ECCM 2024 Conference
The 2nd International Electronic Conference on Clinical Medicine
13–15 November 2024 | Online
MDPI

Integrative RNA-Seq and DEG Analysis for the Identification of Clinical Biomarkers in Tuberculosis Infectious Disease

Javed Aalam (javedalambioinformatics@gmail.com) 1*, Rafat Parveen (rparveen@jmi.ac.in) 1
1Department of Computer Science, Jamia Millia Islamia, New Delhi, India – 110025

INTRODUCTION & AIM

Tuberculosis (TB) remains a major global health threat with rising cases. The World Health Organization's 2023 Global Tuberculosis Report reveals a substantial increase in TB cases, with 8.2 million new diagnoses, up from 7.5 million in 2022 **[1]**. Accurate differentiation between Active TB (ATB), characterised by symptoms and high transmission risk, and Latent TB Infection (LTBI), a dormant state, is essential for effective treatment. However, traditional diagnostics often lack precision in distinguishing these stages. This study seeks to identify diagnostic biomarkers to improve TB staging accuracy using advanced transcriptomic methods. By analysing differential gene expression through RNA sequencing and bioinformatics, the research focuses on genes linked to inflammatory responses in ATB and regulatory pathways in LTBI, aiming to enhance diagnostic methods and patient care **[2]**.

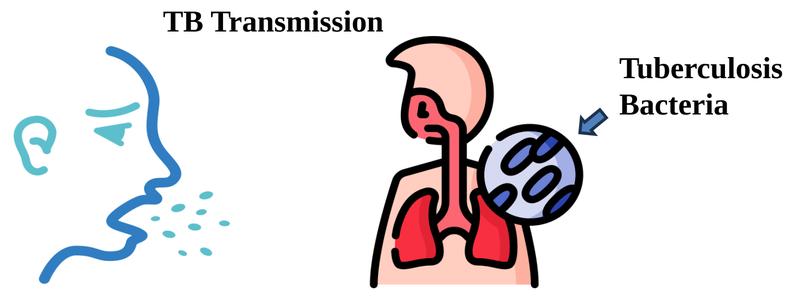

Figure 1. Major routes of Tuberculosis transmission: A Global Perspective

METHOD

Sample Collection Samples were obtained from the NCBI database with accession number GSE41055, including individuals diagnosed with active tuberculosis (ATB), latent tuberculosis infection (LTBI), and healthy controls.

High-Throughput Transcriptome Analysis: RNA sequencing (RNA-Seq) was employed to examine gene expression profiles, with differential gene expression analysis identifying distinct gene profiles in ATB and LTBI. Proteomic techniques were used to support the identification of biomarker candidates.

Data Processing and Analysis: Preprocessing- Raw RNA-Seq data were quality-checked, trimmed, and aligned to a reference genome.

Differential Gene Expression (DEG) Analysis: DESeq2 in R was used to normalise and analyse expression across groups, identifying statistically significant genes with distinct expression in ATB versus LTBI.

Biomarker Identification: Significant DEGs were identified as potential diagnostic biomarkers for distinguishing ATB and LTBI.

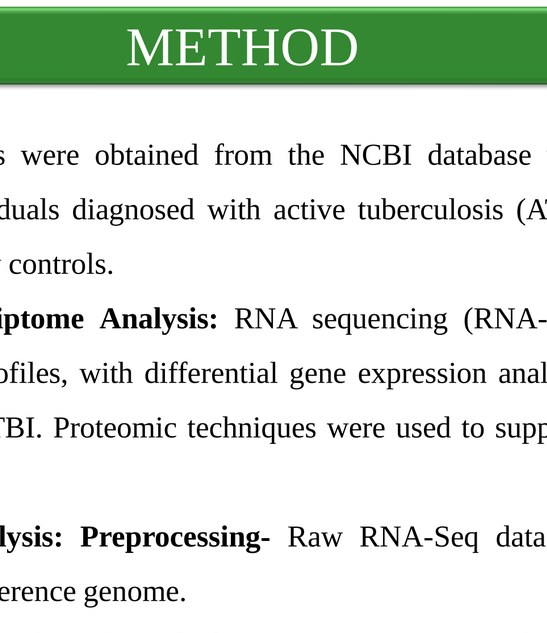

Figure 2. Demonstrates workflow for biomarkers methodology

RESULTS & DISCUSSION

Type of TB	Category	S/No	ID	adj.P.Val	P.Value	t	B	logFC	GB_LIST	Gene Symbols
ATB	Upregulated	1.	2328990	0.823	0.0092213	2.91335426	-3.54	1.31	NM_005610	RBBP4
		2.	3444525	0.823	0.0117107	2.80261108	-3.62	3.21	NM_176887	TAS2R46
		3.	2565935	0.823	0.0135246	2.73538174	-3.66	4.58	NM_025190	ANKRD36B
		4.	3019158	0.823	0.0237513	2.46806061	-3.84	1.04	NM_001099660	LRRN3
		5.	3935486	0.823	0.0304327	2.34763831	-3.91	1.07	NM_006272	S100B
	Downregulated	6.	3026216	0.823	0.0002905	-4.47061788	-2.59	-1.26	NM_001006626	CHRM2
		7.	2742109	0.823	0.0004813	-4.24354007	-2.72	-4.85E-01	NM_002006	FGF2
		8.	3197955	0.823	0.0008935	-3.96643147	-2.88	-3.06E-01	NM_000170	GLDC
		9.	3835814	0.823	0.001126	-3.86301112	-2.94	-9.41E-01	NM_001042724	NECTIN2
		10.	2663295	0.823	0.0014465	-3.75099755	-3.01	-7.86E-01	NM_018306	TMEM40
LTBI	Upregulated	1.	2717078	0.511	0.0001776	4.695006	-2.19	2.1494231	NM_005980	S100P
		2.	4048241	0.511	0.0075366	3.006615	-3.35	6.7526135	NM_002125	HLA-DRB5
		3.	3061268	0.511	0.0097059	2.890128	-3.43	1.0154671	NM_001040057	FAM133B
		4.	2921374	0.511	0.0124237	2.775456	-3.52	1.1434911	NM_032194	RPF2
		5.	4048265	0.511	0.0128802	2.758602	-3.53	3.9713506	NM_002124	HLA-DRB1
	Downregulated	6.	2326774	0.511	0.000109	-4.917991	-2.06	-0.3533251	NM_006142	SFN
		7.	3447798	0.511	0.0003033	4.452585	-2.34	0.3830332	NM_001082972	DNAI7
		8.	2721633	0.511	0.0003062	-4.448279	-2.34	-0.2944936	NM_003102	SOD3
		9.	3455516	0.511	0.0005591	-4.17742	-2.52	-0.3289351	NM_002273	KRT8
		10.	2544565	0.511	0.0008018	4.015893	-2.63	0.3636033	NM_016544	DNAJC27

- In active tuberculosis (ATB), six genes were significantly upregulated (e.g., RBP4, S100B) and five genes were downregulated (e.g., FGF2, GLDC) based on criteria of p-value < 0.05 and log fold change ≥1 for upregulation or ≤-1 for downregulation.
- In latent tuberculosis infection (LTBI), five genes showed significant upregulation (e.g., S100P) and five showed downregulation (e.g., SFN), suggesting potential biomarkers to differentiate between ATB and LTBI.

CONCLUSION AND FUTURE WORK

Conclusion: Differential expression of specific genes in ATB and LTBI suggests potential biomarkers for distinguishing between these tuberculosis states.

Future Work: Further validation of these biomarkers in clinical settings will help refine their diagnostic accuracy for TB differentiation.

REFERENCES

1. World Health Organization, “Global Tuberculosis Report 2023,” WHO, Geneva, 2023. Available: https://www.who.int/teams/global-tuberculosis-programme/tb-reports.
2. Tailleux, L., Waddell, S. J., Pelizzola, M., Mortellaro, A., Withers, M., Tanne, A., Butcher, P. D., ... & Neyrolles, O. (2008). Probing host-pathogen cross-talk by transcriptional profiling of both *Mycobacterium tuberculosis* and infected human dendritic cells and macrophages. *PLoS ONE*, *3*(1), e1403. https://doi.org/10.1371/journal.pone.0001403.
3. Huang, Z., Luo, Q., Guo, Y., Chen, J., Xiong, G., Peng, Y., & Su, B. (2014). RNA-Seq analysis reveals distinct gene expression profiles in response to *Mycobacterium tuberculosis* infection, contributing to biomarker discovery for tuberculosis. *Inflammation*, *37*(2), 531-541. https://doi.org/10.1007/s10753-013-9765-2.