to the duration of generation time. The transcription of rRNA (*rrn*) operon is the rate-determining step of novel protein synthesis, the slow growing pathogens *M. tuberculosis* and *M. leprae* have a single *rrn* operon, while most fast growing Mycobacterium have at least two *rrn* operons [Liesack et al., 1991; Domenech et al., 1994; Menendez et al., 2002]. In addition, the longer distance of *M. tuberculosis rrn* operon to its *ori*C [Cole et al., 1998] might have a role in slow growth. The capability to utilize carbon source might also have a profound effect on this. There are at least three GO enrichments in *M. smegmatis* relevant to carbon metabolism.

GO: 0044262 contains carbohydrate metabolism genes. Carbohydrates are preferable short-term fuel for organisms due to more readily metabolism than fats or amino acid. Therefore, enhanced carbohydrate metabolism of *M. smegmatis* serves to guarantee sufficient energy for fast-grow. Transposon site hybridization (TraSH) is a powerful tool to identify *M. tuberculosis* genes essential for survival during infection. Mutation of several genes predicted to metabolize carbohydrates and several disaccharide importers at the early infection phase were found by TraSH to be able to limit the growth at the early of infection [Sassetti and Rubin, 2003]. Genes within this GO of *M. smegmatis* not only involve in glycolysis and citric acid cycle, but also other carbohydrate, such as pentoses, arabinose, tagatose, xylose, rhamnose, and triose.

M. smegmatis enriched GO:0018189 handles with the pyrroloquinoline quinone (PQQ) biosynthetic process. PQQ is a redox cofactor, enzymes containing PQQ are quinoproteins including dehydrogenases, oxidases, oxygenases, hydratases, and decarboxylases [Duine and Jongejan, 1989]. *M. smegmatis* possesses a full complements of PQQ biosynthesis genes, providing enough PQQ as cofactor binding to corresponding apoenzymes that include glucose dehydrogenase (GDH), a sensor of external carbohydrate concentration and be regarded as a star molecule of glucose metabolism. *M. smegmatis* produces enough PQQ to furnish GDH with full potential to utilize the energy substance-glucose required for fast growing. Nitriles hydratase is also a quinoprotein, and the enough PQQ guarantee the enzyme activity, ensure the enhanced nitriles degradation ability. PQQ also has a growth stimulating effect by the reduction of the lag time, these results indicated that PQQ must have an important role in the initiation of cell reproduction [Ameyama et al., 1988]. It is remarkable that no other Mycobacterium species harbored enrichment genes belongs to this GO (Table II).

In addition to carbon metabolism, *M. smegmatis* genome also harbors genes for nitrogen metabolism to support its fast-growth. GO: 0031975 and GO: 0044462 have nearly the same genes that most of which are amino acid transporters, including glutamine, ectione, and function unknown transporters. Glutamine is a source of nitrogen for the synthesis of purines, pyrimidines, amino acids, glucosamine, and benzoate [Reitzer, 2003], and it can be regarded as a universal ''currency'' in nitrogen compounds biosynthesis. Enhanced glutamine transport might implicate *M. smegmatis* possess robust glutamine capture ability of up taking ready-made glutamine from external media. To accelerate growth, *M. smegmatis* genome is evolutionally geared to handle diverse substrates.

Taken together, the remarkable feature of the carbohydrates metabolism and amino acid acquisition ability of *M. smegmatis*, can meet the energy demand for fast growth. The enrichment gene list of these GO was provided in table S2.

TABLE II. The Distribution of Enrichment GO in Different Mycobacterium Species

	GO:0016810	GO:0044262	GO:0018189	GO:0031975	GO:0010468 GO:0003700 GO:0010556 GO:0034961 GO:0006350 GO:0016070 GO:0019219
M. tuberculosis H37Rv	0	0	0	0	0
M. tuberculosis H37Ra	0	0	0	0	2
M. tuberculosis F11	0	0	0	0	0
M. bovis BCG	3	0	0	0	2
M. Bovis	0	0	0	0	1
M. tuberculosis CDC1551	0	0	0	0	2
M. lepare	0	0	0	0	2
M. ulcer	0	0	0	0	0
M. marinum	1	0	0	0	2
M. avium	0	3	0	0	13
M. subavium	0	2	0	0	7
M. abscessus	5	6	0	3	53
M. vanbaalenii	5	1	0	6	36
M. gilvum	4	5	0	0	16
M. JLS	2	1	0	2	14
M. KMS	0	0	0	0	1
M. MCS	0	0	0	0	0
M. smegmatis	36	19	5	19	101

TABLE III. The Blastp Result of Randomly Chosen 10 Proteins Belonging to Enrichment GO: 0010468 in *M. smegmatis*

Gene ID	Homologus in other bacteria	Function annotation
MSMEG_2682	*Kineococcus radiotolerans* Krad_0551	GntR family transcriptional regulator
MSMEG_6700	*Thermobispora bispora* Tbis_1733	Regulatory Protein Gntr
MSMEG_2368	*Rhodococcus erythropolis* RER_04170	TetR family transcriptional regulator
MSMEG_4964	*Nocardiopsis dassonvillei* Ndas_1233	TetR family transcriptional regulator
MSMEG_6300	*Pseudonocardia dioxanivorans* Psed_4176	GntR family transcriptional regulator
MSMEG_2807	*Streptomyces hygroscopicus*	Two-component system response regulator
MSMEG_5575	*Rhodococcus jostii* RHA1_ro02808	Transcriptional regulator
MSMEG_1082	*Rhodococcus equi* REQ_11490	LuxR family transcriptional regulator
MSMEG_0120	*Nocardia brasiliensis*	LuxR family transcriptional regulator
MSMEG_1391	*Rhodococcus imtechensis*	LysR family transcriptional regulator

Genes are reported as close homologues if their corresponding protein sequences show more than 30% identity over 80 % of the length of the protein.

SALT RESISTANCE

Bacteria have evolved complex stress management strategies to sense and response to the external environment [Sleator and Hill, 2002]. *M. smegmatis* can tolerate higher salt concentration than pathogenic species of the very same genus [Kubica Vas et al., 1970]. Intracellular accumulation of compatible solutes as a strategy for adaptation to high environmental osmolality, and the enrichment of ectoine transporter of *M. smegmatis* may be related to its salt resistant phenotype. Some amino acid transporters within GO: 0031975 and GO: 0044462 are responsible for ectoine transporting. Ectoine (1, 4, 5, 6-tetrahydro-2-methyl-4-pyrimidinecarboxylic acid) is a subset of neutral non-ionic organic molecules of low molecular mass, called compatible solutes [Bursy et al., 2007]. Most halophilic eubacteria, yeasts, fungi, algae, and plants accumulate compatible solutes to maintain the osmotic equilibrium to confer resistance toward salt and temperature stress. Exogenous ectoine can stimulate growth of *Escherichia coli* in media of inhibitory salinity condition [Jebbar et al., 1992]. Ectoine is readily available from saprophytic niche [Mauchline et al., 2006]. And the genome of *M. smegmatis* posses two ectoine transport systems (MSMEG_3547-3548-3549-3550, MSMEG_5368-5369-5370-5371) that may endow this specie high level ectione transport ability. It is presumably that the enhanced ectoine transportation contributing to *M. smegmatis* high salt concentration resistance. Further experiments are needed to confirm these bioinformatic results.

TRANSCRIPTION FACTOR

It is logical that *M. smegmatis* genome also enriches transcriptional factors crucial for the adaptation, such as GO: 0003700: Transcription factor activity, GO: 0010468: Regulation of gene expression, GO: 0010556: Regulation of macromolecule biosynthesis process, GO: 0034961: Cellular biopolymer biosynthetic process, GO: 0006350: transcription, GO: 0016070: RNA metabolic process, GO: 0019219: Regulation of nucleobase nucleoside, nucleotide, and nucleic acid metabolic process. These GO have almost the same genes, most of which are TetR family protein, GntR-family protein transcriptional regulator, LacI-family protein transcriptional regulator, LysR family protein, MarR family protein, Two-component regulator, IclR family protein transcriptional regulator, HTH-type transcriptional regulator DegA, AsnC-family protein transcriptional regulator, sigma factor, DeoR-family protein transcriptional regulator, and LuxR family protein. These transcriptional factors are superfluous compare other cogants, may contribute for its special niche adaptation. Randomly choose 10 proteins in this GO to process blastp, the best blastp hits were predominantly with *Rhodococcus*, *Nocardia*, *Streptomyces*, etc. [Bentley et al., 2002; Eschbach et al., 2003], revealed genes in these genes were acquired from the environment by lateral gene transfer (LGT) from other saprophytic bacterial (Table III). Further investigation into the regulatory role of these transcriptional factors will help to elucidate their role.

CONCLUSION

Several genomic hallmarks representative of the saprophyte lifestyle of *M. smegmatis* have been found by gene ontology enrichment analysis, and some special physiological characteristics were also deciphered, such as the high salt resistance and shorter generation time. These basic physiological differences between *M. smegmatis* and pathogenic *Mycobacterium* should be cautioned for *M. smegmatis* as a model to address the pathogenic properties of *Mycobacterium*.

AUTHORS' CONTRIBUTION

Prof. Jianping Xie and Prof. Wen Wang designed research; Quanxin Long, Qi Zhou and Lei Ji performed research; Quanxin Long and Jun Wu analyzed data; Quanxin Long, Lei Ji, and Jianping Xie wrote the paper.

REFERENCES

Alexeyenko A, Tamas I, Liu G, Sonnhammer EL. 2006. Automatic clustering of orthologs and inparalogs shared by multiple proteomes. Bioinformatics 22:e9–e15.

Ameyama M, Matsushita K, Shinagawa E, Hayashi M, Adachi O. 1988. Pyrroloquinoline quinone: Excretion by methylotrophs and growth stimulation for microorganisms. Biofactors 1:51–53.

Bange FC, Collins FM, Jacobs WR Jr. 1999. Survival of mice infected with Mycobacterium smegmatis containing large DNA fragments from Mycobacterium tuberculosis. Tuber Lung Dis 79:171–180.

Bauer S, Grossmann S, Vingron M, Robinson PN. 2008. Ontologizer 2.0–a multifunctional tool for GO term enrichment analysis and data exploration. Bioinformatics 24:1650–1651.

Bentley SD, Chater KF, Cerdeno-Tarraga AM, Challis GL, Thomson NR, James KD, Harris DE, Quail MA, Kieser H, Harper D, Bateman A, Brown S, Chandra G, Chen CW, Collins M, Cronin A, Fraser A, Goble A, Hidalgo J, Hornsby T, Howarth S, Huang CH, Kieser T, Larke L, Murphy L, Oliver K, O'Neil S, Rabbinowitsch E, Rajandream MA, Rutherford K, Rutter S, Seeger K, Saunders D, Sharp S, Squares R, Squares S, Taylor K, Warren T, Wietzorrek A, Woodward J, Barrell BG, Parkhill J, Hopwood DA. 2002. Complete genome sequence of the model actinomycete Streptomyces coelicolor A3(2). Nature 417:141–147.

Bercovier H, Vincent V. 2001. Mycobacterial infections in domestic and wild animals due to Mycobacterium marinum, M. fortuitum, M. chelonae, M. porcinum, M. farcinogenes, M. smegmatis, M. scrofulaceum, M. xenopi, M. kansasii, M. simiae and M. genavense. Rev Sci Tech 20:265–290.

Bursy J, Pierik AJ, Pica N, Bremer E. 2007. Osmotically induced synthesis of the compatible solute hydroxyectoine is mediated by an evolutionarily conserved ectoine hydroxylase. J Biol Chem 282:31147–31155.

Cole ST, Brosch R, Parkhill J, Garnier T, Churcher C, Harris D, Gordon SV, Eiglmeier K, Gas S, Barry CE 3rd, Tekaia F, Badcock K, Basham D, Brown D, Chillingworth T, Connor R, Davies R, Devlin K, Feltwell T, Gentles S, Hamlin N, Holroyd S, Hornsby T, Jagels K, Krogh A, McLean J, Moule S, Murphy L, Oliver K, Osborne J, Quail MA, Rajandream MA, Rogers J, Rutter S, Seeger K, Skelton J, Squares R, Squares S, Sulston JE, Taylor K, Whitehead S, Barrell BG. 1998. Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. Nature 393:537–544.

Cole ST, Eiglmeier K, Parkhill J, James KD, Thomson NR, Wheeler PR, Honore N, Garnier T, Churcher C, Harris D, Mungall K, Basham D, Brown D, Chillingworth T, Connor R, Davies RM, Devlin K, Duthoy S, Feltwell T, Fraser A, Hamlin N, Holroyd S, Hornsby T, Jagels K, Lacroix C, Maclean J, Moule S, Murphy L, Oliver K, Quail MA, Rajandream MA, Rutherford KM, Rutter S, Seeger K, Simon S, Simmonds M, Skelton J, Squares R, Squares S, Stevens K, Taylor K, Whitehead S, Woodward JR, Barrell BG. 2001. Massive gene decay in the leprosy bacillus. Nature 409:1007–1011.

Domenech P, Menendez MC, Garcia MJ. 1994. Restriction fragment length polymorphisms of 16S rRNA genes in the differentiation of fast-growing mycobacterial species. FEMS Microbiol Lett 116:19–24.

Duine JA, Jongejan JA. 1989. Quinoproteins, enzymes with pyrrolo-quinoline quinone as cofactor. Annu Rev Biochem 58:403–426.

Eschbach M, Mobitz H, Rompf A, Jahn D. 2003. Members of the genus Arthrobacter grow anaerobically using nitrate ammonification and fermentative processes: Anaerobic adaptation of aerobic bacteria abundant in soil. FEMS Microbiol Lett 223:227–230.

Kubica Vas GP, Kilburn JO, Smithwick RW, Beam RE, Jones WD JR, Stottmeier KD. 1970. Differential identification of mycobacteria VI. Mycobacterium triviale Kubica sp. nov. Int J Syst Bacteriol 20:161.

Jebbar M, Talibart R, Gloux K, Bernard T, Blanco C. 1992. Osmoprotection of Escherichia coli by ectoine: Uptake and accumulation characteristics. J Bacteriol 174:5027–5035.

Lagier B, Pelicic V, Lecossier D, Prod'hom G, Rauzier J, Guilhot C, Gicquel B, Hance AJ. 1998. Identification of genetic loci implicated in the survival of Mycobacterium smegmatis in human mononuclear phagocytes. Mol Microbiol 29:465–475.

Liesack W, Sela S, Bercovier H, Pitulle C, Stackebrandt E. 1991. Complete nucleotide sequence of the Mycobacterium leprae 23 S and 5 S rRNA genes plus flanking regions and their potential in designing diagnostic oligonucleotide probes. FEBS Lett 281:114–118.

Mauchline TH, Fowler JE, East AK, Sartor AL, Zaheer R, Hosie AH, Poole PS, Finan TM. 2006. Mapping the Sinorhizobium meliloti 1021 solute-binding protein-dependent transportome. Proc Natl Acad Sci USA 103:17933–17938.

Menendez MC, Garcia MJ, Navarro MC, Gonzalez-y-Merchand JA, Rivera-Gutierrez S, Garcia-Sanchez L, Cox RA. 2002. Characterization of an rRNA operon (rrnB) of Mycobacterium fortuitum and other mycobacterial species: Implications for the classification of mycobacteria. J Bacteriol 184:1078–1088.

O'Brien KP, Remm M, Sonnhammer EL. 2005. Inparanoid: A comprehensive database of eukaryotic orthologs. Nucleic Acids Res 33:D476–D480.

Pelicic V, Reyrat JM, Gicquel B. 1998. Genetic advances for studying Mycobacterium tuberculosis pathogenicity. Mol Microbiol 28:413–420.

Reitzer L. 2003. Nitrogen assimilation and global regulation in Escherichia coli. Annu Rev Microbiol 57:155–176.

Remm M, Storm CE, Sonnhammer EL. 2001. Automatic clustering of orthologs and in-paralogs from pairwise species comparisons. J Mol Biol 314:1041–1052.

Reyrat JM, Kahn D. 2001. Mycobacterium smegmatis: An absurd model for tuberculosis? Trends Microbiol 9:472–474.

Sassetti CM, Rubin EJ. 2003. Genetic requirements for mycobacterial survival during infection. Proc Natl Acad Sci USA 100:12989–12994.

Sleator RD, Hill C. 2002. Bacterial osmoadaptation: The role of osmolytes in bacterial stress and virulence. FEMS Microbiol Rev 26:49–71.

Stinear TP, Seemann T, Pidot S, Frigui W, Reysset G, Garnier T, Meurice G, Simon D, Bouchier C, Ma L, Tichit M, Porter JL, Ryan J, Johnson PD, Davies JK, Jenkin GA, Small PL, Jones LM, Tekaia F, Laval F, Daffe M, Parkhill J, Cole ST. 2007. Reductive evolution and niche adaptation inferred from the genome of Mycobacterium ulcerans, the causative agent of Buruli ulcer. Genome Res 17:192–200.

Tyagi JS, Sharma D. 2002. Mycobacterium smegmatis and tuberculosis. Trends Microbiol 10:68–69.