Supplementary Material

Nyawo GR, *et al. Thorax* 2023; 78:297–308. doi: 10.1136/thorax-2022-219103

Methods

FNAB collection and TB microbiology

Needle passes were done on the largest (using surface area recorded in cm^2) distinct node with a 23-gauge needle and a 10 mL syringe as described [1]. The first two passes were used to prepare standard microscope slides for cytological examination using Rapidiff and Papanicolaou staining. A flush of the needle was collected in 1.5 mL of TB transport medium [2] media and sent to the National Health Laboratory Services (NHLS) microbiology laboratory for Xpert MTB/RIF (Xpert) or Xpert MTB/RIF Ultra (Ultra), Mycobacteria Growth Indicator Tube 960 liquid culture (MGIT960; BD), and acid-fast bacilli (AFB) staining.

Definitions

We used a reference standard to designate patients as definite-TBLs (dTBLs), probable-TBLs (pTBLs), or non-TBLs (nTBLs) based on bacteriological or cytological evidence of TB as previously described [1]. dTBLs had at least one *Mycobacterium tuberculosis* complex (MTBC)-positive extrapulmonary or pulmonary specimen by acid-fast bacilli (AFB) staining microscopy, Xpert MTB/RIF (Xpert) and/or Xpert MTB/RIF Ultra (Ultra), or Mycobacteria Growth Indicator Tube (MGIT) 960 liquid culture (culture). pTBLs did not meet dTBL criteria but commenced treatment empirically. nTBLs had no microbiological TB, were not placed on treatment, and/or had an alternative diagnosis.

Clustering

We then evaluated for presence of distinct groups of samples based on identification of distinct microbial communities in lymph nodes which we called lymphotypes. Dirichlet multinomial mixture modelling (DMM) was performed using the R package *DirichletMultinomial* to establish clustering within groups [3]. Using genus tables, the number of clusters was determined by selecting the number of Dirichlet components that reduced the Laplace approximation of the model [3] (i.e. lower values indicate better fits). Clustering profiles indicate unique groupings, interpreted as “lymphotypes”.

Nyawo GR, *et al. Thorax* 2023; 78:297–308. doi: 10.1136/thorax-2022-219103

Results

Environmental and background controls

It is important to evaluate possible sources of microbial DNA contamination in low biomass samples such lymph fluid. Given the concern for potential carry-over of skin commensals and DNA present on medical apparatus (i.e., syringe used for biopsy), microbiome readouts from patients' lymph fluid were analysed in parallel with that from skin (sampled at puncture site) and saline flush of the syringe used for aspiration (done for 1 in 5 participants). Pairwise comparisons of α-diversity were similar between saline and skin, and lymph fluid and saline (**Figure S1A**). β-diversity was different between the three fluid types (*p=0.001*; **Figure S1B**), with lymph enriched in the respiratory pathogen *Mycobacterium* (**Figure S1C**) vs. skin, and vs. saline (**Figure S1D**). Skin was enriched with *Psychrobacter* and *Corynebacterium* vs. lymph and saline, respectively (**Figures S1C and S1E**), whilst no taxa were enriched in saline (**Figures S1D-E**).

Cohort characteristics

We had 89 dTBLs, 61 nTBLs (**Figure 1**) and 8 pTBLs (latter subsequently excluded due to small numbers), from which we collected FNABs from the head, neck or thorax regions. 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.

Microbial comparisons including probable TBL patients

We compared the microbiome of pTBLs (n=9) to dTBLs and nTBLs. There were no differences in α-diversity (**Figure S2A**). β-Diversity differed between pTBLs and dTBLs, but not between pTBLs and nTBLs (**Figure S2B**). We excluded the pTBLs from the primary analysis due to few patients meeting this definition.

*α- and β-diversities according to demographic, clinical, and microbiological characteristics **(Table S2)***

Overall: Females had a higher α-diversity than males (*p=0.016*), patients who used antibiotics within a year, and at recruitment had a lower α-diversity than those who did not (*p=0.042; p=0.003 respectively*), and patients with smaller lymph nodes had a higher α-diversity than those with larger nodes (*p=0.001*). β–diversity was different in patients with antibiotic use at recruitment versus none (*p=0.032*) and antibiotic use within one year versus later use (*p=0.020*). Furthermore, within PLHIV, β–diversity differed by ART status (*p=0.042*) and CD4 count stratum (*p=0.038*). In those patients who received non-TB antibiotics (i.e., current or within one year of enrolment), α-diversities by TB status were similar but β-diversity differed (*p=0.001*). In patients without antibiotics (n=109), α-diversity was lower in dTBLs vs. nTBLs (*p=0.035*) and β-diversity differed (*p=0.035*).

dTBLs: α-diversity was decreased with antibiotic use at recruitment (*p=0.025*) and within one year (*p=0.007*) as well as in larger nodes lymph node size (*p=0.034*). β-diversity also differed by antibiotics usage (current and within one year) and CD4 count stratum in PLHIV (*p=0.034*).

nTBLs: α-diversity was less in males than females (*p=0.003*) and in smokers than non-smokers (*p=0.002*). β-diversity was only associated with specimen appearance in nTBLs (*p=0.047*).

Correlation between 16S rRNA gene sequencing and TB diagnostic tests

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 S4A**), and culture days-to-positivity; r_s=-0.684, 95% CI [-0.859, -0.367.], *p=0.003*; **Figure S4B**). However, there was no correlation between lymph node size and mycobacterial load (relative abundance) (**Figure S4C**).

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, however, the relative abundance of *Mycobacterium* was higher in PLHIV (**Figure S5C**).

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 by HIV status within dTBLs (*p=0.017,* **Figure 3C**) but not within nTBLs. HIV-positive dTBLs had higher *Mycobacterium* relative abundance than HIV-negative dTBLs **(Figure S5C)**.

Comparisons within HIV-positives or -negatives by TB status: In people of 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 3D**) where dTBLs were enriched in *Mycobacterium* (**Figure S5D**). In HIV-negatives, there were no differences between dTBLs and nTBLs (**Figure S5B**).

Lymphotype identification and their associations with clinical characteristics

We further explored this data using DMM to identify potential clusters in the TBL microbiome. These clusters were termed “lymphotypes”. We then looked for associations between each lymphotype(s) and patients’ clinical characteristics.

Overall: We first examined whether all patients could be grouped into distinct lymphotypes; these were termed overall lymphotypes (oLs). Five oLs with differing α- and β-diversities were identified (**Figure 4A-C, Table S3**). oL1 had no dominant taxa (**Figure 4D**), whilst oL4 was *Mycobacterium*-dominated and had the least α-diversity, and oL2, oL3 and oL5 were *Corynebacterium*-, *Prevotella*- and *Streptococcus*-dominated, respectively. While no taxa were differentially abundant in oL1 vs. other oLs (**Figure S6A-C**), oL2, oL3, and oL5 were enriched relative to oL4 in *Corynebacterium*, *Prevotella*, and *Streptococcus,* respectively (**Figure 4E-G**). The proportions of dTBLs in oL1, oL2, oL3, oL4, and oL5 were 35% (17/48), 63% (28/44), 57% (12/21), 100% (21/21), and 69% (11/16), respectively (**Table S4**). The patients in these lymphotypes are associated with distinct clinical characteristics. The majority of nTBLs occurred in highly diverse oLs with a heterogenous mixtures of taxa; likely reflecting the spectrum of pathologies in people with TBL ruled out. oL1 was associated with characteristics indicative of less severe lymphadenitis. Compared separately to oL2, oL4, and oL5, oL1s were less likely to have dTBL. Furthermore, oL1s were less likely to be HIV-positive vs. oL4s but, oL1 PLHIVs had lower CD4 counts vs. oL2 and oL3 PLHIVs. In contrast, oL4 was associated with characteristics resembling more severe lymphadenitis. oL4 was more likely to contain dTBL patients than each other oL. Furthermore, compared with oL2s, oL4s were more likely to have a bigger lymph node, chylous FNABs and, of PLHIV, a smaller proportion on ART. Compared with oL3s, oL4s were more likely to have previous TB and HIV, and those with HIV were more likely to have lower CD4 counts. Compared with oL5s, oL3s with HIV had lower CD4 counts. Therefore, in summary, oL1 appears to be associated with less severe lymphadenitis forms, whereas oL4 was associated more severe forms (**Table S4**).

Nyawo GR, *et al. Thorax* 2023; 78:297–308. doi: 10.1136/thorax-2022-219103

Within patients of the same TB status: We then examined whether patients within each TB group could be grouped into distinct lymphotypes. Within dTBLs, three lymphotypes (termed definite TBL lymphotypes; dL) with differing β-diversities were identified (**Figure 5A-B**). dL1 was abundant in *Prevotella* and *Corynebacterium*, dL2 in *Prevotella* and *Streptococcus*, and dL3 in *Mycobacterium* (**Figure 5C**) and these taxa were differentially abundant (**Figure 5D-F**). These lymphotypes were termed *Prevotella-Corynebacterium*, *Prevotella-Streptococcus* and *Mycobacterium,* respectively. dL3s were more likely to be HIV-positive, with larger lymph nodes, compared with dL1s. In addition, dL3s were more likely to have larger lymph nodes than dL2s. Lastly, dL2s are more likely to be female than dL1s (**Table S5**). Together, these differences suggest dL3 is associated with more severe TBL than other lymphotypes. Within nTBLs, no lymphotypes were identified (**Figure S7**).

Predictive metagenome profiling shows increased short chain fatty acid metabolism

We further predicted the bacterial metagenome content and made functional inferences of the microbiome using the PICRUSt algorithm. The overall differences among pathways between groups were evaluated and visualised by DESeq2 analysis.

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 S8**).

In different lymphotypes: When comparing inferred pathways in oLs, a similar core of pathways was enriched in oL4. 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. oL4, oL1 was enriched in “epithelial cell signalling in *Helicobacter pylori* infection”, oL2 was enriched in “carbohydrate digestion and absorption”, oL3 was enriched in “dioxin degradation”, and oL5 was enriched in “carbohydrate digestion and absorption” (**Figure S9A-H)**. When comparing the three dTBL lymphotypes, *Mycobacterium*-dominated dL3 was, compared with each other dLs, enriched in the similar core pathways seen for the *Mycobacterium*-dominated oL4 in all patients (**Figure 8C; Figure S10**).

Table S1: Reference standard definition used in the study. Due a small number of pTBLs, they were excluded from analyses.

	dTBLs	nTBLs	pTBLs
	Site-of-disease fluid		
Xpert	✓	×	×
Ultra	✓	×	×
MGIT960 Culture	✓	×	×
Smear microscopy	✓	×	×
Cytology	✓	×	×
	Non-site-of-disease fluid		
Smear microscopy	✓	×	×
Xpert	✓	×	×
Ultra	✓	×	×
MGIT960	✓	×	×
	Treatment information		
TB treatment initiated	×	×	✓
Response to treatment self-reported by patient	×	×	✓

Abbreviations: dTBLs: definite tuberculous lymphadenitis; nTBLs: non-tuberculous lymphadenitis; pTBLs: probable-tuberculous lymphadenitis; Xpert: Xpert MTB/RIF; Ultra: Xpert MTB/RIF Ultra; MGIT960 Culture: Mycobacteria Growth Indicator Tube 960 liquid culture.

Nyawo GR, *et al. Thorax* 2023; 78:297–308. doi: 10.1136/thorax-2022-219103

Table S2: α- and β-diversities in presumptive TBL patients when patients with different demographic and clinical characteristics were compared. Several characteristics, described in the Supplementary Results text, were associated with differing diversities.

204

Characteristics	Overall (n=150)			dTBLs (n=89)			nTBLs (n=61)		
	α-diversity *p-value* (Shannon's Index)	β-diversity		α-diversity *p-value* (Shannon's Index	β-diversity		α-diversity *p-value* (Shannon's Index	β-diversity	
		p-value (PERMANOVA)	R^2 *value*		*p-value* (PERMANOVA)	R^2 *value*		*p-value* (PERMANOVA)	R^2 *value*
dTBL	0.110	**0.001**	0.037	-	-	-	-	-	-
Sex	**0.016**	0.121	0.010	0.406	0.616	0.035	**0.003**	**0.012**	0.008
HIV	0.860	**0.004**	0.023	0.179	**0.008**	0.043	0.312	0.731	0.432
CD4+ <200 cells/µl	0.459	**0.038**	0.032	0.053	**0.034**	0.055	0.140	0.455	0.045
On ART	0.662	**0.042**	0.030	0.267	0.344	0.022	0.306	0.267	0.055
Previous TB	0.337	0.072	0.012	0.426	0.141	0.018	0.501	0.603	0.015
Tobacco smoking	0.084	0.189	0.009	0.636	0.658	0.008	**0.002**	0.276	0.020
Antibiotic use within 1 year of recruitment	**0.042**	**0.020**	0.015	**0.025**	**0.012**	0.036	0.547	0.212	0.022
Antibiotic use at recruitment	**0.003**	**0.032**	0.061	**0.007**	**0.025**	0.141	0.062	0.064	0.115
Site (neck vs. thorax)	0.220	0.134	0.010	0.128	0.142	0.018	0.809	0.830	0.011
Specimen appearance (bloody vs. chylous)	0.213	0.068	0.012	0.771	0.198	0.016	**0.020**	**0.047**	0.778
Lymph node characteristics: size, cm^2	**0.011**	0.128	0.012	**0.034**	0.065	0.265	0.197	0.612	0.017

*R^2 provides the proportion of variation explained (e.g., a factor that has a $R^2 = 0.037$, explains 3.7% of the variation in community composition) by β-diversity.

Abbreviations: TB: tuberculosis; TBL: tuberculous lymphadenitis; ART: antiretroviral therapy; dTBLs: definite tuberculous lymphadenitis; nTBLs: non-tuberculous lymphadenitis

Table S3: Adjusted p-values for α-diversity comparisons between lymphotypes measured by Shannon's diversity index.

Comparison	Lymphotype with highest α-diversity	Adjusted *p-value*
Lymphotype comparisons in overall lymphotypes		
oL1 vs. oL2	oL2	**<0.0001**
oL1 vs. oL3	oL3	**0.0012**
oL1 vs. oL4	oL1	>0.9999
oL1 vs. oL5	oL5	**<0.0001**
oL2 vs. oL3	oL2	>0.9999
oL2 vs. oL4	oL2	**<0.0001**
oL2 vs. oL5	oL5	**0.0329**
oL3 vs. oL4	oL3	**<0.0001**
oL3 vs. oL5	oL5	0.1088
oL4 vs. oL5	oL5	**<0.0001**
Lymphotype comparisons in all dTBL lymphotypess		
dL1 vs. dL2	oL2	**0.001**
dL1 vs. oL3	oL3	**<0.0001**
dL2 vs. oL3	oL3	**0.001**

Definition of abbreviations: oL: overall lymphotype;; dL: dTBL lymphotype.

205
206

Table S4: Demographic, clinical, and microbiological differences in each lymphotype (overall in all patients) showing oL1 is likely associated with less severe forms of lymphadenitis whereas oL4 is associated with more severe forms. Amongst other differences, oL1s were less likely to have dTBL than oL2s, oL4s, and oL5s. Furthermore, oL1s were less likely to be HIV-positive vs. oL4s. oL1 PLHIV had lower CD4 counts vs. oL2 and oL3 PLHIVs. In contrast, oL4s were more likely to be dTBLs than other lymphotypes. Furthermore, compared to oL2, oL4s had bigger lymph nodes and were more likely to have chylous FNABs and a smaller proportion of PLHIVs on ART. Compared to oL3, oL4s were more likely to have previous TB and HIV, and oL3 PLHIVs were more likely to have lower CD4 counts. Compared to oL5, oL3 PLHIV had lower CD4 counts.

Characteristic[¶]	Total (n=150)	L1 (n=48) (No dominant taxa)	L2 (n=44) *Corynebacterium*)	L3 (n=21) (*Prevotella*)	L4 (n=21) (*Mycobacterium*)	L5 (n=16) (*Streptococcus*)	*p-value* (L1 vs. L2)	*p-value* (L1 vs. L3)	*p-value* (L1 vs. L4)	*p-value* (L1 vs. L5)	*p-value* (L2 vs. L3)	*p-value* (L2 vs. L4)	*p-value* (L2 vs. L5)	*p-value* (L3 vs. L4)	*p-value* (L3 vs. L5)	*p-value* (L4 vs. L5)
Age, years	36 (30-45)	35 (29-47)	37 (32-47)	31 (28-46)	37 (34-43)	36 (28-45)	>0.999	>0.999	>0.999	>0.999	>0.999	>0.999	>0.999	>0.999	>0.999	>0.999
dTBLs	89/150 (59)	17/48 (35)	28/44 (64)	12/21 (57)	21/21 (100)	11/16 (69)	**0.007**	0.093	**<0.001**	**0.020**	0.615	**0.001**	0.713	**0.001**	0.471	**0.006**
Female	83/150 (55)	25/48 (52)	26/44 (59)	8/21 (38)	12/21 (57)	12/16 (75)	0.499	0.284	0.698	0.108	0.113	0.882	0.258	0.217	**0.026**	0.260
HIV	72/148 (49)	20/48 (42)	24/42 (57)	6/21 (29)	15/21 (71)	7/16 (44)	0.143	0.302	**0.023**	0.884	**0.032**	0.271	0.361	**0.006**	0.338	0.089
CD4+	166 (90-308)	35 (29-48)	171 (86-332)	255 (154-387)	83 (17-163)	136 (54-334)	**<0.0001**	**0.001**	0.733	0.103	>0.999	0.620	>0.999	0.223	>0.999	>0.999
CD4+ <200 cells/µl	43/72 (60)	10/20 (50)	14/24 (58)	2/6 (33)	12/15 (80)	5/15 (33)	0.580	0.473	0.069	0.324	0.272	0.163	0.129	**0.040**	>0.999	**0.010**
On ART	35/71 (49)	9/20 (45)	16/24 (67)	3/6 (50)	5/15 (33)	2/6 (33)	0.149	0.829	0.486	0.612	0.449	**0.042**	0.136	0.477	0.558	>0.999
Previous TB	33/148 (22)	10/48 (21)	9/42 (21)	2/21 (10)	9/21 (43)	3/16 (19)	0.945	0.254	0.060	0.858	0.241	0.076	0.822	**0.014**	0.416	0.121
Tobacco smoking	43/149 (29)	18/48 (38)	12/44 (27)	8/21 (38)	3/21 (14)	2/15 (13)	0.296	0.936	0.054	0.062	0.377	0.245	0.232	0.079	0.082	0.875
Antibiotic use within 1 year of recruitment	38/147 (26)	11/47 (23)	9/9 (100)	4/20 (20)	9/20 (45)	5/16 (31)	**<0.001**	0.760	0.077	0.533	**<0.001**	**0.005**	**0.001**	0.091	0.439	0.400
At recruitment	21/38 (55)	8/11 (73)	5/9 (56)	1/4 (25)	6/9 (67)	1/5 (20)	0.423	0.095	0.769	**0.049**	0.308	0.629	0.198	0.164	0.858	0.094
Lymph node characteristics: sites																
Neck	133/150 (89)	46/48 (96)	37/44 (84)	18/21 (86)	20/21 (95)	12/16 (75)	0.058	0.136	0.911	**0.013**	0.865	0.201	0.421	0.293	0.410	0.074
Deep anterior cervical	60/133 (45)	19/46 (41)	16/37 (43)	8/18 (44)	13/20 (65)	4/12 (33)	0.859	0.819	0.077	0.615	0.933	0.117	0.544	0.203	0.543	0.082
Deep lateral cervical	25/133 (19)	13/46 (28)	8/37 (22)	2/18 (11)	2/20 (10)	0/12 (0)	0.489	0.145	0.104	**0.037**	0.343	0.271	0.078	0.911	0.232	0.258
Superficial	15/133 (11)	8/46 (17)	2/37 (5)	3/18 (17)	2/20 (10)	0/12 (0)	0.095	0.945	0.442	0.120	0.173	0.517	0.411	0.544	0.136	0.258
Supraclavicular	20/133 (15)	2/46 (4)	7/37 (19)	3/18 (17)	3/20 (15)	5/12 (42)	**0.034**	0.099	0.133	**<0.001**	0.839	0.710	0.111	0.888	0.129	0.092
Head	13/133 (10)	4/46 (9)	4/37 (11)	2/18 (11)	0/20 (0)	3/12 (25)	0.746	0.766	0.174	0.123	0.973	0.127	0.222	0.126	0.317	**0.019**
Thorax	17/150 (11)	2/48 (4)	7/44 (16)	3/21 (14)	1/21 (5)	4/16 (25)	0.058	0.136	0.911	**0.013**	0.865	0.201	0.421	0.293	0.410	0.074
Axillary (vs. breast)	13/17 (76)	1/2 (50)	7/7 (100)	3/3 (100)	1/1 (100)	1/4 (25)	**0.047**	0.171	0.387	0.540	-	-	**0.007**	-	**0.047**	0.171
Lymph node characteristics: size, cm^2	4 (2-9)	4 (4-9)	3 (1-4)	5 (3-10)	6 (4-29)	4 (1-9)	**0.0288**	>0.9999	>0.9999	>0.9999	0.1069	**0.002**	>0.9999	>0.9999	>0.9999	0.420
Specimen appearance																
Bloody (vs. chylous)	123/150 (82)	40/48 (83)	39/44 (89)	19/21 (90)	14/21 (67)	11/16 (69)	0.466	0.438	0.122	0.209	0.823	**0.033**	0.068	0.060	0.095	0.893

Abbreviations: TB: tuberculosis; TBLs: tuberculous lymphadenitis; HIV: human immunodeficiency virus; ART: antiretroviral therapy; L: lymphotype; oL: overall lymphotype; dTBLs: definite tuberculous lymphadenitis; nTBLs: non-tuberculous lyphadenitis;

Table S5: Demographic, clinical, and microbiological differences between dTBL lymphotypes. dL3s had characteristics associated with more severe TBL. dL3s were more likely to have HIV and larger lymph nodes compared to dL1s and dL2s. dL2s were more likely to be female than dL1s.

Characteristic	Total (n=89)	dL1 (n=48) (*Prevotella*-*Corynebacterium*)	dL2 (n=21) (Prevotella-*Streptococcus*)	dL3 (n=20) (*Mycobacterium*)	*p-value* (dL1 vs. dL2)	*p-value* (dL1 vs. dL3)	*p-value* (dL2 vs. dL3)
Age, years	35 (29-40)	33 (28-38)	36 (28-46)	37 (34-44)	>0.999	0.197	0.873
Female	48/89 (54)	22/48 (46)	15/21 (71)	11/20 (55)	**0.050**	0.491	0.275
HIV	49/89 (55)	23/48 (48)	11/21 (52)	15/20 (75)	0.733	**0.040**	0.133
CD4+	155 (76-251)	157 (106-250)	212 (64-385)	92 (17-226)	>0.999	0.254	0.172
CD4+ <200 cells/µl	32/49 (65)	16/23 (70)	5/11 (45)	11/15 (73)	0.180	0.800	0.150
On ART	21/49 (43)	11/23 (48)	5/11 (45)	5/15 (33)	0.900	0.380	0.530
Previous TB	24/88 (27)	11/47 (23)	4/21 (19)	9/20 (45)	0.689	0.077	0.074
Tobacco smoking	21/89 (24)	13/48 (27)	4/21 (19)	4/20 (20)	0.480	0.540	0.940
within 1 year of recruitment	22/87 (25)	7/47 (15)	6/20 (30)	9/20 (45)	0.153	**0.008**	0.327
At recruitment	10/22 (45)	2/7 (29)	2/6 (33)	6/9 (67)	0.850	0.130	0.200
Lymph node characteristics: sites							
Neck	78/89 (88)	42/48 (88)	17/21 (81)	19/20 (95)	0.480	0.350	0.170
Deep anterior cervical	36/78 (46)	16/42 (38)	7/17 (41)	13/19 (68)	0.826	**0.028**	0.101
Deep lateral cervical	15/78 (19)	11/42 (26)	2/17 (12)	2/19 (11)	0.230	0.170	0.910
Superficial	6/78 (8)	5/42 (12)	0/17 (0)	1/19 (5)	0.140	0.420	0.340
Supraclavicular	17/78 (22)	7/42 (17)	7/17 (41)	3/19 (16)	**0.045**	0.920	0.090
Head	4/78 (5)	3/42 (7)	1/17 (6)	0/19 (0)	0.860	0.230	0.280
Thorax	11/89 (12)	6/48 (13)	4/21 (19)	1/20 (5)	0.480	0.350	0.170
Axillary (vs. breast)	9/11 (82)	6/6 (100)	2/4 (50)	1/1 (100)	0.053	-	0.361
:haracteristics: size, cm^2	4 (2-9)	4 (1-7)	4 (1-4)	8 (4-12)	0.827	**0.030**	**0.005**
Specimen appearance							
Bloody (vs. chylous)	66/89 (74)	38/48 (79)	14/21 (67)	14/20 (70)	0.270	0.420	0.820

Abbreviations: TB: tuberculosis; TBLs: tuberculous lymphadenitis; HIV: human immunodeficiency virus; ART: antiretroviral therapy; dTBLs: definite tuberculous lymphadenitis; dL: dTBL lymphotype.

Figure S1: Paired analysis of controls and lymph fluid (n=33) indicates that environmental cross contamination is highly unlikely. (A) α-diversity analyses show skin has higher diversity than lymph fluid. (B) β-diversity of lymph fluid differs to saline and skin. DESeq2 volcano plots depicting differentially abundant taxa show that (C) lymph was enriched in *Mycobacterium* vs. (C) skin and (D) saline, and there were more differentially abundant taxa in skin vs. (C) lymph and (E) saline. Significantly more discriminatory taxa appear closer to the left or right, and higher above the threshold (red dotted line, FDR=0.2) as the degree of significance increases. Relative taxa abundance is indicated by circle size.

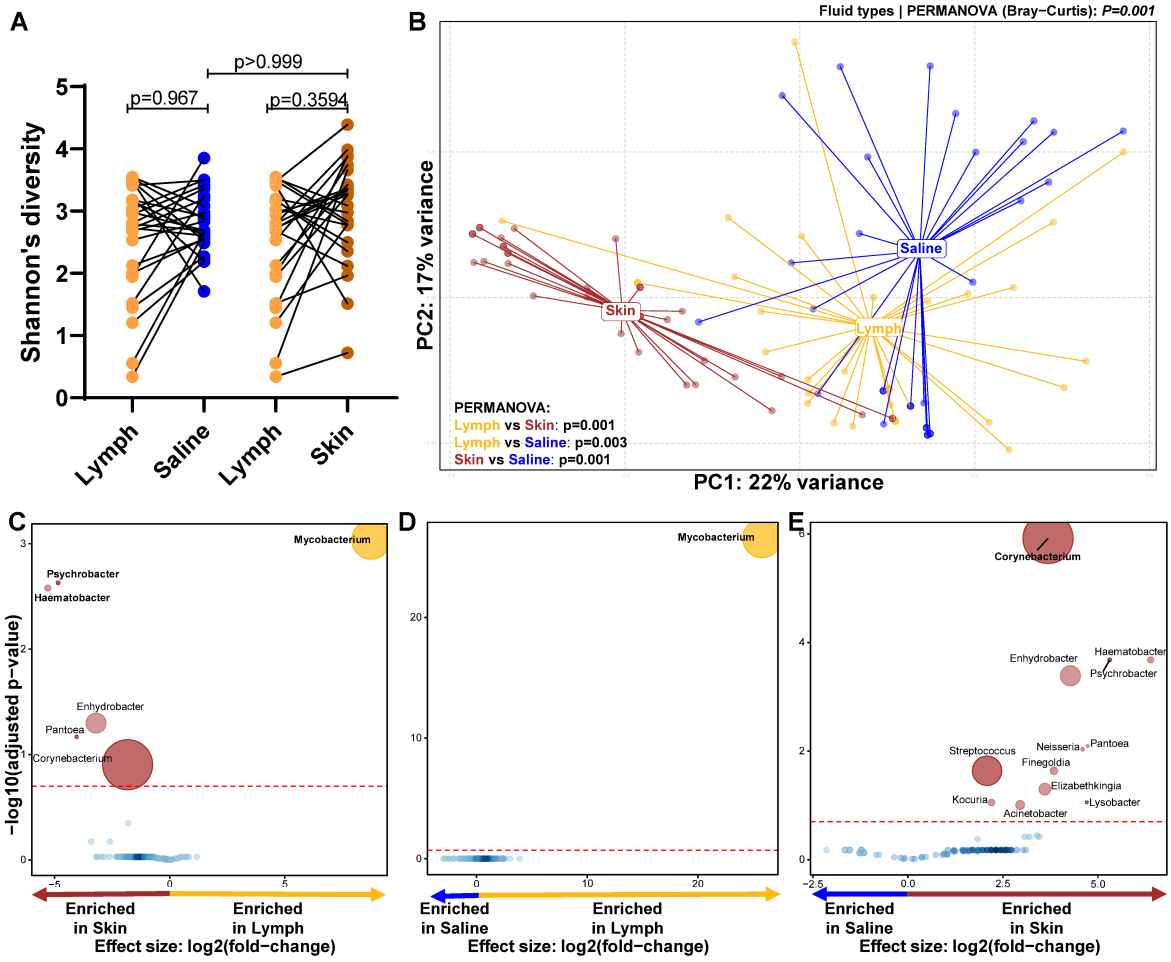
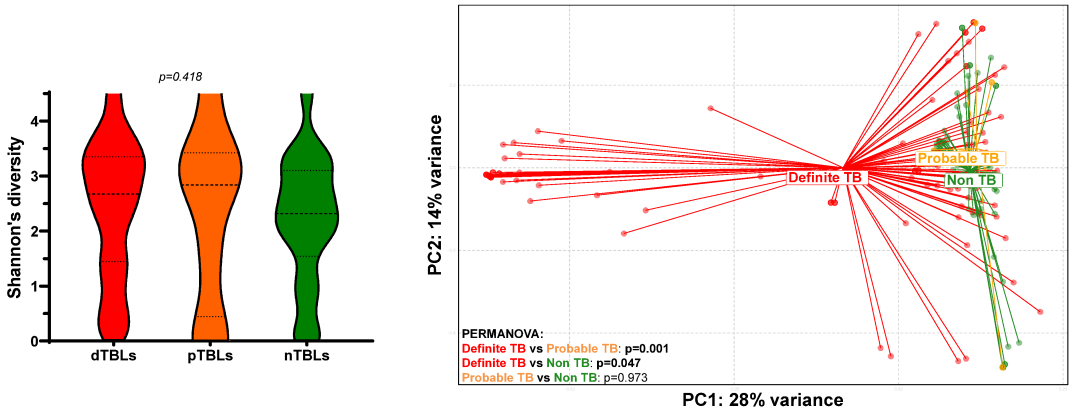


Figure S2: dTBLs have a distinct microbiome to pTBLs and nTBLs s. (A) Microbial diversity is similar, but (B) microbial composition of dTBLs is different from both pTBLs and nTBLs.



Nyawo GR, *et al. Thorax* 2023; 78:297–308. doi: 10.1136/thorax-2022-219103

Figure S3: *Mycobacterium* reads in FNABs of participants showing some nTBLs with Mycobacterium reads. Relative abundance of *Mycobacterium* per participant stratified by TB status shows *Mycobacterium* in some nTBLs. Furthermore, not all dTBLs had detected *Mycobacterium* reads.

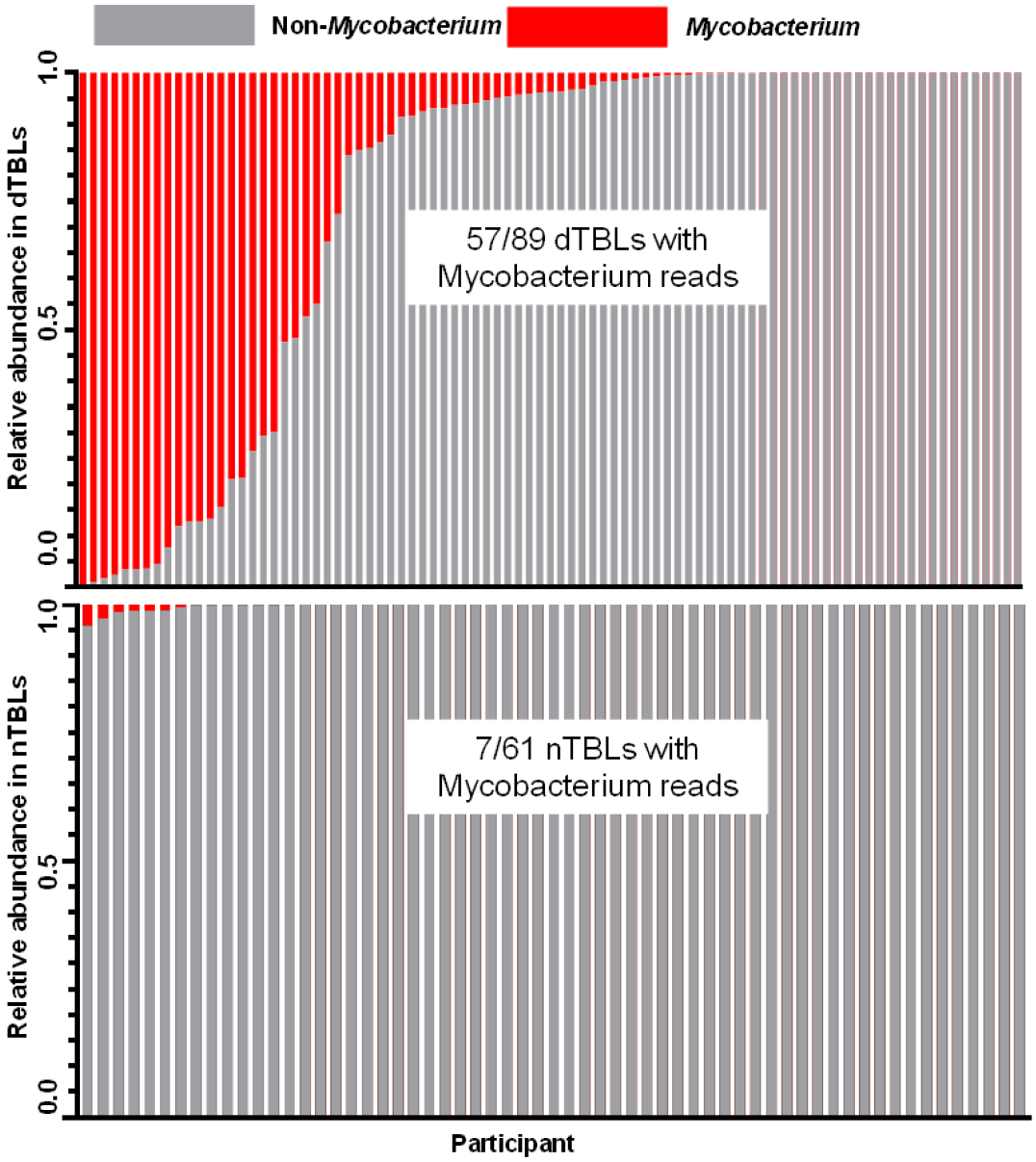
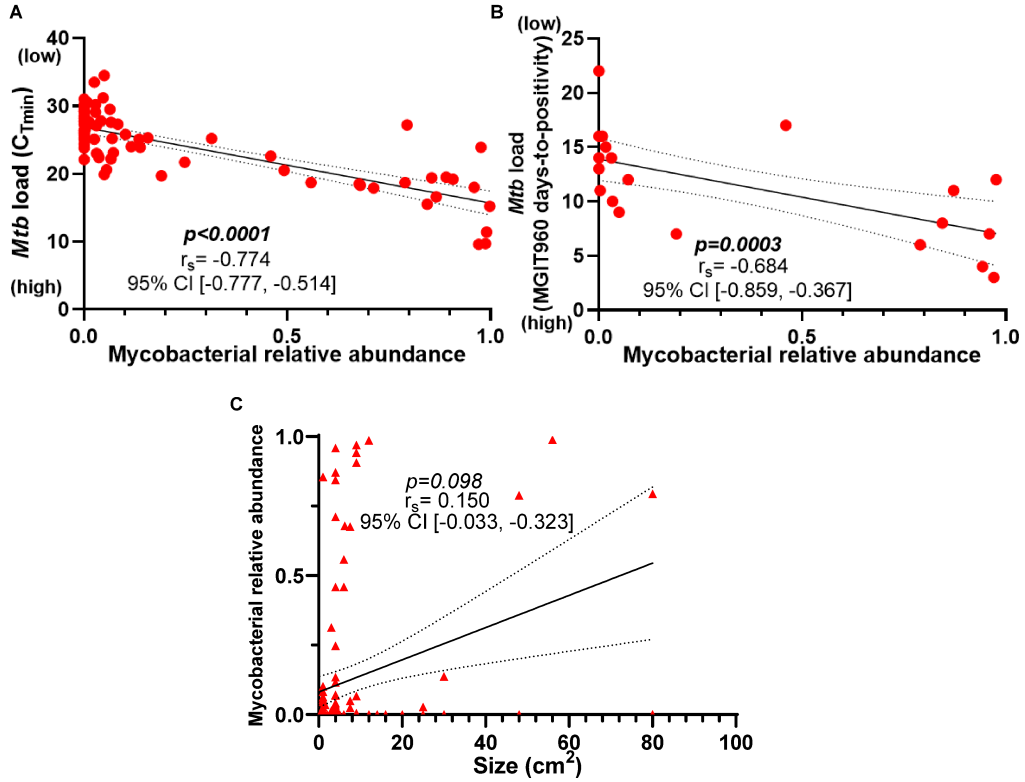


Figure S4: 16S rRNA gene sequencing positively correlated to TB diagnostic tests. Mycobacterial reads positively correlated with *Mtb* load: (A) Xpert and Ultra and (B) culture (days-to-positivity); but not with (C) lymph node size. Xpert: Xpert MTB/RIF; Ultra: Xpert MTB/RIF Ultra; MGIT960: Mycobacteria Growth Indicator Tube 960 liquid culture; rs: Spearman correlation coefficient.



Nyawo GR, *et al. Thorax* 2023; 78:297–308. doi: 10.1136/thorax-2022-219103

Figure S5: HIV has a greater effect on the microbiome in patients co-infected with TB. β-diversity did not differ (A) by HIV status in nTBLs, or by TBL status in (B) HIV-negatives. Circle sizes represent relative abundances. dTBLs: definite tuberculous lymphadenitis; nTBLs: non-tuberculous lymphadenitis**. (C)** *Mycobacterium* abundance did not differ by HIV status within dTBLs and within nTBLs, and (D) HIV-positive dTBLs were enriched in *Mycobacterium* compared to nTBLs. Circle sizes represent relative abundances. dTBLs: definite tuberculous lymphadenitis; nTBLs: nontuberculous lymphadenitis.

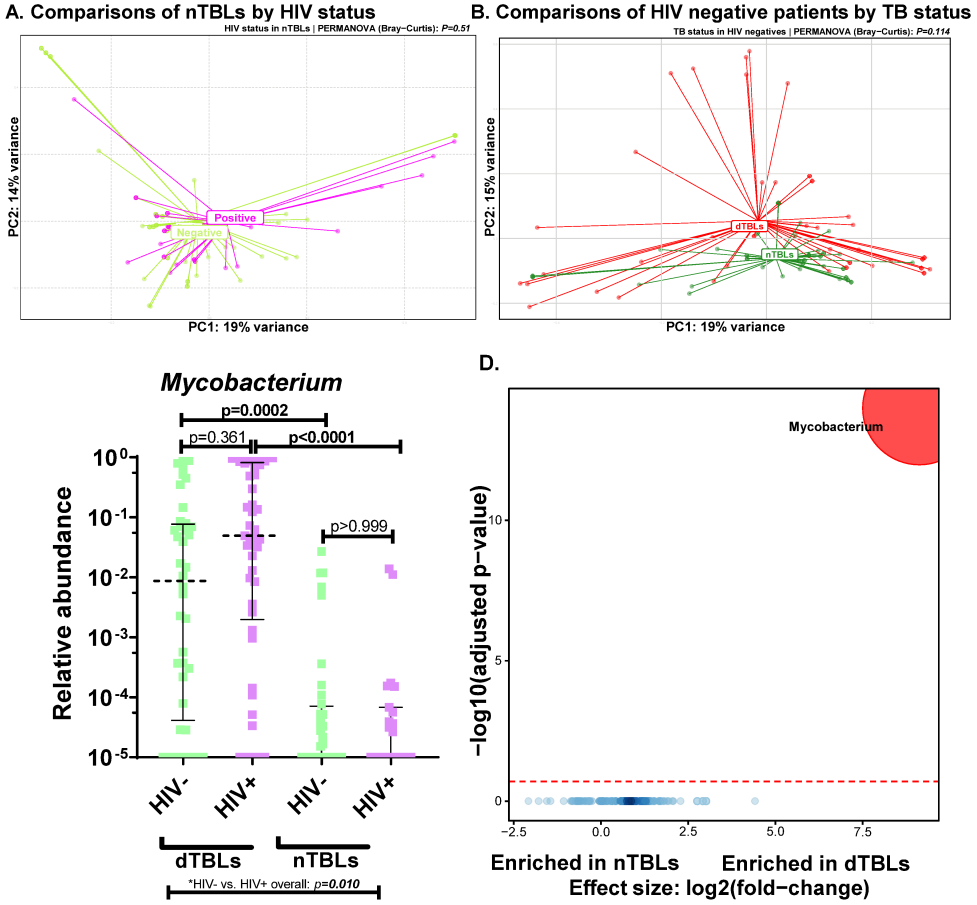


Figure S6: Five microbial community states observed in presumptive TBL patients are enriched with distinct taxa. oL1 had no enriched taxa, and was depleted in (A) *Enhydrobacter*, (B) *Mycobacterium*, and (C) *Streptococcus*, *Anaerosinus*, *Neisseria* and *Kocuria*. oL3 was enriched in (D) *Acinetobacter* and depleted of *Prevotella*. oL5 was enriched in *Streptococcus* accompanied with (E) *Anaerosinus*, *Neisseria*, *Kocuria* and *Prevotella* vs. oL2, and with (F) *Bacteroides* and *Kocuria* vs. oL3. 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 abundance of taxa is indicated by circle size. oL: Lymphotype.

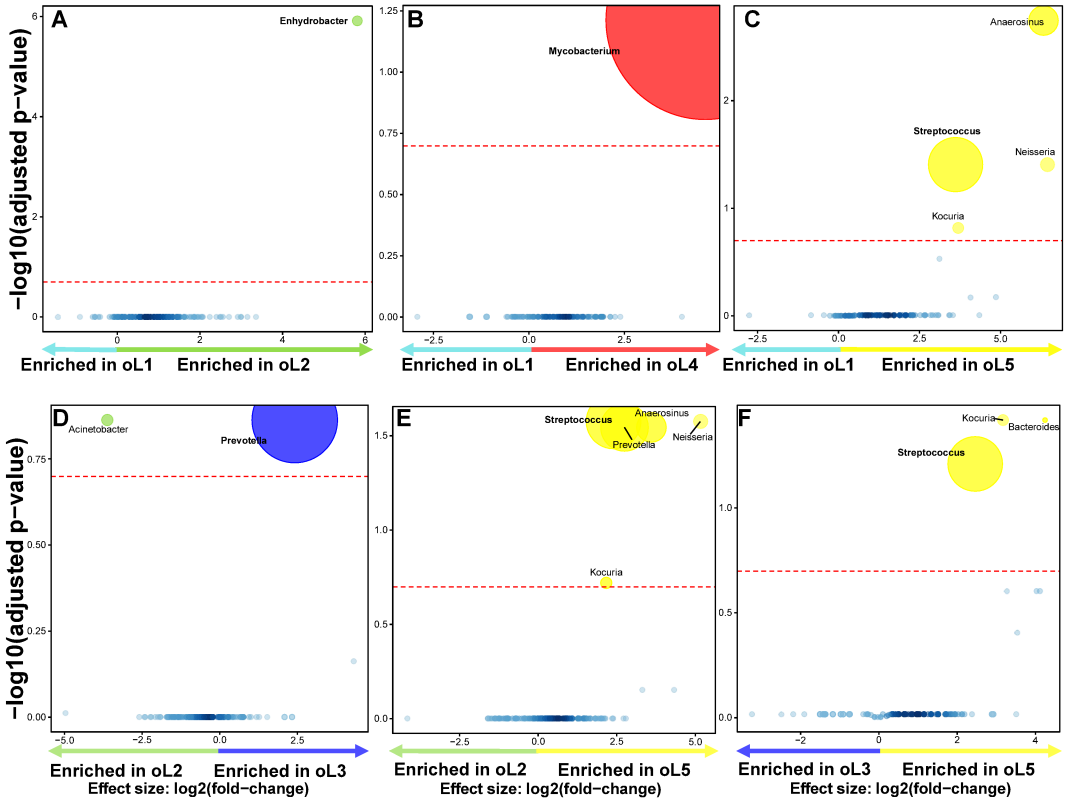


Figure S7: The Laplace approximation of model evidence is a measure of the model fit. Laplace approximation predicts no clustering for nTBL patients. Lower values indicate better fit. nTBLs: non-tuberculous lymphadenitis.

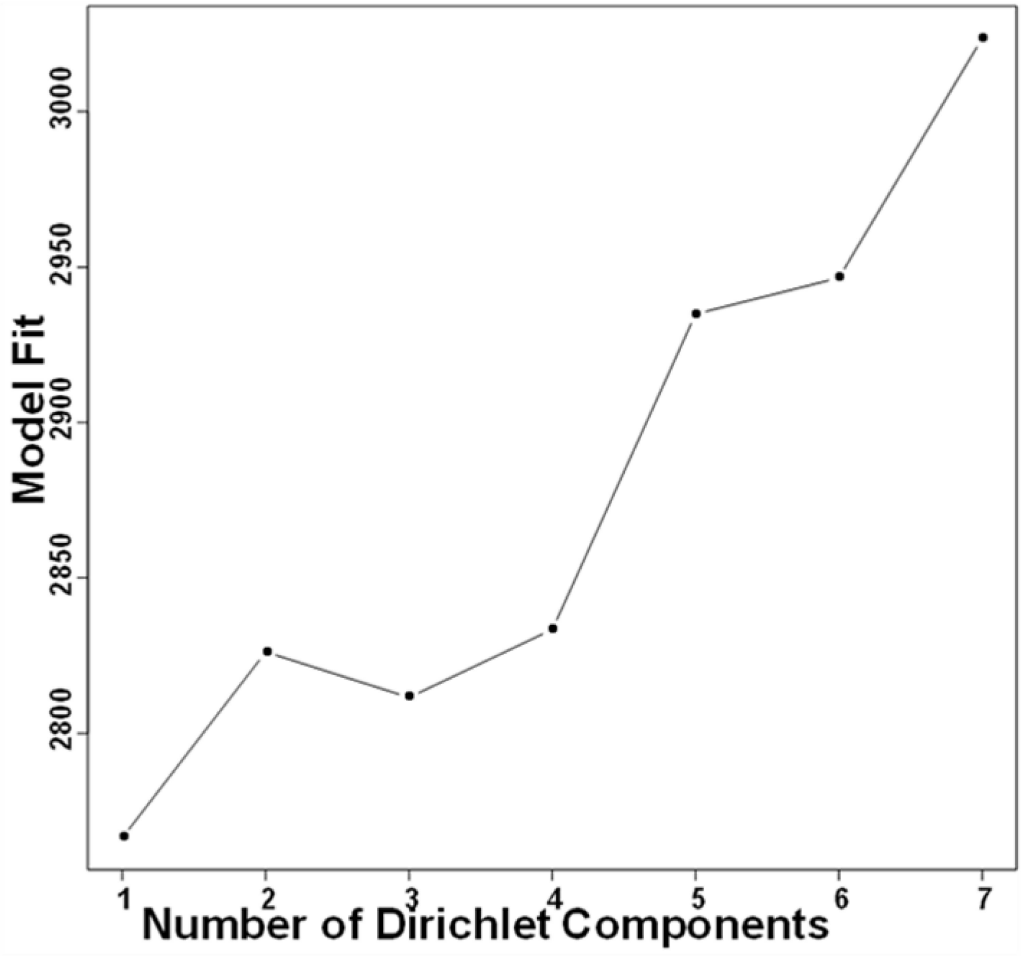
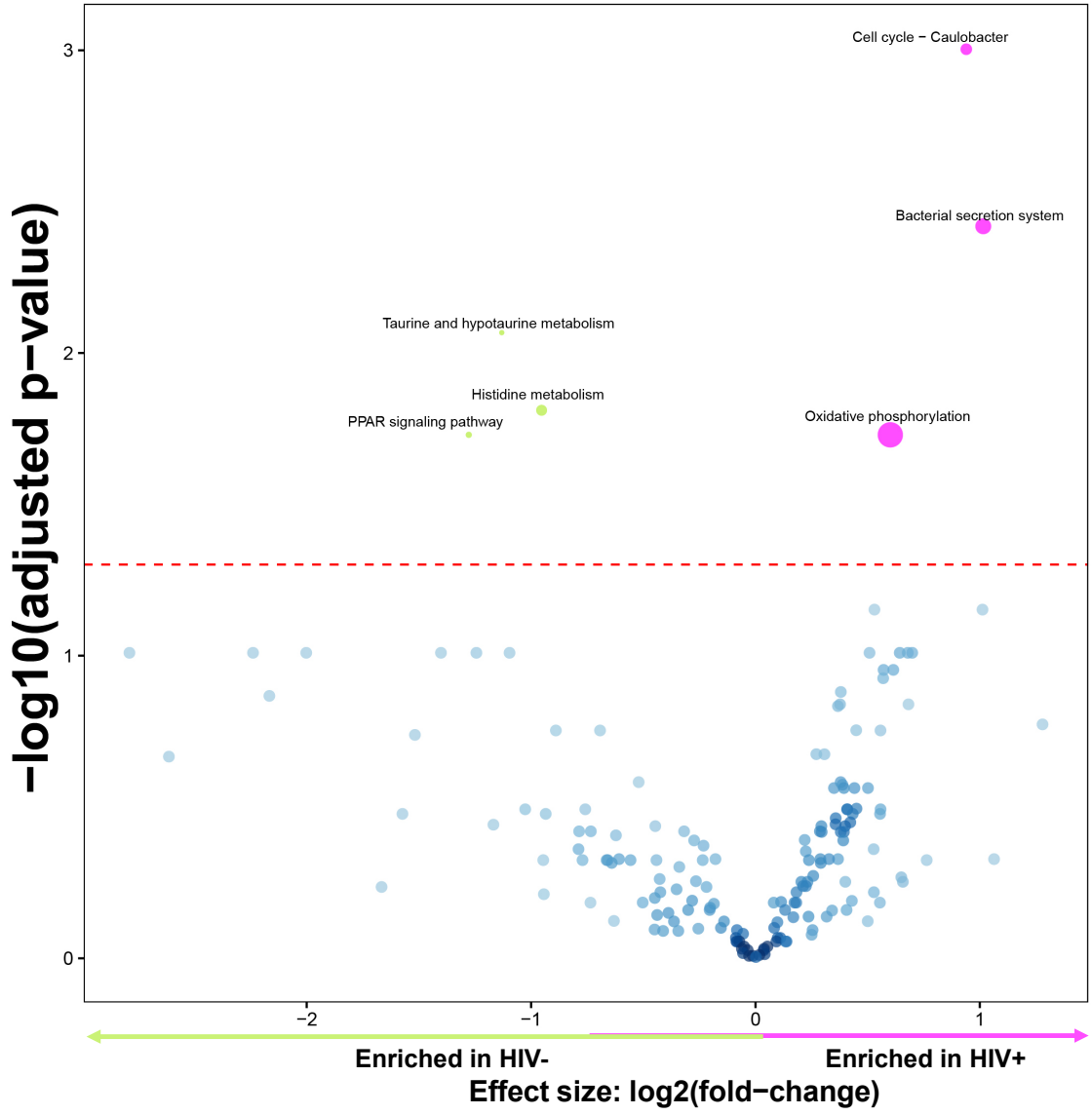


Figure S8: Predicted metagenome function in HIV-positive nTBLs versus HIV-negative nTBLs. Volcano plot depicting functional pathways differing between HIV-positive and HIV-negative nTBLs. Significantly more discriminatory pathways appear closer to the left or right, and higher above the threshold (red dotted line, FDR=0.05). Key pathways of interest include "cell cycle - *Caulabacter*", "bacterial secretion system", "taurine and hypotaurine metabolism", and "histidine metabolism". Relative gene abundance is indicated by circle size.



Nyawo GR, *et al. Thorax* 2023; 78:297–308. doi: 10.1136/thorax-2022-219103

Figure S9: Inferred metagenomes of lymphotypes in all patients. Volcano plot depicting differentially enriched pathways in oL4 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. 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. oL: overall Lymphotype.

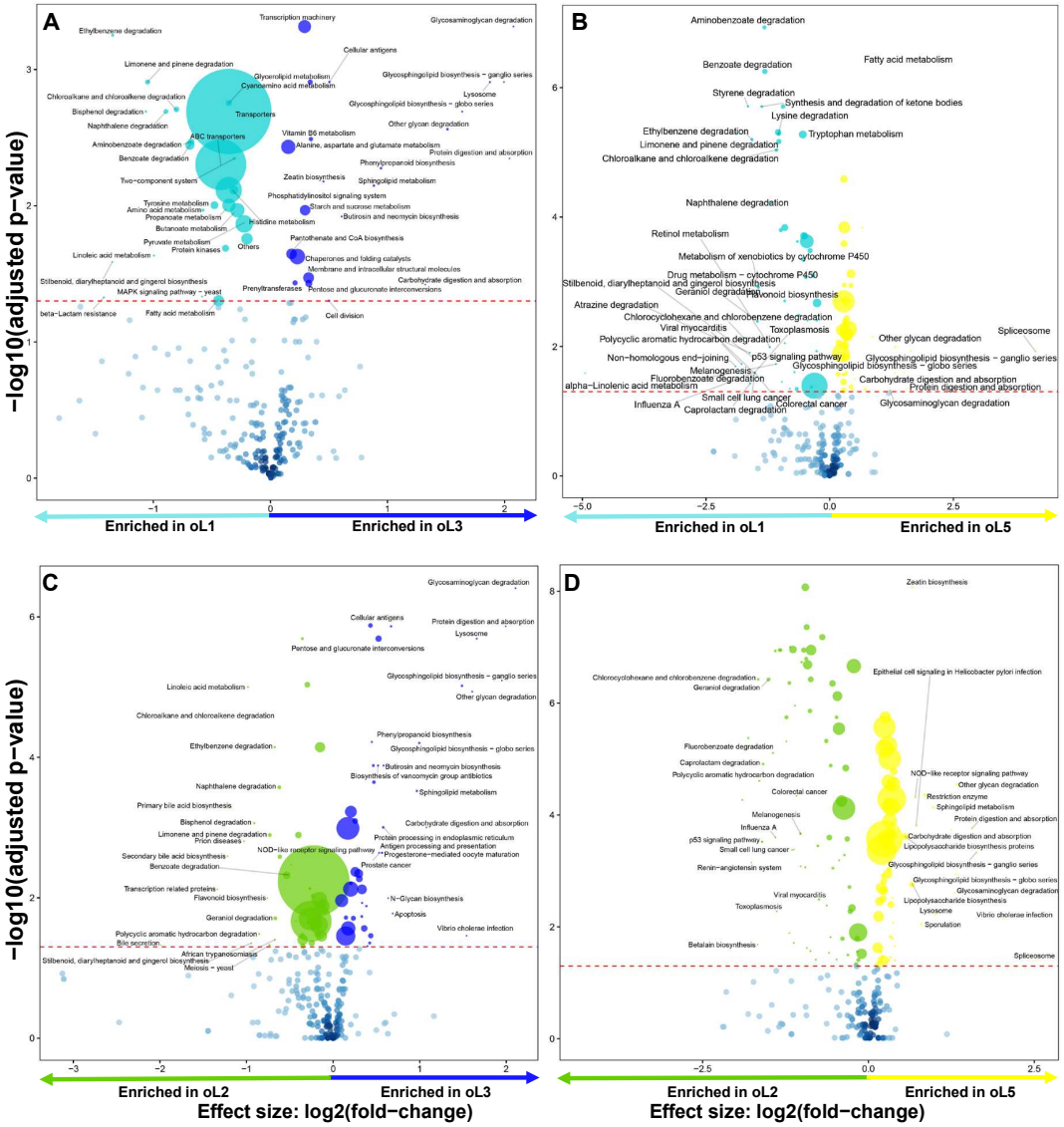


Figure S9 cont.

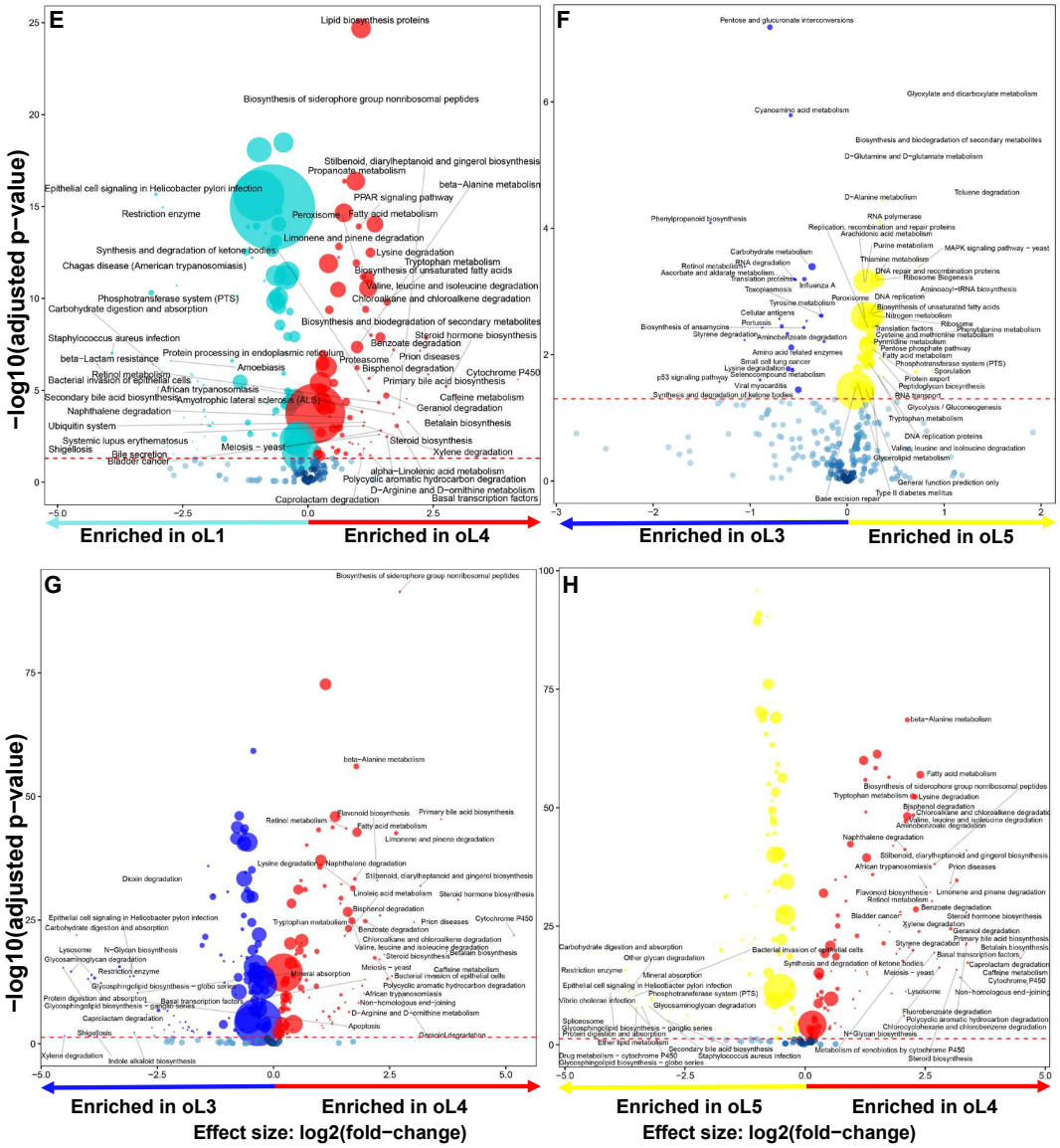
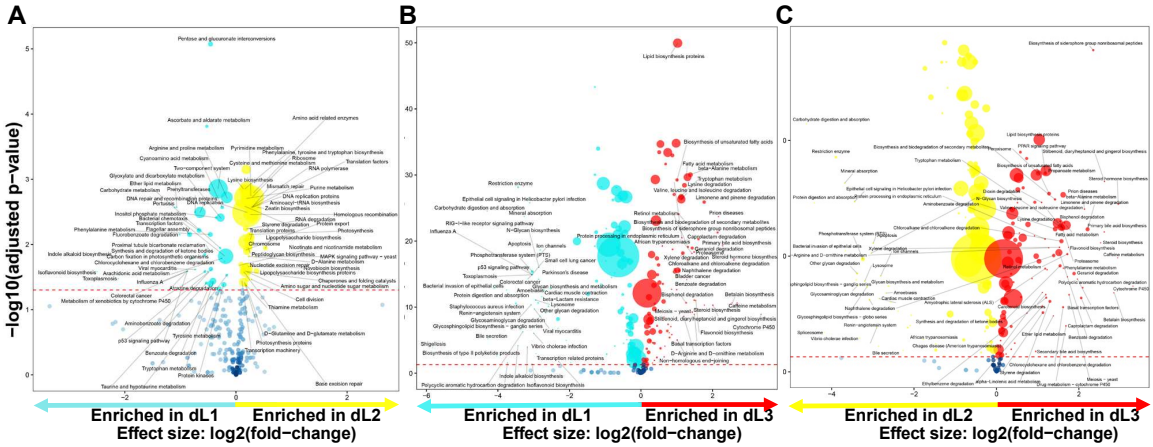


Figure S10: Inferred metagenomes of lymphotypes in dTBLs. Volcano plot depicting differentially enriched pathways in dL3 included pathways involving lipid biosynthesis, fatty acids, and SCFA metabolism i.e. lipid biosynthesis proteins, propanoate metabolism, benzoate degradation, and valine, leucine and isoleucine degradation. 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. dL: dTBL Lymphotype.



1
2

References

1. Minnies S, Reeve BW, Rockman L, et al. Xpert MTB/RIF Ultra Is Highly Sensitive for the Diagnosis of Tuberculosis Lymphadenitis in a High-HIV Setting. *Journal of clinical microbiology* 2021;59(12):e01316-21.
2. Wright CA, Bamford C, Prince Y, et al. Mycobacterial transport medium for routine culture of fine needle aspiration biopsies. *Archives of disease in childhood* 2010;95(1):48-50.
3. Holmes I, Harris K, Quince C. Dirichlet multinomial mixtures: generative models for microbial metagenomics. *PLoS ONE* 2012;7(2):e30126.