

A potential link between tuberculosis and lung cancer through non-coding RNAs

Debmalya Barh^{1,2*}, Sandeep Tiwari², Ranjith N. Kumavath³, Vasco Azevedo²

¹Centre for Genomics and Applied Gene Technology, Institute of Integrative Omics and Applied Biotechnology (IIOAB), Nonakuri, Purba Medinipur, West Bengal, India

²Laboratório de Genética Celular e Molecular, Departamento de Biologia Geral, Instituto de Ciências Biológicas (ICB), Universidade Federal de Minas Gerais, Pampulha, Belo Horizonte, Minas Gerais, Brazil

³Department of Genomic Sciences, School of Biological Sciences, Central University of Kerala, Kasaragod, India

***Corresponding author:** Email: dr.barh@gmail.com; Tel/Fax: +91 944 955 0032

ABSTRACT

Pulmonary tuberculosis caused by *Mycobacterium* and lung cancer are two major causes of deaths worldwide and the former increases the risk of developing lung cancer. However, the precise molecular mechanism of *Mycobacterium* associated increased risk of lung cancer is not entirely understood. Here, using *in silico* approaches, we show that hsa-mir-21 and *M. tuberculosis* sRNA_1096 and sRNA_1414 could play important roles in the pathogenesis of both these diseases. Further, we postulated a “*Genetic remittance*” hypothesis where these sRNAs may play important roles. The sRNA_1096 could be involved in tuberculosis through multiple infectious processes, and if transferred to the host, it may activate the TLR8 mediated pro-metastatic inflammatory pathway by acting as a ligand to TLR8 similar to the mir-21 leading to lung tumorigenesis and chemo-resistance. Analogous to SH3GL1, it may also regulate cell cycle. On the other hand, sRNA_1414 is probably involved in survivability and drug response of the pathogen. However, it may be a metastatic factor for lung cancer providing EPS8L1 and SORBS1 like functions upon remittance. Further, all these three non-coding RNAs are predicted to act in rifampicin resistance in *Mycobacterium*. Currently, we are applying robust bioinformatics strategies and conducting experimental validations to confirm our *in-silico* findings and hypothesis.

Key words: Genetic predisposition; lung cancer risk; microRNA; *Mycobacterium tuberculosis*, sRNA

INTRODUCTION

Viral involvements and their causal roles in oncology are well accepted for various cancers including ovarian neoplasms [1], hepatocellular carcinoma [2], and lung cancer [3] among others. Although, bacterial infections are not considered as major threats to cancer, yet a number of bacterial pathogens are reported to be associated with several cancers. Some examples include: *Mycoplasma* in prostate malignancy [4], *Robinsoniella* in pancreatic cancer [5], *S. typhi*, *H. bilis*, *H. hepaticus*, and *E. coli* in carcinoma of the gallbladder [6], *Chlamydia* in cervical cancer [7], and *Mycobacterium* in lung cancer[8-14] .

Lung cancer is the leading cause of all cancer related deaths with recently estimated 1.6 million deaths worldwide [15]. Similar to lung cancer, pulmonary tuberculosis caused by *M. tuberculosis* is a global health problem. It is one of the major causes of death amongst infectious diseases and according to WHO 2013 report, it is estimated that 9 million people are infected and 1.5 million are died from tuberculosis in 2012 [16]. Several reports have documented co-existence of tuberculosis and lung cancer [14,17-20] and pulmonary tuberculosis is a risk factor for developing lung cancer [17-20]. However, it is not yet fully established at molecular level, how the *Mycobacterium* increases susceptibility to lung cancer. Some report say that *M. tuberculosis* induces ROS mediated DNA damage pathway and produces epiregulin growth factor to induce cell proliferation [21]; while other study indicates mechanisms along with COX-2 mediated activation of inflammatory pathway in *M. tuberculosis* associated carcinogenesis [22].

Bacterial small regulatory RNAs (sRNAs) are a class of small non-coding RNAs of 40-500 nt in length that regulate various essential patho-physiologies in bacteria such as outer membrane protein biogenesis, virulence, quorum sensing etc. sRNA functions through complementary base-pairing with 3'- or 5'-UTRs of target mRNAs to inhibit translation, alters activity of a protein by directly binding, and by mimicking RNA and DNA structures [23-25]. Several sRNAs have been identified or predicted from *M. tuberculosis* [26-28] having probable role in pulmonary tuberculosis pathogenesis [26,29,30]. However, no report so far is available on sRNA from *M. tuberculosis* having role in lung cancer. On the other hand, human micro RNAs (miRNAs) are small non-coding RNAs of 20-25 nt length that inhibit post-transcriptional gene regulation by complementary base pairing at the 3' -UTRs of target mRNAs and regulate various patho-physiological conditions including cell cycle

regulation, cell differentiation, development, metabolism, aging, different types of cancers, metabolic disorders, and neuronal diseases etc. [31,32]. Several miRNAs have been implemented to be associated with lung cancer having causative roles and diagnostic potentials. Similarly, a number of miRNAs are found deregulated in pulmonary tuberculosis patients [33-35].

Since the sRNAs and miRNAs are similar in structure, their mode of actions, and the miRNAs are associated with both pulmonary tuberculosis and lung cancer; we hypothesized that *Mycobacterium* sRNAs may be associated with lung cancer tumorigenesis. Since *Mycobacterium* infection is a risk factor in developing lung cancer, here we postulated a “*Genetic remittance*” hypothesis which presumes that *M. tuberculosis* sRNAs having similarity with human miRNAs (that are associated with either pulmonary tuberculosis or lung cancer or both these two diseases) are transmitted to the host during *Mycobacterium* infection, remained within the human, and act as a predisposition factor to increase the risk of lung cancer.

In this study, using *in silico* strategies, we aimed to understand (i) the role of *M. tuberculosis* sRNAs in lung carcinogenesis; (ii) if there are structural and functional similar *M. tuberculosis* sRNAs to human miRNAs responsible for pathogenesis in both pulmonary tuberculosis and lung cancer and their probable functions in these diseases; and (iii) proof of concept of our “*Genetic remittance*” hypothesis.

METHODS

Collection of human miRNAs and M. tuberculosis siRNAs

We used PubMed (<http://www.ncbi.nlm.nih.gov/pubmed>), miRegulome [32], miR2Disease [36], TUMIR [37], sRNAdb [38], and BSRD [39] databases to get the data. The data deposited in these databases during January 2006 to November 2015 were searched. In the first step, we collected all the validated deregulated miRNAs associated with lung cancer and pulmonary tuberculosis from published literature indexed in PubMed (<http://www.ncbi.nlm.nih.gov/pubmed>) and from databases such as miRegulome [32], miR2Disease [36], and TUMIR [37]. The common miRNAs associated with both the diseases were manually identified and listed out separately. Similar to miRNA, we collected all

reported and novel *M. tuberculosis* sRNAs by means of PubMed literature mining and searching sRNA databases such as sRNAdb [38] and BSRD [39]. For PubMed search, specific key words such as “miRNA + *M. tuberculosis*”, “miRNA + tuberculosis”, “miRNA + lung cancer”, sRNA + *M. tuberculosis*”, “sRNA + tuberculosis” etc. were used.

Prediction of miRNAs that may function as sRNA or vice versa

To achieve this, we used a simple strategy. Since the sRNAs and miRNAs are of very short in sequence, we presumed that small sequence motifs in these non-coding RNAs are very important for their specific functions. Therefore, we used comparative BLASTn (miRNA against sRNA) with default parameters in order to identify if there is any human miRNA (from the group of miRNAs common to both lung cancer and pulmonary tuberculosis) having sequence similarity to any *M. tuberculosis* sRNA, so that, they may function similarly. From the BLASTn results, the specific motif sequences that are common in sRNAs and miRNAs were noted down.

Functional annotation of M. tuberculosis sRNAs

The functional annotation of the sRNAs was carried out using target based reverse annotation approaches following a modified protocol as described by Barh et al., 2013 [40,41]. In brief, we identified the validated targets of sRNA using sRNATarBase [42] database and RNApredator [43] was used to predict putative targets in *M. tuberculosis*, H37Rv. Top 100 targets were used for functional annotation of using DAVID functional annotation tool [44]. Further, we presumed that, if there is a coding gene that has identity with sRNA; the function of the sRNA could be similar to that gene. Therefore, we performed sRNA BLASTn against *M. tuberculosis* genome using NCBI BLASTn server (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) to identify if there is any coding sequence present in *M. tuberculosis* genome having sequence identity with the sRNA. To check if the sRNA is targeting an essential gene of *M. tuberculosis*, H37Rv, we used Database of Essential Genes (DEG) [45] BLASTn and to check if the target could be a drug target, we used the strategy as described by Barh et al., 2013 [41]. Further, we did miRNA (that is having sequence similarity with sRNA) BLASTn against *M. tuberculosis* genome in NCBI BLASTn server with default parameters to check if any *M. tuberculosis* coding sequence is matching with the miRNA sequence. Since, in this

way identified miRNA shares sRNA sequence, if there is a coding sequence that matches with the miRNA; we may assume that, the matched coding sequence of the sRNA and the miRNA may have similar function.

Testing “Genetic remittance” hypothesis

To test our “*Genetic remittance*” hypothesis, we performed BLASTn of the *M. tuberculosis* sRNAs having sequence similarity with human miRNA (that are associated with both pulmonary tuberculosis and lung cancer) against the human genome on NCBI BLASTn server in order to identify if there is any human coding sequence having identity with the sRNAs. If there is a sequence match and that human gene sequence is responsible for any tumorigenesis process, we assumed that, there may be possibility that these sRNAs may transfer to the host during *Mycobacterial* infection, and with a yet unknown mechanism regulate those human genes to increase the risk of lung tumorigenesis. Further, as those sRNAs are sharing human miRNA that are associated with lung cancer, there could be possibility that the predisposition of sRNA may have synergetic effects in presence of miRNA that leads to lung cancer.

RESULTS

Common miRNAs in pulmonary tuberculosis and lung cancer

From various literature and databases, we collected differential expression of 186 human miRNAs in pulmonary tuberculosis and 242 miRNAs in lung cancer patients. However, while we checked for the common miRNAs that are deregulated both in pulmonary tuberculosis and lung cancer, we found the number is only 45 (**Supplementary Table S1**).

M. tuberculosis sRNA_1096 and sRNA_1414 shares hsa-mir-21 sequence

From the sRNadb [38] and BSRD [39] databases and literature [27,46] we collected 120 reported *M. tuberculosis* sRNAs and their sequences. The comparative BLASTn between the 120 sRNAs and the common 45 human miRNAs in pulmonary tuberculosis and lung cancer revealed that the mature human miRNA hsa-mir-21-5p having very short sequence similarity with *M. tuberculosis* sRNA_1096 and sRNA_1414. Both these sRNAs are experimentally validated in *M. tuberculosis* [27]. Human hsa-mir-21-5p and *M. tuberculosis* sRNA_1096 share a common sequence of GTTG/ GUUG, while the common sequence between hsa-mir-21-5p and *M. tuberculosis* sRNA_1414 is ATCAG/ AUCAG (**Supplementary Table S2**).

Functional annotation of sRNA_1096 and sRNA_1414

To understand the functions of sRNA_1096 and sRNA_1414 in *M. tuberculosis*, first we used sRNATarBase [42] to search validated targets of these sRNAs. However, we did not get any target from this database. Therefore, we used RNApredator [43] to predict the targets of these sRNAs. The top 100 targets based on the Z-score of RNApredator were further used for DAVID functional annotation.

For sRNA_1096, phosphate-binding protein pstS 2 (Z-score: -11.29) was found to be the best target. pstS2 is an inorganic phosphate transmembrane transporter that is involved in two-component system and ABC transporters pathways. Further, among the top 100 targets, fourteen targets are PE PGRS family proteins (**Supplementary Table S3**). DAVID functional annotation analysis of the top 100 targets of sRNA_1096 shows that most targets are membrane located and the two-component system is the top annotation cluster.

The top target of sRNA_1414 is found to be glycyl-tRNA synthetase / glycine--tRNA ligase (glyS) (Z-score: -11.41) by RNApredator [43] (**Supplementary Table S3**). The drug transporter activity is ranked as the first annotation cluster and most targets are transmembrane proteins as observed through DAVID functional annotation analysis. According to the Database of Essential Genes (DEG) [45], the *M. tuberculosis* glyS is an essential gene in the pathogen but has 31% identity at protein level with human glycine--tRNA ligase as per NCBI human BLASTp.

When we performed sRNA_1096 BLASTn against *M. tuberculosis* genome, a very short similarity “CCGTCACCGTTG” was observed with *M. tuberculosis* Arabinosyltransferase

EmbC (*embC*) gene. The BLASTn of sRNA_1414 with *M. tuberculosis* genome did not show any hit with any specific protein coding gene of the pathogen.

Common function of sRNA_1096, sRNA_1414, and hsa-mir-21

As in previous analysis we found there are sequence similarities among hsa-mir-21 and *M. tuberculosis* sRNA_1096 and sRNA_1414; we performed a NCBI BLASTn of hsa-mir-21-5p against the *M. tuberculosis* genome. We observed that the *M. tuberculosis* *rpoB* gene that provides rifampin/ rifampicin resistant in *M. tuberculosis* has sequence similarity with mir-21 and having both the GTTG and ATCAG short-stretch sequences that are present in *M. tuberculosis* sRNA_1096 and sRNA_1414, respectively. Therefore, all these non-coding may be involved in rifampicin resistance in pulmonary tuberculosis.

Support to “Genetic remittance” hypothesis

To support our “Genetic remittance” hypothesis, we tried to identify if there is any sequence match of sRNA_1096 and sRNA_1414 in human coding sequence. The BLASTn of sRNA_1096 against human genome hits with SH3GL1 (SH3-domain GRB2-like 1) and some sequence of sRNA_1414 matches with human EPS8L1 (EPS8-like 1) and SORBS1 (Sorbin and SH3 domain containing 1). All these human genes are associated with tumorigenesis, thus providing a preliminary support to our hypothesis.

DISCUSSION

In this study, we found that there could be correlations between lung cancer and pulmonary tuberculosis at non-coding RNA level. *M. tuberculosis* sRNA_1096 and sRNA_1414, and human hsa-mir-21-5p could probably are the links to explain why the pulmonary tuberculosis is a risk factor in developing lung cancer. Among the 45 human miRNAs that are deregulated both in both the diseases (**Supplementary Table S1**) and 120 reported *M. tuberculosis* sRNAs *M. tuberculosis* genome, we found that there are sequence similarities among human hsa-mir-21-5p and *M. tuberculosis* sRNA_1096 and sRNA_1414 (**Supplementary Table S2**).

The onco-miR hsa-mir-21 is frequently upregulated in lung cancer [47-55] while it is downregulated in CD4⁺ T cells in tuberculosis patients [56]. However, hsa-mir-21 is found upregulated in the host during *M. bovis* BCG infection [57] (**Supplementary Table S1**). It is also observed that mir-21 plays a role in T-cell immunity against *M. tuberculosis* [56] and found to affect the anti-mycobacterial T- cell response through targeting IL12 and BCL2 [57].

Reports suggest that the bacteria induced carcinogenesis occurred in multiple ways. These mechanisms include induction or interference of chronic inflammatory and other signalling cascades including TLRs (Toll-like receptors) signalling and acetaldehyde metabolism pathways in various cancers [58-60]. TLRs signalling are generally involved in innate and adaptive immune responses. However, activation of TLRs signalling promotes tumor cell proliferation, growth, invasion, and metastasis [61]. In lung cancer, activation of TLR7 and TLR8 increases survival and chemo-resistance of the tumor cells, therefore these two TLRs could be targets for tumor immunotherapy [62]. In *Mycobacterial* infection, TLRs signalling regulates host innate and inflammatory responses and determines the disease outcome [63]. Polymorphisms and over expression of TLR8 is associated with pulmonary tuberculosis susceptibility and infection, respectively [64]. However, the precise mechanism of TLR8 in pulmonary tuberculosis is not yet known [65].

Fabbri *et al* in 2012 first reported the novel mechanism of onco-miR mir-21 that can act as a ligand for TLR8 to induce inflammatory response leading to tumor growth and metastasis [66]. This study also showed that the “GUUG” motif miR-21 directly binds to TLR8 to induce the TLR-mediated pro-metastatic inflammatory response.

***M. tuberculosis* sRNA_1096**

In our analysis, we found that the GTTG/ GUUG sequence is conserved in hsa-miR-21 and experimentally validated *M. tuberculosis* sRNA_1096 (**Supplementary Table S2**). Therefore, it may be implemented that the *M. tuberculosis* sRNA_1096 may bind to TLR8 in a similar way the miR-21 does through its GUUG motif and thus may activate immune and inflammatory responses and explains a role of TLR8 in pulmonary tuberculosis.

Further, as per our analysis, *M. tuberculosis* sRNA_1096 may play a role in two-component system pathway and may regulate PE PGRS family and membrane located proteins. PE PGRS family proteins act as variable surface antigens and are involved in multiple levels of the infectious process, modulation of innate immune responses [67], and virulence [68] in *M. tuberculosis*. Similarly, the two-component system is crucial in *Mycobacterial* survival and pathogenicity [69]. In *M. tuberculosis* genome, “CCGTCACCGTTG” sequence of the sRNA_1096 is present in Arabinosyltransferase EmbC (embC) gene which is involved in bacteria-host interactions and also modulates immune response in *M. tuberculosis* [70]. Therefore, *M. tuberculosis* sRNA_1096 could play an important role in *M. tuberculosis* pathogenesis leading to pulmonary tuberculosis.

In support to our “*Genetic remittance*” hypothesis, we predicted that if the *M. tuberculosis* sRNA_1096 is present or predisposes to human, it may modulate human SH3GL1 (SH3-domain GRB2-like 1) which plays a role in endocytosis [71], regulation of cell cycle in leukaemia [72], positive regulation of cell proliferation and inhibition of apoptosis in multiple myeloma [73], and oncogenesis in gliomas [74]. Hence, sRNA_1096 may play a critical role in lung cancer risk. Further, if the sRNA_1096 acts as a ligand to TLR8 similar to mir-21, upon remittance to the host, it may activate TLR8 pathway and along with increasing risk it may also regulate chemo-resistance [62] in lung cancer individually or in combination with mir-21.

Therefore, in summary, we postulate that, the *M. tuberculosis* sRNA_1096 is involved in pulmonary tuberculosis pathogenesis through multiple infectious processes including TLR8 mediated pathway. The sRNA_1096 may be transported to host and predisposed during *M. tuberculosis* infection and later acts as ligand to TLR8 through its GTTG/ GUUG sequence similar to mir-21 to activate TLR8 mediated pro-metastatic pathways and chemo-resistance in lung cancer. Similar to SH3GL1, it may also regulate tumorigenic inflammatory response and cell cycle, respectively leading to lung cancer. Hence, sRNA_1096 may be an emerging marker for tuberculosis and lung cancer risk and chemo-resistance in combination with mir-21.

***M. tuberculosis* sRNA_1414**

On the other hand, we observed ATCAG/ AUCAG as the common sequence between hsa-mir-21-5p and *M. tuberculosis* sRNA_1414 (**Supplementary Table S2**). The sRNA_1414 is predicted to target glycyl-tRNA synthetase (glyS) which is an essential gene in *M. tuberculosis*. Further, we observed that sRNA_1414 may regulate drug transporter activity. Aspartyl-tRNA synthetase and tyrosyl-tRNA synthetase are important drug targets in *M. tuberculosis*, [75,76] and polymorphisms in aspartyl-tRNA synthetase is associated with drug resistance mechanism in this pathogen [77]. Although, we predicted glyS is as an essential gene in *M. tuberculosis*, being a human homolog, it is not a suitable target. Thus sRNA_1414 probably be involved in regulating the survivability and drug response of the pathogen.

To evaluate our “*Genetic remittance*” hypothesis, while we did the BLASTn of sRNA_1414 against human genome, we found sequence matches with human EPS8L1 (EPS8-like 1) and SORBS1 (Sorbin and SH3 domain containing 1). EPS8L1 encodes a protein that is related to epidermal growth factor receptor pathway substrate 8 and is involved in regulation of Rho protein signal transduction [78] which is associated with small cell lung cancer migration [79]. Similarly, SORBS1 (Sorbin and SH3 domain containing 1) plays an important role in cell-matrix adhesion [80] a key process in cell migration. Therefore, if the *M. tuberculosis* sRNA_1414 is transferred to the host during the infection, it may lead to lung cancer metastasis in later stage functioning similar to EPS8L1 and SORBS1.

Common functions of M. tuberculosis sRNA_1096 and sRNA_1414

Since the hsa-mir-21 sequence GTTG /GUUA is shared by sRNA_1096 and ATCAG/ AUCAG by sRNA_1414 (**Supplementary Table S2**), we tried to predict what could common role of these non-coding RNAs. Our hsa-mir-21 BLASTn against *M. tuberculosis* shows that these GTTG /GUUA and ATCAG/ AUCAG sequences are present in *M. tuberculosis* *rpoB* gene that provides rifampin/ rifampicin resistant in *M. tuberculosis* [81-84]. Therefore, we presume that all these non-coding RNAs: hsa-mir-21, sRNA_1096, and sRNA_1414 could be involved in rifampicin resistance and an up regulation of mir-21 in tuberculosis patient may be a marker of rifampicin resistance.

CONCLUSIONS

In this work, we hypothesise that *M. tuberculosis* sRNA_1096 and sRNA_1414 may be involved in increased risk of developing lung cancer metastasis and chemo-resistance provided they are transferred to human and during *Mycobacterial* infection. These sRNAs may have synergetic effects in presence of human mir-21 that is associated with lung cancer. This analysis also provides an insight on why the *Mycobacterial* infection increases the risk of developing lung cancer and the possible TLR8 mediated molecular mechanisms. Further, all these non-coding RNAs could be involved in rifampicin resistance in *Mycobacterium*. Therefore, all these non-coding RNAs are potential biomarkers for assessing rifampicin resistance and lung cancer risk in pulmonary tuberculosis patients. A robust bioinformatics approach as outlined in **Figure-1** including domain/ motif analysis along with experimental validations are required to establish this proof of concept.

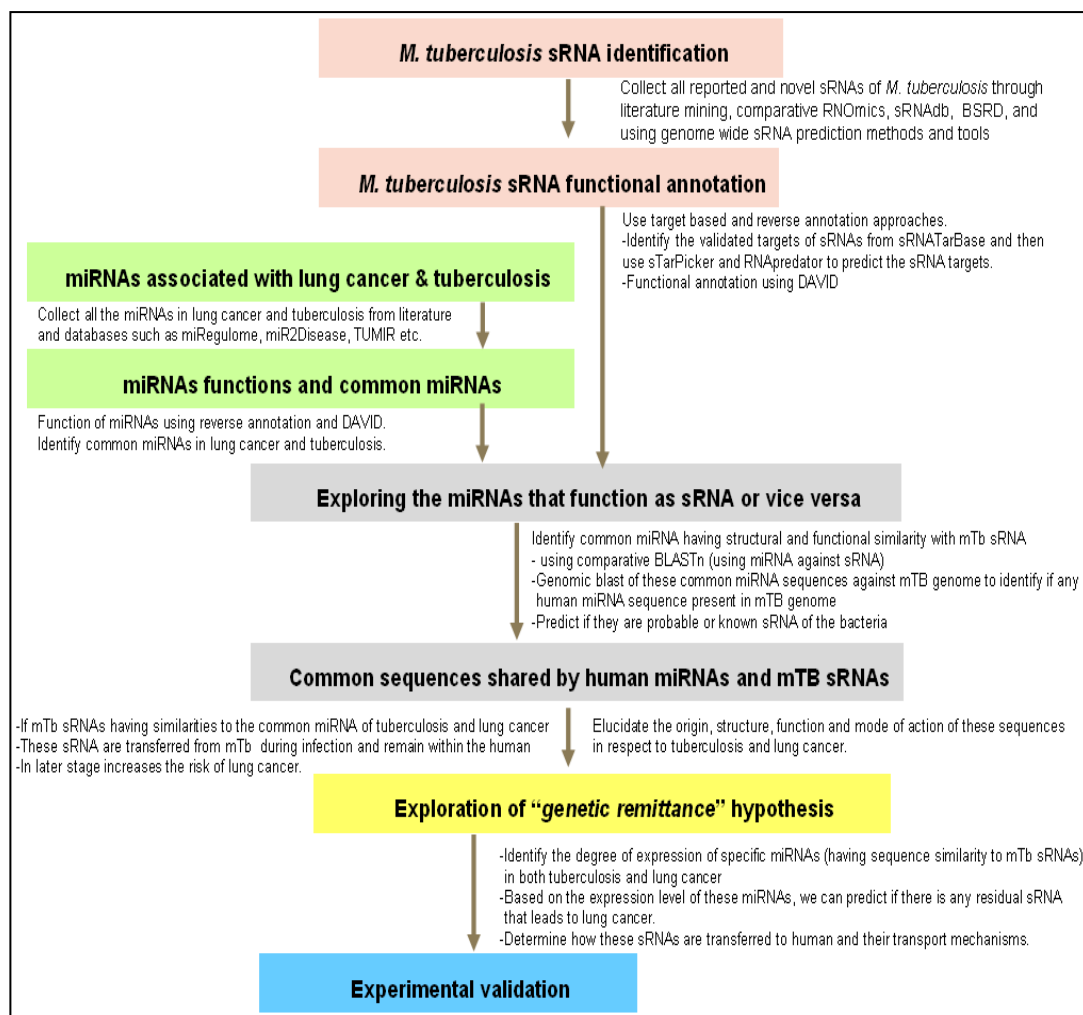


Figure 1: The integrated bioinformatics and experimental validation strategy to elucidate the relationship between tuberculosis and lung cancer linked through non-coding RNAs.

Funding

The authors have no support or funding to report.

Competing interests

The authors have declared that no competing interests exist.

Acknowledgments

Sandeep Tiwari acknowledges the “TWAS-CNPq Postgraduate Fellowship Programme” for granting a fellowship for doctoral studies. Debmalya Barh acknowledges the “TWAS-CNPq Postdoctoral Fellowship Programme” for granting a fellowship for postdoctoral studies.

Author Contributions

Conceived, designed the experiment, collected and analysed initial data, coordinated and led the entire project: DB, performed all *in-silico* analysis: DB, ST. Cross-checked all data, analysis: RNK, VA. Wrote the paper: DB. All authors read and approved the manuscript

REFERENCES

1. Giordano, G.; Azzoni, C.; D'Adda, T.; Rocco, A.; Gnetti, L.; Froio, E.; Merisio, C.; Melpignano, M. Human papilloma virus (hpv) status, p16ink4a, and p53 overexpression in epithelial malignant and borderline ovarian neoplasms. *Pathol Res Pract* **2008**, *204*, 163-174.
2. Ji, X.; Zhang, Q.; Du, Y.; Liu, W.; Li, Z.; Hou, X.; Cao, G. Somatic mutations, viral integration and epigenetic modification in the evolution of hepatitis b virus-induced hepatocellular carcinoma. *Curr Genomics* **2014**, *15*, 469-480.
3. Peng, J.W.; Liu, D.Y.; Lin, G.N.; Xiao, J.J.; Xia, Z.J. Hepatitis b virus infection is associated with poor prognosis in patients with advanced non small cell lung cancer. *Asian Pac J Cancer Prev* **2015**, *16*, 5285-5288.
4. Rogers, M.B. Mycoplasma and cancer: In search of the link. *Oncotarget* **2011**, *2*, 271-273.
5. Shen, D.; Chen, R.; Ye, L.; Luo, Y.; Tang, Y.W. Robinsoniella peoriensis bacteremia in a patient with pancreatic cancer. *J Clin Microbiol* **2010**, *48*, 3448-3450.
6. Nath, G.; Gulati, A.K.; Shukla, V.K. Role of bacteria in carcinogenesis, with special reference to carcinoma of the gallbladder. *World J Gastroenterol* **2010**, *16*, 5395-5404.
7. Ursu, A.; Sen, A.; Ruffin, M. Impact of cervical cancer screening guidelines on screening for chlamydia. *Ann Fam Med* **2015**, *13*, 361-363.
8. El-Sharif, A.; Afifi, S.; El-Dahshan, R.; Rafeh, N.; Eissa, S. Characterization of mycobacterium tuberculosis isolated from cancer patients with suspected tuberculosis infection in egypt: Identification, prevalence, risk factors and resistance pattern. *Clin Microbiol Infect* **2012**, *18*, E438-445.
9. Lande, L.; Peterson, D.D.; Gogoi, R.; Daum, G.; Stampler, K.; Kwait, R.; Yankowski, C.; Hauler, K.; Danley, J.; Sawicki, K., *et al.* Association between pulmonary mycobacterium avium complex infection and lung cancer. *J Thorac Oncol* **2012**, *7*, 1345-1351.
10. Cadinanos Loidi, J.; Abad Perez, D.; de Miguel Buckley, R. Endobronchial tuberculosis simulating lung cancer. *Arch Bronconeumol* **2014**, *50*, 126-127.
11. Christopoulos, A.; Saif, M.W.; Sarris, E.G.; Syrigos, K.N. Epidemiology of active tuberculosis in lung cancer patients: A systematic review. *Clin Respir J* **2014**, *8*, 375-381.
12. Doshi, V.K.; Kulkarni, S.R.; Kham, N.M.; Kapitan, K.S. A case of lung cancer originating from cavitary mycobacterium xenopi infection. *Respir Care* **2015**, *60*, e56-58.
13. Li, J.; Pan, Y.; Zhang, B.; Chen, Q. Macrophages are needed in the progression of tuberculosis into lung cancer. *Tumour Biol* **2015**, *36*, 6063-6066.
14. Skowronski, M.; Iwanik, K.; Halicka, A.; Barinow-Wojewodzki, A. Squamous cell lung cancer in a male with pulmonary tuberculosis. *Pneumonol Alergol Pol* **2015**, *83*, 298-302.
15. Jemal, A.; Bray, F.; Center, M.M.; Ferlay, J.; Ward, E.; Forman, D. Global cancer statistics. *CA Cancer J Clin* **2011**, *61*, 69-90.

16. Glaziou, P.; Sismanidis, C.; Floyd, K.; Raviglione, M. Global epidemiology of tuberculosis. *Cold Spring Harb Perspect Med* **2014**, *5*, a017798.
17. Yu, Y.Y.; Pinsky, P.F.; Caporaso, N.E.; Chatterjee, N.; Baumgarten, M.; Langenberg, P.; Furuno, J.P.; Lan, Q.; Engels, E.A. Lung cancer risk following detection of pulmonary scarring by chest radiography in the prostate, lung, colorectal, and ovarian cancer screening trial. *Arch Intern Med* **2008**, *168*, 2326-2332; discussion 2332.
18. Liang, H.Y.; Li, X.L.; Yu, X.S.; Guan, P.; Yin, Z.H.; He, Q.C.; Zhou, B.S. Facts and fiction of the relationship between preexisting tuberculosis and lung cancer risk: A systematic review. *Int J Cancer* **2009**, *125*, 2936-2944.
19. Wu, C.Y.; Hu, H.Y.; Pu, C.Y.; Huang, N.; Shen, H.C.; Li, C.P.; Chou, Y.J. Pulmonary tuberculosis increases the risk of lung cancer: A population-based cohort study. *Cancer* **2011**, *117*, 618-624.
20. Yu, Y.H.; Liao, C.C.; Hsu, W.H.; Chen, H.J.; Liao, W.C.; Muo, C.H.; Sung, F.C.; Chen, C.Y. Increased lung cancer risk among patients with pulmonary tuberculosis: A population cohort study. *J Thorac Oncol* **2011**, *6*, 32-37.
21. Nalbandian, A.; Yan, B.S.; Pichugin, A.; Bronson, R.T.; Kramnik, I. Lung carcinogenesis induced by chronic tuberculosis infection: The experimental model and genetic control. *Oncogene* **2009**, *28*, 1928-1938.
22. Ardies, C.M. Inflammation as cause for scar cancers of the lung. *Integr Cancer Ther* **2003**, *2*, 238-246.
23. Gottesman, S.; Storz, G. Bacterial small rna regulators: Versatile roles and rapidly evolving variations. *Cold Spring Harb Perspect Biol* **2011**, *3*.
24. Ying, X.; Cao, Y.; Wu, J.; Liu, Q.; Cha, L.; Li, W. Starpicker: A method for efficient prediction of bacterial srna targets based on a two-step model for hybridization. *PLoS One* **2011**, *6*, e22705.
25. Oliva, G.; Sahr, T.; Buchrieser, C. Small rnas, 5' utr elements and rna-binding proteins in intracellular bacteria: Impact on metabolism and virulence. *FEMS Microbiol Rev* **2015**, *39*, 331-349.
26. Arnvig, K.B.; Comas, I.; Thomson, N.R.; Houghton, J.; Boshoff, H.I.; Croucher, N.J.; Rose, G.; Perkins, T.T.; Parkhill, J.; Dougan, G., *et al.* Sequence-based analysis uncovers an abundance of non-coding rna in the total transcriptome of mycobacterium tuberculosis. *PLoS Pathog* **2011**, *7*, e1002342.
27. Miotto, P.; Forti, F.; Ambrosi, A.; Pellin, D.; Veiga, D.F.; Balazsi, G.; Gennaro, M.L.; Di Serio, C.; Ghisotti, D.; Cirillo, D.M. Genome-wide discovery of small rnas in mycobacterium tuberculosis. *PLoS One* **2012**, *7*, e51950.
28. Tsai, C.H.; Baranowski, C.; Livny, J.; McDonough, K.A.; Wade, J.T.; Contreras, L.M. Identification of novel srnas in mycobacterial species. *PLoS One* **2013**, *8*, e79411.
29. DiChiara, J.M.; Contreras-Martinez, L.M.; Livny, J.; Smith, D.; McDonough, K.A.; Belfort, M. Multiple small rnas identified in mycobacterium bovis bcg are also expressed in mycobacterium tuberculosis and mycobacterium smegmatis. *Nucleic Acids Res* **2010**, *38*, 4067-4078.
30. Pelly, S.; Bishai, W.R.; Lamichhane, G. A screen for non-coding rna in mycobacterium tuberculosis reveals a camp-responsive rna that is expressed during infection. *Gene* **2012**, *500*, 85-92.
31. Barh, D.; Bhat, D.; Viero, C. Mireg: A resource for microRNA regulation. *J Integr Bioinform* **2010**, *7*.

32. Barh, D.; Kamapantula, B.; Jain, N.; Nalluri, J.; Bhattacharya, A.; Juneja, L.; Barve, N.; Tiwari, S.; Miyoshi, A.; Azevedo, V., *et al.* Miregulome: A knowledge-base of mirna regulomics and analysis. *Sci Rep* **2015**, *5*, 12832.
33. Latorre, I.; Leidinger, P.; Backes, C.; Dominguez, J.; de Souza-Galvao, M.L.; Maldonado, J.; Prat, C.; Ruiz-Manzano, J.; Sanchez, F.; Casas, I., *et al.* A novel whole-blood mirna signature for a rapid diagnosis of pulmonary tuberculosis. *Eur Respir J* **2015**, *45*, 1173-1176.
34. Xu, Z.; Zhou, A.; Ni, J.; Zhang, Q.; Wang, Y.; Lu, J.; Wu, W.; Karakousis, P.C.; Lu, S.; Yao, Y. Differential expression of mirnas and their relation to active tuberculosis. *Tuberculosis (Edinb)* **2015**, *95*, 395-403.
35. Zheng, L.; Leung, E.; Lee, N.; Lui, G.; To, K.F.; Chan, R.C.; Ip, M. Differential microrna expression in human macrophages with mycobacterium tuberculosis infection of beijing/w and non-beijing/w strain types. *PLoS One* **2015**, *10*, e0126018.
36. Jiang, Q.; Wang, Y.; Hao, Y.; Juan, L.; Teng, M.; Zhang, X.; Li, M.; Wang, G.; Liu, Y. Mir2disease: A manually curated database for microrna deregulation in human disease. *Nucleic Acids Res* **2009**, *37*, D98-104.
37. Dong, L.; Luo, M.; Wang, F.; Zhang, J.; Li, T.; Yu, J. Tumir: An experimentally supported database of microrna deregulation in various cancers. *J Clin Bioinforma* **2013**, *3*, 7.
38. Pischmarov, J.; Kuenne, C.; Billion, A.; Hemberger, J.; Cemic, F.; Chakraborty, T.; Hain, T. Srnadb: A small non-coding rna database for gram-positive bacteria. *BMC Genomics* **2012**, *13*, 384.
39. Li, L.; Huang, D.; Cheung, M.K.; Nong, W.; Huang, Q.; Kwan, H.S. Bsrdb: A repository for bacterial small regulatory rna. *Nucleic Acids Res* **2013**, *41*, D233-238.
40. Barh, D.; Jain, N.; Tiwari, S.; Field, J.K.; Padin-Iruegas, E.; Ruibal, A.; Lopez, R.; Herranz, M.; Bhattacharya, A.; Juneja, L., *et al.* A novel in silico reverse-transcriptomics-based identification and blood-based validation of a panel of sub-type specific biomarkers in lung cancer. *BMC Genomics* **2013**, *14 Suppl 6*, S5.
41. Barh, D.; Gupta, K.; Jain, N.; Khatri, G.; Leon-Sicairos, N.; Canizalez-Roman, A.; Tiwari, S.; Verma, A.; Rahangdale, S.; Shah Hassan, S., *et al.* Conserved host-pathogen ppis. Globally conserved inter-species bacterial ppis based conserved host-pathogen interactome derived novel target in c. Pseudotuberculosis, c. Diphtheriae, m. Tuberculosis, c. Ulcerans, y. Pestis, and e. Coli targeted by piper betel compounds. *Integr Biol (Camb)* **2013**, *5*, 495-509.
42. Wang, J.; Liu, T.; Zhao, B.; Lu, Q.; Wang, Z.; Cao, Y.; Li, W. Srnatarbase 3.0: An updated database for srna-target interactions in bacteria. *Nucleic Acids Res* **2016**, *44*, D248-253.
43. Eggenhofer, F.; Tafer, H.; Stadler, P.F.; Hofacker, I.L. Rnapredator: Fast accessibility-based prediction of srna targets. *Nucleic Acids Res* **2011**, *39*, W149-154.
44. Huang da, W.; Sherman, B.T.; Lempicki, R.A. Systematic and integrative analysis of large gene lists using david bioinformatics resources. *Nat Protoc* **2009**, *4*, 44-57.
45. Luo, H.; Lin, Y.; Gao, F.; Zhang, C.T.; Zhang, R. Deg 10, an update of the database of essential genes that includes both protein-coding genes and noncoding genomic elements. *Nucleic Acids Res* **2014**, *42*, D574-580.
46. Haning, K.; Cho, S.H.; Contreras, L.M. Small rnas in mycobacteria: An unfolding story. *Front Cell Infect Microbiol* **2014**, *4*, 96.
47. Liu, X.; Sempere, L.F.; Galimberti, F.; Freemantle, S.J.; Black, C.; Dragnev, K.H.; Ma, Y.; Fiering, S.; Memoli, V.; Li, H., *et al.* Uncovering growth-suppressive micromas in lung cancer. *Clin Cancer Res* **2009**, *15*, 1177-1183.

48. Lebanony, D.; Benjamin, H.; Gilad, S.; Ezagouri, M.; Dov, A.; Ashkenazi, K.; Gefen, N.; Izraeli, S.; Rechavi, G.; Pass, H., *et al.* Diagnostic assay based on hsa-mir-205 expression distinguishes squamous from nonsquamous non-small-cell lung carcinoma. *J Clin Oncol* **2009**, *27*, 2030-2037.
49. Chen, X.; Ba, Y.; Ma, L.; Cai, X.; Yin, Y.; Wang, K.; Guo, J.; Zhang, Y.; Chen, J.; Guo, X., *et al.* Characterization of micrnas in serum: A novel class of biomarkers for diagnosis of cancer and other diseases. *Cell Res* **2008**, *18*, 997-1006.
50. Markou, A.; Tsaroucha, E.G.; Kaklamanis, L.; Fotinou, M.; Georgoulas, V.; Lianidou, E.S. Prognostic value of mature microrna-21 and microrna-205 overexpression in non-small cell lung cancer by quantitative real-time rt-pcr. *Clin Chem* **2008**, *54*, 1696-1704.
51. Shen, J.; Todd, N.W.; Zhang, H.; Yu, L.; Lingxiao, X.; Mei, Y.; Guarnera, M.; Liao, J.; Chou, A.; Lu, C.L., *et al.* Plasma micrnas as potential biomarkers for non-small-cell lung cancer. *Lab Invest* **2011**, *91*, 579-587.
52. Cho, W.C.; Chow, A.S.; Au, J.S. Restoration of tumour suppressor hsa-mir-145 inhibits cancer cell growth in lung adenocarcinoma patients with epidermal growth factor receptor mutation. *Eur J Cancer* **2009**, *45*, 2197-2206.
53. Yanaihara, N.; Caplen, N.; Bowman, E.; Seike, M.; Kumamoto, K.; Yi, M.; Stephens, R.M.; Okamoto, A.; Yokota, J.; Tanaka, T., *et al.* Unique microrna molecular profiles in lung cancer diagnosis and prognosis. *Cancer Cell* **2006**, *9*, 189-198.
54. Volinia, S.; Calin, G.A.; Liu, C.G.; Ambs, S.; Cimmino, A.; Petrocca, F.; Visone, R.; Iorio, M.; Roldo, C.; Ferracin, M., *et al.* A microrna expression signature of human solid tumors defines cancer gene targets. *Proc Natl Acad Sci U S A* **2006**, *103*, 2257-2261.
55. Yu, L.; Todd, N.W.; Xing, L.; Xie, Y.; Zhang, H.; Liu, Z.; Fang, H.; Zhang, J.; Katz, R.L.; Jiang, F. Early detection of lung adenocarcinoma in sputum by a panel of microrna markers. *Int J Cancer* **2010**, *127*, 2870-2878.
56. Kleinstaub, K.; Heesch, K.; Schattling, S.; Kohns, M.; Sander-Julch, C.; Walzl, G.; Hesseling, A.; Mayatepek, E.; Fleischer, B.; Marx, F.M., *et al.* Decreased expression of mir-21, mir-26a, mir-29a, and mir-142-3p in cd4(+) t cells and peripheral blood from tuberculosis patients. *PLoS One* **2013**, *8*, e61609.
57. Wu, Z.; Lu, H.; Sheng, J.; Li, L. Inductive microrna-21 impairs anti-mycobacterial responses by targeting il-12 and bcl-2. *FEBS Lett* **2012**, *586*, 2459-2467.
58. Li, T.T.; Ogino, S.; Qian, Z.R. Toll-like receptor signaling in colorectal cancer: Carcinogenesis to cancer therapy. *World J Gastroenterol* **2014**, *20*, 17699-17708.
59. Hoste, E.; Arwert, E.N.; Lal, R.; South, A.P.; Salas-Alanis, J.C.; Murrell, D.F.; Donati, G.; Watt, F.M. Innate sensing of microbial products promotes wound-induced skin cancer. *Nat Commun* **2015**, *6*, 5932.
60. Khajuria, N.; Metgud, R. Role of bacteria in oral carcinogenesis. *Indian J Dent* **2015**, *6*, 37-43.
61. Huang, B.; Zhao, J.; Unkeless, J.C.; Feng, Z.H.; Xiong, H. Tlr signaling by tumor and immune cells: A double-edged sword. *Oncogene* **2008**, *27*, 218-224.
62. Cherfils-Vicini, J.; Platonova, S.; Gillard, M.; Laurans, L.; Validire, P.; Caliandro, R.; Magdeleinat, P.; Mami-Chouaib, F.; Dieu-Nosjean, M.C.; Fridman, W.H., *et al.* Triggering of tlr7 and tlr8 expressed by human lung cancer cells induces cell survival and chemoresistance. *J Clin Invest* **2010**, *120*, 1285-1297.
63. Basu, J.; Shin, D.M.; Jo, E.K. Mycobacterial signaling through toll-like receptors. *Front Cell Infect Microbiol* **2012**, *2*, 145.

64. Davila, S.; Hibberd, M.L.; Hari Dass, R.; Wong, H.E.; Sahiratmadja, E.; Bonnard, C.; Alisjahbana, B.; Szeszko, J.S.; Balabanova, Y.; Drobniowski, F., *et al.* Genetic association and expression studies indicate a role of toll-like receptor 8 in pulmonary tuberculosis. *PLoS Genet* **2008**, *4*, e1000218.
65. Kleinnijenhuis, J.; Oosting, M.; Joosten, L.A.; Netea, M.G.; Van Crevel, R. Innate immune recognition of mycobacterium tuberculosis. *Clin Dev Immunol* **2011**, *2011*, 405310.
66. Fabbri, M.; Paone, A.; Calore, F.; Galli, R.; Gaudio, E.; Santhanam, R.; Lovat, F.; Fadda, P.; Mao, C.; Nuovo, G.J., *et al.* Micrnas bind to toll-like receptors to induce prometastatic inflammatory response. *Proc Natl Acad Sci U S A* **2012**, *109*, E2110-2116.
67. Sampson, S.L. Mycobacterial pe/ppe proteins at the host-pathogen interface. *Clin Dev Immunol* **2011**, *2011*, 497203.
68. Kohli, S.; Singh, Y.; Sharma, K.; Mittal, A.; Ehtesham, N.Z.; Hasnain, S.E. Comparative genomic and proteomic analyses of pe/ppe multigene family of mycobacterium tuberculosis h(3)(7)rv and h(3)(7)ra reveal novel and interesting differences with implications in virulence. *Nucleic Acids Res* **2012**, *40*, 7113-7122.
69. Parish, T. Two-component regulatory systems of mycobacteria. *Microbiol Spectr* **2014**, *2*, MGM2-0010-2013.
70. Korkegian, A.; Roberts, D.M.; Blair, R.; Parish, T. Mutations in the essential arabinosyltransferase embc lead to alterations in mycobacterium tuberculosis lipoarabinomannan. *J Biol Chem* **2014**, *289*, 35172-35181.
71. Renard, H.F.; Simunovic, M.; Lemiere, J.; Boucrot, E.; Garcia-Castillo, M.D.; Arumugam, S.; Chambon, V.; Lamaze, C.; Wunder, C.; Kenworthy, A.K., *et al.* Endophilin-a2 functions in membrane scission in clathrin-independent endocytosis. *Nature* **2015**, *517*, 493-496.
72. Cheung, N.; So, C.W.; Yam, J.W.; So, C.K.; Poon, R.Y.; Jin, D.Y.; Chan, L.C. Subcellular localization of een/endophilin a2, a fusion partner gene in leukaemia. *Biochem J* **2004**, *383*, 27-35.
73. Huang, E.W.; Xue, S.J.; Li, X.Y.; Xu, S.W.; Cheng, J.D.; Zheng, J.X.; Shi, H.; Lv, G.L.; Li, Z.G.; Li, Y., *et al.* Een regulates the proliferation and survival of multiple myeloma cells by potentiating igf-1 secretion. *Biochem Biophys Res Commun* **2014**, *447*, 271-277.
74. Matsutani, T.; Hiwasa, T.; Takiguchi, M.; Oide, T.; Kunimatsu, M.; Saeki, N.; Iwadate, Y. Autologous antibody to src-homology 3-domain grb2-like 1 specifically increases in the sera of patients with low-grade gliomas. *J Exp Clin Cancer Res* **2012**, *31*, 85.
75. Gurucha, S.S.; Usha, V.; Cox, J.A.; Futterer, K.; Abrahams, K.A.; Bhatt, A.; Alderwick, L.J.; Reynolds, R.C.; Loman, N.J.; Nataraj, V., *et al.* Biochemical and structural characterization of mycobacterial aspartyl-trna synthetase asps, a promising tb drug target. *PLoS One* **2014**, *9*, e113568.
76. Zhu, N.; Lin, Y.; Li, D.; Gao, N.; Liu, C.; You, X.; Jiang, J.; Jiang, W.; Si, S. Identification of an anti-tb compound targeting the tyrosyl-trna synthetase. *J Antimicrob Chemother* **2015**, *70*, 2287-2294.
77. Ioerger, T.R.; O'Malley, T.; Liao, R.; Guinn, K.M.; Hickey, M.J.; Mohaideen, N.; Murphy, K.C.; Boshoff, H.I.; Mizrahi, V.; Rubin, E.J., *et al.* Identification of new drug targets and resistance mechanisms in mycobacterium tuberculosis. *PLoS One* **2013**, *8*, e75245.

78. Offenhauser, N.; Borgonovo, A.; Disanza, A.; Romano, P.; Ponzanelli, I.; Iannolo, G.; Di Fiore, P.P.; Scita, G. The eps8 family of proteins links growth factor stimulation to actin reorganization generating functional redundancy in the ras/rac pathway. *Mol Biol Cell* **2004**, *15*, 91-98.
79. Li, B.; Zhao, W.D.; Tan, Z.M.; Fang, W.G.; Zhu, L.; Chen, Y.H. Involvement of rho/rock signalling in small cell lung cancer migration through human brain microvascular endothelial cells. *FEBS Lett* **2006**, *580*, 4252-4260.
80. Mandai, K.; Nakanishi, H.; Satoh, A.; Takahashi, K.; Satoh, K.; Nishioka, H.; Mizoguchi, A.; Takai, Y. Ponsin/sh3p12: An I-afadin- and vinculin-binding protein localized at cell-cell and cell-matrix adherens junctions. *J Cell Biol* **1999**, *144*, 1001-1017.
81. Telenti, A.; Imboden, P.; Marchesi, F.; Lowrie, D.; Cole, S.; Colston, M.J.; Matter, L.; Schopfer, K.; Bodmer, T. Detection of rifampicin-resistance mutations in mycobacterium tuberculosis. *Lancet* **1993**, *341*, 647-650.
82. Chaves, F.; Alonso-Sanz, M.; Rebollo, M.J.; Tercero, J.C.; Jimenez, M.S.; Noriega, A.R. RpoB mutations as an epidemiologic marker in rifampin-resistant mycobacterium tuberculosis. *Int J Tuberc Lung Dis* **2000**, *4*, 765-770.
83. Bobadilla-del-Valle, M.; Ponce-de-Leon, A.; Arenas-Huertero, C.; Vargas-Alarcon, G.; Kato-Maeda, M.; Small, P.M.; Couary, P.; Ruiz-Palacios, G.M.; Sifuentes-Osornio, J. RpoB gene mutations in rifampin-resistant mycobacterium tuberculosis identified by polymerase chain reaction single-stranded conformational polymorphism. *Emerg Infect Dis* **2001**, *7*, 1010-1013.
84. Zaczek, A.; Brzostek, A.; Augustynowicz-Kopec, E.; Zwolska, Z.; Dziadek, J. Genetic evaluation of relationship between mutations in rpoB and resistance of mycobacterium tuberculosis to rifampin. *BMC Microbiol* **2009**, *9*, 10.

Supplementary table legends

Supplementary Table S1

Common miRNAs that are deregulated both in pulmonary tuberculosis and lung cancer.

Supplementary Table S2

Sequences and similarities among of *M. tuberculosis* sRNA_1096, sRNA_1414, and human hsa-miR-21.

Supplementary Table S3

RNApredator based predicted targets of *M. tuberculosis* sRNA_1096 and sRNA_1414.

II. Issues about DSM identification

What evidence is there to support the validity of DSM identification?
How many conditions, how many DSM, and how many are interrelated conditions?

III. Issues about DSM functional attribution

How many DSM conditions are there?
How do DSM conditions relate to the DSM and to the DSM conditions?
How many are interrelated conditions?

DSM conditions are related to DSM conditions

How many DSM conditions are related to DSM conditions?
How many are interrelated conditions?

DSM conditions are related to DSM conditions

How many DSM conditions are related to DSM conditions?
How many are interrelated conditions?

Exploring the DSM conditions as DSM in the DSM

How many DSM conditions are related to DSM conditions?
How many are interrelated conditions?
How many are interrelated conditions?
How many are interrelated conditions?

How many DSM conditions are related to DSM conditions?

How many DSM conditions are related to DSM conditions?
How many are interrelated conditions?
How many are interrelated conditions?

How many DSM conditions are related to DSM conditions?
How many are interrelated conditions?
How many are interrelated conditions?

Exploring the DSM conditions' hypothesis

How many DSM conditions are related to DSM conditions?
How many are interrelated conditions?
How many are interrelated conditions?
How many are interrelated conditions?

Experimental studies