STATISTICAL ANALYSIS OF HUMAN TUBERCULOSIS MICROARRAY GENE EXPRESSION DATA IN THE BIOCONDUCTOR R PACKAGE

UMAR SHITTU

A dissertation submitted in partial fulfilment of the
Requirements for the award of
Master of Science (Biotechnology)

Faculty of Biosciences and Medical Engineering
Universiti Teknologi Malaysia

JANUARY 2015

DEDICATION

I dedicated this thesis to my beloved parents
Late Shittu Umar and Late Hauwa'u Lawal.

ACKNOWLEDGEMENT

Alhamdulillah, in the name of ALLAH, the most Gracious, the most Merciful, for His guidance and blessing throughout my master's programme. I am very grateful and would like to express my sincere gratitude to my supervisor Dr. Mohammed Abu Naser for his invaluable guidance, continuous encouragement and constant support in making this research possible.

I am highly grateful to my lovely wife Maryam Abdullahi and my son Muhammad Kabir Umar who gave me courage and support to achieve success in my chosen career. My sincere appreciation to all my colleagues and others who have provided assistance at various occasions. Their views and tips were useful indeed. Unfortunately, it is not possible to list all of them in this limited space. To you all i said JAZAKUMULLAHU KHAIRAN.

ABSTRACT

Tuberculosis is an intracellular bacterial infection that attack organs of human body system, it is a worldwide disease with high estimated number of death rate every year. Microarray technology produces large amount of disease gene expression data and provides opportunities mine the data to understand disease mechanisms at molecular level. The aim of this study is to explore the usage of some tools available for analysing human TB microarray gene expression data. The control stimulated samples with phosphate buffer saline (PBS) and experimental unstimulated samples of three different clinical forms of human TB microarray gene expression data such as latent TB (LTB), pulmonary TB (PTB) and meningeal TB (TBM) were collected from GEO-NCBI database and all analysis were performed by using Bioconductor R packages. The results of this study, explore the use of affycoretool for microarray TB image visualization analysis, AffyQCReport tool for TB microarray data quality assessment, GCRMA method for TB microarray data normalization and LIMMA as a statistical tool for the identification of significantly expressed genes of human TB. According to LIMMA, there was a significant different between stimulated and unstimulated tuberculosis and majority of the significantly expressed genes identified were genes responsible for cellular immune response. The regulated genes identified from the LIMMA analysis using Venn diagram indicated more decrease in rate of gene expression than the increase in stimulated tuberculosis while show more increase in rate of gene expression than decrease in unstimulated tuberculosis. Hierarchical clustering (hclust) method was used to determine common expression pattern among the three different clinical forms of human TB infection, it suggested that, hierarchical clustering analysis distinguish different clinical forms of human TB infection. This study recommended that the results generated from these findings can be used in further analysis for detection and control of human TB infection.

ABSTRAK

Batuk kering akibat daripada jangkitan bakteria intrasel yang menyerang organ-organ sistem badan manusia, merupakan penyakit di seluruh dunia kadar kematian yang tinggi setiap tahun. Teknologi microarray menghasilkan sejumlah besar penyakit gen data ungkapan dan memberi peluang melombong data untuk memahami mekanisme penyakit pada peringkat molekul. Tujuan kajian ini adalah untuk meneroka penggunaan beberapa alatan yang disediakan untuk menganalisis microarray TB data ekspresi gen manusia. Kawalan dirangsang sampel dengan penimbal fosfat masin (PBS) dan ujikaji sampel tidak dirangsang daripada tiga bentuk klinikal yang berbeza data ungkapan TB manusia microarray gen seperti TB terpendam (LTB), TB paru-paru (PTB) dan TB meningeal (TBM) telah diambil dari GEO pangkalan data -NCBI dan semua analisis telah dijalankan dengan menggunakan Bioconductor R pakej. Hasil kajian ini, meneroka penggunaan affycoretool untuk microarray TB analisis visualisasi imej, alat AffyQCReport untuk penilaian kualiti data microarray TB, kaedah GCRMA untuk normalisasi data microarray TB dan LIMMA zdcvcsebagai alat statistik untuk mengenal pasti gen yang ketara daripada TB manusia . Menurut LIMMA, terdapat perbezaan yang signifikan antara dirangsang dan unstimulated batuk kering dan majoriti daripada gen ketara yang dikenal pasti ialah gen yang bertanggungjawab untuk tindak balas imun selular. Gen dikawal selia dikenal pasti daripada analisis LIMMA dengan menggunakan gambar rajah Venn menunjukkan penurunan lebih dalam kadar ekspresi gen daripada peningkatan batuk kering dirangsang manakala persembahan lebih meningkat pada kadar ekspresi gen daripada penurunan batuk kering tidakdirangsang. Hierarki kelompok (hclust) kaedah telah digunakan untuk menentukan corak ungkapan biasa di kalangan tiga bentuk berbeza klinikal jangkitan TB manusia, ia mencadangkan bahawa, analisis pengelompokan hierarki membezakan bentuk klinikal berbeza jangkitan TB manusia. Kajian ini mencadangkan bahawa hasil yang dijana daripada hasil kajian ini boleh digunakan dalam analisis selanjutnya untuk pengesanan dan kawalan jangkitan TB manusia.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABRAVIATION/ SYMBOLS

AIDS	Acquired immune deficiency syndrome
BCG	Bacillus calmette Guerin
CDNA	Colour deoxyribonucleic acid
DNA	Deoxyribonucleic acid
HIV	Human immune deficiency virus
LTB	Latent tuberculosis
MDMs	Monocyte-derived macrophages
MM	Mismatch
PBMCs	Peripheral blood mononuclear cells
PBS	Phosphate buffer saline
PM	Perfect match
PMA	Present margin absent
PTB	Pulmonary tuberculosis
RNA	Ribonucleic acid
RT-PCR	Reverse transcription-polymerase chain reaction
Stm	Stimulated
TBM	Meningeal tuberculosis
Unstm	Unstimulated
WHO	World health organization
=	Equal sign
%	Percentage

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1 Background of the Research

Tuberculosis is an intracellular bacterial infection that usually attacks the lungs in the human body system, but it can infect other parts of the human body. The causative agent of this infection is known as *Mycobacterium tuberculosis*. TB disease is a communicable disease that can easily be transmitted between the human being via the inhalation process, when a person with pulmonary TB infection exhaled the bacteria. The individuals nearby or those who live with others having active TB infections, or smoking cigarettes, young children, alcoholics and intravenous drug users, malnutrition, Patients with HIV/AIDS or other immune system deficiencies can become infected. Tuberculosis is a major cause of illness and death worldwide, but most of the TB infections of about 90% to 95% are latent TB infections which does not show any symptom of the disease. At 2012, 8.6 million meningeal TB cases were estimated all over the world (WHO, 2013). Many of these infections occurred in those with HIV infections and also most of these cases are occurring in developing countries where the majority of people were having very weak immune system in their bodies due to the lack of proper nutrition. The abilities of immune responses in the body against facultative intracellular pathogens are dependent on the rich food items taken by the individual (Kaufmann, 2002).

Microarray is a new method for detection of thousands of DNA or RNA at one time. Microarray technology is one of the technologies that contributes toward the development of studies in the field of Molecular Biology. This technology deals with the collections of DNA or RNA samples from microarray experiments and uses these samples to measure the expression levels of large numbers of genes or the whole genome of an organism, or to identify differential expressed genes or to determine common expression pattern among the genes at a time. Microarray experiment involves the growing of an organism, collecting of tissues, extraction of RNA or DNA, hybridization and scanning process to generate microarray data. (Miller *et al.,* 2009). Microarray data analysis involves different steps and a number of variables because the data sets are commonly very large.

Microarray is the combination of two complementary single-stranded DNA or RNA to form a single double-stranded through base pairing to a very large number of genes in which each gene is normally represented by more than one probe. .Microarray is used to determine gene sequence or to identify variations among genes or expression. Microarray technology is a molecular technology that used genetic materials of an organism to identify new genes for a particular character with their expression levels and functions under different environmental conditions. This technology is very useful in providing information about a particular disease and possible way of controlling it. The first issue reported concerning the use of microarray technology was that, 378 arrays of bacterial colonies was generated for gene expression in normal and tumour tissues. With development more than 4000 human gene sequences was generated using scanning process and image processing of computer system to carry out quantitative analysis of human tumours and normal tissues. The same process is also used to compare tumour tissues at different genetic risk (Augenlicht *et al.,* 1982; 1991). In 1995, microarray data for gene expression profiling was first generated and used, and after which in 1997, a complete microarray data of eukaryotic genome (*Saccharomyces cerevisiae*) were generated and published (Schena *et al.,* 1995; Lashkari *et al.,* 1997).

Different microarray experiments were designed to analyse gene expression and variations among the genes of an organism. This development in microarray technology help medical personnel to diagnose different infections. Microarray is currently used to analyse many different systems, including the classification of microbes and human microbial pathogens, how body mechanisms of an infected individual response to infections, drug and toxic exposures, tumour classification, single nucleotide polymorphism detection, the detection of gene fusions, comparative genomic hybridization, alternative splicing detection (exon junction array/exon arrays) and gene expression profiling (Miller *et al.,* 2009).

1.2 Statement of the Research Problem

Microarray technology is a technique that has evolved to allow the study of genes that are being used in each particular cell throughout the body of an organism. This technology allows the use genetic materials to analyse gene expression and genetic variations among the living organisms, and also it contributed significantly to the understanding of pathogenic mechanisms of different microbial infections and provide knowledge of how drugs for treating such infections can be discovered. In addition, microarray technology has been used to detect and monitor how body cellular immune system of an organism's response to different infections.

Different studies of microarray indicated that, the whole-genome of *Mycobacterium tuberculosis* was first studied and described using the amplicon arrays (Wilson *et al*., 1999). Advancement of chronic TB infection using whole genome was studied and later microarrays study is used to monitor the gene expression of *Mycobacterium bacterium* responses to a different environmental conditions and to measure exposure of these genes to antibiotics (Betts *et al.,* 2003; Kaushal *et al.,* 2002; Rengarajan *et al.,* 2005; Stewart *et al.,* 2004; Voskuil *et al.,* 2003). Samples of tuberculosis were collected from the infected human lungs and transcriptional profile of that sample was studied, the result shows that, number of genes are actively

transcribed to prevent the attack from host defence systems **(**Rachman *et al.*, 2006**)**. Microarray analysis of host immune response to TB infection and analysis of microarray meningeal TB infection were performed (Gonzalez-Juarrero *et al.*, 2009). Microarray study on early lung immune responses to TB infections using mouse model was also carried out (Dongwan *et al*., 2011). However, with all the significant contributions of microarray studies to tuberculosis infection, still the major issues now concern with this infection are detection and control. On the other hand also the challenge with microarray technology is that, due to the large dimensionalities in microarray data, it is very difficult to analyze microarray data despite availabilities of many tools for analyzing microarray data, but the usage of these tools depends on the objectives of the study.

In view of the above, this study explore the usage of some tools and a common statistical tool available in Bioconductor R packages for analysing microarray gene expression data of human tuberculosis infection under different conditions with different objectives.

1.3 Significance of the Research

This research is immensely important to live all over the world in the following ways: The study will aid in understanding molecular mechanism of tuberculosis required of controlling the disease. It can be used as a source of information's for future studies.

1.4 Aims and Objectives of the Research

The aim of this research has been use already published microarray expression data of three different clinical forms of human tuberculosis infection under different conditions with different objectives as:

1. To process the microarray raw data of human tuberculosis infection
2. To explore a common statistical tool available for analysing microarray gene expression data of human tuberculosis infection.
3. To determine common expression pattern of human tuberculosis infection.

1.5 Scope and Limitations

The research is exclusively theoretical in nature and no microarray experiments were performed. Instead, only one set of microarray data was used of three different clinical forms of human tuberculosis infection collected from GEO-NCBI database. The research was limited to explore some of the tools and a common statistical tool for analysing microarray gene expression data. The duration of the research was 5 months. FBME computer facilities were used to perform the analysis.

LIST OF REFERENCES

American Thoracic Society (ATS), Centres for Disease Control and Prevention, Infectious Diseases Society of America (2003). Treatment of tuberculosis. American Journal of Respiratory and Critical Care Medicine, 167(4): 603–662.

Augenlicht LH, Kbrin D (1982). "Cloning and screening of sequences expressed in a mouse colon tumour". Cancer Research 42(3): 1088-1093. PMID 7059971.

Augenlicht (1991). " *Patterns of Gene Expression that Characterize the Colonic Mucosa in Patients at Genetic Risk for Colonic Cancer*" Proceeding of the National Academy of Sciences of the United States of America 88(8): 3286- 28 doi:10.1073/pnas.88.8,3286.

Beresford B, Sadoff JC (2010). Update on research and development pipeline: tuberculosis vaccines. Clin Infect Dis. 2010; 50 (Supply 3):S178–83. [PubMed]

Betts J. C., A. McLaren, M. G. Lennon, F. M. Kelly, P. T. Lukey, S. J. Blakemore, and K. Duncan *(*2003). *"Signature gene expression profiles discriminate between isoniazid-, thiolactomycin-, and triclosan-treated Mycobacterium tuberculosis"*. Antimicrobial. Agents Chemother. 47:290 2913. PMID: 12936993 [PubMed - indexed for MEDLINE] PMCID: PMC182614.

Bolstad BM, Irizarry RA, Astrand M, Speed TP (2003). A comparison of normalization methods for high density oligonucleotide array data based on Variance and bias. Bioinformatics 19(2): 185- 93. Doi: 10. 1093/bioinformatics/95.2.185 PMID 12538238

Brenude Ed. Wilmoth Lerner and K. Lee Lerner. Robert Koch (2006)." *World of* Microbiology *and Immunology.* Detroit: Gale, 2006. Biography in Context. Web. 14 Apr. 2013.

Bretz, F., Land grebe, J., and Brunner, E. (2005). Design and analysis of two-colour Microarray experiments using linear models. Methods Info Med, 44(3):423– 430.

Carl Zimmer, (2014). "Tuberculosis Is Newer than Thought, Study says", New York Times, 21 August 2014. Centers for Disease Control and Prevention

CDCP (Centre for disease control and prevention, 2011). Recommendations for use of an isoniazid-rifapentine. Regimen with direct observation to treat latent Mycobacterium tuberculosis Infection. MMWR, 60(48): 1650–1653. Available online.

Churchill, GA (2002). "Fundamental of experimental design for cDNA microarrays Nature Genetics supplement 32: 490-5. doi:10.1038/ng1031.PMID 12454643. Archived from the original on 2005-05-08. Retrieved 12 December 2013.

Cohn DL, Catlin BJ, Peterson KL, Judson FN, Sbarbaro JA (1990). "A 62-dose, 6-month therapy for pulmonary and extra pulmonary tuberculosis. A twice-weekly, directly observed, and cost-effective regime "Annals of Internal Medicine 112(6):407-15. Doi: 10.7326/0003-4819-76-3-112-6-407. PMID 2106816.

Craig Parman and Conrad Halling (2014). AffyQCReport: A Package to Generate QC Reports for Affymetrix Array Data. April 11, 2014

Dongwan D. Kang, Yinyao Lin, Javier-Rangel Moreno, Troy D. Randall (2011). "*Profiling Early Lung Immune Responses in the Mouse Model of*

Tuberculosis". Published online Jan 13, 2011. Do: 10.1371/journal.pone.0016161 PMCID: PMC3020951.

Dudoit, S., J.P. Shaffer, and J.C. Boldrick (2002). "*Multiple hypothesis testing in Microarray experiments. Technical Report*" 110, Division of Biostatistics, University of California, Berkeley. URL http://www.bepress.com/ucbbio-stat/paper110/.

Dunlap (2013). Pulmonary and critical care medicine. University of Virginia School of Medicine in Charlottesville.

Eisen, M., Spellman, P., Brown, P. and Botstein, D (1998). Clusterana- lysis and display of genome-wide expression patterns. Proc. Natl Acad. Sci. USA, 95, 14863- 14868.

Elzinga G, Raviglione MC, Maher D (2004). "Scale up: meeting targets in global tuberculosis control "Lancet 363(9411): 814-19. Doi: 10.1016/S0140-6736(04)15698-5. PMID 15016493.

Gentleman R, Carey V, Huber W (2005). Bioinformatics and computational Biology Solutions using R and Bioconductor. USA, Springer Science Business Media, inch, 2005.

George Bodington (2010). **"**An essay on the treatment and cure of pulmonary Consumption: On Principles Natural, Rational". *First Published* 1840 by Longman. Last edited by Open Library Bot.

Golden MP and Vikram HR (2005). Extra pulmonary tuberculosis: an overview. American Family Physician, 72 (9), 1761–8.

Golub, T., Slonim, D., Tamayo, P., Huard, C., Gaasenbeek, M., Mesirov, J., Coller, H., Loh, M (1999). Molecular classification of cancer: class

discovery and class prediction by gene expression monitoring. Science 286:531-537.

Gonzalez-Juarrero M., Kingry LC, Ordway DJ, Henao-Tamayo M., Harton M. (2009). "Immune response to Mycobacterium tuberculosis and identification of molecular markers of disease". Am J Respir Cell Mol Biol. 2009; 40:398– 409. PMCID: PMC2660559.

Herrero J., A. Valencia, J. Dopazo (2001). A hierarchical unsupervised growing neural Network for clustering gene expression patterns. *Bioinformatics 17(2):126 136.*

Huber W., Robert G., Vincent J. Carey, Rafael A. Irizarry, Sandrine Duodoit. (2006). Bioinformatics and computational Biology Solution Using R. books.google.com.my/books?isbn=0387293620.

Irizarry, RA; Hobbs, B; Collin F; Beazer-Barclay, YD; Antonellis, KJ; Scherf, U; Speed TP (2003). "*Exploration, normalization, and summaries of high density oligonucleotide array probe level data*." Biostatistics 4 (2): 249 64. Doi:10.1093/biostatistics/4.2.249.PMID12925520.

Jarno Tuimala (2008). DNA microarray data analysis using Bioconductor. CSC, the Finnish IT centre for Science, CSC – IT Center for Science Ltd. 2008 ISBN 978-9525520-34-7http://www.csc.fi/oppaat/R.

Jennifer Villaseñor-Park and Alex G Ortega-Loayza (2013). Microarray Technique, Analysis, and Applications in Dermatology. Journal of Investigative Dermatology (2013) 133, e7. doi:10.1038/jid.2013.64 Jones and Bartlett (2010). Series in biomedical informatics, R for Medicine and Biology, 2010 ISBN 978-0-7637-5808-0.

Kauffmann A., R. Gentleman, W. Huber (2009). Array Quality Metrics a Bioconductor package for quality assessment of microarray data, Bioinformatics 25 (2009)415–416.

Kaufmann S. H. E. (2002). "*Protection against tuberculosis: cytokines, T cells, and Macrophages"*. Department of Immunology, Schumannstrasse 21-22, D-10117 Berlin, Germany. PMCID: PMC1766701.

Kaushal D., B. G. Schroeder, S. Tyagi, T. Yoshimatsu, C. Scott, C. Ko, L. Carpenter, J. Mehrotra, Y. C. Manabe, R. D. Fleischmann, and W. R. Bishai (2002). *"Reduced immunopathology and mortality despite tissue persistence in a Mycobacterium tuberculosis mutant lacking alternative sigma factor".* Sigh. Proc. Natl. Acad. Sci. USA 99**:**8330-8335.

Kendziorski et al. (2005). On the utility of pooling biological samples in microarray Experiments, PNAS, 102(12):4252- 4257.

Kentucky (2006). Mammoth Cave long on history. *CNN*. 27 February 2004. Accessed 8 October, 2006.

Kulesh DA, Clive DR, Zarlega DS, Greene JJ (1987). "Identification of interferon-Modulated proliferation-related cDNA sequences". Proc. Natl. Acad. Sci. U.S.A.84 (23):8453-8457.doi:10.1073/pnas.84.23.8453.PMC 299562. PMID. 2446323.

Lashkari DA, DeRisi JL, McCusker JH, Namath AF, Gentile C, Hwang SY, Brown PO, Davis RW (1997). "Yeast microarrays for genome wide parallel genetic and gene expression analysis". Proc. Natl. Acad. Sci. U.S.A.94 (24): 13057-13062. doi:10.1073/pnas.94.24.13057.PMC 24262. PMID 9371799.

Leming Shi, National Centres for Toxicological Research (2007). "Microarray

Quality Control (MAQC) Project". U.S.Food and Drug Administration. Retrieved 2007-12- 26.

Lim WK., Wang K Lefebvre C., Califano A. (2007). "Comparative analysis Microarray normalization procedures: effect on reverse engineering gene Networks". Bioinformatics 23 (13): i282-8. doi:10.1093/bioinformatics/btm201 PMID 1764630

Lindblad-Toh K, Tanenbaum DM, Daly MJ, Winchester E, Lui WO, Villapakkam A, (2000). "*Loss of heterozygosity analysis of small-cell lung carcinomas using Single nucleotide polymorphism arrays*". Nat Biotechnology 2000; 18:1001-5.

Matthew E. Ritchie, Jeremy Silver, Alicia Oshlack Dileepa Diyagama and Gordon K. Smyth (2007). *A comparison of background correction methods for two-colour microarrays.* Oxford Journals of Science & Mathematics Bioinformatics Volume 23, Issue 20 Pp. 2700- 2707.

Matthew N. McCall, Benjamin M. Bolstad, Rafael A. Irizarry (2009*). Frozen robust Multiarray analysis (fRMA)*. Oxford Journals Science Mathematics Volume 11, Issue 2 Pp. 242-253.

Michael A and Peter GG (2005). Evidence-based respiratory medicine. Oxford, Blackwell. Pp. 321. ISBN 978-0- 7279-1605-1.

Miller M. B., Tang Y. W. (2009). **"***Basic concepts of microarrays and potential Applications in clinical microbiology".* Clinical Microbiology Review, v.22 (4); PMC2772365.

Mishra, G (2013). "Tuberculosis Prescription practices In Private and Public Sector in India". NJIRM 4(2):71-78.

Mitchell, MA (2012). "Mycobacterial infections in reptiles". The civics. Veterinary clinics of North America. Exotic animal practice 15(1): 101-11, vii. Do: 10.1016/j. civic. 2011.10.002. PMID 22244116.

Nguyen Thuy Thuong Thuong, Sarah J. Dunstan,Tran Thi Hong Chau, Vesteinn Thorsson, Cameron P. Simmons, Nguyen Than Ha Quyen, Guy E. Thwaites,[5]Nguyen Thi Ngoc Lan, Martin Hibberd,Yik Y. Teo, Mark Seielstad, Alan Aderem, Jeremy J. Farrar, and Thomas R. Hawn (2008). "*Identification of Tuberculosis Susceptibility Genes with Human Macrophage Gene Expression Profiles*". PLoS Pathog. Dec 2008; 4(12): e1000229. Published online Dec 5, 2008. Doi: 10.1371/journal.ppat.1000229 PMCID: PMC2585058.

Page (2006). The Power Atlas: a power and sample size atlas for microarray Experimental design and research, BMC Bioinformatics, 7:84.

Paul Kielstra (30 June 2014). "Ancient enemy, modern imperative-A time for greater action against tuberculosis". In Zoe Tabary. Economist Insights (The Economist Group). Retrieved 1 August 2014.

Pekeliling Kawalan Penyakit Bil.1 (2002). Pemberhentian Pelalian Dos Tambahan (Booster) *BCG Dalam Program Kawalan Penyakit Tibi.*

Philip Lobue (2013). "Tuberculosis Elimination". Centres for Disease Control and Prevention (CDC) Atlanta GA 30333 TB Notes No. 4, 2013.

Rachman H., M. Strong, T. Ulrichs, L. Grode, J. Schuchhardt, H. Mollenkopf, G. A. Kosmiadi, D. Eisenberg, and S. H. Kaufmann *(*2006). *"Unique Transcriptome signature of Mycobacterium tuberculosis in pulmonary Tuberculosis".* Infect. Immun. p. 1233–1242 0019-9567. doi:10.1128/IAI.Vol. 74.2.

Rengarajan J., B. R. Bloom, and E. J. Rubin (2005). *"Genome-wide requirements for Mycobacterium tuberculosis adaptation and survival in macrophages".* Proc. Natl. Acad. Sci. USA 102**:**8327-8332. Do: 10.1073/pnas.0503272102. vol. 102 no. 23.

Sakula A. (1983). Carlo Forlanini, inventor of artificial pneumothorax for treatment of pulmonary tuberculosis. Thorax, 38, 326–32.

Schena M, Shalon D, Davis RW, Brown PO (1995). "Quantitatives monitoring of Gene expression patterns with a complementary DNA microarray". Science 270(5235): 467-470. doi:10.1126/science.270.5235.467. PMID 7569999.

Shivaprasad, HL; Palmieri, C (2012). "Pathology of mycobacteriosis in birds". The civics. Veterinary clinics of North America. Exotic animal practice 15(1): 41 55, v-VI. Do: 10.1016/j. 2011.11.004. PMID 22244112.

Smyth GK (2004). Linear models and empirical Bayes methods for assessing Differential expression in microarray experiments. *Stat Appl Genet Mol Biol* 2004, 3: Article 3.

Stewart G. R., B. D. Robertson, and D. B. Young *(*2004). *"Analysis of the function of mycobacterial DNA J proteins by overexpression and microarray profiling".* Tuberculosis (Edinburgh) 84:180-*187*. PMID: 15207487 [PubMed - indexed for MEDLINE].

Thomas Girke, UC Riverside (2010). R & Bioconductor Manual. Institute for Integrative, genome Biology.

Voskuil M. I., D. Schnappinger, K. C. Visconti, M. I. Harrell, G. M. Dolganov, D. R. Sherman, and G. K. Schoolnik (2003). *"Inhibition of respiration by nitric Oxide induces a Mycobacterium tuberculosis dormancy program".* J. Exp. Med. 198**:**705-713. PMID: 12953092 [PubMed - indexed for MEDLINE] PMCID: PMC2194188.

Wang JY, Hsueh PR, Jan IS, (2006). "Empirical treatment with a fluoroquinolone Delays the treatment for tuberculosis and is associated with a poor prognosis in endemic areas". Thorax 61(10):903-08.doi:10.1136/thx.2005.056887.PMC 2104756. PMID 16809417.

Wilson CL, Miller CJ (2005). "Simpleaffy: a Bioconductor package for Affymetrix Quality Control and data analysis". Bioinformatics 21(18):3683-5 doi:10.1093/bioinformatics/bti605. PMID 16076888.

Wilson CL, SD Pepper, Y Hey, and CJ Miller (2004). Amplification protocols Introduce systematic but reproducible errors into gene expression studies. Biotechniques, 36:498–506, 2004.

Wilson M., J. DeRisi, H. H. Kristensen, P. Imboden, S. Rane, P. O. Brown, and G.K. Schoolnik (1999). "*Exploring drug-induced alterations in gene Expression in Mycobacterium tuberculosis by microarray hybridization"*. Proc. Natl. Acad. Sci. USA 96:12833-12838. PMID: 10536008 [PubMed – Indexed for MEDLINE] PMCID: PMC23119.

World Health Organization (2005). Global tuberculosis control: surveillance, planning, financing. Geneva: World Health Organization; 2005.

World Health Organization (2009). "Epidemiology". *Global tuberculosis control*: epidemiology, strategy, financing. Pp. 6–33. ISBN 978-92-4-156380-2.

WHO REPORT (2011) GLOBAL TUBERCULOSIS CONTROL?

WHO (2012) Tuberculosis Fact Sheet. World Health Organization, Geneva, Switzerland.

World Health Organization (WHO) (2013). Tuberculosis fact sheet. Retrieved from http://www.who.int/mediacentre/factsheets/fs104/en/.

WHO (2013). "*Global tuberculosis report*". World Health Organization.

Zumla A, Raviglione M, Hafner R, & Fordham VR. (2013). Current concepts: Tu berculosis. New England Journal of Medicine, 368(8), 74555. Retrieved From http://search.proquest.com docview/1295104636? Accounted=35812.