More than *Mycobacterium tuberculosis:* specific site-of-disease microbial communities, functional capacities, and their distinct clinical profiles in tuberculous lymphadenitis

Georgina Nyawo[1,2], Charissa Naidoo[1,2], Benjamin Wu[3], Imran Sulaiman[3], Jose Clemente[4] Yonghua Li[3], Stephanie Minnies[1], Byron Reeve[1], Suventha Moodley[1,2], Cornelia Rautenbach[5,6], Colleen Wright[7], Shivani Singh[3], Andrew Whitelaw[5,6], Pawel Schubert[5,7], Robin Warren[1], Leopoldo Segal[3], Grant Theron[1,2]

Affiliations:

[1]DSI-NRF Centre of Excellence for Biomedical Tuberculosis Research; South African Medical Research Council Centre for Tuberculosis Research; Division of Molecular Biology and Human Genetics, Faculty of Medicine and Health Sciences, Stellenbosch University, Cape Town, South Africa;

[2]African Microbiome Institute, Faculty of Medicine and Health Sciences, Stellenbosch University, Cape Town, South Africa;

[3]Division of Pulmonary, Critical Care, and Sleep Medicine, Department of Medicine, New York University School of Medicine, New York, USA;

[4]Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY, USA;

[5]National Health Laboratory Service, Tygerberg Hospital, Cape Town, South Africa;

[6]Division of Medical Microbiology, Faculty of Medicine and Health Sciences, Stellenbosch University, Cape Town, South Africa;

[7]Division of Anatomical Pathology, Faculty of Medicine and Health Sciences, Stellenbosch University, Cape Town, South Africa.

Corresponding author:

Professor Grant Theron, 2nd Floor BMRI Building, Division of Molecular Biology and Human Genetics, Faculty of Medicine and Health Sciences, Francie Van Zijl Drive, Tygerberg, South Africa, 7550.

E-mail: gtheron@sun.ac.za, Telephone: +27 21 938 9693

29 **Keywords:** Microbiome, Extrapulmonary Tuberculosis, Tuberculous Lymphadenitis**,** HIV

30 **Word count:** 2995

ABSTRACT

Background: Lymphadenitis is the most common extrapulmonary tuberculosis (EPTB) manifestation and a major cause of death. The microbiome is important to human health but uninvestigated in EPTB. We profiled the site-of-disease lymph node microbiome in tuberculosis lymphadenitis (TBL).

Methods: Fine needle aspiration biopsies (FNABs) were collected from 159 pre-treatment presumptive TBL patients in Cape Town, South Africa. 16S Illumina MiSeq rRNA gene sequencing was done.

Results: We analysed 89 definite TBLs (dTBLs) and 61 non-TBLs (nTBLs), which had similar α- but different β-diversities (*p=0.001*). Clustering identified five lymphotypes prior to TB status stratification: *Mycobacterium-, Prevotella-* and *Streptococcus*-dominant lymphotypes were more frequent in dTBLs whereas a *Corynebacterium*-dominant lymphotype and a fifth lymphotype (no dominant taxon) were more frequent in nTBLs. When restricted to dTBLs, clustering identified a *Mycobacterium*-dominant lymphotype with low α-diversity and other non-*Mycobacterium*-dominated lymphotypes (termed *Prevotella-Corynebacterium* and *Prevotella-Streptococcus*). The *Mycobacterium* dTBL lymphotype was associated with HIV-positivity and clinical features characteristic of severe lymphadenitis (e.g., node size). dTBL microbial communities were enriched with potentially proinflammatory microbial short chain fatty acid metabolic pathways (propanoate, butanoate) vs. those in nTBLs. 11% (7/61) of nTBLs had *Mycobacterium* reads.

Conclusions: TBL at the site-of-disease is not microbially homogenous and distinct microbial community clusters exist that are associated with different immunomodulatory potentials and clinical characteristics. Non-*Mycobacterium*-dominated dTBL lymphotypes, which contain taxa potentially targeted by TB treatment, represent less severe potentially earlier stage disease. These investigations lay foundations for studying the microbiome's role in lymphatic TB and the long-term clinical significance of lymphotypes requires prospective evaluation.

250/250

INTRODUCTION

Tuberculosis (TB), which kills 1.5 million people globally each year (including 214 000 people with HIV), causes extrapulmonary tuberculosis (EPTB) [1]. EPTB accounts for ~15% of all TB, and as much as half of all TB in in people living with HIV (PLHIV) in some settings [2]. EPTB is difficult to diagnose [3] and has high mortality.

TB lymphadenitis (TBL) is the most common EPTB manifestation, accounting for 70% of EPTB and most frequently affects peripheral and cervical lymph nodes [4]. TBL occurs after *Mtb* enters the airways, is taken up by phagocytic cells, and transported to a thoracic lymph node where granulomas may form. These steps are also necessary for priming T-cells to generate adaptive immune responses for microbial killing mediated by cytokines and other effector mechanisms [5 6].

Lymph nodes have an important role in TB pathogenesis: enlargement has been documented following exposure, even if only a fraction of patients with enlarged nodes develop active disease [7]. Furthermore, *Mtb* DNA is often found in the lymph nodes of exposed yet healthy people. Lymph nodes are therefore hypothesised to serve as a *Mtb* growth and persistence niche [7] that can spread to bodily sites [8] (in animals lymph node infection almost always accompanies infection in the lungs [9 10]); suggesting that TB may primarily be a lymphatic rather than pulmonary disease [11]. For example, prior to development of active TB, lymph nodes demonstrate enhanced metabolic activity on PET-CT scans [12]. Together these studies show the lymph nodes have an important role in TB pathogenesis, however, the determinants of why *Mtb* sometimes successfully establishes itself in the lymph nodes and subsequently proliferates, including the potential role of other microbes, is understudied. Key to understanding this is characterizing the local site-of-disease.

The human microbiome influences immune function [13]. Two studies assessed lymph node microbial content [14 15], both in mesenteric lymph nodes in Crohn's disease where reduced diversity was observed. The site-of-disease microbiome in TB is underexamined [16]: in bronchoalveolar lavage fluid (BALF), active pulmonary TB was associated with *Mycobacterium* enrichment and *Streptococcus* depletion [17 18].

The site-of-disease microbiome in TBL (including in HIV-endemic settings where TB is common) remains uncharacterised. Therefore, given the apparent role of the lymph nodes in TB pathogenesis, and the importance of the microbiome as a modulator of immunity, we characterised the site-of-disease lymph microbiome in presumptive TBL patients from a high HIV burden setting before the potentially confounding effects of antibiotic-based TB treatment.

METHODS

Patient recruitment and follow-up

Presumptive TBL participants (≥18 years) were recruited from Tygerberg Academic Hospital in Cape Town, SA (25 January 2017-11 December 2018). Participants were programmatically referred for a routine fine needle aspiration biopsy (FNAB) for the investigation of lymphadenopathy as described [19]. Eligible participants were not on TB treatment within six months. Clinical and demographic data were collected by interview and medical record review. Patients with TBL were programmatically diagnosed with TB, initiated on treatment, and study staff assessed treatment response by telephonic follow-up ≥12 weeks. The study had no role in patient management.

Ethics

Patients provided written informed consent. The study was approved by the Health Research and Ethical Committee of Stellenbosch University (N16/04/050), Tygerberg Hospital (Project ID:4134), and the Western Cape Department of Health (WC_2016RP15_762).

Specimen collection and processing

For each patient, two background DNA sampling controls were collected in microcentrifuge tubes prior to lymph node aspiration: a skin swab (collected into saline; Ysterplaat Medical Supplies, Cape Town, South Africa) of the site to be punctured, followed by a saline flush of the syringe to be used for aspiration. Aspiration and microbiological procedures are in the Supplement. Aspirated material from the third pass was collected into 500µL sterile saline and stored at −80 °C until batched DNA extraction.

Routine specimen testing

Patients were categorised based on lymphatic or non-lymphatic mycobacteriological evidence, provided by the government programmatic laboratory (National Health Laboratory

Service [NHLS]), and/or clinical decision to start treatment (**Figure 1**) by responsible clinician thereafter. Case definitions are described in the Supplement (**Table S1**).

Microbial DNA extraction and sequencing

DNA was extracted from specimens and controls using the PureLink Microbiome DNA Purification Kit (Invitrogen, Carlsbad, USA). The 16S rRNA gene V4 hypervariable region (150 bp read length) was amplified and sequenced (paired-ends) on the Illumina MiSeq platform. Lymph, skin swab, and 1 in 5 saline flushes were extracted and sequenced.

Microbiome data analysis

16S rRNA gene sequences (Sequence Read Archive PRJNA738676) were processed, denoised and analysed in Quantitative Insights Into Microbial Ecology (QIIME 2, v2020.8) [20] and DADA2 [21] using closed-reference picking by assigning taxonomy at a 97% similarity against representative sequences in Greengenes (v13.8) [22]. QIIME2 outputs (phylogenetic tree, feature table, taxonomy) and metadata imported into R (v3.5.2) and analyses done using *phyloseq* [23]. Shannon's index was calculated using *vegan* [24] to measure α-diversity (within-sample diversity). Bray-Curtis distances were used to measure β-diversity (between-sample diversity) and construct Principal Coordinate Analysis (PCoA) plots. Dirichlet-Multinomial Mixtures (DMM) modelling was done to estimate the optimal number of clusters (microbial community states known as lymphotypes) based on compositional similarity [25].

Inferred metagenome

Phylogenetic Investigation of Communities by Reconstruction of Unobserved States (PICRUSt) v.2.1.3-b [26] was used to predict gene family abundance with PICRUSt2 default options (picrust2_pipeline.py). The resulting gene table was mapped against the Kyoto Encyclopaedia of Genes and Genomes (KEGG) database, and pathway abundances were inferred from predicted KEGG ORTHOLOGY (KO) abundances.

Differential abundance analysis

Differentially abundant taxa and pathways were identified using *DESeq2* (v1.22.2) [27] with Benjamini-Hochberg multiple testing correction. Feature tables were pruned to taxa with ≥5% relative abundance in 0.5% of samples [28]. Adjusted p-values <0.2 and <0.05 was considered significant for taxa and pathways, respectively (DESq2 tables). PICRUSt *DeSeq2* outputs were used to identify common pathways in L4 vs. each lymphotype (overall patients), and L3 vs. each other lymphotype (dTBLs).

Statistical analyses

Statistical analysis was done in GraphPad Prism Version 7.00 (GraphPad Software, USA) and R. For α-diversity comparisons, the non-parametric Mann-Whitney test or the Kruskal-Wallis ANOVA with Dunn's multiple comparisons was used. For paired comparisons, the Wilcoxon signed rank test was used. Permutational multivariate analysis of variance (PERMANOVA) was computed with 999 permutations to test β-diversity differences and R^2 used to measure the proportion variation explained by a variable. For correlation analysis, the non-parametric Spearman ranking and parametric Pearson ranking tests were used. A p-value <0.05 was significant for all comparisons.

RESULTS

Cohort characteristics

We had 89 dTBLs, 61 nTBLs (**Figure 1**) and 9 pTBLs (latter subsequently excluded due to small numbers). dTBLs were more likely to have supraclavicular or head lymph node involvement than nTBLs, if HIV-positive were more likely to have a lower CD4 count (**Table 1**) and were more likely to have a FNAB that appeared bloody rather than chylous.

Table 1: Demographic and clinical characteristics of patients with presumptive TBL. dTBLs were more likely to have HIV and a lower CD4 count if HIV-positive, supraclavicular lymph node involvement, and a bloody FNAB. Data are n/N (%) or median (IQR).

	Patients with presumptive TB (n=158)			*p-value*
	Total (n=158)*	**dTBL (n=89)**	**nTBL (n=61)**	
Age, years	36 (21-44)	35 (29-40)	38 (30-49)	0.053
Female	85/159 (53)	48/89 (54)	35/61 (57)	0.677
HIV	77/156 (49)	49/89 (55)	23/59 (39)	0.055
CD4+	166 (90-308)	155 (76-251)	250 (139-458)	**0.027**
CD4+ <200 cells/μl	47/77 (61)	32/49 (65)	11/23 (48)	0.159
On ART	38/76 (50)	21/49 (43)	14/22 (64)	0.105
Previous TB	36/156 (23)	24/88 (27)	9/60 (15)	0.078
Tobacco smoking	44/157 (28)	21/89 (24)	22/60 (37)	0.084
Antibiotic use within 1 year of recruitment	41/155 (26)	22/87 (25)	16/60 (27)	0.851
At recruitment	24/41 (59)	10/22 (45)	11/16 (69)	0.154
Lymph node characteristics: sites				
Neck	138/158 (87)	78/89 (88)	55/61 (90)	0.632
Deep anterior cervical	65/138 (47)	36/78 (46)	24/55 (47)	0.774
Deep lateral cervical	25/138 (18)	15/78 (19)	10/55 (18)	0.879
Superficial	15/138 (11)	6/78 (8)	9/55 (16)	0.120
Supraclavicular	19/138 (14)	16/78 (21)	3/55 (5)	**0.015**
Head	13/138 (10)	4/78 (5)	9/55 (16)	**0.032**
Thorax	20/158 (13)	11/89 (12)	6/61 (10)	0.632
Axillary (vs. breast)	16/21 (81)	9/11 (82)	3/5 (60)	0.350
Lymph node characteristics: size, cm^2	4 (2-9)	4 (2-9)	4 (4-9)	0.150
Specimen appearance				
Bloody (vs. chylous)	130/158 (82)	66/89 (72)	57/61 (93)	**0.003**

Bolded items indicate that *p* values are significant at $p<0.05$
*8 Probable TBLs (pTBLs) excluded from table
Abbreviations: dTBLs: definite tuberculous lymphadenitis; nTBLs: non-tuberculous lymphadenitis; pTBLs: probable-tuberculous; ART: Antiretroviral therapy.
¶Missing data: HIV (n=2); Specimen appearance (n=3)

Lymph microbiome is distinct from background sampling controls.

Lymph fluid had similar α-diversity to background controls (skin, saline) but different β-diversity resulting from an enrichment of *Mycobacterium* **(Figure S1A-D)**, suggesting environmental contamination unlikely.

Mycobacterium enrichment in dTBLs drives differences with nTBLs

α-Diversity was similar in dTBLs and nTBLs (**Figure 2A**) and, in β-diversity analyses, *Mycobacterium* was the most discriminatory taxon (**Figure 2B-C**) appearing at several fold higher frequencies than in nTBLs (**Figure 2D**). Bray distances within nTBLs were greater than within dTBLs (**Figure 2E**), thus dTBLs were more like each other than nTBLs to each other (likely reflecting the mixture of different disease pathologies in the nTBLs and relative homogeneity of dTBLs). *Mycobacterium* reads were present in 64% (57/89) of dTBLs and 11% (7/61; *p<0.0001*) of nTBLs (**Figure S2**) and, when sequences underwent BLAST, all reads matched with *Mtb.*

Correlation between 16S rRNA gene sequencing and TB diagnostic test results

As expected, there was a higher relative abundance of *Mycobacterium* reads in dTBLs [median 0.034 (IQR 0.001-0.460) vs. 0.001 (0.001-0.001), *p<0.0001;* **Figure 2D**]. *Mycobacterium* relative abundance in dTBLs showed a positive correlation with bacillary load (based on Xpert and Ultra cycle threshold values; r_s=-0.774, 95% CI [-0.777, -0.514], *p<0.0001*; **Figure 2F**), and culture days-to-positivity; r_p=-0.642, [-0.833, -0.31.], *p=0.001*; **Figure 2G**). Furthermore, a non-significant trend towards a positive correlation between lymph node size and mycobacterial load (relative abundance, Xpert/Ultra C_{Tmin} values) was observed (**Figure S3A-B**).

Differences by HIV status

Overall: α-diversity did not differ by HIV status (**Figure 3A**) and although β-diversity did (**Figure 3B**) no differentially enriched taxa were found (not shown), however, the relative abundance of *Mycobacterium* was higher in PLHIV (**Figure S4A**).

Comparisons within dTBLs or nTBLs by HIV status: There were 55% (49/89) and 39% (23/59) HIV-positive dTBLs and nTBLs, respectively. Within dTBLs or nTBLs, α-diversities did not differ by HIV status (**Figure 3A**) and β-diversity differed within dTBLs by HIV status (*p=0.017,* **Figure 3C**) but not within nTBLs. HIV-positive dTBLs had higher *Mycobacterium* relative abundance than HIV-negative dTBLs **(Figure S4A)**.

Comparisons within HIV-positives or -negatives by TB status: In people with the same HIV status, α-diversity did not differ by TB status (**Figure 3A**) and β-diversity only differed between dTBLs vs. nTBLs in HIV-positives (*p=0.009*, **Figure 3E**) where dTBLs were enriched in *Mycobacterium* (**Figure S4B**). In HIV-negatives, there were no differences between dTBLs and nTBLs (**Figure 3F**).

Lymphotype identification and their associations with clinical characteristics

Overall: Five lymphotypes with differing α- and β-diversities were identified (**Figure 4A-C, Table S3**). L1 had no dominant taxa (**Figure 4D**), whilst L4 was *Mycobacterium*-dominated and had the least α-diversity, and L2, L3 and L5 were *Corynebacterium*-, *Prevotella*- and *Streptococcus*-dominated, respectively. While no taxa were differentially abundant in L1 vs. other lymphotypes (**Figure S5A-C**), L2, L3, and L5 were enriched relative to L4 in *Corynebacterium*, *Prevotella*, and *Streptococcus,* respectively (**Figure 4E-G**). The proportions of dTBLs in L1, L2, L3, L4, and L5 were 35% (17/48), 63% (28/44), 57% (12/21), 100% (21/21), and 69% (11/16), respectively. The patients in these lymphotypes are associated with distinct clinical characteristics. The majority of nTBLs occurred in highly diverse lymphotypes with a heterogenous mixtures of taxa; likely reflecting the spectrum of pathologies in people with TBL ruled out. L1 was associated with characteristics indicative of less severe lymphadenitis. Compared separately to L2, L4, and L5, L1s were less likely to have dTBL. Furthermore, L1s were less likely to be HIV-positive vs. L4s but, L1 PLHIVs had

lower CD4 counts vs. L2 and L3 PLHIVs. In contrast, L4 was associated with characteristics resembling more severe lymphadenitis. L4 was more likely to contain dTBL patients than each other lymphotype. Furthermore, compared to L2s, L4s were more likely to have a bigger lymph node, chylous FNABs and, of PLHIV, a smaller proportion on ART. Compared to L3s, L4s were more likely to have previous TB and HIV, and those with HIV were more likely to have lower CD4 counts. Compared to L5s, L3s with HIV had lower CD4 counts. Therefore, in summary, L1 appears to be associated with less severe forms of lymphadenitis, whereas L4 was associated more severe forms (**Table S4**).

Within patients of the same TB status: Within dTBLs, three lymphotypes with differing β-diversities were identified (**Figure 5A-B**). L1 was abundant in *Prevotella* and *Corynebacterium*, L2 in *Prevotella* and *Streptococcus*, and L3 in *Mycobacterium* (**Figure 5C**) and these taxa were differentially abundant (**Figure 5D-F**). These lymphotypes were termed *Prevotella-Corynebacterium*, *Prevotella-Streptococcus* and *Mycobacterium,* respectively. L3s were more likely to be HIV-positive, with larger lymph nodes, compared to L1s. In addition, L3s were more likely to have larger lymph nodes than L2s. Lastly, L2s are more likely to be female than L1s (**Table S5**). Together, these differences suggest L3 is associated with more severe TBL than other lymphotypes. Within nTBLs, no lymphotypes were identified (**Figure S6**).

Predictive metagenome profiling shows increased short chain fatty acid metabolism

dTBLs vs. nTBLs: 139 inferred microbial metabolic pathways were differentially enriched (75 in dTBLs, 64 in nTBLs). In dTBLs, “fatty acid metabolism”, “benzoate degradation”, “propanoate metabolism” and “butanoate metabolism” were enriched, suggesting increased SCFA production (**Figure 6**).

HIV-positive vs. negatives: The above SCFA-related pathways were enriched in HIV-positive vs. -negative patients overall and, within dTBLs, in HIV-positives vs. -negatives

(**Figure 7A-B**). Within nTBLs, HIV-positives were enriched in the "cell cycle – *Caulobacter*", "bacterial secretion system" and "oxidative phosphorylation" vs. -negatives (**Figure S7**).

In different lymphotypes: When comparing lymphotypes' inferred pathways in all patients (overall including dTBLs and nTBLs), a similar core of pathways was enriched in lymphotype 4. These included the "propanoate metabolism", "tuberculosis", "lipid biosynthesis", "butanoate metabolism", "fatty acid metabolism" and "PPAR signalling pathway" (most-to-least enriched) (**Figure 8A-B)**. In contrast, vs. lymphotype 4, lymphotype 1 was enriched in "epithelial cell signalling in *Helicobacter pylori* infection", lymphotype 2 was enriched in "carbohydrate digestion and absorption", lymphotype 3 was enriched in "dioxin degradation", and lymphotype 5 was enriched in "carbohydrate digestion and absorption" (**Figure S8A-H)**. When comparing the three dTBL lymphotypes, *Mycobacterium*-dominated lymphotype 3 was, compared to each other dTBL lymphotypes, enriched in the similar core pathways seen for the *Mycobacterium*-dominated lymphotype 4 overall in all patients (**Figure 8C; Figure S9**).

DISCUSSION

We characterised the local microbial environment in patients with lymphadenitis undergoing investigation for TB in a HIV-endemic setting. Our key findings are: 1) lymphatic microbial communities in dTBLs clustered into three distinct "lymphotypes" we termed "*Prevotella-Corynebacterium", "Prevotella-Streptococcus"*, and "*Mycobacterium"*, 2) the *Mycobacterium* dTBL lymphotype was associated with HIV-positivity and other clinical features characteristic of severe lymphadenitis, and 3) dTBLs relative to nTBLs were functionally enriched in fatty acid-, amino acid-, and SCFA-related microbial metabolic pathways with known immunomodulatory effects (the *Mycobacterium* lymphotype was most enriched in these pathways than other dTBL lymphotypes). Finally, 4) dTBLs without *Mycobacterium* reads and nTBLs with *Mycobacterium* reads were identified. These data show TBL at the site-of-disease is not microbially homogenous and that distinct clusters of microbial communities exist associated with different clinical characteristics. The long-term significance and importance of these lymphotypes requires prospective evaluation.

We identified three lymphotypes within dTBLs termed "*Prevotella-Corynebacterium", "Prevotella-Streptococcus"*, and "*Mycobacterium"*, distinguished by different relative abundances of these taxa (*Prevotella* co-occurred in the first two lymphotypes). These individual taxa are enriched in respiratory secretions from pulmonary TB cases [29] [30]. Furthermore, within dTBLs , *Streptococcus* is associated with low BMI and extent of lung damage [30]. *Prevotella* in bronchoalveolar lavage fluid also positively correlates with SCFA concentrations and independently predicts incident TB in people without co-prevalent TB [31]. Compared to the other dTBL lymphotypes, "*Mycobacterium"* was associated with severe disease and most frequently occurred in PLHIV, agreeing with diagnostics studies that show stronger baseline mycobacterial PCR test readouts predict long term clinical outcomes in pulmonary [32] and extrapulmonary TB [33]. Together, these data show distinct lymphotypes are associated with different clinical characteristics and suggests that patients with the most severe *Mycobacterium*-dominated lymphotype may initially progress through different site-of-

312 disease microbial states characterised by *Corynebacterium*-, *Streptococcus*- and/or
313 *Prevotella*-domination. Studies with longitudinal follow-up and repeat sampling are required
314 to examine whether these lymphotypes have potential for clinical staging.
315
316 Importantly, *Corynebacterium* and *Streptococcus* often dominated in dTBL patients.
317 Members of both taxa are causative agents of lymphadenitis and, even though these
318 patients have TB lymphadenitis confirmed via conventional diagnostics, *Corynebacterium*
319 and *Streptococcus* may therefore co-contribute to pathology and symptoms [34-36].
320 Coincidently, these taxa fall within the anti-microbial spectrum of first-line TB treatment [16],
321 meaning that this regimen may, in part, cure lymphadenitis by killing *Corynebacterium* and
322 *Streptococcus* in addition to *Mycobacterium.*
323
324 Microbial pathways predicted to be most enriched in dTBLs involved fatty acid, amino acid,
325 and SCFAs (benzoate, propanoate) metabolism; all of which are associated with pulmonary
326 TB disease compared to sick patients without TB [37 38]. SCFAs in particular suppress immune
327 pathways involved in IFN-γ and IL-17A production and, *ex vivo*, limit macrophage-mediated
328 kill of *Mtb*. SCFA concentrations hence predict incident TB in patients [31]. Our research
329 therefore suggests that the inflammation associated with lymphadenopathy is in part caused
330 by the presence of microbes including but limited to *Mycobacterium* that are able to produce
331 SCFAs that interfere with these immunological pathways; revealing potentially new
332 therapeutic targets to reduce lymphadenopathy.
333
334 We detected *Mycobacterium* in nTBLs. These reads could be from previous TB exposure or
335 disease. *Mtb* DNA has been found in the lymph nodes of healthy individuals and primates
336 exposed to TB, where the sites are hypothesised to serve as a *Mtb* growth and persistence
337 niche [7]. dTBLs without *Mycobacterium* reads were also documented, however, 16S rRNA
338 sequencing has well known sub-optimal sensitivity for *Mycobacterium*, in part due to low 16s
339 RNA gene copy number [39].

Our study has strengths and limitations. Patients were sampled once, as close as possible to treatment initiation; animal models might permit repeat invasive sampling especially if treatment is withheld. We did not co-analyse host immune signatures but plan to do so. The programmatic context enabled large numbers of patients to be recruited, however, detailed long-term follow-up, which could include imaging of lymph nodes and more detailed measurements of differential responses to treatment, was not possible.

In conclusion, we show dTBL patients have a distinct microbiome at the site of disease, characterized by three lymphotypes (*Mycobacterium, Prevotella-Corynebacterium, Prevotella-Streptococcus)*. This dysbiosis of the lymphatic microbiome likely contributes to pathophysiology, including inflammatory state and clinical severity, which itself may reflect the chronicity of TB disease. TB lymphadenitis does therefore not appear to be a microbially homogenesis disease, and this reveals potentially new diagnosis, therapeutic, and prognostic targets.

ACKNOWLEDGMENTS:

The authors thank study participants and Tygerberg Hospital FNA clinic staff especially Sr Cupido. Additionally, we thank CLIME research group staff, especially Sr Ruth Wilson and Roxanne Higgit. GRN acknowledges funding from L'Oréal-UNESCO For Women in Science Sub-Saharan Africa Young Talents Award, and the International Rising Talents Award. The content is the solely the responsibility of the authors and does not necessarily represent the official views of the funders. Computations were performed using facilities provided by the University of Cape Town's ICTS High Performance Computing team: *hpc.uct.ac.za*

CONTRIBUTORSHIP: GN, CCN and GT contributed to conceptualisation and design of the study and supervised the study, funding acquisition, data collection, and wrote the manuscript. GRN, CCN, IS, BGW, IS, JCC, LNS, and GT contributed to data and statistical analysis, and figures. All authors contributed to interpretation of data and editing of the manuscript. As the study guarantor, GT is responsible for the overall content of this manuscript.

FUNDING: This work and authors were supported by the European & Developing Countries Clinical Trials Partnership (EDCTP; project numbers SF1041, TMA2017CDF-1914-MOSAIC and TMA2019CDF-2738-ESKAPE-TB), National Research Foundation (NRF), the South African Medical Research Council (SAMRC), the Harry Crossley Foundation and Stellenbosch University Faculty of Health Sciences, and the National Institute of Allergy and Infectious Diseases of the National Institutes of Health under award numbers (R01AI136894; U01AI152087; U54EB027049; D43TW010350).

COMPETING INTRESTS: None

DATA AVAILABILITY: Data are available upon reasonable request De-identified patient data, the study protocol, informed consent, and datasets generated in this study may be requested from the corresponding author.

REFERENCES

1. Organization WH. Global tuberculosis report 2020: executive summary. 2020
2. Gupta RK, Lawn SD, Bekker L-G, et al. Impact of HIV and CD4 count on tuberculosis diagnosis: analysis of citywide data from Cape Town, South Africa. *The international journal of tuberculosis and lung disease: the official journal of the International Union against Tuberculosis and Lung Disease* 2013;17(8):1014.
3. Theron G, Peter J, Calligaro G, et al. Determinants of PCR performance (Xpert MTB/RIF), including bacterial load and inhibition, for TB diagnosis using specimens from different body compartments. *Scientific reports* 2014;4:5658. doi: 10.1038/srep05658
4. Sharma S, Mohan AJIJoMR. Extrapulmonary tuberculosis. 2004;120:316-53.
5. Zhang Z, Liu Y, Wang W, et al. Identification of differentially expressed genes associated with lymph node tuberculosis by the bioinformatic analysis based on a microarray. *Journal of Computational Biology* 2020;27(1):121-30.
6. Maji A, Misra R, Mondal AK, et al. Expression profiling of lymph nodes in tuberculosis patients reveal inflammatory milieu at site of infection. *Scientific reports* 2015;5(1):1-10.
7. Ganchua SKC, White AG, Klein EC, et al. Lymph nodes—The neglected battlefield in tuberculosis. *PLoS pathogens* 2020;16(8):e1008632.
8. Ganchua SKC, Cadena AM, Maiello P, et al. Lymph nodes are sites of prolonged bacterial persistence during Mycobacterium tuberculosis infection in macaques. *PLoS pathogens* 2018;14(11):e1007337.
9. Lin PL, Coleman T, Carney JP, et al. Radiologic Responses in Cynomolgus Macaques for Assessing Tuberculosis Chemotherapy Regimens. *Antimicrob Agents Chemother* 2013;57(9):4237-44. doi: 10.1128/aac.00277-13 [published Online First: 2013/06/26]
10. Martin CJ, Cadena AM, Leung VW, et al. Digitally Barcoding Mycobacterium tuberculosis Reveals In Vivo Infection Dynamics in the Macaque Model of Tuberculosis. *mBio* 2017;8(3) doi: 10.1128/mBio.00312-17 [published Online First: 2017/05/11]
11. Behr MA, Waters WR. Is tuberculosis a lymphatic disease with a pulmonary portal? *The Lancet infectious diseases* 2014;14(3):250-55.
12. Esmail H, Lai RP, Lesosky M, et al. Characterization of progressive HIV-associated tuberculosis using 2-deoxy-2-[18 F] fluoro-D-glucose positron emission and computed tomography. *Nature medicine* 2016;22(10):1090-93.
13. Young VB. The role of the microbiome in human health and disease: an introduction for clinicians. *Bmj* 2017;356:j831. doi: 10.1136/bmj.j831
14. Kiernan MG, Coffey JC, McDermott K, et al. The human mesenteric lymph node microbiome differentiates between Crohn's disease and ulcerative colitis. *Journal of Crohn's and Colitis* 2018;13(1):58-66.
15. O'Brien CL, Pavli P, Gordon DM, et al. Detection of bacterial DNA in lymph nodes of Crohn's disease patients using high throughput sequencing. *Gut* 2014;63(10):1596-606.
16. Naidoo CC, Nyawo GR, Wu BG, et al. The microbiome and tuberculosis: state of the art, potential applications, and defining the clinical research agenda. 2019
17. Vázquez-Pérez JA, Carrillo CO, Iñiguez-García MA, et al. Alveolar microbiota profile in patients with human pulmonary tuberculosis and interstitial pneumonia. *Microbial pathogenesis* 2020;139:103851.
18. Zhou Y, Lin F, Cui Z, et al. Correlation between either Cupriavidus or Porphyromonas and primary pulmonary tuberculosis found by analysing the microbiota in patients' bronchoalveolar lavage fluid. *PLoS ONE* 2015;10(5):e0124194.
19. Minnies S, Reeve BW, Rockman L, et al. Xpert MTB/RIF Ultra is highly sensitive for the diagnosis of tuberculosis lymphadenitis in an HIV-endemic setting. *medRxiv* 2021
20. Bolyen E, Rideout JR, Dillon MR, et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nature biotechnology* 2019;37(8):852-57.

21. Callahan BJ, McMurdie PJ, Rosen MJ, et al. DADA2: High-resolution sample inference from Illumina amplicon data. *Nature methods* 2016;13(7):581-83.
22. McDonald D, Price MN, Goodrich J, et al. An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *The ISME journal* 2012;6(3):610.
23. McMurdie PJ, Holmes S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PloS one* 2013;8(4):e61217.
24. Dixon P. VEGAN, a package of R functions for community ecology. *Journal of Vegetation Science* 2003;14(6):927-30.
25. Holmes I, Harris K, Quince C. Dirichlet multinomial mixtures: generative models for microbial metagenomics. *PLoS ONE* 2012;7(2):e30126.
26. Douglas GM, Maffei VJ, Zaneveld J, et al. PICRUSt2: An improved and customizable approach for metagenome inference. *BioRxiv* 2020:672295.
27. McMurdie PJ, Holmes S. Waste not, want not: why rarefying microbiome data is inadmissible. *PLoS computational biology* 2014;10(4)
28. Reiner A, Yekutieli D, Benjamini Y. Identifying differentially expressed genes using false discovery rate controlling procedures. *Bioinformatics* 2003;19(3):368-75.
29. Krishna P, Jain A, Bisen PS. Microbiome diversity in the sputum of patients with pulmonary tuberculosis. *Eur J Clin Microbiol Infect Dis* 2016;35(7):1205-10. doi: 10.1007/s10096-016-2654-4
30. Ticlla MR, Hella J, Hiza H, et al. The Sputum Microbiome in Pulmonary Tuberculosis and Its Association With Disease Manifestations: A Cross-Sectional Study. *Front Microbiol* 2021;12
31. Segal LN, Clemente JC, Li Y, et al. Anaerobic Bacterial Fermentation Products Increase Tuberculosis Risk in Antiretroviral-Drug-Treated HIV Patients. *Cell host & microbe* 2017;21(4):530-37 e4. doi: 10.1016/j.chom.2017.03.003
32. Pires MdM, Pereira GR, Barbosa MS, et al. Association of Xpert MTB/RIF cycle threshold values with tuberculosis treatment outcomes. *Lung* 2020;198(6):985-89.
33. Martyn EM, Bangdiwala AS, Kagimu E, et al. Cerebrospinal fluid bacillary load by Xpert MTB/RIF Ultra polymerase chain reaction cycle threshold value predicts 2-week mortality in human immunodeficiency virus–associated tuberculous meningitis. *Clinical Infectious Diseases* 2021;73(9):e3505-e10.
34. Peel MM, Palmer GG, Stacpoole AM, et al. Human lymphadenitis due to Corynebacterium pseudotuberculosis: report of ten cases from Australia and review. *Clinical Infectious Diseases* 1997;24(2):185-91.
35. Barroso LF, Pegram SP. Clinical manifestations, diagnosis, and treatment of diphtheria. In: Post TW, ed. UpToDate. Waltham, MA: UpToDate 2022.
36. Denis Spelman M, Baddour LM. Cellulitis and skin abscess: Epidemiology, microbiology, clinical manifestations, and diagnosis. In: Post TW, ed. UpToDate. Waltham, MA: UpToDate 2021.
37. Naidoo CC, Nyawo GR, Sulaiman I, et al. Anaerobe-enriched gut microbiota predicts pro-inflammatory responses in pulmonary tuberculosis. *EBioMedicine* 2021;67:103374.
38. Somboro AM, Diallo D, Holl JL, et al. The role of the microbiome in inflammation during tuberculosis. *EBioMedicine* 2021;68
39. Sulaiman I, Wu BG, Li Y, et al. Evaluation of the airway microbiome in nontuberculous mycobacteria disease. *European Respiratory Journal* 2018;52(4)

FIGURE LEGENDS

Figure 1: Study flow chart. Fine needle aspirates, skin and saline controls were collected from presumptive TBL patients. dTBLs: definite tuberculous lymphadenitis; nTBLs: non-tuberculous lymphadenitis; pTBLs: probable tuberculous lymphadenitis; Smear: Smear microscopy; MGIT960 Culture: Mycobacteria Growth Indicator Tube 960 liquid culture; Xpert: Xpert MTB/RIF; Ultra: Xpert MTB/RIF Ultra.

Figure 2: dTBLs have a distinct microbiome to nTBLs with *Mycobacterium* enrichment. (A) Although α-diversity was similar, **(B)** β-diversity differed. *Mycobacterium* was enriched in dTBLs compared to nTBLs based on **(C)** differential abundance testing and **(D)** relative abundance. Discriminatory taxa appear above the threshold (red dotted line, FDR=0.2). **(E)** dTBLs were more compositionally similar to each other than nTBLs. Mycobacterial reads positively correlated with *Mtb* load: **(F)** Xpert and Ultra and **(G)** culture (days-to-positivity). dTBLs: definite tuberculous lymphadenitis; nTBLs: non-tuberculous lymphadenitis; Xpert: Xpert MTB/RIF; Ultra: Xpert MTB/RIF Ultra; MGIT960: Mycobacteria Growth Indicator Tube 960 liquid culture; r_s: Spearman correlation coefficient; r_p: Pearson correlation coefficient.

Figure 3: Microbiome differences in HIV-positive dTBLs versus nTBLs but not in HIV-negative dTBLs versus nTBLs. (A) α-Diversity did not differ by HIV or TBL statuses, **(B)** however, β-diversity differed between HIV-positives and -negatives overall (shaded circles are dTBLs, empty circles are nTBLs). β-diversity differed by HIV status in **(C)** dTBLs but not in **(D)** nTBLs. β-diversity differed by TBL status in **(E)** HIV-positives and **(F)** HIV-negatives.

Figure 4: Five lymphotypes observed in presumptive TBL. (A) LaPlace approximation identified five clusters. **(B)** L5 had the highest α-diversity. **(C)** β-diversity differed between each lymphotype (shaded circles dTBLs, empty circles nTBLs). **(D)** Stacked bar plots

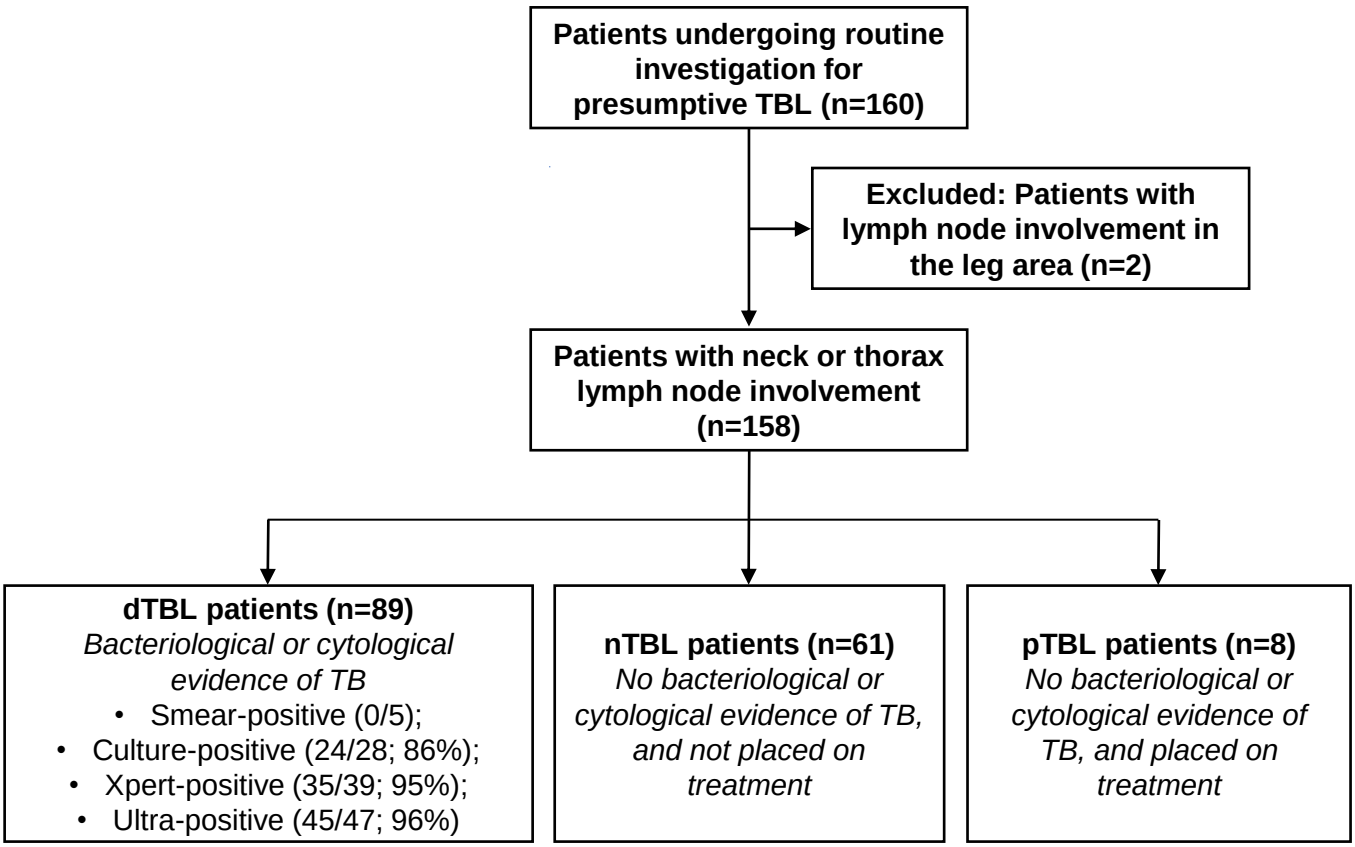
showing L1 with a heterogenous mixture of genera, L2 dominated by *Corynebacterium*, L3 dominated by *Prevotella*, L4 dominated by *Mycobacterium*, and L4 dominated by *Streptococcus*. Bolded taxa represent dominating taxa. **(E)** *Corynebacterium* was enriched in L2; **(F)** *Prevotella* enriched in L3, **(G)** *Mycobacterium* enriched in L4, and *Streptococcus* enriched in L5. Significantly more discriminatory taxa (bolded) appear closer to the left or right, and higher above the threshold (red dotted line, FDR=0.2) as significance increases. Relative taxa abundance is indicated by circle size. dTBLs: definite tuberculous lymphadenitis; nTBLs: non-tuberculous lymphadenitis; L: lymphotype.

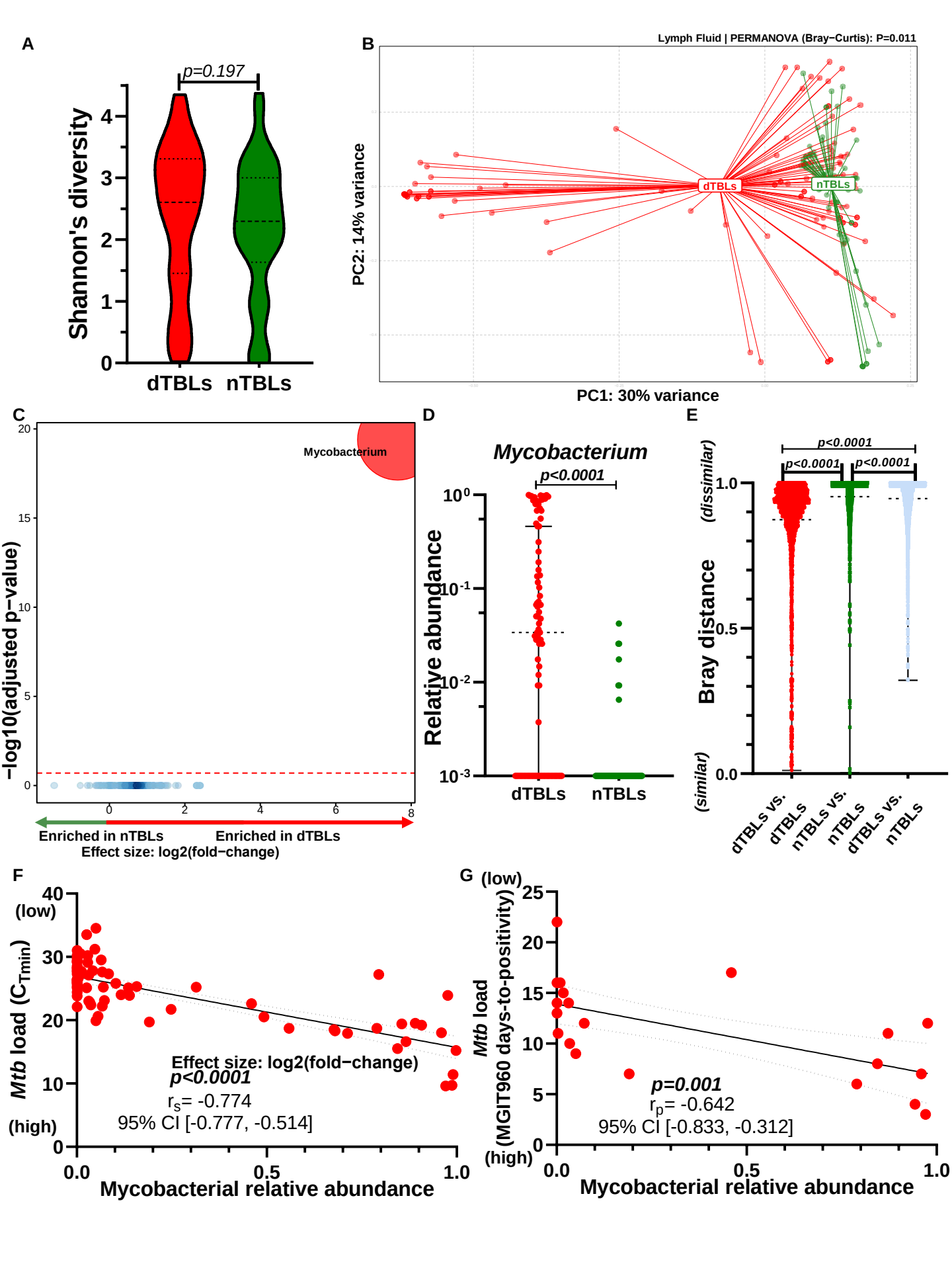
Figure 5: Three lymphotypes identified in dTBLs. (A) Best model fit based on LaPlace approximation identified three clusters within dTBLs. **(B)** β-diversity differed between lymphotypes. **(D)** Stacked bar plots showing L1 comprised of *Mycobacterium* and accompanying heterogenous taxa, L2 dominated by *Prevotella* and *Streptococus*, and L3 dominated by *Mycobacterium*. Bolded taxa represent dominating taxa. **(D)** No taxa were enriched in L1, **(E)** L2 was enriched in *Streptococcus,* **(F)** and *Mycobaterium* was enriched in L3. Significantly more discriminatory taxa (bolded) appear closer to the left or right, and higher above the threshold (red dotted line, FDR=0·2) as significance increases. Relative taxa abundance is indicated by circle size. dTBLs: definite tuberculous lymphadenitis; L: lymphotype.

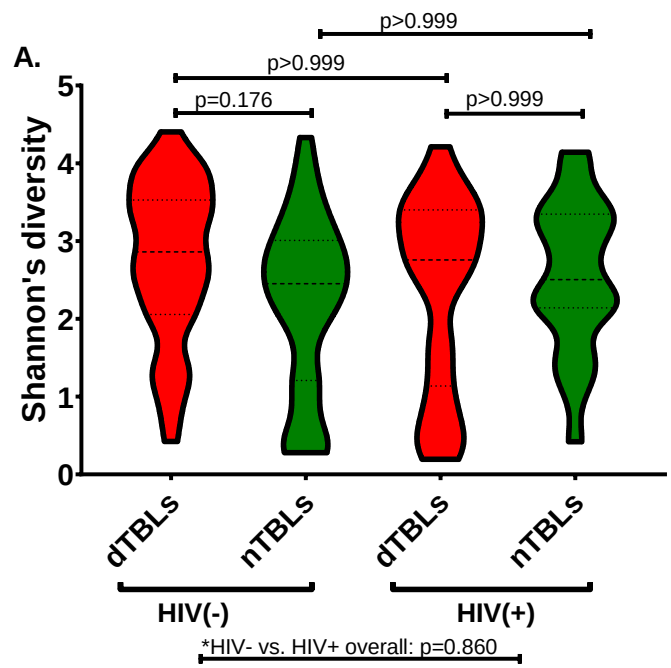
Figure 6: Enriched microbial capacity for SCFA pathways in dTBLs versus nTBLs. Volcano plot depicting differentially abundant microbial pathways in dTBLs vs. nTBLs inferred by PICRUSt2. Key pathways of interest are bolded including aminobenzoate degradation, benzoate degradation, and propanoate degradation. Significantly more discriminatory pathways appear closer to the left or right, and higher above the threshold (red dotted line, FDR=0.05) as significance increases. Relative gene abundance is indicated by circle size. SCFA: short chain fatty acids; dTBLs: definite tuberculous lymphadenitis; nTBLs: non-tuberculous lymphadenitis.

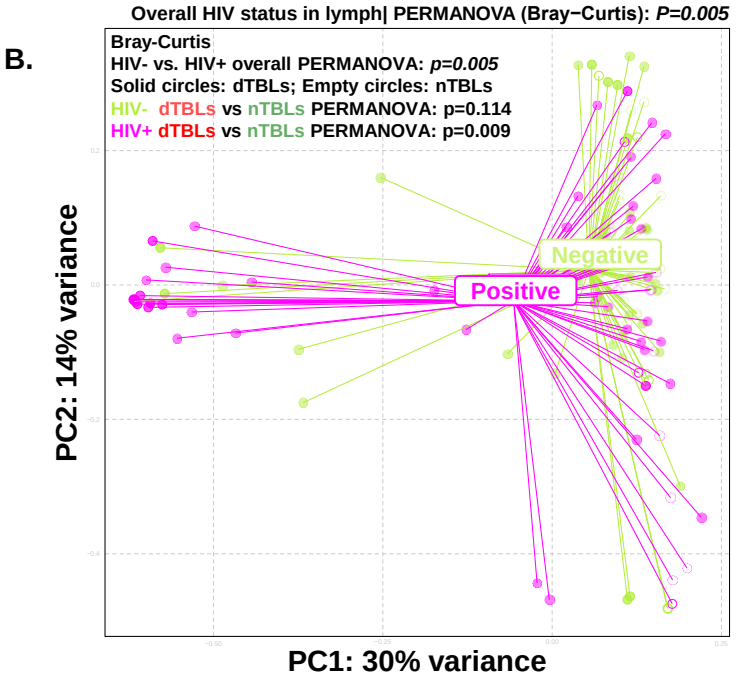
Figure 7: Predicted metagenome function reveals increased capacity for SCFA production in HIV-positive versus HIV-negative patients overall, and in dTBLs. Volcano plot depicting functional pathways differing between **(A)** HIV-positive and HIV-negative patients with presumptive TBL, and **(B)** in dTBLs. Key pathways of interest include butanoate metabolism, propanoate metabolism and benzoate degradation. Significantly more discriminatory pathways appear closer to the left or right, and higher above the threshold (red dotted line, FDR=0.05). Relative gene abundance is indicated by circle size. SCFA: short chain fatty acids.

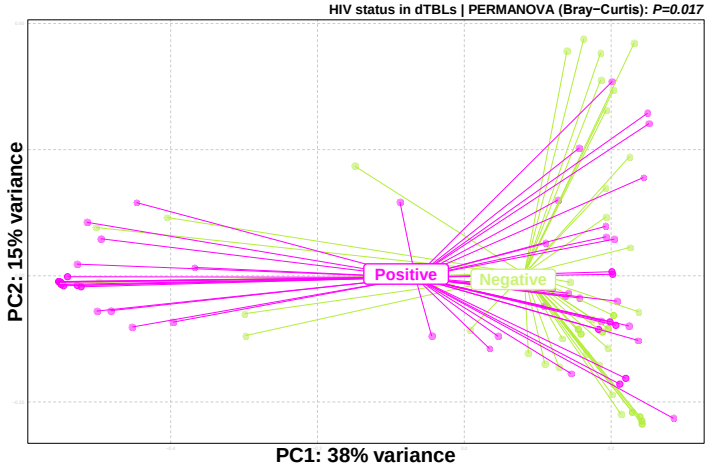
Figure 8: Differential microbial pathways between lymphotypes showing similar core pathways enriched in the *Mycobacterium*-dominated lymphotype.. (A) Volcano plot showing differentially abundant microbial pathways inferred by PICRUSt2 in lymphotype 2 vs. lymphotype 4 representing pathways enriched in lymphotype 4 compared to every other lymphotype in all patients (overall including dTBLs and nTBLs). Significantly more discriminatory pathways appear closer to the left or right, and higher above the threshold (red dotted line, FDR=0.05) as significance increases. Relative gene abundance is indicated by circle size. **(B)** 65.5% of all inferred pathways enriched L4 compared to each other overall lymphotypes were common, whilst **(C)** 85.8% were common in L3 compared to each other dTBL lymphotypes.Differentially enriched pathways common in all comparisons with the *Mycobacterium* dominant lymphotype included pathways involving lipid biosynthesis, fatty acids, and SCFA metabolism i.e. lipid biosynthesis proteins, propanoate metabolism, benzoate degradation, and valine, leucine and isoleucine degradation. SCFA: short chain fatty acids; dTBLs: definite tuberculous lymphadenitis; nTBLs: non-tuberculous lymphadenitis; L: lymphotype.

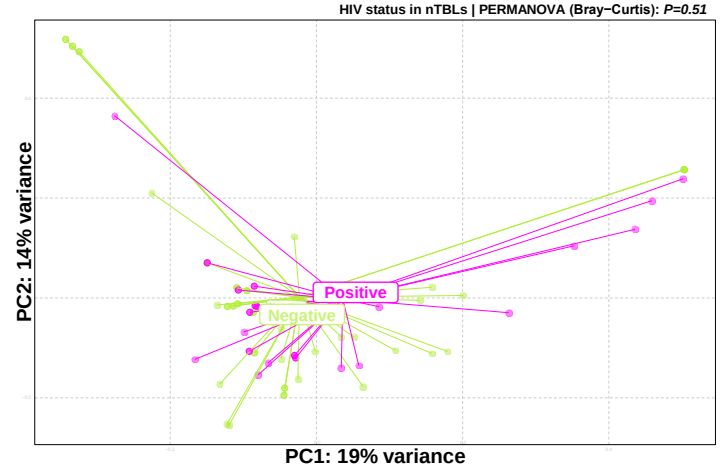

Patients undergoing routine investigation for presumptive TBL (n=160)
Excluded: Patients with lymph node involvement in the leg area (n=2)
Patients with neck or thorax lymph node involvement (n=158)
dTBL patients (n=89)
Bacteriological or cytological evidence of TB
• Smear-positive (0/5);
• Culture-positive (24/28; 86%);
• Xpert-positive (35/39; 95%);
• Ultra-positive (45/47; 96%)
nTBL patients (n=61)
No bacteriological or cytological evidence of TB, and not placed on treatment
pTBL patients (n=8)
No bacteriological or cytological evidence of TB, and placed on treatment

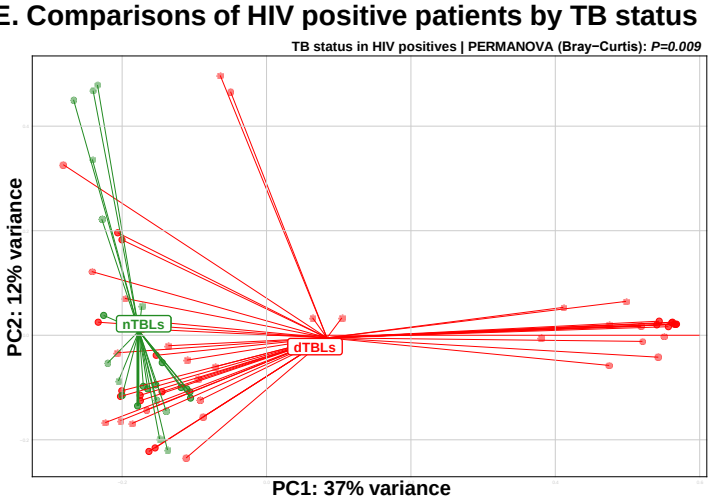

A
p=0.197
Shannon's diversity
dTBLs
nTBLs
B
Lymph Fluid | PERMANOVA (Bray−Curtis): P=0.011
PC2: 14% variance
PC1: 30% variance
dTBLs
nTBLs
C
Mycobacterium
−log10(adjusted p−value)
Enriched in nTBLs
Enriched in dTBLs
Effect size: log2(fold−change)
D
Mycobacterium
p<0.0001
Relative abundance
dTBLs
nTBLs
E
p<0.0001
p<0.0001
p<0.0001
(dissimilar)
Bray distance
(similar)
dTBLs vs. dTBLs
nTBLs vs. nTBLs
dTBLs vs. nTBLs
F
(low)
(high)
Mtb load (CTmin)
Effect size: log2(fold−change)
p<0.0001
rs= -0.774
95% CI [-0.777, -0.514]
Mycobacterial relative abundance
G
(low)
(high)
Mtb load (MGIT960 days-to-positivity)
p=0.001
rp= -0.642
95% CI [-0.833, -0.312]
Mycobacterial relative abundance

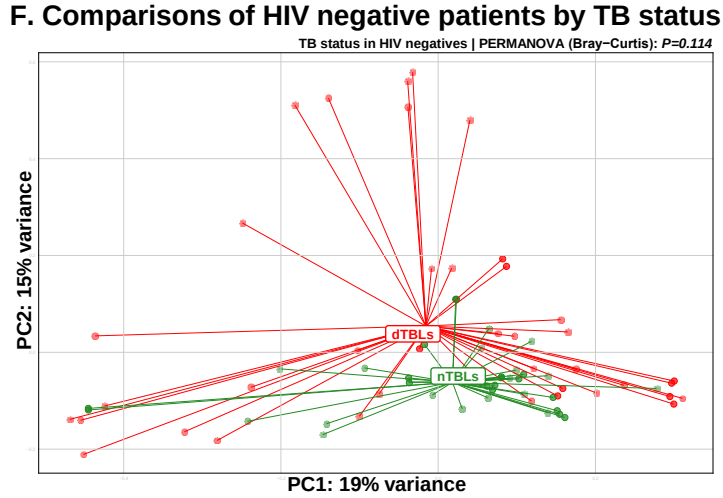

A.
p>0.999
p>0.999
p=0.176
p>0.999
Shannon's diversity
5
4
3
2
1
0
dTBLs
nTBLs
dTBLs
nTBLs
HIV(-)
HIV(+)
*HIV- vs. HIV+ overall: p=0.860

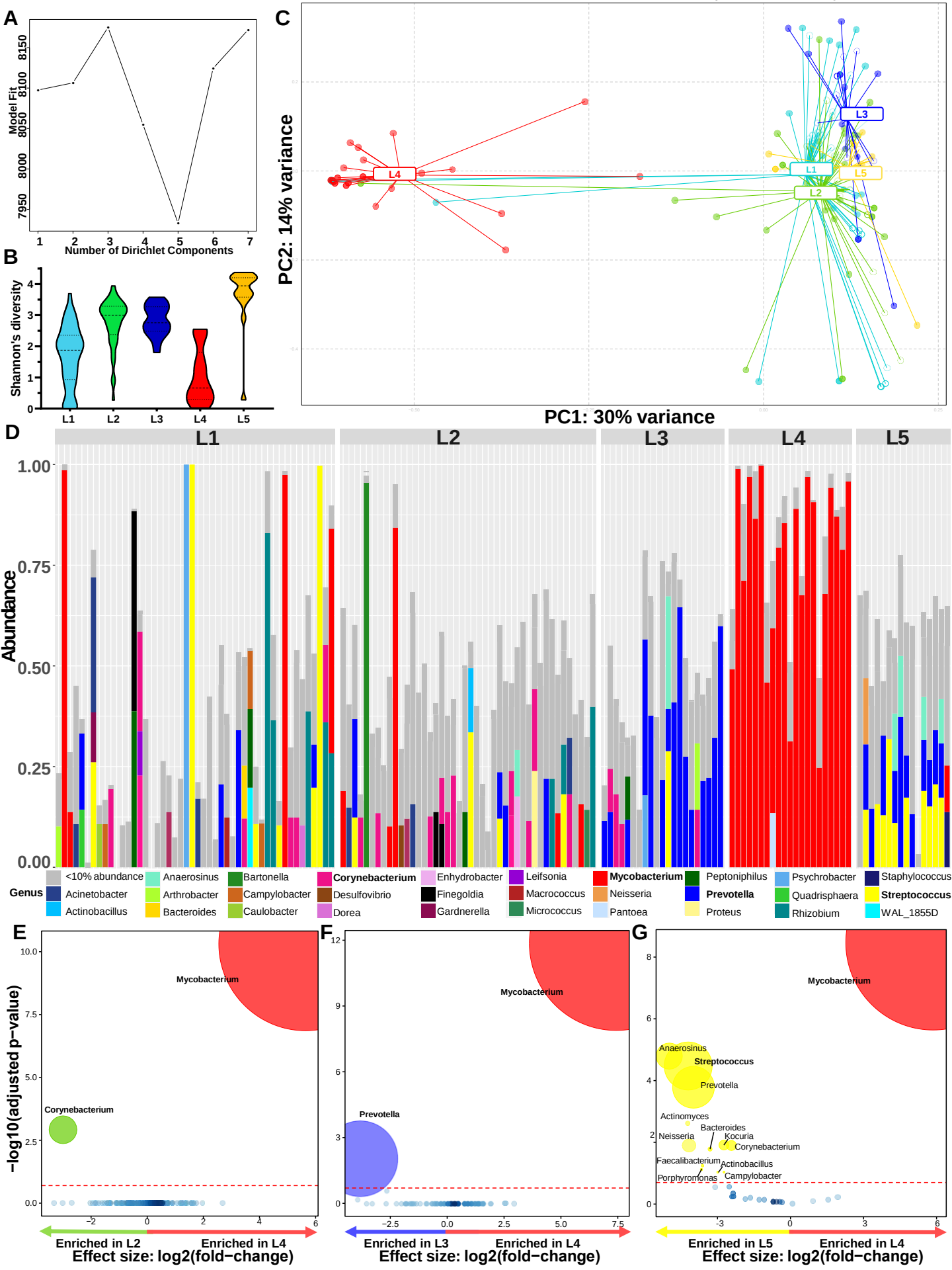

B.
Overall HIV status in lymph| PERMANOVA (Bray−Curtis): P=0.005
Bray-Curtis
HIV- vs. HIV+ overall PERMANOVA: p=0.005
Solid circles: dTBLs; Empty circles: nTBLs
HIV- dTBLs vs nTBLs PERMANOVA: p=0.114
HIV+ dTBLs vs nTBLs PERMANOVA: p=0.009
Negative
Positive
PC2: 14% variance
PC1: 30% variance

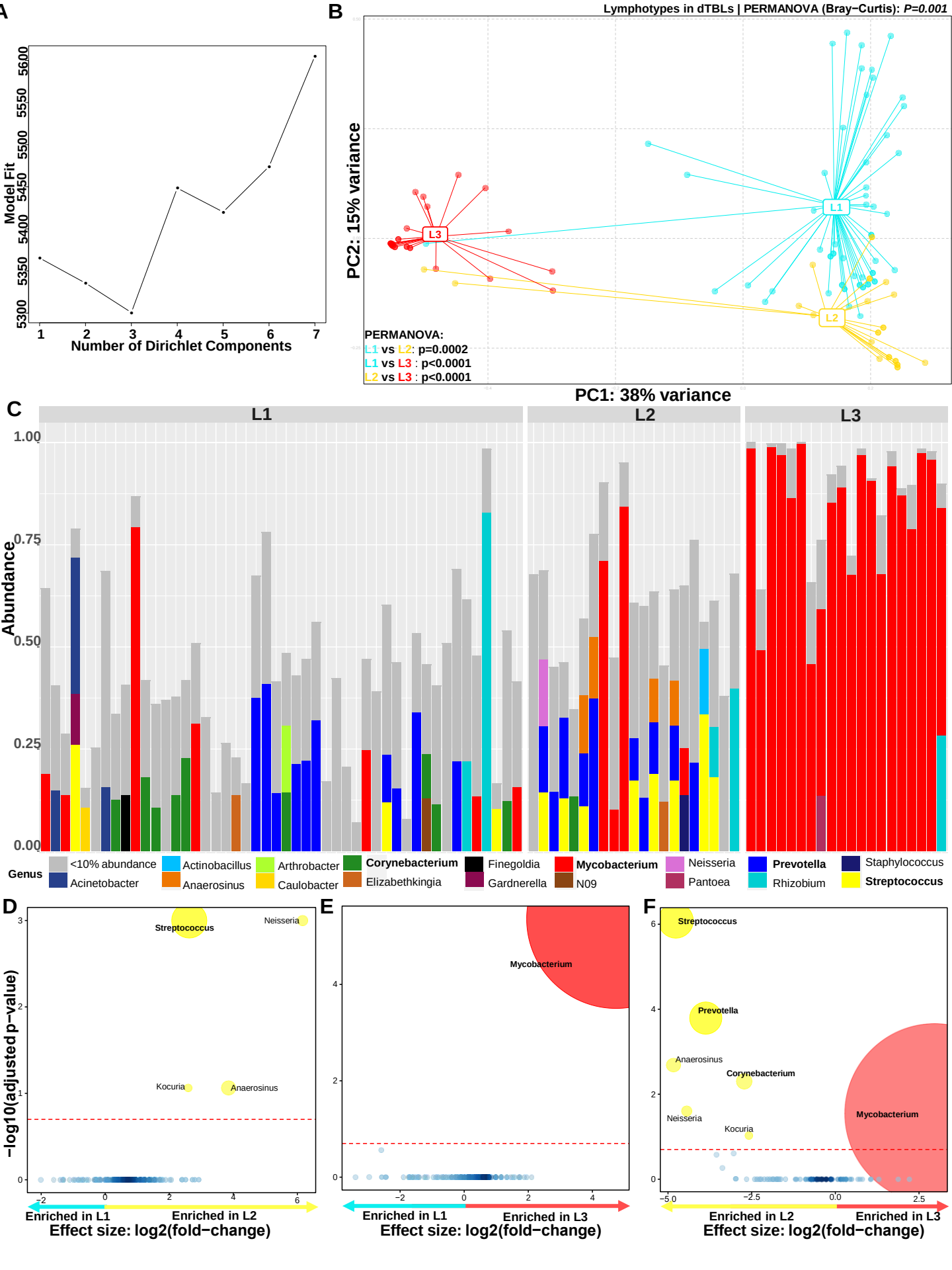

C. Comparisons of dTBLs by HIV status
HIV status in dTBLs | PERMANOVA (Bray−Curtis): P=0.017
Positive
Negative
PC2: 15% variance
PC1: 38% variance

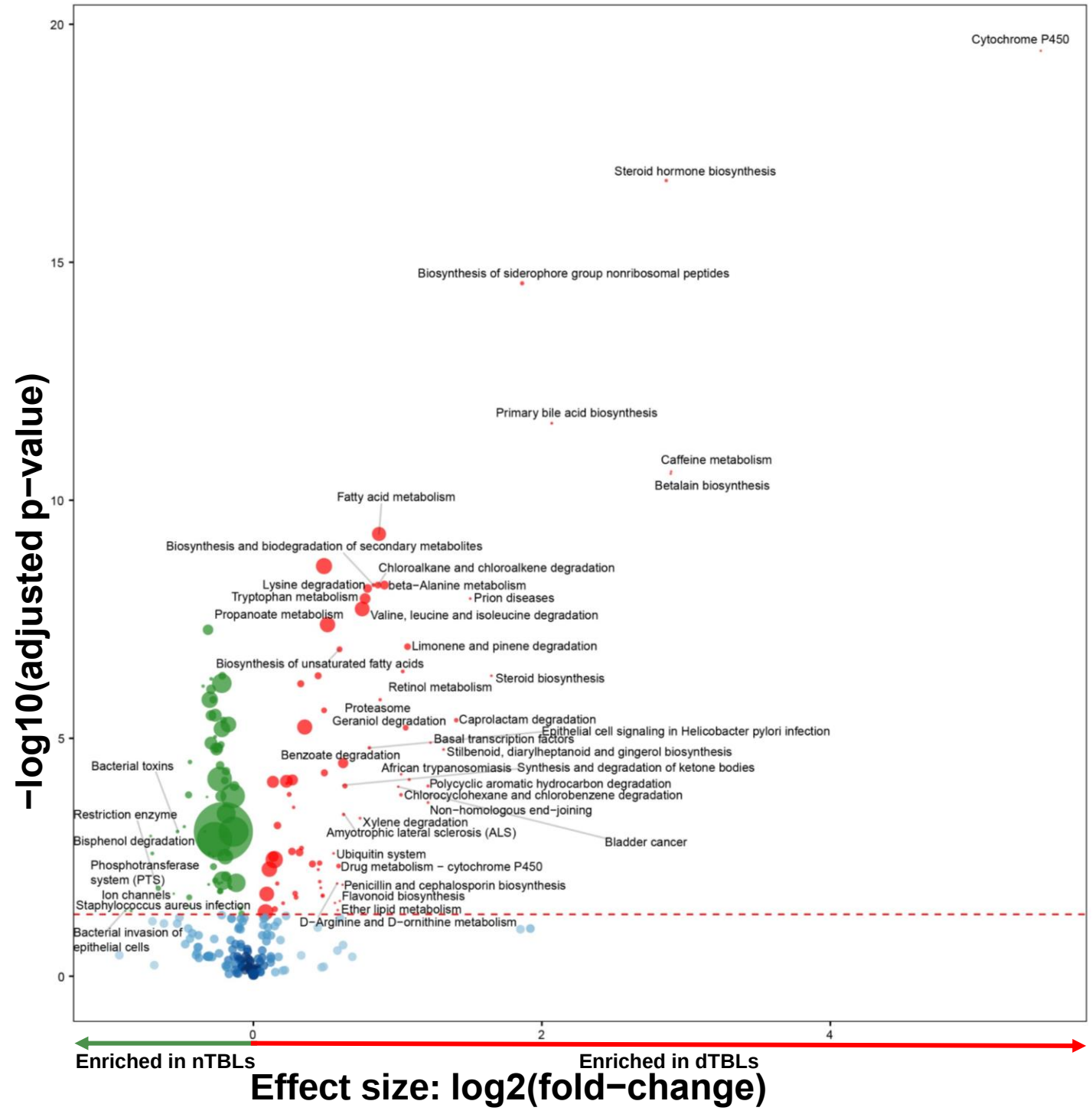

D. Comparisons of nTBLs by HIV status
HIV status in nTBLs | PERMANOVA (Bray−Curtis): P=0.51
Positive
Negative
PC2: 14% variance
PC1: 19% variance

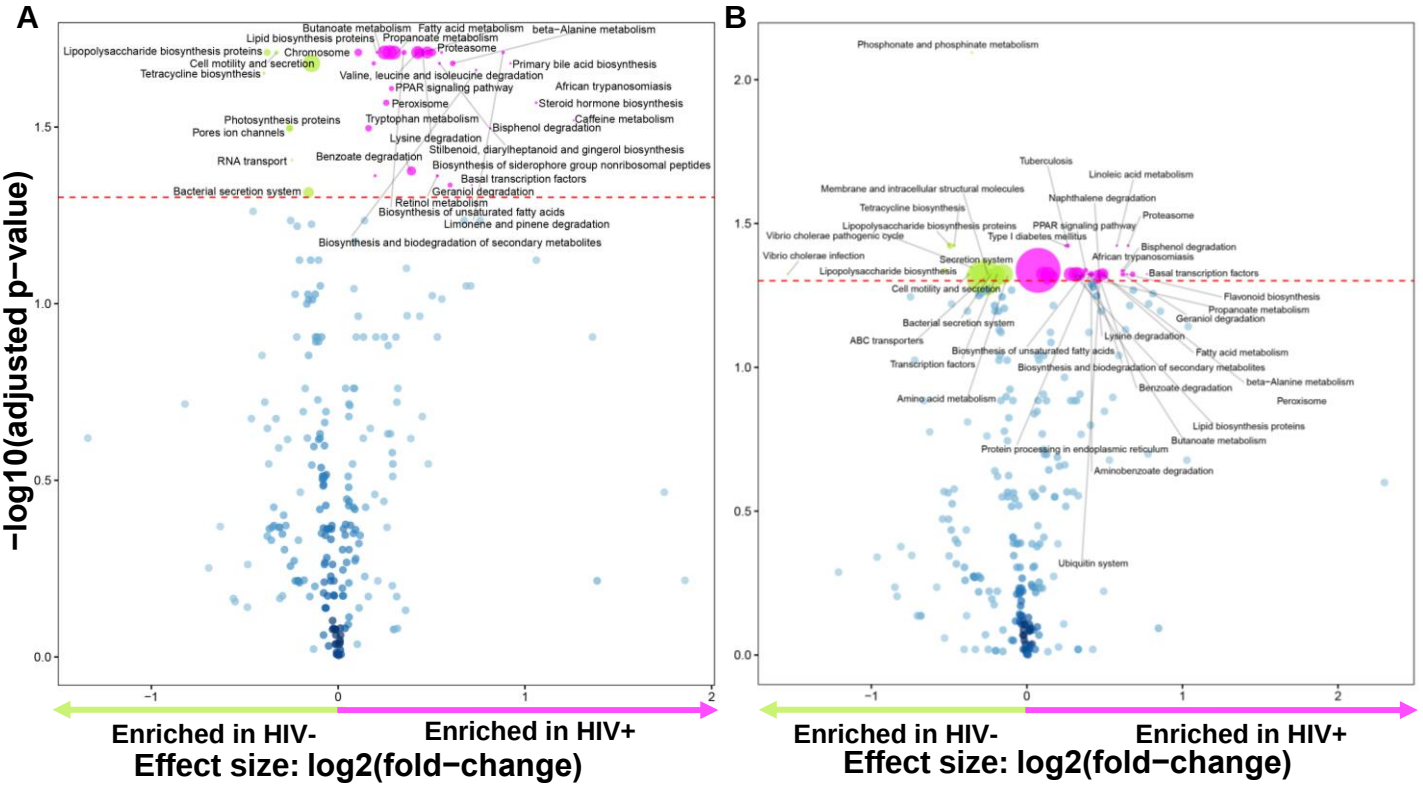

E. Comparisons of HIV positive patients by TB status
TB status in HIV positives | PERMANOVA (Bray−Curtis): P=0.009
nTBLs
dTBLs
PC2: 12% variance
PC1: 37% variance

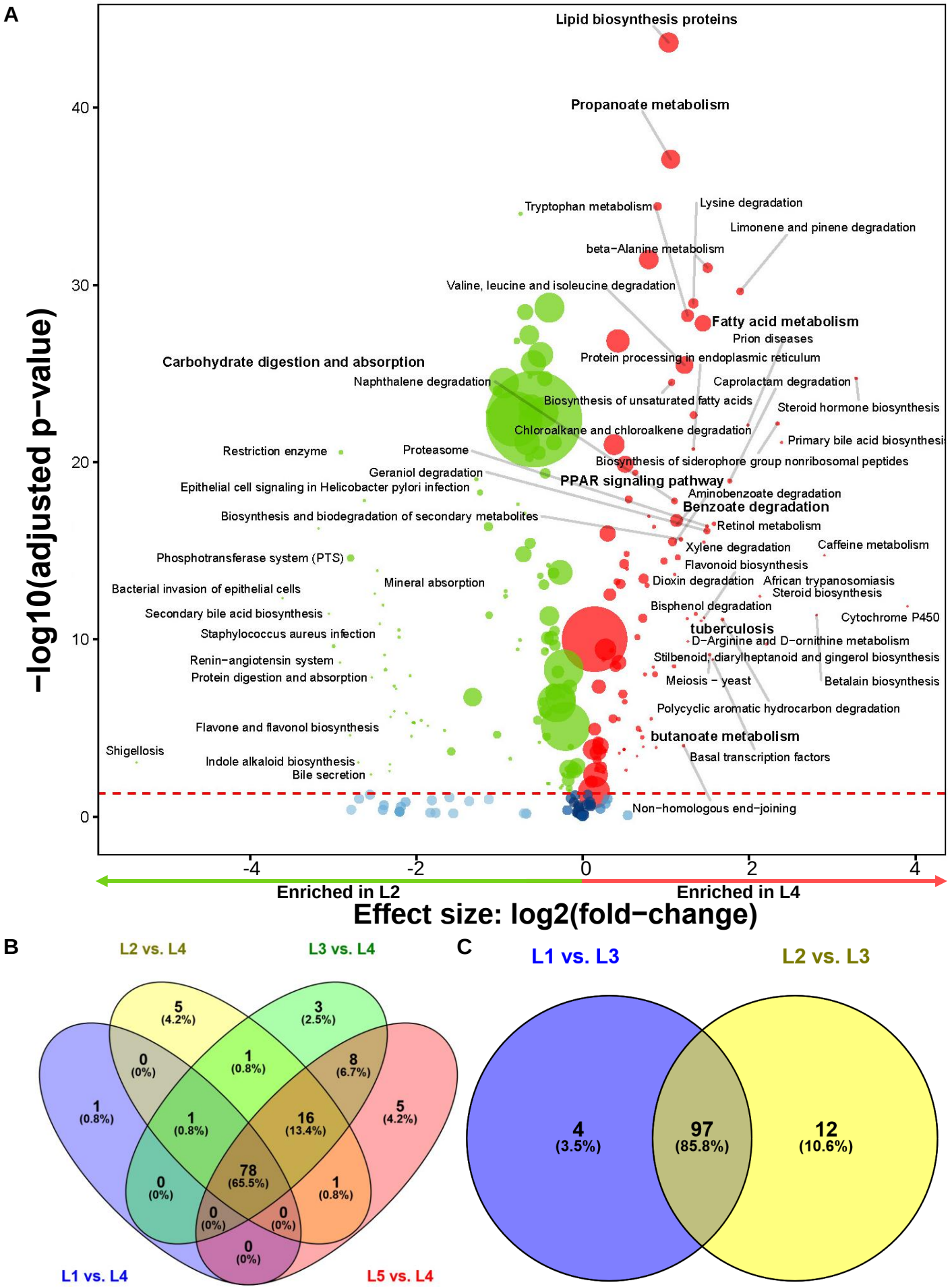

F. Comparisons of HIV negative patients by TB status
TB status in HIV negatives | PERMANOVA (Bray−Curtis): P=0.114
dTBLs
nTBLs
PC2: 15% variance
PC1: 19% variance


A
Model Fit
Number of Dirichlet Components
B
Shannon's diversity
L1
L2
L3
L4
L5
C
PC2: 14% variance
PC1: 30% variance
D
Abundance
Genus
<10% abundance
Acinetobacter
Actinobacillus
Anaerosinus
Arthrobacter
Bacteroides
Bartonella
Campylobacter
Caulobacter
Corynebacterium
Desulfovibrio
Dorea
Enhydrobacter
Finegoldia
Gardnerella
Leifsonia
Macrococcus
Micrococcus
Mycobacterium
Neisseria
Pantoea
Peptoniphilus
Prevotella
Proteus
Psychrobacter
Quadrisphaera
Rhizobium
Staphylococcus
Streptococcus
WAL_1855D
E
F
G
−log10(adjusted p−value)
Mycobacterium
Corynebacterium
Prevotella
Anaerosinus
Streptococcus
Actinomyces
Bacteroides
Neisseria
Kocuria
Faecalibacterium
Actinobacillus
Porphyromonas
Campylobacter
Enriched in L2
Enriched in L4
Enriched in L3
Enriched in L5
Effect size: log2(fold−change)


A
Model Fit
5300
5350
5400
5450
5500
5550
5600
1
2
3
4
5
6
7
Number of Dirichlet Components
B
Lymphotypes in dTBLs | PERMANOVA (Bray−Curtis): P=0.001
PC2: 15% variance
PC1: 38% variance
L1
L2
L3
PERMANOVA:
L1 vs L2: p=0.0002
L1 vs L3 : p<0.0001
L2 vs L3 : p<0.0001
C
L1
L2
L3
Abundance
1.00
0.75
0.50
0.25
0.00
Genus
<10% abundance
Acinetobacter
Actinobacillus
Anaerosinus
Arthrobacter
Caulobacter
Corynebacterium
Elizabethkingia
Finegoldia
Gardnerella
Mycobacterium
N09
Neisseria
Pantoea
Prevotella
Rhizobium
Staphylococcus
Streptococcus
D
−log10(adjusted p−value)
Streptococcus
Neisseria
Kocuria
Anaerosinus
Enriched in L1
Enriched in L2
Effect size: log2(fold−change)
E
Mycobacterium
Enriched in L1
Enriched in L3
Effect size: log2(fold−change)
F
Streptococcus
Prevotella
Anaerosinus
Corynebacterium
Neisseria
Kocuria
Mycobacterium
Enriched in L2
Enriched in L3
Effect size: log2(fold−change)


Cytochrome P450
Steroid hormone biosynthesis
Biosynthesis of siderophore group nonribosomal peptides
Primary bile acid biosynthesis
Caffeine metabolism
Betalain biosynthesis
Fatty acid metabolism
Biosynthesis and biodegradation of secondary metabolites
Chloroalkane and chloroalkene degradation
Lysine degradation
beta−Alanine metabolism
Tryptophan metabolism
Prion diseases
Propanoate metabolism
Valine, leucine and isoleucine degradation
Limonene and pinene degradation
Biosynthesis of unsaturated fatty acids
Steroid biosynthesis
Retinol metabolism
Proteasome
Geraniol degradation
Caprolactam degradation
Epithelial cell signaling in Helicobacter pylori infection
Basal transcription factors
Benzoate degradation
Stilbenoid, diarylheptanoid and gingerol biosynthesis
Bacterial toxins
African trypanosomiasis
Synthesis and degradation of ketone bodies
Polycyclic aromatic hydrocarbon degradation
Chlorocyclohexane and chlorobenzene degradation
Restriction enzyme
Non−homologous end−joining
Xylene degradation
Amyotrophic lateral sclerosis (ALS)
Bisphenol degradation
Bladder cancer
Ubiquitin system
Phosphotransferase system (PTS)
Drug metabolism − cytochrome P450
Penicillin and cephalosporin biosynthesis
Ion channels
Flavonoid biosynthesis
Staphylococcus aureus infection
Ether lipid metabolism
D−Arginine and D−ornithine metabolism
Bacterial invasion of epithelial cells
20
15
10
5
0
0
2
4
−log10(adjusted p−value)
Enriched in nTBLs
Enriched in dTBLs
Effect size: log2(fold−change)


A
B
−log10(adjusted p−value)
Butanoate metabolism
Fatty acid metabolism
beta−Alanine metabolism
Lipid biosynthesis proteins
Propanoate metabolism
Lipopolysaccharide biosynthesis proteins
Chromosome
Proteasome
Cell motility and secretion
Primary bile acid biosynthesis
Tetracycline biosynthesis
Valine, leucine and isoleucine degradation
PPAR signaling pathway
African trypanosomiasis
Peroxisome
Steroid hormone biosynthesis
Photosynthesis proteins
Tryptophan metabolism
Caffeine metabolism
Pores ion channels
Bisphenol degradation
Lysine degradation
Stilbenoid, diarylheptanoid and gingerol biosynthesis
RNA transport
Benzoate degradation
Biosynthesis of siderophore group nonribosomal peptides
Basal transcription factors
Bacterial secretion system
Geraniol degradation
Retinol metabolism
Biosynthesis of unsaturated fatty acids
Limonene and pinene degradation
Biosynthesis and biodegradation of secondary metabolites
Enriched in HIV-
Enriched in HIV+
Effect size: log2(fold−change)
Phosphonate and phosphinate metabolism
Tuberculosis
Linoleic acid metabolism
Membrane and intracellular structural molecules
Naphthalene degradation
Tetracycline biosynthesis
Proteasome
Lipopolysaccharide biosynthesis proteins
PPAR signaling pathway
Vibrio cholerae pathogenic cycle
Type I diabetes mellitus
Bisphenol degradation
Vibrio cholerae infection
Secretion system
African trypanosomiasis
Lipopolysaccharide biosynthesis
Basal transcription factors
Cell motility and secretion
Flavonoid biosynthesis
Propanoate metabolism
Bacterial secretion system
Geraniol degradation
ABC transporters
Lysine degradation
Biosynthesis of unsaturated fatty acids
Fatty acid metabolism
Transcription factors
Biosynthesis and biodegradation of secondary metabolites
Benzoate degradation
beta−Alanine metabolism
Amino acid metabolism
Peroxisome
Lipid biosynthesis proteins
Butanoate metabolism
Protein processing in endoplasmic reticulum
Aminobenzoate degradation
Ubiquitin system
Enriched in HIV-
Enriched in HIV+
Effect size: log2(fold−change)


A
Lipid biosynthesis proteins
Propanoate metabolism
Tryptophan metabolism
Lysine degradation
Limonene and pinene degradation
beta−Alanine metabolism
Valine, leucine and isoleucine degradation
Fatty acid metabolism
Prion diseases
Protein processing in endoplasmic reticulum
Carbohydrate digestion and absorption
Naphthalene degradation
Caprolactam degradation
Biosynthesis of unsaturated fatty acids
Steroid hormone biosynthesis
Chloroalkane and chloroalkene degradation
Primary bile acid biosynthesis
Restriction enzyme
Proteasome
Biosynthesis of siderophore group nonribosomal peptides
Geraniol degradation
PPAR signaling pathway
Epithelial cell signaling in Helicobacter pylori infection
Aminobenzoate degradation
Benzoate degradation
Biosynthesis and biodegradation of secondary metabolites
Retinol metabolism
Xylene degradation
Caffeine metabolism
Phosphotransferase system (PTS)
Flavonoid biosynthesis
Mineral absorption
Dioxin degradation
African trypanosomiasis
Bacterial invasion of epithelial cells
Steroid biosynthesis
Bisphenol degradation
Secondary bile acid biosynthesis
Cytochrome P450
tuberculosis
Staphylococcus aureus infection
D−Arginine and D−ornithine metabolism
Renin−angiotensin system
Stilbenoid, diarylheptanoid and gingerol biosynthesis
Protein digestion and absorption
Meiosis − yeast
Betalain biosynthesis
Polycyclic aromatic hydrocarbon degradation
Flavone and flavonol biosynthesis
butanoate metabolism
Shigellosis
Basal transcription factors
Indole alkaloid biosynthesis
Bile secretion
Non−homologous end−joining
−log10(adjusted p−value)
40
30
20
10
0
-4
-2
0
2
4
Enriched in L2
Enriched in L4
Effect size: log2(fold−change)
B
L2 vs. L4
L3 vs. L4
5 (4.2%)
3 (2.5%)
0 (0%)
1 (0.8%)
8 (6.7%)
1 (0.8%)
1 (0.8%)
16 (13.4%)
5 (4.2%)
78 (65.5%)
0 (0%)
1 (0.8%)
0 (0%)
0 (0%)
0 (0%)
L1 vs. L4
L5 vs. L4
C
L1 vs. L3
L2 vs. L3
4 (3.5%)
97 (85.8%)
12 (10.6%)