Systematic Pipeline for the analysis of microRNA-gene interactions in active and latent TB infection

Julia Tzu-Ya Weng[1], Kai-Yao Huang[1], Shih-Wei Lee[2], Lawrence Shih-Hsin Wu[3], Yi-Cheng Chen[4], and Tzong-Yi Lee[1]
Email: julweng@saturn.yzu.edu.tw, kaiyao.tw@gmail.com, chestman9@gmail.com, lshwu@hotmail.com, ycchen@mail.tku.edu.tw, francis@saturn.yzu.edu.tw
[1]Department of Computer Science and Engineering, Yuan Ze University, Taoyuan, Taiwan
[2]Taoyuan General Hospital, Ministry of Health and Welfare, Taoyuan, Taiwan
[3]Institute of Medical Sciences, Tzu Chi University, Hualien, Taiwan
[4]Department of Computer Science & Information Engineering, Tamkang University, New Taipei City, Taiwan

***Abstract*—Tuberculosis (TB) is the second most common cause of death from infectious diseases. About 90% of those infected are asymptomatic—the so-called latent TB infections (LTBI), with a 10% lifetime chance of progressing to active TB. Several gene expression studies have compared healthy controls versus active TB or LTBI patients. The results vary due to diverse genetic background, study designs, and the inherent complexity of the disease process. Thus, developing a sensitive and efficient method for the detection of LTBI is both crucial and challenging. Our objective was to establish an efficient and cost-effective pipeline for gene and microRNA expression profiling in TB and LTBI. We attempted to investigate the interaction between these two types of molecular signatures as biomarkers for a more sensitive and specific differentiation among active TB, LTBI, and healthy individuals. Following our systematic pipeline, we have uncovered novel differences specific to the Taiwanese population. Differentially expressed microRNAs and their interactions with the corresponding target genes will serve as potential molecular signatures to enhance our understanding of the underlying mechanisms of TB and facilitate the development for a more sensitive diagnostic assay for LTBI.**

Keywords—tuberculosis, latent infection, microRNA, gene expression

I. INTRODUCTION

Tuberculosis (TB) is an infectious disease usually caused by Mycobacterium tuberculosis (Mtb) [1]. Approximately one-third of the world's population is estimated to be latently infected with Mtb [2]. About 90% of those infected with Mtb are asymptomatic—the so-called latent TB infections (LTBI), with a 10% lifetime chance of progressing to active TB [3].

The only currently available vaccine is bacillus Calmette–Guérin (BCG), which shows decreased effectiveness after about ten years [4]. Developing a sensitive and efficient method for the detection of LTBI represents one of the major challenges in TB prevention. In primary active TB, the bacteria overcome the immune system defense and begin to multiply soon after the initial infection [1]. However, in LTBI, the bacteria remain dormant for many years before progressing to active TB. Even after treatment, there is still the risk of reactivation due to immunosuppression, or multiple-drug resistant TB bacteria [5].

To date, several studies have compared gene expression profiles between healthy individuals and active TB or LTBI patients [6-8]. These findings reveal important transcriptionally regulated markers of key biological processes, including genes involved in inflammatory responses, immune defense, cell activation, homeostatic processes, regulation of cell proliferation and apoptosis, etc. It appears that TB and LTBI share similar affected pathways, in which specific molecular markers may be able to discriminate the two disease statuses.

More recent evidence suggests the use of microRNAs as biomarkers for active TB. MicroRNAs (miRNAs) are small, noncoding, single-stranded RNAs that actively circulate in bodily fluids, regulating the expression of other genes, and thus are thought to represent a more direct indicator of altered physiology [9]. Gene and miRNA expression array study results agree on the involvement of cytokine and chemokine responses in the progression from latent infection to active disease [10-12]. Yet, many of the identified molecular markers vary due to diverse genetic background of the study population, differences in the study design, and the inherent complexity of the disease process.

With the latest advances in technology and bioinformatics, we believe utilizing complementary platforms will aid in improving the discrimination between TB and LTBI. Here, we present a systematic approach of combining gene and miRNA expression profiling to uncover the complex networks of molecular interactions associated with TB and LTBI. Our study began with the analysis of gene expression among active TB, LTBI, and healthy individuals. This was followed by a series of bioinformatics analyses to predict the corresponding regulatory miRNAs. Subsequent miRNA expression array experiment allowed us to validate our predictions, construct a comprehensive gene-miRNA interaction network, and reveal novel molecular signatures to help improve our understanding and the diagnostic differentiation of TB and LTBI.

This study was supported by the Ministry of Health and Welfare and Taoyuan General Hospital, Taiwan, and the National Science Council of the Republic of China for financially supporting this research under Contract Number of NSC 101-2628-E-155-002-MY2.

978-1-4799-2770-8/14/$31.00 ©2014 IEEE

II. MATERIALS AND METHODS

A. System Flow

Fig. 1 depicts the system flow of our study. Gene expression microarray experiment was performed to identify differentially expressed transcripts among 7 healthy control, 7 active TB, and 7 LTBI individuals. Differentially expressed candidates were categorized into up- and down-regulated genes, and divided into TB- and LTBI-specific groups. Putative miRNAs targeting the differentially expressed transcripts were predicted using five miRNA knowledge bases. Validation for our analysis was approached in two ways. Experimentally

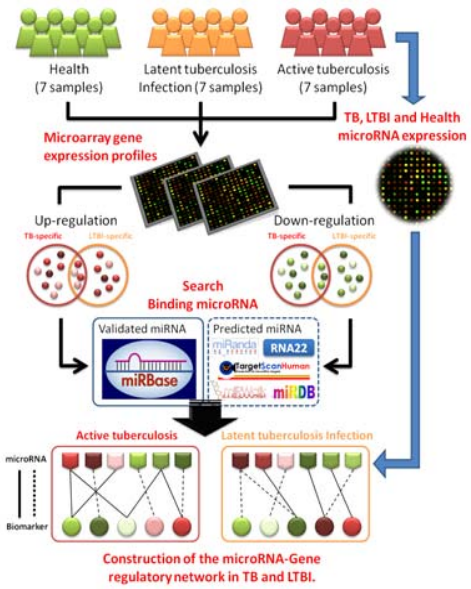

Fig. 1. System flow of our wet bench and bioinformatics analyses.

verified miRNA-gene interactions curated by miRWalk [13] represent one level of confirmation. Differentially expressed miRNAs in TB and LTBI as identified from our subsequent miRNA expression array study is the more direct validation.

B. Clinical sample collection

All procedures were reviewed and approved by the Institutional Review Board of Taoyuan General Hospital, Department of Health, Taoyuan, Taiwan. Written informed consents were obtained from all participants. Eligibility for entry into the study was based on clinical signs and symptoms of Mtb infection. LTBI subjects were recruited from close contacts with active TB patients, with positive T-SPOT TB test, and negative chest radiograph, but without clinical evidence of active TB. Healthy controls were individuals who had not been in close contacts with TB or LTBI patients and showed no clinical signs of TB or LTBI. Individuals with allergic diseases, diabetes, cancer, immune-compromised conditions, and co-infections with any types of infectious diseases were excluded.

C. RNA isolation

RNA was isolated from peripheral blood mononuclear cells. RNA quality was determined by an optical density (OD) 260/280 ratio ≥1.8, and OD 260/230 ratio ≥1.5 on a spectrophotometer and by the intensity of the 18S and 28S rRNA bands on a 1% formaldehyde-agarose gel. RNA quantity was detected by a spectrophotometer. RNA integrity was examined on an Agilent Bioanalyzer. RNA with a RNA integrity number (RIN) ≥6.0 and 28S/18S >0.7 was subjected to microarray analysis.

D. Gene expression microarray analysis

RNA samples were subjected to Human OneArray ® v6 and each gene expression profile was run in duplicates (Phalanx Biotech, Hsinchu, Taiwan). Data were analyzed using R /Bioconductor. Active TB and LTBI expression profiles were normalized against those of the healthy controls before comparisons were made between active TB and LTBI. Genes showing significant differential expression between normalized active TB and LTBI (absolute value of log2 ratio ≥ 1.5, False discovery rate or FDR <0.05) were categorized into TB-specific down- and up-regulated genes, and LTBI-specific down- and up-regulated genes. The uncentered Pearson correlation coefficient was calculated to cluster genes and disease statuses based on the relative expression of these differentially expressed genes using the hierarchical clustering algorithm in Cluster3 [14]. Differentially expressed genes were used as input for the Database for Annotation, Visualization and Integrated Discovery (DAVID) bioinformatics resource for gene ontology (GO) enrichment analysis [15] (Benjamini adjusted p<0.05).

E. microRNA arget prediction

We employed the features in miRWalk [13] to find miRNAs that may regulate the candidate genes. The cross-validation function in miRWalk was used to compare the prediction results with four other miRNA resources: miRanda [16], miRDB [17], RNA22 [18], and TargetScan [19]. A predicted miRNA was accepted as a potential regulatory candidate when four out of the five miRNA resources agreed

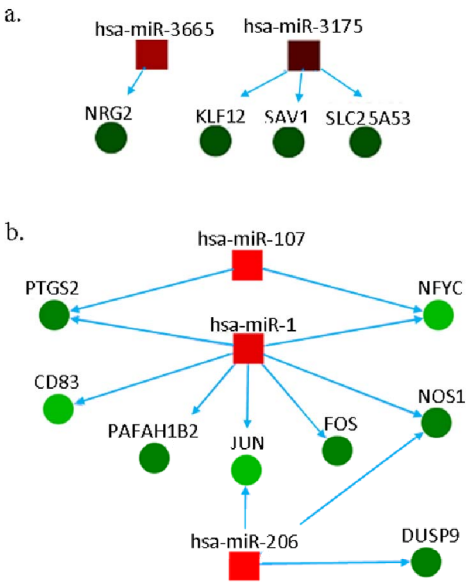

Fig. 2. Predicted and validated miRNA-gene interactions for TB and LTBI

that it may interact with the differentially expressed genes. Finally, we used miRWalk and miRTarBase [20] to further validate miRNA-target interactions.

F. MicroRNA expression analysis

To verify our computational predictions, RNA samples from the same participants were processed for microRNA array analysis. Total RNA samples were subjected to Human miRNA OneArray ® v5 and each sample was run in duplicates (Phalanx Biotech, Hsinchu, Taiwan). Data were analyzed with R/Bioconductor. Active TB and LTBI microRNA expression profiles were normalized against those of the healthy controls. MicroRNAs showing significant differential expression between normalized active TB and LTBI (absolute value of log2 ratio $\geq$0.5, FDR <0.05) were categorized into TB-specific down- and up-regulated microRNAs, and LTBI-specific down- and up-regulated microRNAs. These candidate miRNAs were analyzed for correlation with the candidate genes to reveal potential miRNA-gene interactions, in which decreased miRNA expression may be associated with increased target gene expression and vice versa.

G. MicroRNA-gene interaction visualization

To visualize the relationship between the selected candidate miRNAs and transcripts, we have built TB- and LTBI-specific interaction networks with Cytoscape [21]. We incorporated these candidates' differential expression patterns in our network construction. The interactions include predicted, as well as our microarray- and other experimentally verified miRNA-gene associations found on miRNA knowledge bases.

III. RESULTS

A. Differential gene expression profiles discriminate between TB and LTBI

In total, there are 248 and 251 specifically down-regulated genes in TB, and 64 and 144 down-regulated genes in LTBI. GO enrichment analysis indicated that the candidate genes in both TB and LTBI are associated with chemokine activity and immune response. Our findings are consistent with the pathology of TB infection, in which cytokines, chemokines, and molecules of the immune system are actively involved [22].

B. Differential miRNA expression profiles discriminate between TB and LTBI

To understand the mechanisms regulating the differentially expressed genes in TB and LTBI, we utilized five miRNA target prediction databases and cross-validated the results to find putative miRNA-gene interactions. Then, we performed miRNA expression profiling on the same TB, LTBI, and healthy individuals. There are 3 up-regulated and 8 down-regulated microRNAs specific to the TB group, and 33 up-regulated and 305 down-regulated microRNAs in LTBI. We analyzed the significance of correlation between the differentially expressed miRNAs and their target genes, based on bioinformatics predictions and their corresponding patterns of expression. Tables 1 and 2 present the association between up-regulated miRNAs and down-regulated candidate genes in TB and LTBI, respectively. Note that currently, the three up-regulated miRNAs in TB have no known target genes. Hence, the miRNA-gene relationships in TB are based on the observed correlation of expression patterns between the two types of molecules in our study.

C. TB- and LTBI-specific miRNA-gene interactions

Graphical representations of the miRNA-gene interactions for TB and LTBI are shown in Fig. 2a and 2b, respectively. Fig. 2a illustrates the up-regulated miRNAs whose expression patterns correlate with those of their predicted target genes in TB. These miRNA-gene relationships were neither predicted nor experimentally verified except in our array study. Fig. 2b diagrams the validated and predicted miRNA-gene interactions that are also observed in the LTBI expression profile in our experiment.

IV. DISCUSSION

In this study, we integrated the available gene and miRNA expression array technology and bioinformatics tools to investigate the possibility of uncovering molecular events indicative of TB and LTBI. Our system flow allowed us to assess the results of gene expression analysis before proceeding with the miRNA expression array study. Given the limited sample size (7 healthy control, 7 TB, and 7 LTBI participants), we could still differentiate between TB and LTBI patients. Thus, our systematic approach is both efficient and cost-effective.

Our findings do not correspond entirely with those of other groups, though the identified miRNA-gene associations are biologically relevant in the context of TB infection. For instance, many of the up-regulated microRNAs in LTBI are associated with established markers of TB, including chemokine CXCL1 [23], interleukin IL-8 [24], and interferon-induced protein IFIT1B [25]. Our results suggest that the regulation of these host defense genes may involve microRNAs.

Many of the differentially expressed microRNAs from our results are novel markers of LTBI and TB, but have potential roles in the disease process. For example, in LTBI, hsa-mir-16-5p is up-regulated. The microRNA has-mir-16 is a known inducer of apoptosis [26] and is associated with several tuberculosis candidate genes: SLC11A2 [27], PGM1 [28], and TAP2 [29]. Our miRNA-gene interaction network reveals more genes that may modulate the disease progression from latent to active TB infection through the interaction of hsa-mir-16-5p with genes such as nitric oxide synthase 1 (NOS1), platelet activating factor acetylhydrolase (PAFA1B2), and sorting nexin-18 (SNX18). The potential interaction between the differentially expressed miRNAs and their target genes may represent important underlying mechanisms of LTBI.

In conclusion, we have performed a complementary analysis of gene and miRNA expression profiling, establishing a comprehensive analytical pipeline to identify genes, microRNAs, as well as the interaction between these two types of molecules in TB and LTBI. Our approach has allowed us to identify new molecular interactions that may help differentiate TB from LTBI, further our understanding of TB pathogenesis,

978-1-4799-2770-8/14/$31.00 ©2014 IEEE

and facilitate the development of a molecular diagnostic platform for LTBI detection.

TABLE I. TB-SPECIFIC DOWN-REGULATED GENES AND THEIR CORRESPONDING BINDING MICRORNAS AS PREDICTED BY THE CORRELATION OF THEIR EXPRESSION PATTERNS.

Gene symbol	*Fold Change*	*microRNA*	*Fold Change*
KLF12	-1.91	hsa-miR-3175	0.55
NRG2	-1.87	hsa-miR-3665	0.64
SAV1	-1.51	hsa-miR-3175	0.55
SLC25A53	-1.65	hsa-miR-3175	0.55

TABLE II. LTBI-SPECIFIC DOWN-REGULATED GENES AND THEIR CORRESPONDING BINDING MICRORNAS AS PREDICTED BY THE CORRELATION OF THEIR EXPRESSION PATTERNS.

Gene symbol	*Fold Change*	*microRNA*	*Fold Change*
FOS	-1.77	hsa-miR-1	0.88
JUN	-2.55	hsa-miR-1 hsa-miR-206	0.88 1.04
NFYC	-2.93	hsa-miR-1 hsa-miR-107	0.88 1.09
NOS1	-2.93	hsa-miR-1 hsa-miR-206	0.88 1.04
PAFAH1B2	-2.25	hsa-miR-1	0.88
PTGS2	-1.95	hsa-miR-1 hsa-miR-107	0.88 1.09

REFERENCES

[1] V. Kumar, S.L. Robbins, Robbins basic pathology, 8th ed. Philadelphia, PA: Saunders/Elsevier, 2007.

[2] C. Dye, B.G. Williams, "The population dynamics and control of tuberculosis," Science, 328(5980), pp. 856-861, 2010.

[3] R.L. Skolnik, Global health 101, 2nd edn. Burlington, MA: Jones & Bartlett Learning, 2012.

[4] E.C. Cole, C.E. Cook, "Characterization of infectious aerosols in health care facilities: an aid to effective engineering controls and preventive strategies," American journal of infection control, 26(4), pp.453-464, 1998.

[5] S.D. Lawn, A.I. Zumla, "Tuberculosis," Lancet, 2011, 378(9785):57-72.

[6] P. Constantoulakis, E. Filiou, N. Rovina, G. Chras, A. Hamhougia, S. Karabela, A. Sotiriou, C. Roussos, N. Poulakis, "In vivo expression of innate immunity markers in patients with Mycobacterium tuberculosis infection," BMC infectious diseases, 10, pp. 243, 2010.

[7] J. Maertzdorf, D. Repsilber, S.K. Parida, K. Stanley, T. Roberts, G. Black, G. Walzl, S.H. Kaufmann, "Human gene expression profiles of susceptibility and resistance in tuberculosis," Genes and immunity, 12(1), pp. 15-22, 2011.

[8] L.M. Verhagen, A. Zomer, M. Maes, J.A. Villalba, B.D. Nogal, M. Eleveld, S.A. van Hijum, J.H. de Waard, P.W. Hermans, "A predictive signature gene set for discriminating active from latent tuberculosis in Warao Amerindian children," BMC genomics, 14, pp. 74, 2013.

[9] S. Gilad, E. Meiri, Y. Yogev, S. Benjamin, D. Lebanony, N. Yerushalmi, H. Benjamin, M. Kushnir, H. Cholakh, N. Melamed et al, "Serum microRNAs are promising novel biomarkers," PLoS One, 3(9), pp. 3148, 2008.

[10] C. Lu, J. Wu, H. Wang, S. Wang, N. Diao, F. Wang, Y. Gao, J. Chen, L. Shao, X. Weng et al, "Novel biomarkers distinguishing active tuberculosis from latent infection identified by gene expression profile of peripheral blood mononuclear cells," PLoS One, 6(8), pp. 24290, 2011.

[11] Y. Fu, Z. Yi, X. Wu, J. Li, F. Xu, "Circulating microRNAs in patients with active pulmonary tuberculosis," Journal of clinical microbiology, 49(12), pp. 4246-4251, 2011.

[12] C. Wang, S. Yang, G. Sun, X. Tang, S. Lu, O. Neyrolles, Q. Gao, "Comparative miRNA expression profiles in individuals with latent and active tuberculosis," PLoS One, 6(10), pp. 25832, 2011.

[13] H. Dweep H, C. Sticht, P. Pandey, N. Gretz, "miRWalk--database: prediction of possible miRNA binding sites by "walking" the genes of three genomes," Journal of biomedical informatics, 44(5), pp. 839-847, 2011.

[14] M.J. de Hoon, S. Imoto, J. Nolan, S. Miyano, "Open source clustering software," Bioinformatics, 20(9), pp. 1453-1454, 2004.

[15] D.W. Huang, B.T. Sherman, Q. Tan, J. Kir, D. Liu, D. Bryant, Y. Guo, R. Stephens, M.W. Baseler, H.C. Lane et al, "DAVID Bioinformatics Resources: expanded annotation database and novel algorithms to better extract biology from large gene lists," Nucleic acids research, 35, W169-175, 2007.

[16] K.C. Miranda, T. Huynh, Y. Tay, Y.S. Ang, W.L. Tam, A.M. Thomson, B. Lim, I. Rigoutsos, "A pattern-based method for the identification of microRNA binding sites and their corresponding heteroduplexes," Cell, 126(6), pp. 1203-1217, 2006.

[17] X. Wang, I.M. El Naqa, "Prediction of both conserved and nonconserved microRNA targets in animals," Bioinformatics, 24(3), pp. 325-332, 2008.

[18] L. Phillipe, I. Rigoutsos, "Interactive Exploration of RNA22 MicroRNA Target Predictions," Bioinformatics, 15, 28(24), pp. 3322-3, 2012.

[19] A. Grimson, K.K.H. Farh, W.K. Johnston, P. Garrett-Engele, L.P. Lim, D.P. Bartel, "MicroRNA targeting specificity in mammals: Determinants beyond seed pairing," Mol Cell, 27(1), pp. 91-105, 2007.

[20] S.D. Hsu, F.M. Lin, W.Y. Wu, C. Liang, W.C. Huang, W.L. Chan, W.T. Tsai, G.Z. Chen, C.J. Lee, C.M. Chiu et al, "miRTarBase: a database curates experimentally validated microRNA-target interactions," Nucleic Acids Res, 39(Database issue), D163-169, 2011.

[21] P. Shannon P, A. Markiel, O. Ozier, N.S. Baliga, J.T. Wang, D. Ramage, N. Amin, B. Schwikowski, T. Ideker, "Cytoscape: a software environment for integrated models of biomolecular interaction networks," Genome Res, 13(11), pp. 2498-2504, 2003.

[22] G. Walzl, K. Ronacher, W. Hanekom, T.J. Scriba, A. Zumla, "Immunological biomarkers of tuberculosis," Nature reviews Immunology, 11(5), pp. 343-354, 2011.

[23] D.D. Riedel, S.H. Kaufmann, "Chemokine secretion by human polymorphonuclear granulocytes after stimulation with Mycobacterium tuberculosis and lipoarabinomannan," Infection and immunity, 65(11), pp. 4620-4623, 1997.

[24] E. Ferrero, P. Biswas, K. Vettoretto, M. Ferrarini, M. Uguccioni, L. Piali, B.E. Leone, B. Moser, C. Rugarli, R. Pardi, "Macrophages exposed to Mycobacterium tuberculosis release chemokines able to recruit selected leucocyte subpopulations: focus on gammadelta cells," Immunology, 108(3), pp. 365-374, 2003.

[25] P.S. Manzanillo, M.U. Shiloh, D.A. Portnoy, J.S. Cox, "Mycobacterium tuberculosis activates the DNA-dependent cytosolic surveillance pathway within macrophages," Cell host & microbe, 11(5), pp. 469-480, 2012.

[26] J. Shen, R. Wan, G. Hu, L. Yang, J. Xiong, F. Wang, S. He, X. Guo, J. Ni, C. Guo et al, "miR-15b and miR-16 induce the apoptosis of rat activated pancreatic stellate cells by targeting Bcl-2 in vitro," Pancreatology, 12(2), pp. 91-99, 2012.

[27] X.W. Li, Y. Yang, F. Zhou, Y.Z. Zhang, H.Z. Lu, Q. Jin, L. Gao, "SLC11A1 (NRAMP1) Polymorphisms and Tuberculosis Susceptibility: Updated Systematic Review and Meta-Analysis," PLoS One, 6(1), e15831, 2011.

[28] A.S. Sergeyev, I.V. Bogadelnikova, R.K. Agapova, M.I. Perelman, "Analysis of heterozygosity at the PI, TF, PGM1, ACP1, HP, GC, GLO1, C3, and ESD loci in pulmonary tuberculosis patients differing in response to treatment," Russ J Genet, 37(12), pp. 1409-1414, 2001.

[29] E.Y. Roh, M.H. Park, P.A. Yoon, M.Y. Kim, S.Y. Kwon, K.J. Yim, S.K. Han, Y.S. Shim, "Association of TAP1 and TAP2 gene polymorphisms with tuberculosis in Koreans," Hum Immunol, 65, S35-S35, 2004.

978-1-4799-2770-8/14/$31.00 ©2014 IEEE