

Phylogenetic Analysis and Transcriptional Expression Study of *RV2787* Gene in *Mycobacterium tuberculosis*

การวิเคราะห์วิวัฒนาการชาติพันธุ์และการศึกษาการแสดงออกของ ยีน *RV2787* ในเชื้อ *Mycobacterium tuberculosis*

Arunnee Sangka (อรุณณี สังกา) ^{1*} Akkpon Hanrin (อรรคพล หาญรินทร์)**

Dr.Wises Namwat (ดร.วิเศษ นามวาท)**

ABSTRACT

Mycobacterium tuberculosis (Mtb) is an important infectious agent and the causative agent of tuberculosis in human. Mtb is thought to infect one third of the world's population. Most of them are asymptomatic and latent infection. A group of genes involving in survival and adaptation in latent infection has been studied. *RV2787* gene was hypothesized to involve in adaptation under stress condition such as in phagosome of macrophage or granuloma of infected pulmonary tissue. This gene posses ATPase conserved domain and express in iron-dependent manner. Phylogenetic analysis revealed that *RV2787* closely related genes were found in *M. bovis* and *M. canettii* with 99 % identities. Its homologue gene in *M. leprae* was found to be longer relationship with 76 % identities. Whereas other related genes found in non-pathogenic organisms showed less than 50 % identities. This result suggested that function of *RV2787* should be specific and distinctive for pathogenic related process. Transcriptional analysis was performed to investigate the relative expression level of the *RV2787* in nutrient-, oxygen- and iron-depleted condition. The gene was expressed to a certain extent under normal growth condition and slightly decreased in nutrient-depleted growth condition. The expression was dramatically decreased in both iron-depleted condition but increased in oxygen-depleted condition. This result suggested that this gene possibly involves in adaptation and/or survival under stress condition.

¹ Correspondent author : arusan@kku.ac.th

* Student, Doctor of Philosophy Program in Biomedical Sciences, Graduate School, Khon Kaen University.

** Departments of Microbiology, Faculty of Medicine, and Research and Diagnostic Center for Emerging Infectious Disease, Khon Kaen University, Khon Kaen 40002, Thailand.

บทคัดย่อ

Mycobacterium tuberculosis เป็นเชื้อแบคทีเรียก่อโรควัณโรคในมนุษย์ ซึ่งก่อปัญหาในด้านสาธารณสุขที่สำคัญของโลก ในปัจจุบันมีประชากรโลกจำนวนถึง 1 ใน 3 ที่มีการติดเชื้อนี้โดยไม่แสดงอาการและเป็นการติดเชื้อแฝงจากการศึกษาก่อนหน้านี้ พบยีนหลายชนิดมีความเกี่ยวข้องกับการดำรงชีพและการปรับตัวในระหว่างการเจริญของเชื้อในระยะแฝง กลุ่มยีน *RV2787* มีส่วนของยีนที่แสดงคุณสมบัติเป็นเอนไซม์ *AT-Pase* และมีข้อมูลเบื้องต้นว่าการแสดงออกของยีนขึ้นอยู่กับระดับของธาตุเหล็กในสิ่งแวดล้อม จึงมีสมมุติฐานว่ายีนนี้มีความเกี่ยวข้องกับการเจริญของเชื้อภายใต้สภาวะกดดัน เช่นการเจริญในเซลล์ฟาโกโซมของมาโครฟาจหรืออาศัยอยู่ในแกรนูลโอมาของเซลล์เนื้อเยื่อปอดที่ติดเชื้อ ซึ่งเป็นกลไกที่เชื้อใช้หลบหลีกไม่ให้ถูกทำลายจากระบบภูมิคุ้มกันของโฮสต์ จากการวิเคราะห์วิวัฒนาการชาติพันธุ์ของยีน *RV2787* พบยีนที่มีความสัมพันธ์ใกล้ชิดในเชื้อ *M. bovis* และ *M. canettii* โดยมีความเหมือนร้อยละ 99 และพบยีนที่มีความสัมพันธ์ใกล้ชิดรองลงมาในเชื้อ *M. leprae* โดยมีความเหมือนร้อยละ 76 ส่วนยีนที่พบในกลุ่มเชื้อไม่ก่อโรควัณโรคมีความเหมือนน้อยกว่าร้อยละ 50 ผลการวิเคราะห์ชี้ให้เห็นว่าบทบาทของยีน *RV2787* น่าจะมีความจำเพาะและเกี่ยวข้องกับการก่อโรคของเชื้อ จากการศึกษาการแสดงออกของยีนโดยวิธี RT-PCR พบว่าในสภาวะการเจริญเติบโตปกติ ยีนมีการแสดงออกระดับหนึ่ง ในสภาวะขาดสารอาหารยีนมีระดับการแสดงออกลดลงแต่ไม่มาก ในสภาวะขาดแคลนธาตุเหล็ก พบว่ามีการแสดงออกน้อยกว่าสภาวะปกติมาก ขณะที่สภาวะขาดแคลนออกซิเจนมีการแสดงออกมากกว่าสภาวะปกติ ผลการทดลองนี้บ่งชี้ว่ายีนนี้น่าจะเกี่ยวข้องกับการปรับตัวและการดำรงชีวิตในสภาวะกดดัน

Key Words : Tuberculosis, *RV2787*, *Mycobacteria*

คำสำคัญ : วัณโรค ยีนก่อโรค

1. Introduction

Mycobacterium tuberculosis (Mtb) is a significant infectious agent of the human population and causative agent of tuberculosis. It is thought to infect one third of the world's population. Thailand ranks 18th in the list of 22 high TB burden countries. The prevalence and incidence of TB was estimated at 192 and 62 per 100,000 population, respectively, [1]. Most of the infected people are asymptomatic and latent infection. A Group of genes involve survival and adaptation during growth in latent infection has been studied. The complete genome sequence of *M. tuberculosis* H37Rv has been

determined and analyzed [2, 3]. Several genes were characterized and indicated as the potential candidates of drug targets leading to the more effective treatment and control of this organism [4]. The virulence related genes were analyzed for their functions, e.g. the carboxyl esterase *Rv3487c* play a role in resisting of acid stress during infection [5], genes involve in persistent infection in macrophage resulting in latent infection of *M. tuberculosis* [6–9], iron regulated genes, Fur-like proteins (FurA and FurB) and DtxR family (IdeR and SirR) [2].

RV2787 and *RV2788* gene of Mtb was 1764 bp and 687 bp, respectively. They were

found to be co-transcribed as an operon (data not shown). The *RV2787* encoded protein was predicted by bioinformatics analysis as an ATPase enzyme to bind and hydrolyze ATP whereas the *RV2788* encoded protein was predicted as an iron-responsive transcriptional regulator. This operon was hypothesized to involve in adaptation during the growth in low-iron or under stress condition, such as in phagosome of macrophage or granuloma in infected pulmonary tissue. This study aimed to gain further insight into phylogenetic characteristic of *RV2787* and its transcription activity under given stress conditions.

2. Material and Methodology

2.1 Phylogenetics analysis

Homologous proteins of *RV2787* were searched and retrieved by using Basic Local Alignment Search Tool (BLAST) [10]. ClustalW was used to perform multiple alignments of the proteins sequences and constructed phylogram by neighbor-joining method [11]. Phylogenetic analysis of the amino acid sequence of the *Rv2787* protein was performed by using tools and integrated database at Tuberculosis Database (TBDB) homepage (<http://www.tbdb.org/>) [12].

2.2 Bacterial strains, media and growth conditions

Mtb H37Rv was grown in Middlebrook 7H9 liquid medium (BBL) supplemented with 0.2% glycerol, 0.05% Tween 80 and 10% albumin-dextrose-sodium chloride complex (ADC supplement; Sigma) (cM7H9). An inoculum was prepared by culturing the Mtb in M7H9 medium until OD reach 0.1.

The inoculum was aliquot to 10 ml per tube. An oxygen-depleted condition was obtained by culturing Mtb H37Rv in 20 ml cM7H9 at OD of 0.05 at start point. The tube was tightly closed by screw-cap and sealed with paraffin. It was incubated at 37 °C and shaking at 200 rpm for 7 days. A methylene blue was used as oxygen level indicator. In *M. tuberculosis* was also grown in cM7H9 supplemented with 120 µM 2,2'-Bipyridyl (Sigma-Aldrich), an iron chelator, to obtain low iron growth medium. High iron growth medium was obtained by addition of 200 µM FeCl₃ to the low iron growth medium.

2.3 RNA isolation

The cultured Mtb from various conditions were cooled down to 4 °C. The cells were harvested from the culture medium by centrifugation at 5,000 g for 10 min. RNA was isolated from the cell pellets by thoroughly mixing with 1 ml of Trizol RNA isolation (Invitrogen) reagents and sonicating for 45 sec. to disrupt the bacterial cell wall. The lysate was mixed with 200 µl chloroform, vigorously shaken for 10 sec and kept at room temperature for 2 min and then centrifuged at 12,000 g for 30 min at 4 °C to separate the cell debris from the supernatant. The upper aqueous phase containing the total RNA was carefully collected into the new microcentrifuge tube and was mixed with 500 µl isopropanol. The mixture was kept at room temperature for 10 min. and centrifuged at 12,500 g for 15 min at 4 °C. After centrifugation, the RNA pellet was collected, washed once with 75% (v/v)

ethanol, and then resuspended with 50 µl RNase-free water. The RNA concentration was measured at 260 nm (OD₂₆₀) by spectrophotometer. The integrity of RNA was determined by electrophoresis in 2% agarose gel and visualization the ethidium bromide stained-gel under UV transillumination. .

2.4 Reverse transcriptase polymerase chain reaction (RT-PCR)

RT-PCR was performed using SuperScrip III one step RT-PCR (Invitrogen). RT-PCR. The amplification was done with Gene Amp PCR System 2400 (Perkin-Elmer). Forward and reverse primer sequences were 5'-CTAGGCGCCCGCG-CATACCGAA CGATTGG-3' and 5'-GCCG-GTACCCGGGTTTCGCT GTACAGTGC-3', respectively. The cycle profile was performed as follows: 1 cycle at 60°C for 15 min to synthesize cDNA, 1 cycle at 95°C for 5 min, 35 cycles at 95°C for 45 sec, 53°C for 1 min, and 72°C for 1 min. After the last cycle, elongation was performed at 72°C for 7 min. The PCR products were subjected to 2% agarose gel electrophoresis, and visualized under UV transilluminator after staining gel with ethidium bromide.

3. Results and Discussion

3.1 Phylogenetic analysis

Similarity of amino acid sequence of RV2787 protein was retrieved by using protein-protein BLAST. Twelve RV2787 proteins from different microorganisms which show similarity greater than 50 % were

selected to perform phylogenetic analysis. Only three species of Mycobacteria contained RV2787 homologue gene, i.e., *M. bovis*, *M. canetti*, and *M. leprae* (Fig. 1 and 2). *M. canettii* is a member of *M. tuberculosis* complex. It was reported as the causative agent of lymphadenitis in human. A high relationship was found among the three members of *M. tuberculosis* complex. *M. leprae* shared common ancestor with the members of *M. tuberculosis* complex. A group of nontuberculous mycobacteria showed a long- distance relationship including *M. parascrofuraceum*, *M. spp.* MCS-KMS strain and *M. tusciae*. Other genus containing a RV2787 homologue with longer distance relationship are *Corynebacterium amycolatum*, *Rhodococcus apacus* and *Nocardia brasiliensis*.

Phylogenetic analysis revealed that RV2787 closely related genes were found in *M. bovis* and *M. canettii* with 99 % identities. Its homologue gene in *M. leprae* was found to be longer relationship with 76 % identities. Whereas other related genes found in non-pathogenic organisms showed less than 50 % identities.

This result suggested that function of RV2787 should be specific and distinctive for pathogenic related processes or its lifestyles.

3.2 Transcriptional expression

The expression of RV2787 gene in oxygen-, nutrient- and iron-depleted growth condition was determined by RT-PCR (Fig. 2). The gene was expressed to a certain extent under normal growth condition in cM7H9

medium and slightly decreased in nutrient-depleted growth condition in phosphate buffer saline (PBS).

Low iron medium was achieved by addition of 120 μM 2,2'-Bipyridyl which is an iron-chelating substance. The expression during growth in low iron medium was found at a certain level which was lower than those of normal growth condition (51%). The result indicated that the expression of this gene was down-regulated in iron-depleted condition. This was confirmed by addition of 200 μM FeCl_3 into the low iron medium, recovery of the expression level of *RV2787* was found (94%). The result contradicts with a previous study of transcriptomic analysis

by microarray. The average expression level of The *RV2787* gene interpreted from six repeated dots of microarray indicated that the gene was induced in low iron growth condition [13]. In oxygen-depleted condition, the expression showed higher than those of normal growth condition (120%). From this study, the result indicated that the *RV2787* gene was expressed in either oxygen- or iron- dependent manner. Another previous study found an up-regulation of the gene in non-replicating growth condition [14]. This could be concluded that the gene might involve in adaptation and survival in stress conditions.

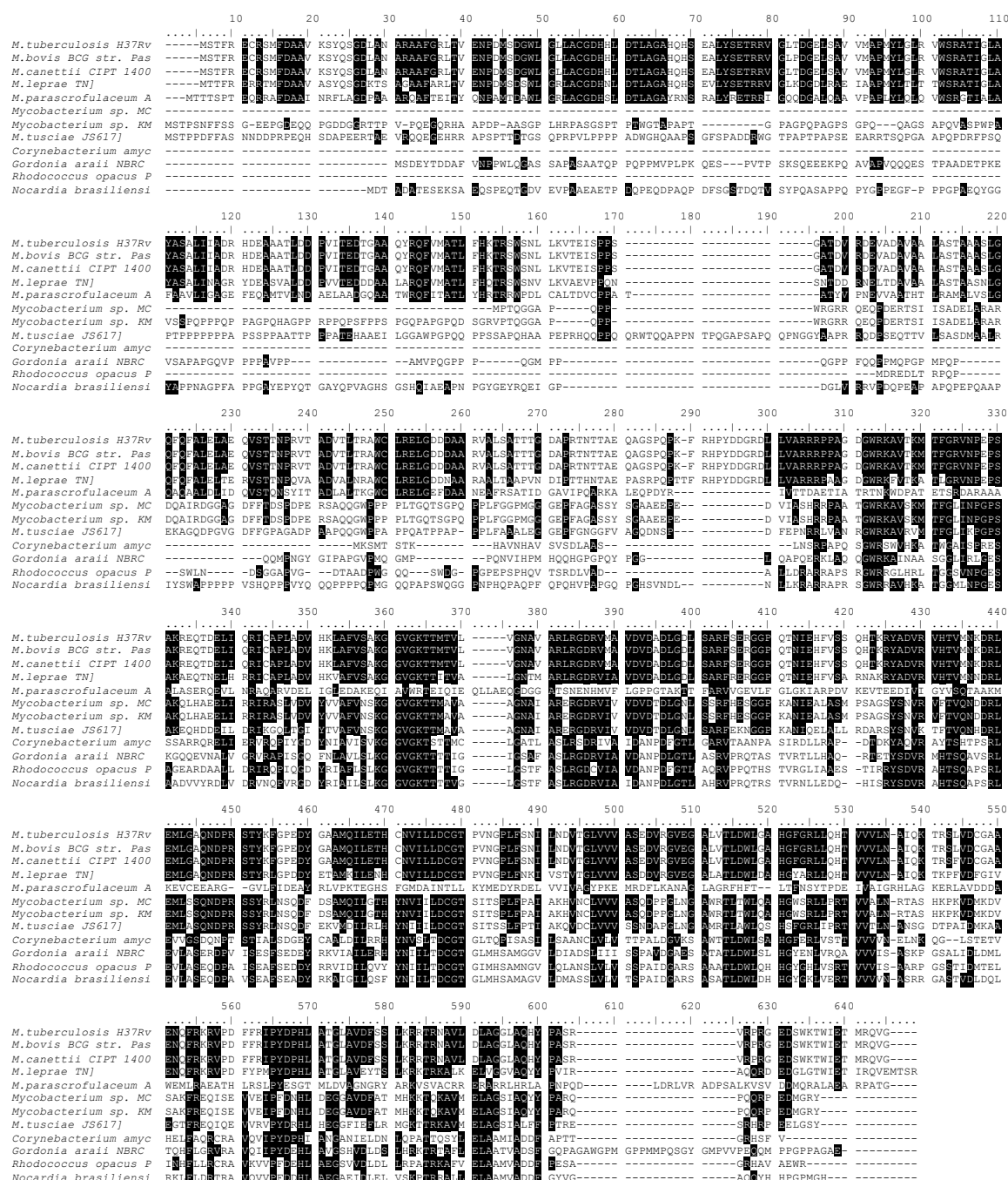


Figure 1 Multiple alignment of amino acid sequences of *Rv2787* and its homologue proteins determined by using ClustalW. Amino acids highlighted in black are conserved in at least five proteins.

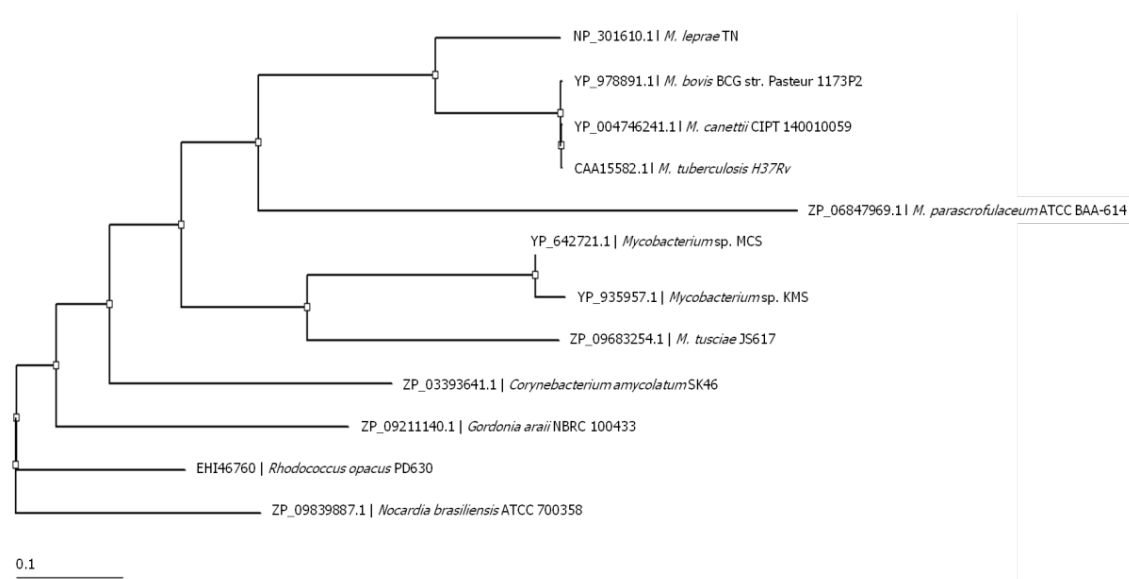


Figure 2 Phylogram based on an alignment of the amino acid sequences of *M. tuberculosis* Rv2787 and its homologue proteins. Multiple alignment was conducted by ClustalW and the phylogram was constructed by the neighbor-joining method. It was derived from 12 proteins of various strains which showed similarity score greater than 50 % (BlastP; NCBI). *M. bovis* BCG str. Pasteur 1173P2 was representative strain of *M. bovis*. The length of the branch is proportional to the amount of amino acid differences between the sequences.

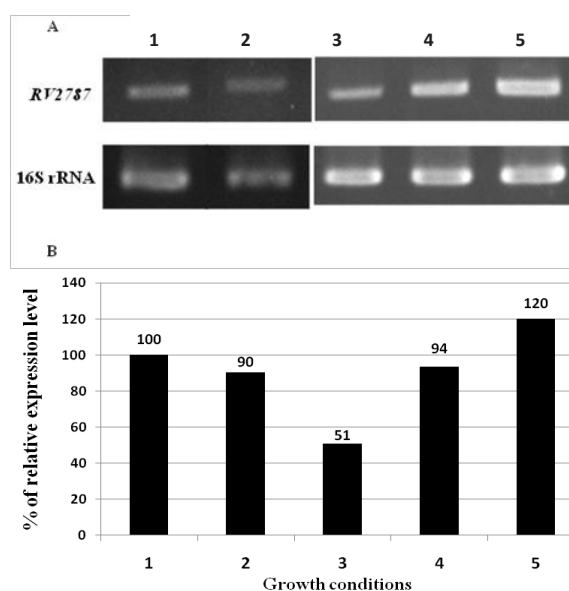


Figure 3 A : Gel electrophoresis of PCR product obtained from RT-PCR used for detection of *Rv2787* expression in O_2 depletion, nutrient starvation and in iron depletion condition; upper panel show integrity of *Rv2787* expression as follow; Lane 1: in cM7H9 [normal growth], Lane 2 : in PBS [nutrient starvation condition], Lane 3 : 2,2'-Bipyridyl treated cM7H9 [iron depletion condition], Lane 4 : 200 μ M added $FeCl_3$ in iron chelating cM7H9, Lane 5 : in cM7H9 with oxygen- depletion; lower panel: Agarose gel profile of the levels and integrity of corresponding 16S rRNA samples. **B :** The relative expression level of *Rv2787* in various growth conditions: The numbers of the growth conditions are represent as same as figure 3A.

Acknowledgements:

This research was supported by Centre for Research and Development of Medical Diagnostic Laboratory (CMDL), the Research and Diagnostic Center for Emerging Infectious Diseases (RCEID), Khon Kaen University, Thailand, and the Higher Education Research Promotion and National Research University Project.

References

1. World Health Organization. Tuberculosis Fact sheet N°104; 2010.
2. Cole ST, Brosch R, Parkhill J, Garnier T, Churcher C, Harris D, et al. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature*. 1998;393:537-44.
3. Camus JC, Pryor MJ, Medigue C, Cole ST. Re-annotation of the genome sequence of *Mycobacterium tuberculosis* H37Rv. *Microbiology*. 2002;148:2967-73.

4. Cole ST. Comparative mycobacterial genomics as a tool for drug target and antigen discovery. *Eur Respir.* 2002;20 (36_suppl):78S-86.
5. Zhang M, Wang J-d, Li Z-f, Xie J, Yang Y-p, Zhong Y, et al. Expression and characterization of the carboxyl esterase Rv3487c from *Mycobacterium tuberculosis*. *Protein Expression and Purification.* 2005 ;42(1):59-66.
6. Pieters J. Entry and survival of pathogenic mycobacteria in macrophages. *Microbes Infect.* 2001;3(3):249-55.
7. Pieters J, Gatfield J. Hijacking the host: survival of pathogenic mycobacteria inside macrophages. *Trends Microbiol.* 2002;10(3):142-6.
8. Kusner DJ. Mechanisms of mycobacterial persistence in tuberculosis. *Clinical Immunology.* 2005 ;114(3):239-47.
9. Nguyen L, Pieters J. The Trojan horse: survival tactics of pathogenic mycobacteria in macrophages. *Trends Cell Biol.* 2005;15(5):269-76.
10. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol.* 1990;215(3):403-10.
11. Thompson JD, Gibson TJ, Higgins DG. Multiple sequence alignment using ClustalW and ClustalX. *Curr Protoc Bioinformatics.* 2002;Chapter 2:Unit 2 3.
12. Reddy TB, Riley R, Wymore F, Montgomery P, DeCaprio D, Engels R, et al. TB database: an integrated platform for tuberculosis research. *Nucleic Acids Res.* 2009;37:D499-508.
13. Rodriguez GM, Voskuil MI, Gold B, Schoolnik GK, Smith I. ideR, an Essential Gene in *Mycobacterium tuberculosis*: Role of IdeR in Iron-Dependent Gene Expression, Iron Metabolism, and Oxidative Stress Response. *Infect Immun.* 2002;70(7): 3371-81.
14. Muttucumaru DG, Roberts G, Hinds J, Stabler RA, Parish T. Gene expression profile of *Mycobacterium tuberculosis* in a non-replicating state. *Tuberculosis.* 2004;84(3-4):239-46.