

Blood transcriptomic diagnosis of active TB with five genes

JK Roe^{1*}, N Thomas¹, E Gil¹, K Best¹, E Tsaliki¹, S Morris-Jones², S Stafford¹, N Simpson¹, KD Witt³, B Chain¹, R Miller⁴, A Martineau³, M Noursadeghi¹

1 Infection & Immunity, University College London, UK
2 Department of Microbiology, University College London Hospitals NHS Trust, UK
3 Blizard Institute, Queen Mary University of London, UK
4 Research Department of Infection and Population Health, University College London, UK

Contact: j.roe@ucl.ac.uk Website: www.innate2adaptive.com



BACKGROUND

- Confirming the diagnosis of active tuberculosis (TB) requires samples from the site of disease. Peripheral blood transcriptional profiling offers a potential alternative. Multiparametric blood transcriptomic signatures of TB have been described as potential diagnostic tests.

AIM

- Identify the minimal set of genes that can accurately diagnose active TB, both pulmonary and extrapulmonary, and evaluate their specificity.

METHODS

- Support vector machine learning (SVM), combined with feature selection, was used to identify a minimum TB-specific signature then classify active TB vs health or fever. Performance shown by receiver operating characteristic (ROC) curves, area under curve (AUC) represents accuracy of classification.

BATF2 DISCRIMINATES ACTIVE TB AND HEALTH

- Whole genome ranked by SVM weightings (Fig 1A)
- Tested genes cumulatively in rank order to determine minimal gene set for discrimination of active TB and health (Fig 1B)

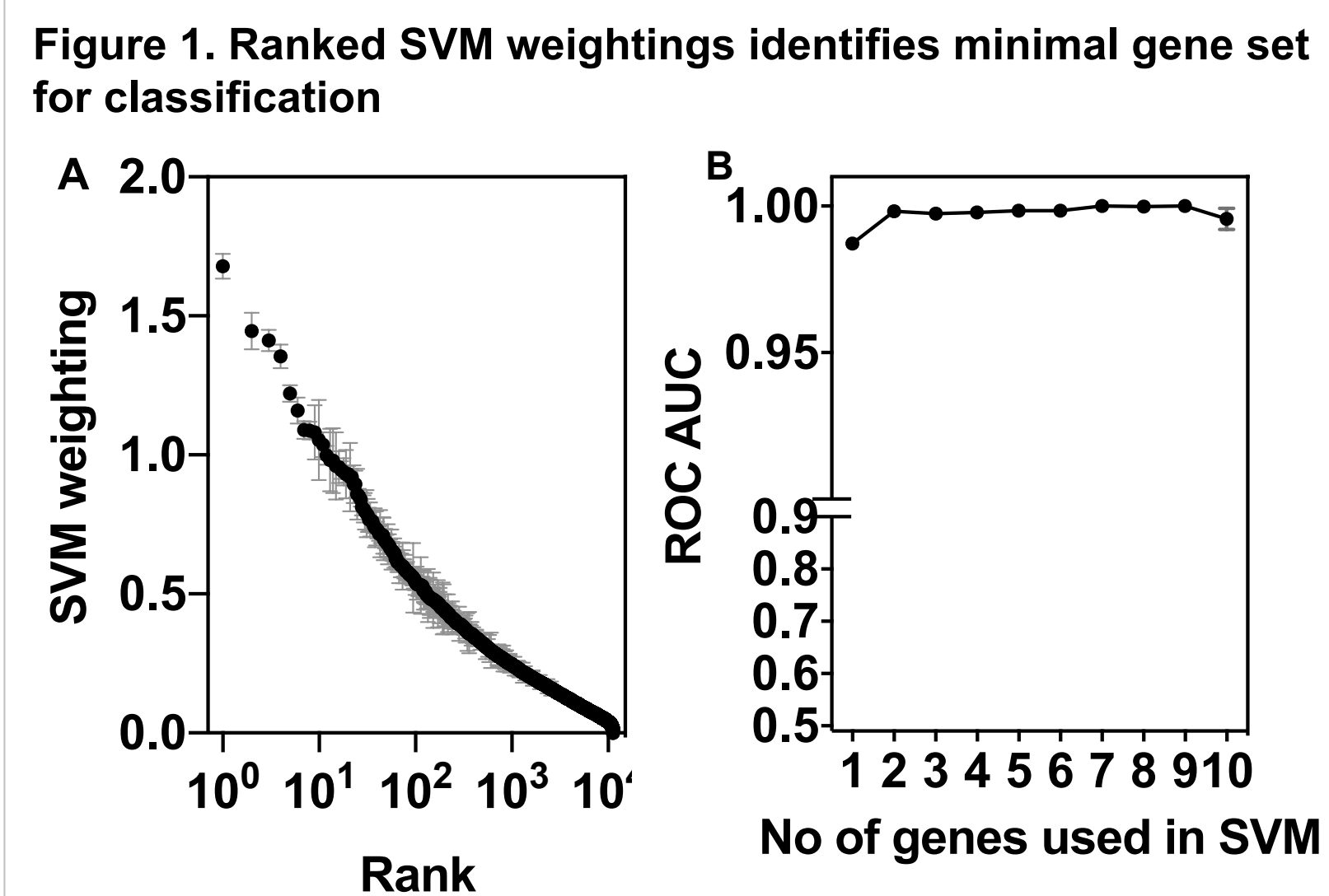
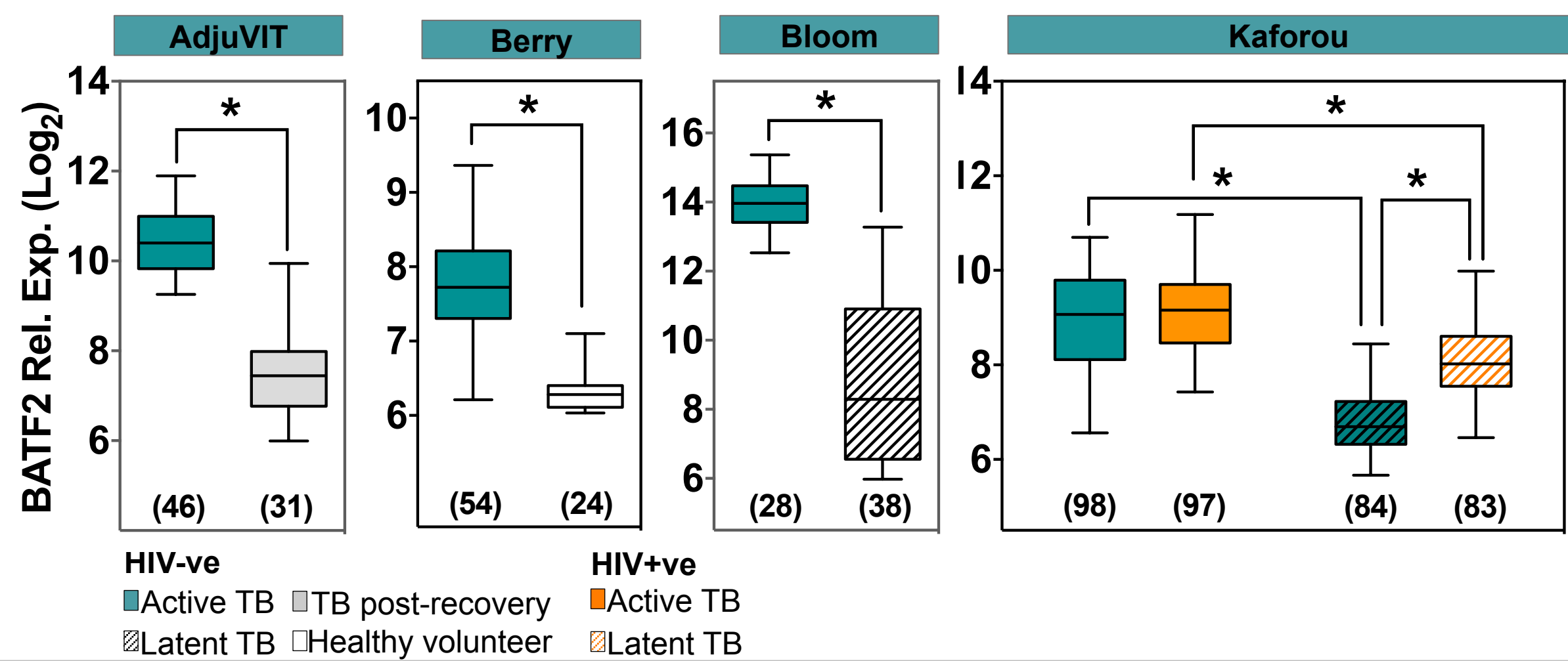
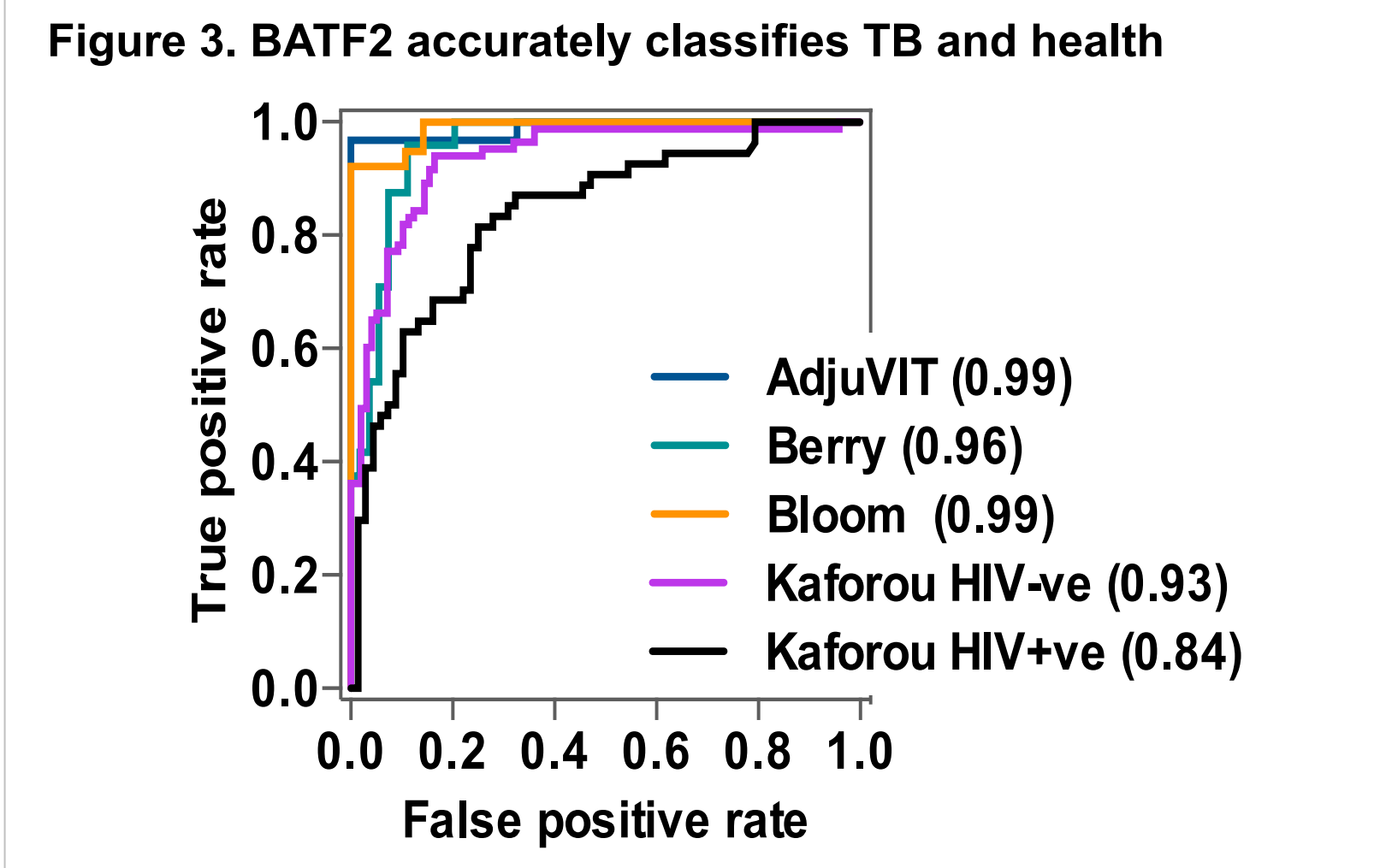


Figure 2. Expression of BATF2 is higher TB than health across 4 independent cohorts



- BATF2 alone discriminates TB from health in 5 cohorts including an HIV+ve cohort (Figs 2 & 3)



CLASSIFICATION OF ACTIVE TB AND FEVER WITH 4 GENES

- Fever cohort = range of community acquired infections, T>38C, attending hospital
- BATF2 not TB-specific (Fig 4A)
- Whole genome SVM of Active TB vs Fever generated ranked weightings

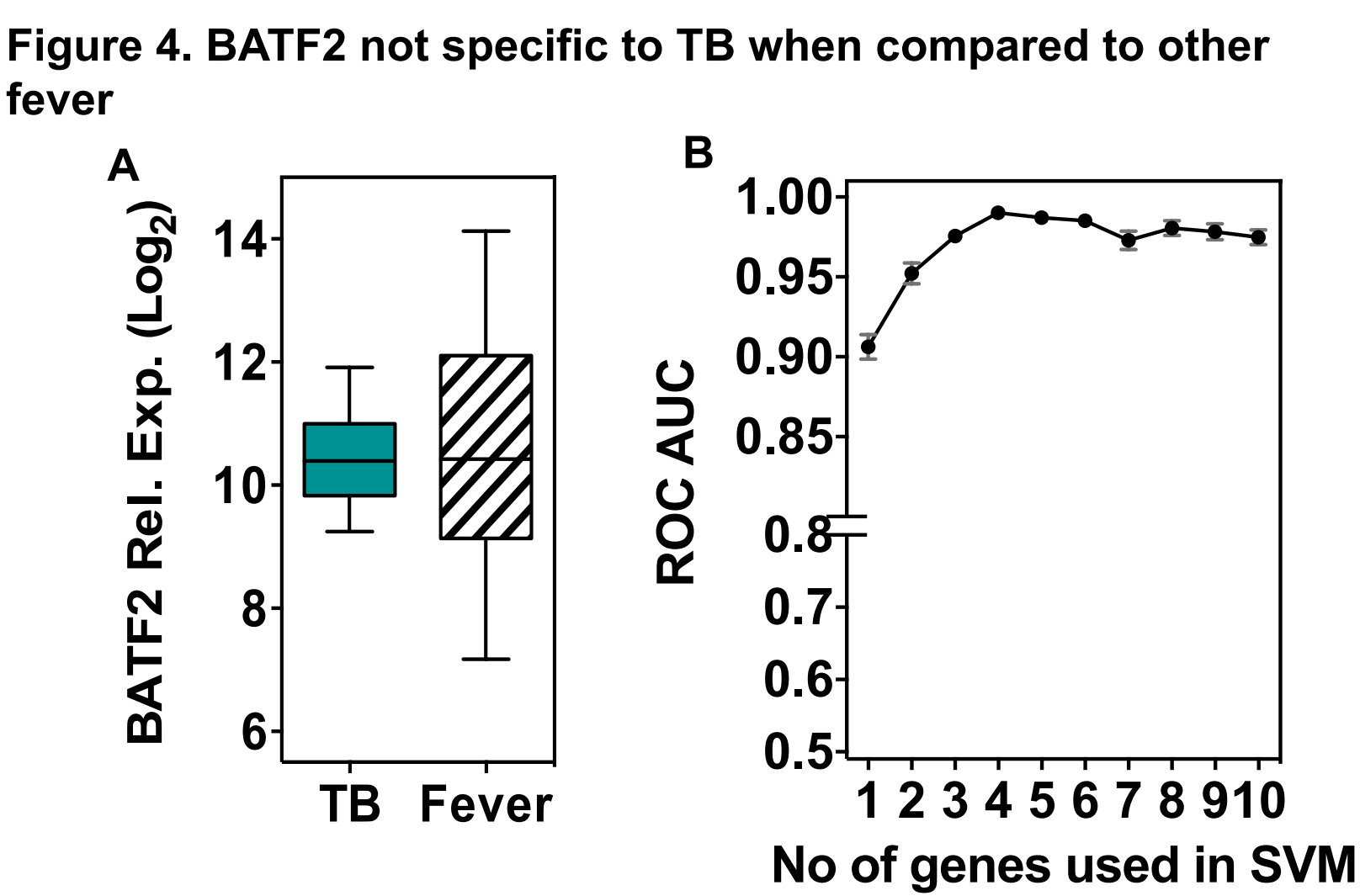
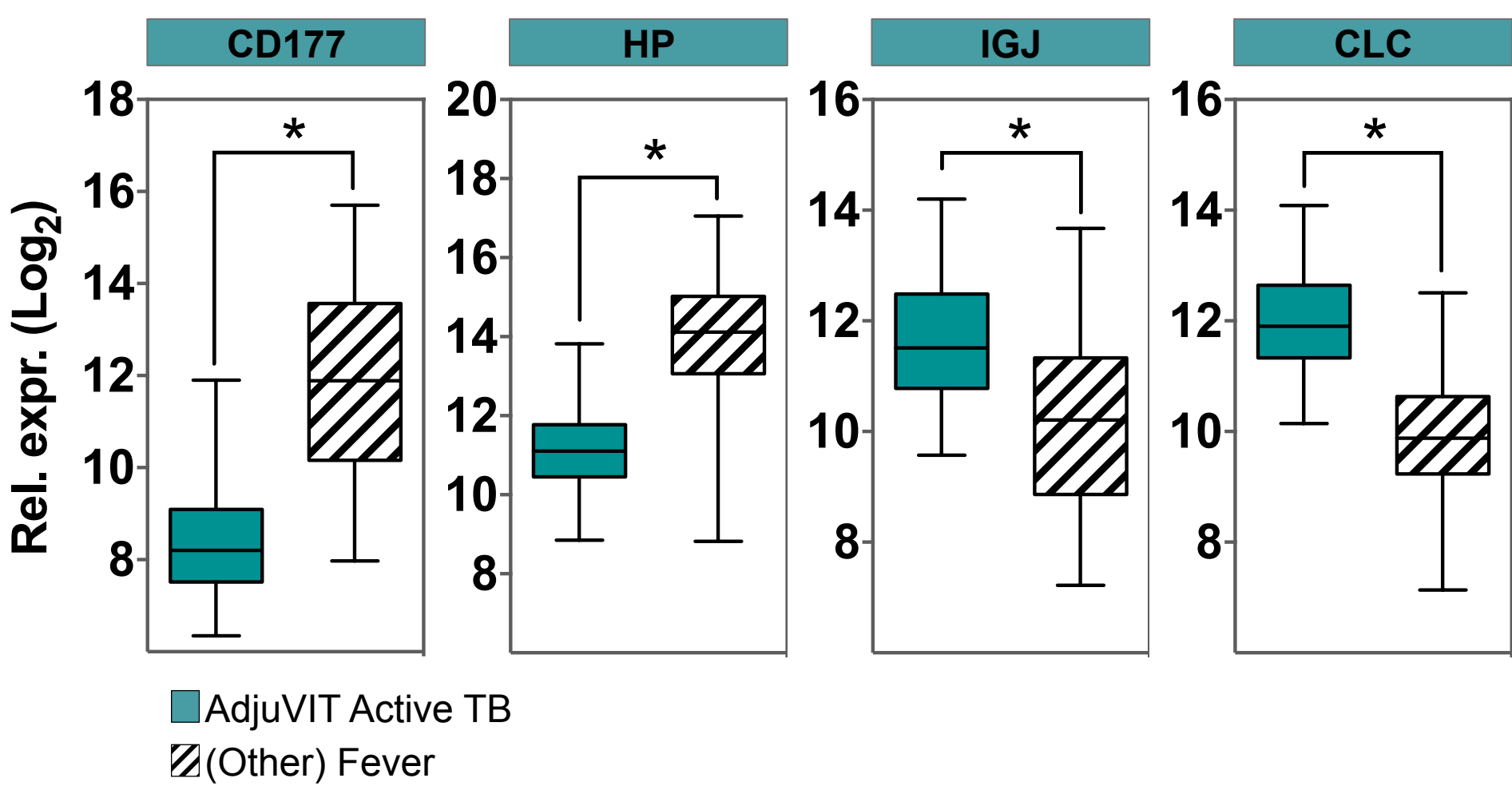
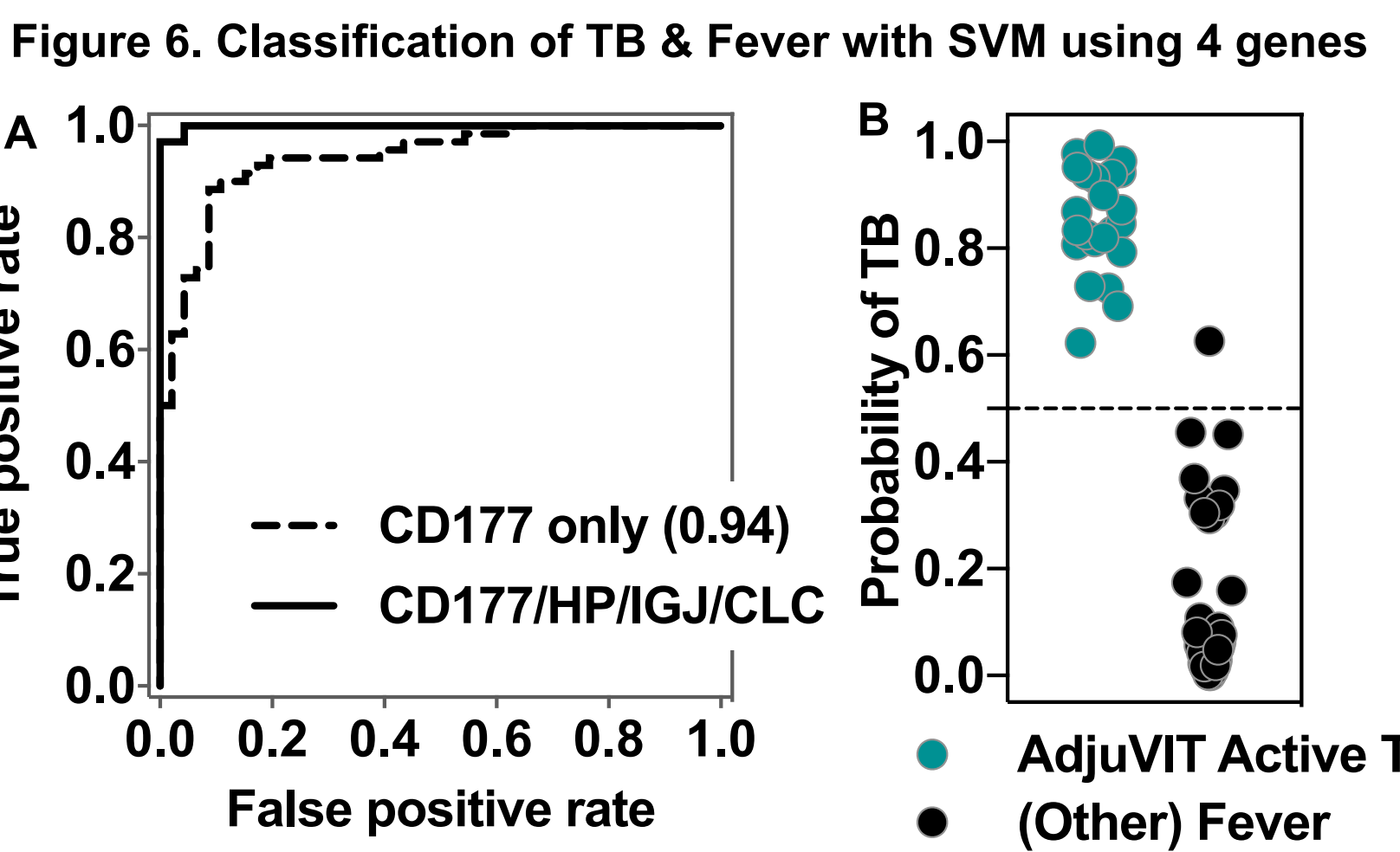


Figure 5. Expression levels of 4 top genes in Active TB and Fever cases



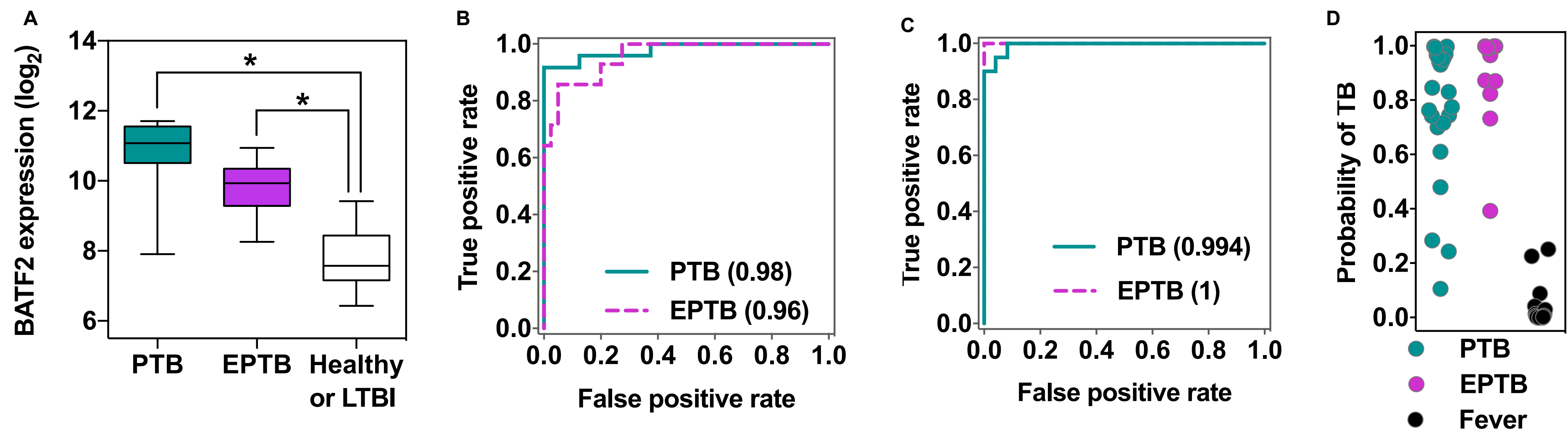
- Cumulative testing strategy identified 4 genes that discriminate TB from Fever (Fig 4B & 5)

- CD177 most discriminatory gene
- AUC improved by inclusion of 3 further genes (Fig 6A)
- SVM used to derive single probability score of active TB for each case (Fig 6B)



BATF2 & 4 GENE SIGNATURE DIAGNOSE PULMONARY & EXTRAPULMONARY TB

Figure 7. BATF2 and 4 gene signature accurately classify Active TB in independent cohorts



- BATF2 expression higher in active TB than health, regardless of site of TB (Fig 7A)
- 4 gene signature accurately classifies cases of PTB or extrapulmonary TB (EPTB) vs healthy cases (Fig 7B) or cases of fever (non-TB pneumonia) (Fig 7C)
- A single probability score can give case by case assessment of confidence in this classification (Fig 7D)

CONCLUSIONS

- Elevated blood BATF2 transcript levels provide a sensitive biomarker discriminating TB from healthy individuals
- A novel 4 gene signature differentiates active TB and other infectious diseases
- These transcriptional biomarkers perform equally well in pulmonary & extrapulmonary TB

FUNDING

