Gene expression signatures in tuberculosis have greater overlap with autoimmune than infectious diseases

Kalum Clayton[1], Marta E. Polak[1], Christopher H. Woelk[1], Paul Elkington[1*]

[1]NIHR Southampton Biomedical Research Centre, Clinical and Experimental Sciences Academic Unit, Faculty of Medicine, University of Southampton, UK

***Address for correspondence:**

Paul Elkington

Professor of Respiratory Medicine

Clinical and Experimental Sciences

University of Southampton

Mailpoint 811, Southampton General Hospital,

Southampton SO16 1YD

Tel 00 44 23 8079 6149

Fax 00 44 20 8383 3394

E-mail p.elkington@soton.ac.uk

Keywords: tuberculosis, host immune response, autoimmunity, pathology

Acknowledgement: This work was supported by a Medical Research Council Global Challenges Research Fund grant MR/P023754/1. MEP is supported by a Wellcome Trust Sir Henry Dale Fellowship

Tuberculosis continues to be depressingly persistent as a global pandemic, killing more than any other infectious agent (1), and the mechanisms underlying pulmonary pathology remain elusive (2). Diverse clinical phenomena and experimental observations have led us to recently hypothesise that part of the pathological process is an infection-initiated autoimmune phenomenon (3), and others have since noted these associations (4, 5). To investigate this hypothesis, we performed an analysis of differentially expressed genes (DEGs) between patients with tuberculosis and those with autoimmune or infectious diseases using within-study control groups. We obtained gene expression datasets from the Gene Expression Omnibus (Accession numbers GSE19444 and GSE22098, Illumina platform) for reanalysis of a major published study (6) to test this new hypothesis.

The tuberculosis cohort all had culture confirmed pulmonary disease. The autoimmune diseases studied were adult and paediatric Systemic Lupus Erythematosus, and adult Stills disease. The infectious diseases were culture-confirmed staphylococcal or streptococcal bacteraemia. All TB patient samples were taken before treatment initiation, while no information was available to the authors regarding the treatment status of the autoimmune cases, which could potentially introduce confounding factors. To control for variability between different microarray experiments, we first defined all DEGs in each condition relative to the healthy control patients for each dataset. DEGs were identified using the *limma* package in R on quantile normalized raw expression data obtained from the GEO datasets using a moderated T-test statistical filter (Benjamini Hochberg [False discovery rate, FDR] adjusted $p<0.05$) for defining differential expression.

Our analysis identified 2,468 DEGs in tuberculosis, 8,134 DEGs in infection and 11,348 DEGs in autoimmune disorders relative to controls, creating disease-specific gene signatures (Figure 1). 1,481 DEGs were common to all three conditions, therefore representing a generic inflammatory response, and this comprised 60.0% of the tuberculosis signature. Only 96 DEGs were shared exclusively between tuberculosis and infection (3.8% of the tuberculosis signature), whereas between tuberculosis and autoimmune disease, there were 810 common DEGs (32.8% of the tuberculosis signature). Notably, once the shared genes with infection and autoimmune diseases were accounted for, only 81

(3.3% of total DEGs) were exclusive to tuberculosis. We further refined our disease DEG lists by quantitative filtering (fold-change >2 or <-2), which reduced the number of genes identified, and the observed disease-overlap relationships was broadly maintained (Figure 1, values in brackets). By this more stringent analysis, 48.0% of the total tuberculosis signature was common to the autoimmune signature, whereas only 3.9% genes were common to the infectious signature. Of the 35 common genes differentially expressed in both TB and the autoimmune conditions, all were found to be modulated in the same direction (i.e. up- or down-regulated) between TB and autoimmunity.

Next, we studied the pathways that were communal between tuberculosis and autoimmune diseases. We used ToppFun functional enrichment in the ToppGene suite to determine gene ontology for disease associations for the tuberculosis-autoimmune overlap genes. Using the FDR-corrected p-value filtered common tuberculosis-autoimmune DEG list, we annotated significant (FDR p-value <0.05) associations to a number of autoimmune diseases, such as systemic lupus erythematosus, Behcet's disease and rheumatoid arthritis. Applying an additional filter for fold-change (fold-change >2 or <-2), the 35 common TB-autoimmune DEGs were also annotated to infer biological processes common between autoimmune disease and tuberculosis (Table 1). These gene ontology associations were dominated the type I interferon signalling pathway.

These analyses support the hypothesis that an autoimmune process contributes to pathology in pulmonary tuberculosis (3). Differentially expressed genes in tuberculosis predominantly overlap with autoimmune disease, and less so with infection. Furthermore, combining infection and autoimmune disease signatures explains 96.7% of the differentially expressed tuberculosis signature, suggesting pathology in tuberculosis results from the interplay between infection and a currently unrecognised autoimmune processes. Our analysis identified a predominant role for type I interferon signalling, which have been known to be critical in autoimmune disease for many years (7) and has recently emerged as key to the gene expression signature in TB (8). Our study is both strengthened but also limited by the fact that gene expression signatures are all derived from samples taken from whole blood, as opposed to the site of disease. The benefit is that this approach permits comparison

between different diseases by sampling the same compartment. The fact that autoimmune disease and tuberculosis have major commonalities in the circulating gene expression signatures remote from the site of disease supports the concept that underlying pathological mechanisms may be shared.

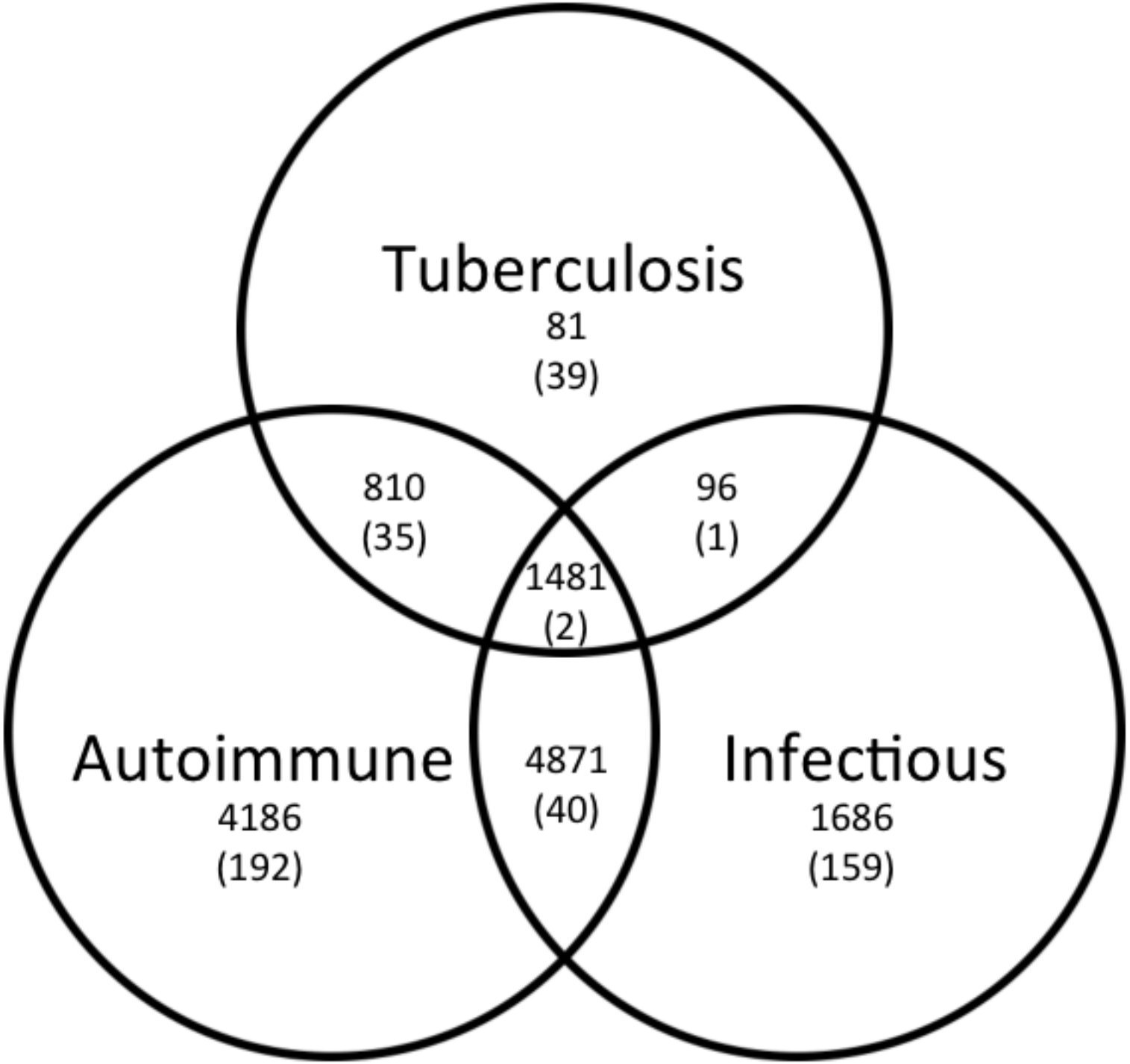
An important next step is to compare expression signatures at the respective sites of disease to identify common pathways that can underpin dissection of the fundamental pathological process. This may elucidate the disease mechanisms in pulmonary tuberculosis, which continues to be poorly understood and for which standard treatment has remained unchanged for 40 years and vaccination for almost 100 years (9). Defining an autoimmune process that exacerbates pathology in tuberculosis would help inform emerging host directed therapies, to limit the lung tissue destruction that results in mortality and to potentially accelerate cure (10), as well as avoid vaccination strategies that may inadvertently exacerbate lung destruction and transmission.

Figure legend: Venn diagram demonstrating overlap in Differentially Expressed Genes (DEGs) relative to healthy controls in tuberculosis, infectious and autoimmune disease. Total Differentially Expressed Genes analyzed by p value are shown, and values further filtered by fold change of greater than +/-2 are shown in brackets.

References

1. Wallis RS, Maeurer M, Mwaba P, Chakaya J, Rustomjee R, Migliori GB, Marais B, Schito M, Churchyard G, Swaminathan S, Hoelscher M, Zumla A. Tuberculosis-advances in development of new drugs, treatment regimens, host-directed therapies, and biomarkers. *Lancet Infect Dis* 2016; 16: e34-46.
2. Ong CW, Elkington PT, Friedland JS. Tuberculosis, pulmonary cavitation, and matrix metalloproteinases. *Am J Respir Crit Care Med* 2014; 190: 9-18.
3. Elkington P, Tebruegge M, Mansour S. Tuberculosis: An Infection-Initiated Autoimmune Disease? *Trends Immunol* 2016; 37: 815-818.
4. Mourik BC, Lubberts E, de Steenwinkel JEM, Ottenhoff THM, Leenen PJM. Interactions between Type 1 Interferons and the Th17 Response in Tuberculosis: Lessons Learned from Autoimmune Diseases. *Frontiers in immunology* 2017; 8: 294.
5. Chao WC, Lin CH, Liao TL, Chen YM, Chen DY, Chen HH. Association between a history of mycobacterial infection and the risk of newly diagnosed Sjogren's syndrome: A nationwide, population-based case-control study. *PLoS One* 2017; 12: e0176549.
6. Berry MP, Graham CM, McNab FW, Xu Z, Bloch SA, Oni T, Wilkinson KA, Banchereau R, Skinner J, Wilkinson RJ, Quinn C, Blankenship D, Dhawan R, Cush JJ, Mejias A, Ramilo O, Kon OM, Pascual V, Banchereau J, Chaussabel D, O'Garra A. An interferon-inducible neutrophil-driven blood transcriptional signature in human tuberculosis. *Nature* 2010; 466: 973-977.
7. Pascual V, Farkas L, Banchereau J. Systemic lupus erythematosus: all roads lead to type I interferons. *Curr Opin Immunol* 2006; 18: 676-682.
8. Cliff JM, Kaufmann SH, McShane H, van Helden P, O'Garra A. The human immune response to tuberculosis and its treatment: a view from the blood. *Immunol Rev* 2015; 264: 88-102.
9. Koenig SP, Furin J. Update in Tuberculosis/Pulmonary Infections 2015. *Am J Respir Crit Care Med* 2016; 194: 142-146.

10. Zumla A, Chakaya J, Hoelscher M, Ntoumi F, Rustomjee R, Vilaplana C, Yeboah-Manu D, Rasolof V, Munderi P, Singh N, Aklillu E, Padayatchi N, Macete E, Kapata N, Mulenga M, Kibiki G, Mfinanga S, Nyirenda T, Maboko L, Garcia-Basteiro A, Rakotosamimanana N, Bates M, Mwaba P, Reither K, Gagneux S, Edwards S, Mfinanga E, Abdulla S, Cardona PJ, Russell JB, Gant V, Noursadeghi M, Elkington P, Bonnet M, Menendez C, Dieye TN, Diarra B, Maiga A, Aseffa A, Parida S, Wejse C, Petersen E, Kaleebu P, Oliver M, Craig G, Corrah T, Tientcheu L, Antonio M, Rao M, McHugh TD, Sheikh A, Ippolito G, Ramjee G, Kaufmann SH, Churchyard G, Steyn A, Grobusch M, Sanne I, Martinson N, Madansein R, Wilkinson RJ, Mayosi B, Schito M, Wallis RS, Maeurer M. Towards host-directed therapies for tuberculosis. *Nat Rev Drug Discov* 2015; 14: 511-512.


Tuberculosis
81
(39)
810
(35)
96
(1)
1481
(2)
Autoimmune
4186
(192)
4871
(40)
Infectious
1686
(159)

Biological Process Gene Ontology analysis: Tuberculosis-Autoimmune Differentially Expressed Gene overlap			
Identity	**Name**	**FDR BH**	**Hits to annotation**
GO:0060337	Type I interferon signalling pathway	1.39E-12	8/81
GO:0071357	Cellular response to type I interferon	1.54E-12	8/82
GO:0034340	Response to type I interferon	2.28E-12	8/86
GO:0034097	Response to cytokine	3.27E-12	15/825
GO:0045087	Innate immune response	4.77E-12	15/847

Table 1: Gene ontology associations of the common differentially expressed genes between tuberculosis and autoimmune diseases. Biological processes were annotated from a Differentially Expressed Gene list filtered statistically and quantitatively (p $p<0.05$, fold-change >2 or <-2). Disease ontology associations were annotated using ToppFun functional enrichment in the ToppGene suite, and we report the Benjamini-Hochberg (BH)-adjusted p-value to account for false discovery rate. FDR: False discovery rate.