



Universiteit
Leiden
The Netherlands

The immune divide: factors influencing immune variation and differences in vaccine responses

Pyuza, J.J.

Citation

Pyuza, J. J. (2025, November 25). *The immune divide: factors influencing immune variation and differences in vaccine responses*. Retrieved from <https://hdl.handle.net/1887/4283867>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/4283867>

Note: To cite this publication please use the final published version (if applicable).

6

Chapter 6

Tanzanian gut microbiota profiles linked to high but rapidly waning yellow fever antibody titers

Jeremia J. Pyuza^{1,2,3,4,12*}, Marloes M.A.R. van Dorst^{1,12*}, David Barnett⁵, Koen Stam¹, Mikhael Manurung¹, Linda Wammes¹, Marion König¹, Yvonne Kruize¹, Nikuntufya Andongolile⁶, Anastazia Ngowi⁶, Elichilia R. Shao^{7,8}, Vesla I. Kullaya^{4,9}, Alex Mremi^{2,10}, Pancras C.W. Hogendoorn¹¹, Sia E. Msuya^{3,6}, Simon P. Jochems¹, John Penders⁵, Maria Yazdanbakhsh^{1,13}, Wouter A.A. de Steenhuijsen Piters^{1,13}.

* Contributed equally

¹² Shared first authors

¹³ Shared last authors

* Corresponding author

Published: NPJ Biofilms Microbiomes. 2025; doi:10.1038/s41522-025-00687-w

Affiliations

¹Leiden University Center for Infectious Diseases (LUCID), Leiden University Medical Center, 2333 ZA Leiden, Netherlands

²Department of Pathology, Kilimanjaro Christian Medical Centre, Moshi, Tanzania

³Institute of Public Health, Kilimanjaro Christian University Medical College (KCMUCo), Moshi, Tanzania

⁴Kilimanjaro Clinical Research Institute (KCRI), Kilimanjaro Christian Medical Centre, Moshi, Tanzania

⁵Department of Medical Microbiology, Infectious Diseases and Infection Prevention, NUTRIM Research Institute on Nutrition and Translational Research in Metabolism, Maastricht University Medical Center, Maastricht, The Netherlands

⁶Department of Community Medicine, Kilimanjaro Christian Medical Centre (KCMC), Moshi, Tanzania

⁷Department of Internal Medicine, Kilimanjaro Christian Medical University College (KCMUCo), Moshi, Tanzania

⁸Department of Internal Medicine, Kilimanjaro Christian Medical Centre (KCMC), Moshi Tanzania

⁹Department of Medical Biochemistry and Molecular Biology, Kilimanjaro Christian Medical College (KCMUCo), Moshi, Tanzania

¹⁰Department of Pathology, Kilimanjaro Christian Medical University College (KCMUCo), Moshi, Tanzania

¹¹Department of Pathology, Leiden University Medical Center, Leiden The Netherlands

Abstract

Vaccine responses vary across populations and are influenced by numerous intrinsic and extrinsic factors, including the gut microbiota. However, studies linking microbiota composition to vaccine immunogenicity in low- and middle-income countries are sparse. In this study, we examined the gut microbiota of 143 healthy rural and urban living Tanzanians who participated in a yellow fever vaccine (YF-17D) trial. We found significant differences in gut microbiota profiles between rural and urban participants. Rural-associated microbiota showed higher diversity and enrichment of taxa like *Prevotella* and *Succinivibrio*, which were linked to dietary intake patterns. Yellow fever neutralizing antibody titers were higher in rural compared to urban participants. Interestingly, a subset of urban individuals with a rural-like microbiota had higher antibody titers and faster waning than those with a more industrialized microbiota. These findings suggest that gut microbiota composition might be linked to vaccine immunogenicity, potentially outweighing the influence of living location.

Keywords: Yellow fever vaccine, YF-17D, microbiota, 16S-rRNA gene sequencing, neutralizing antibody, rural/urban differences, Tanzania.

Introduction

Vaccines play a critical role in the prevention of infectious diseases, especially in high-burden populations[1-5]. However, vaccine-induced immune responses vary across populations. Lower vaccine immunogenicity and efficacy ('hypo-responsiveness') is mainly observed in rural areas in low- and middle-income countries (LMICs), especially when compared to high-income countries (HICs) or to urban areas within the same countries[6-9]. Vaccine responses are impacted by a range of extrinsic and intrinsic factors, including host (age, sex, genetics, co-morbidities), behavioral (e.g. smoking) and environmental (rural/urban living location, season) factors, diet and nutrition, pre-existing immunity and vaccine factors (vaccine type, adjuvants, dose and administration route)[10-12]. Evidence suggests that the gut microbiome influences immune system development and regulation, thereby affecting vaccine responses[13-19]. Like vaccine responsiveness, the gut microbiota is highly variable across age, geographical locations and between people of different lifestyles[20-23]. Broad differences in gut microbiota composition and diversity have been observed between individuals living in rural and urban areas in LMICs[23-27]. Generally, these rural living individuals show a predominance of bacteria capable of polysaccharide degradation and fermentation, including *Prevotella* and other commensal bacteria such as the spirochaete *Treponema succifaciens*, which is enriched in non-industrialized populations[20,23,24,28,29]. In contrast, urban living individuals, who generally consume more processed foods and refined sugars, have a less diverse gut microbiota characterized by high abundance of *Bacteroides*[23,26,30,31].

Gut microbiota variation has been linked to both humoral and cellular vaccine responses⁶ with the strongest associations found in mice receiving non-adjuvanted vaccines (e.g. non-adjuvanted influenza vaccine and inactivated polio vaccine)[32,33]. The mechanisms underlying these associations include the production of immunomodulatory metabolites such as short-chain fatty acids[6,33]. In addition, it has been shown that microbial ligands from the microbiome may serve as natural vaccine adjuvants³⁴. Indeed, studies in knockout mice have shown that innate sensing of

bacterial flagellin in the gut microbiome can provide adjuvant signals enhancing the antibody response to non-adjuvanted vaccines such as trivalent influenza vaccines (TIV) and polio vaccine[32]. The role of the microbiome in live-attenuated (self-adjuvanted) vaccine responses remains contentious. While studies on the yellow fever vaccine show that antibody responses in antibiotic-treated mice are similar to those in non-treated mice, suggesting minimal microbiome influence[32,34], there is also research indicating that the microbiome may impact responses to other live-attenuated vaccines, like BCG[35-37]. These findings so far suggest variation in associations between the host microbiome and the response to specific vaccines. In humans, similar in-human antibiotic intervention studies have not been performed for self-adjuvanted vaccines. Therefore, it is currently unknown whether the microbiota plays a role in potentiating the immune responses to self-adjuvanted vaccines, such as yellow fever vaccine and are generally considered prone to developing poor vaccine responses. Given that the gut microbiota can be modulated through diet and the administration of pre- or probiotics, it may pose an interesting target for future strategies to enhance vaccine responses in these vulnerable populations.

We hypothesize that the gut microbiota profiles in rural and urban living Tanzanians are linked to vaccine responsiveness. To study this, we recruited healthy Tanzanian adults and vaccinated them with yellow fever vaccine (YF-17D). Stool and blood samples were collected over time. In line with previous work, we demonstrate that the gut microbiota composition is profoundly different between rural and urban living individuals and is linked to dietary habits. Variation in yellow fever neutralizing antibody responses may be at least partly explained by differences in microbiota community state types, potentially outweighing the impact of living location.

Results

Baseline characteristics of the study population

We enrolled 185 individuals living in rural and urban locations in Moshi, Northern Tanzania (**Figure 1a-b** and **Supplementary Figure 1**). Individuals were randomized into a vaccination group, including those who received the yellow fever vaccine (YF-

17D; $n = 155$) or a non-vaccinated control group ($n = 30$). All individuals were followed over six months and demographics and lifestyle variables (housing, assets and food history) were collected (**Supplementary Table 1** and **Supplementary Table 2**, **Supplementary Figure 2**). Plasma and/or stool samples were collected at baseline, day 28, 56 and 178 (**Figure 1c**).

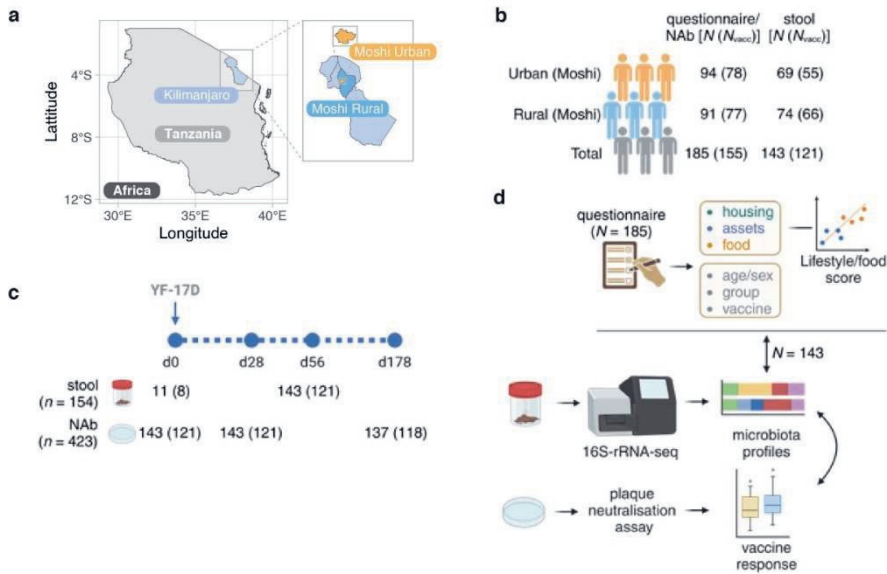


Figure 1 | Study overview.

a) Geographic map of study sites in Tanzania (Moshi Urban and Moshi Rural), within the Kilimanjaro region. b) Graphical representation of the number of urban and rural living participants included in the study and of those from whom stool samples for microbiota assays were available, stratified by the total number of participants (N) and the number of individuals who received the yellow fever (YF-17D) vaccine (N_{vacc}). c) Study design and number of samples at each time point. Numbers depicted as n (n_{vacc}), i.e. total number of samples (number of samples from vaccinated individuals). The number of samples in which neutralizing antibodies were measured only includes those samples with a matching stool sample. d) Graphical summary of study analyses. Questionnaire data from all included individuals ($N = 185$) were used in multiple correspondence analyses (MCA)/principal component analyses (PCA) to derive lifestyle/food scores. The associations between microbiota profiles and demographics and lifestyle factors (summarized in scores) and vaccine responses were assessed. Stool samples were available for microbiota sequencing from 143 of 185 (77.3%) individuals ($n = 74$ rural and 69 urban) (**Table 1**). A total of 154 stool samples were collected from 143 individuals (paired samples before and after vaccination were available from 11 individuals). The median age was 23.1 years (interquartile range [IQR], 21.2-27.1 years) and 43% were female (35% vs 52% in rural

and urban locations, respectively, $p = 0.040$). Baseline characteristics of the microbiota cohort were similar to the overall cohort (**Table 1** and **Supplementary Table 1**). The prevalence of parasitic infections was 4.9% and these infections were detected only in individuals from rural areas (**Table 1**).

Table 1 | Baseline characteristics of the study population ($N = 143$).

Variable	Overall, $N = 143$	Urban Moshi, $N = 69$	Rural Moshi, $N = 74$	p-value
Female sex	62 (43%)	36 (52%)	26 (35%)	0.040
Age	23.1 (21.2, 27.1)	23.0 (21.2, 24.6)	24.0 (21.2, 29.7)	0.209
Age categories				<0.001
18-25	94 (66%)	55 (80%)	39 (53%)	
26-36	49 (34%)	14 (20%)	35 (47%)	
BMI	22.5 (19.8, 25.1)	23.1 (20.4, 26.4)	22.2 (19.8, 24.4)	0.207
BMI classification				0.486
<18.5	16 (11%)	6 (8.7%)	10 (14%)	
18.5-24.9	91 (64%)	42 (61%)	49 (66%)	
25.0-29.9	27 (19%)	16 (23%)	11 (15%)	
>30	9 (6.3%)	5 (7.2%)	4 (5.4%)	
Systolic blood pressure (mmHg)	117 (106, 124)	110 (102, 120)	120 (110, 129)	0.003
Diastolic blood pressure (mmHg)	70 (65, 78)	70 (63, 76)	72 (67, 79)	0.099
Hemoglobin level g/dl	15.00 (13.60, 16.20)	14.30 (12.60, 15.60)	15.65 (14.20, 16.50)	<0.001
Random blood sugar, mmol- ¹	5.20 (4.80, 5.80)	5.30 (4.90, 6.00)	4.95 (4.70, 5.50)	0.008
Highest level of education				<0.001
Primary	43 (30%)	2 (2.9%)	41 (55%)	
Secondary	69 (48%)	38 (55%)	31 (42%)	
College	15 (10%)	14 (20%)	1 (1.4%)	
University	16 (11%)	15 (22%)	1 (1.4%)	
Helminth infection ^a	7 (4.9%)	0 (0%)	7 (9.5%)	0.014
Insurance status	63 (44%)	60 (87%)	3 (4.1%)	<0.001
Occupation				<0.001
Farming	28 (20%)	10 (14%)	18 (24%)	
Elementary occupation	50 (35%)	7 (10%)	43 (58%)	
Employed/business owner	58 (41%)	49 (71%)	9 (12%)	
Other	7 (4.9%)	3 (4.3%)	4 (5.4%)	
Received yellow fever vaccine	121 (85%)	55 (80%)	66 (89%)	0.116

$N = 143$ participants. Values represent number of participants (percentage of total) and median (interquartile range [IQR]) for categorical and continuous variables, respectively. Comparisons between locations were performed using Fisher's exact, chi-squared and Mann-Whitney U-test for categorical and continuous variables, respectively. ^a Stool was tested for helminths using the Kato-Katz method, testing for *Schistosoma mansoni*, *Ascaris lumbricoides*, hookworm and *Trichuris trichuria*. Additionally, urine was tested for *Schistosoma haematobium* and *Schistosoma mansoni* using the POC-CCA method.

Lifestyle and food scores vary with living location

Lifestyle questionnaire data were combined into a single lifestyle score to objectively gauge rural/urban living location-associated lifestyles ($N = 185$). To obtain the lifestyle score, we applied multiple correspondence analysis (MCA, a dimensionality reduction method for categorical data) to 38 questions (118 variable categories; **Supplementary Figure 3a**). MCA separated individuals based on living location, especially across principal component (PC) 1 ('lifestyle score', **Supplementary Figure 3b**), which captured 12.1% of the variation in questionnaire data. Rural individuals showed a larger spread across both PC1 and PC2, indicating they exhibit more diverse lifestyles. All variable categories contributed to the lifestyle score (**Supplementary Figure 3c**), with variables related to possession of assets showing the highest cumulative contribution (57.8%). Variable categories most related to higher lifestyle score (associated with living in urban areas) included possession of household assets (e.g. working television, iron, watch, computer, refrigerator, radio, car/truck or computer). Variables related to housing quality, including the presence of a pit latrine, floors made of soil or sand and walls made of cane/palm/trunks/bamboo contributed to low lifestyle score (related to living in rural areas) (**Supplementary Figure 3d-e**). PC2 scores were additionally driven by livestock-associated variables (**Supplementary Figure 3f-g**). Lifestyle scores (PC1) were similar for the overall ($N = 185$) and microbiota ($N = 143$) cohort (**Supplementary Figure 3h**).

Given the known association between diet and microbiome[38], we additionally developed a food score, based on 11 questions on frequency of consumption of specific food per week (ordinal variables). Using principal component analysis (PCA), variation in food consumption was captured across PC1 ('food score'), again showing a clear separation between rural and urban groups (**Supplementary Figure 4a-b**). In contrast, PC2 covaried with the average reporting frequency across all food variables (**Supplementary Figure 4c**), indicating interindividual differences in how the food questionnaire was filled out. Logistic regression analysis indicated that frequent consumption of fish, locally brewed beer (composed of fermented banana and millet) and green vegetables ($p \leq 0.009$) was significantly associated with rural living, whereas a carbohydrate-rich diet consisting of rice and potatoes was related to living in urban

areas (adjusted for all other food variables and sex, $p \leq 9.0 \times 10^{-4}$; **Supplementary Figure 4d**). Both lifestyle and food score were included in microbiota association tests ($N = 143$; **Figure 1d**).

Characteristics of microbiota data

To characterize the microbiota profiles of rural and urban living individuals, stool samples were subjected to Illumina MiSeq sequencing of V3-V4-region of the 16S-rRNA-gene, resulting in a median of 93,471 reads (range 58,482-124,868) per sample. A median of 68,648 reads (38,118-93,731) per sample remained after bioinformatic processing and quality filtering. We identified 3,354 amplicon sequence variants (ASVs) that occurred in ≥ 2 samples. ASVs could be aggregated into 272 genus-level taxa, 25 of which were part of the ‘core microbiota’ (genera detected in 80% of samples at $\geq 0.1\%$ relative abundance). Highly abundant genera included *Prevotella* (mean relative abundance 27.6%), *Bacteroides* (7.9%), *Faecalibacterium* (7.7%), *Lachnospiraceae* (4.2%), *Oscillospiraceae* (3.5%), *Blautia* (2.6%) and *Succinivibrio* (2.4%). All microbiota analyses are conducted on samples collected at day 56, if not mentioned otherwise.

Microbiota profiles differ between rural and urban living Tanzanian adults

First, we investigated differences in microbiota profiles between rural and urban living Tanzanian adults. The number of observed ASVs was higher in rural compared to urban living Tanzanians (linear model, adjusted for vaccination status, age, sex and sequencing depth, $\beta = 89.0$, $p = 1.6 \times 10^{-8}$). Shannon diversity, a measure of within-sample microbial diversity, was higher in rural compared to urban living adults ($\beta = 0.368$, $p = 1.1 \times 10^{-6}$, respectively; **Figure 2a**). Similarly, a higher number of observed genera was detected in rural compared to urban living individuals ($\beta = 13.2$, $p = 1.0 \times 10^{-4}$), yet genus-level Shannon diversity did not show differences between groups ($\beta = -0.019$, $p = 0.789$; **Figure 2b**).

Next, we investigated the global differences in gut microbial community structure between rural and urban living Tanzanians. Principal component analysis (PCA on a centered-log-ratio [CLR]-transformed genus-level abundance matrix) indicated profound differences linked to rural/urban living location (**Figure 2c**), which was confirmed by Permutational Multivariate Analysis of Variance (PERMANOVA; adjusting for vaccination, age and sex; $R^2 = 9.0\%$, $p < 0.001$). Other important factors driving overall microbial community variation included lifestyle score (PC1 housing, assets and food-related questionnaire data), food score (PC1 food-related questionnaire data), highest level of education, PC2 (lifestyle) and sex (PERMANOVA, $R^2 = 7.9\%$, 6.6% , 6.3% , 2.2% and 1.6% , respectively, $p \leq 0.005$), which are all at least partly collinear with rural/urban living location, which was therefore not considered in our models (**Table 1** and **Supplementary Table 1**, **Supplementary Figure 3b** and **Supplementary Figure 4b**). No statistically significant association between helminth

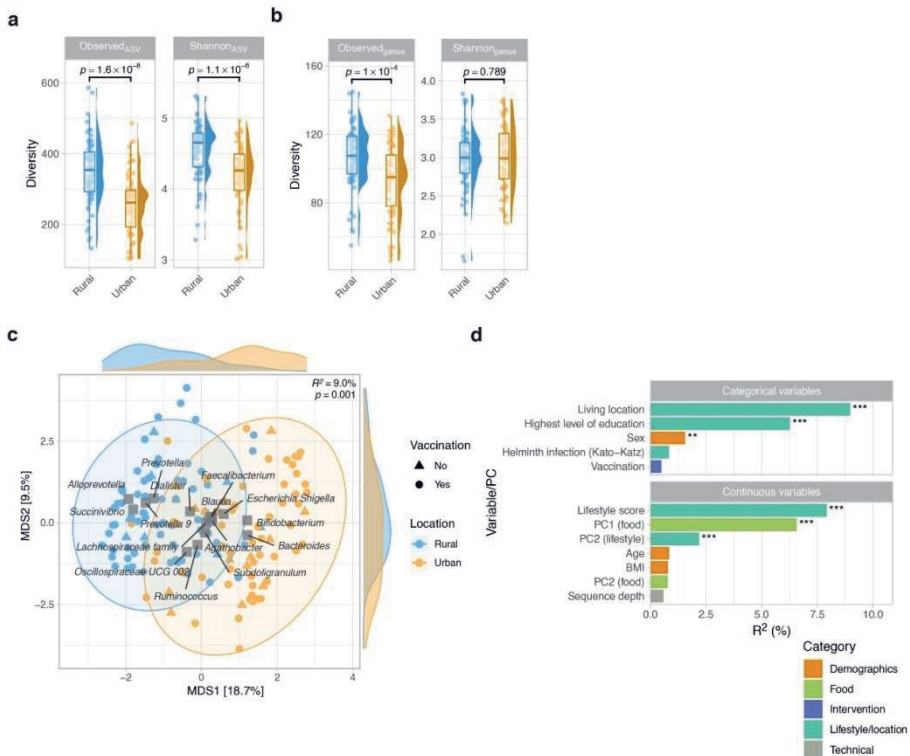


Figure 2 | Microbiota profiles differ across Tanzanians living in rural and urban regions.

a-b) ASV-level (a) and genus-level (b) Shannon diversity between study groups. Box plots represent the 25th and 75th percentiles (lower and upper boundaries of boxes, respectively), the median (middle horizontal line), and measurements that fall within 1.5 times the interquartile range (IQR; distance between 25th and 75th percentiles; whiskers). Density plots were used to visualize the distribution of data points. Statistical significance between groups was assessed using linear models with observed richness or Shannon diversity as outcome, adjusting for vaccination status, age, sex and sequencing depth. c) Principal component analysis (PCA) biplot using CLR-transformed genus-level microbiota features across day 56 samples. Percentages in square brackets denote the total variance explained by the first two principal coordinates. Each data point indicates a stool microbiota sample colored by group (rural/urban living). Ellipses reflect the data spread at a 95% confidence level. Density plots show the distribution of MDS1 (x-axis) and MDS2 (y-axis) score. The 15 highest ranking genera across all day 56 samples were simultaneously visualized (squares). R^2 and statistical significance of the association between group and the overall microbiota composition was assessed using PERMANOVA-test (1,000 permutations), while adjusting for vaccination status, age and sex. d) Bar plots indicating the effect size (R^2) of the association between demographic, technical, lifestyle or food-related variables (or derived scores based on these variables; see Supplementary Figure 3 and Supplementary Figure 4) and the overall microbiota composition. Statistical significance was assessed using PERMANOVA-tests. Each variable was tested separately. All analyses were adjusted for living location (except when assessing the impact of lifestyle/food-related variables, helminth infection status or sex), vaccination status, age and sex. Differential abundance analyses were primarily performed at genus level (MaAsLin2), testing all genera present at >0.1% abundance in $\geq 10\%$ of samples. We found higher abundance of 34 genera in rural living individuals compared to 14 genera enriched in those living in an urban environment (109 genera tested; $p_{adj} < 0.05$ and $\log_2$ -transformed fold change (FC) > 1.5; **Figure 3a**).

infection status and microbiota composition was detected ($R^2 = 0.8\%$, $p = 0.184$). Importantly, no association with vaccination status was detected (PERMANOVA, adjusting for living location, age and sex; $R^2 = 0.5\%$, $p = 0.787$), suggesting that the vaccine had no impact on the gut microbiota composition at day 56 post-vaccination. In addition, no statistically significant effects for age, BMI or sequencing depth were observed (PERMANOVA, adjusting for living location, vaccination age and sex, as appropriate; $R^2 = 0.8\%$, 0.8% and 0.6% , respectively, $p \geq 0.122$; **Figure 2d**).

Specific taxa have previously been associated with industrialized and non-industrialized populations and are referred to as BloSSUM (bloom or selected in societies of urbanization/modernization) and VANISH (volatile and/or associated negatively with industrialized societies of humans) taxa respectively²². We found that urban-associated genera (15/109 genera tested) were significantly enriched for

BloSSUM taxa (9/14 urban-associated genera, compared to 6/95 non-urban-associated genera, Fisher's Exact test $p = 1.8 \times 10^{-6}$; **Figure 3b** and **2c**), whereas rural living individuals were specifically devoid of these genera (1/34 compared to 14/75, $p = 0.034$; **Figure 3d**). In contrast, genera associated with rural living location were not significantly enriched for VANISH taxa (total 18/109; 8/34 rural-associated genera, compared to 10/75 non-rural-associated genera, Fisher's Exact test $p = 0.264$; **Figure 3e**).

For rural living individuals the strongest enriched genus was *Succinivibrio* ($\log_2\text{FC} = 6.76$, $p_{\text{adj}} = 1.3 \times 10^{-12}$). Together with *Treponema* ($\log_2\text{FC} = 1.62$, $p_{\text{adj}} = 0.061$), these genera showed a clear multimodal distribution within rural living individuals, with maximum relative abundance peaks detected at 8.6% and 7.0% for *Succinivibrio* and *Treponema*, respectively (*post-hoc* analysis, **Figure 3f-g**). Six genera belonging to the family of *Prevotellaceae* were strongly associated with rural living, including *Prevotella* (7/9) and *Alloprevotella* ($\log_2\text{FC} \geq 3.92$, $p_{\text{adj}} \leq 7.7 \times 10^{-5}$). Similarly, five genera belonging to the family of *Lachnospiraceae* were enriched in rural living individuals including *Butyrivibrio*, *Eubacterium ruminatum* and *Ruminococcus torques* ($\log_2\text{FC} \geq 1.78$, $p_{\text{adj}} \leq 3.3 \times 10^{-4}$). Other highly significantly enriched genera in rural individuals included *Fournierella*, *Holdemanella*, *Solobacterium*, *Sutterella*, *Anaeroplasma* and *Catenibacterium* ($p_{\text{adj}} \leq 3.4 \times 10^{-5}$; **Figure 3a** and **Supplementary Figure 5**).

In urban living individuals, abundance of *Bacteroides* and *Parabacteroides* was significantly higher compared to rural living ($\log_2\text{FC} = -3.70$ and -2.30 , $p_{\text{adj}} = 8.9 \times 10^{-12}$ and 3.78×10^{-7} , respectively). *Bacteroides* showed a multimodal distribution among individuals living at an urban location with two major peaks at a relative abundance of 1.2% and at 24.2% (**Figure 3g**). Other highly urban-associated genera included *Alistipes*, *Parasutterella*, *Bifidobacterium*, *Odoribacter*, *Bilophila* and *Akkermansia* ($p_{\text{adj}} \leq 2.8 \times 10^{-6}$). The gut microbiome of urban living individuals was further enriched for genera belonging to the family of *Enterobacteriaceae*, including *Enterobacter*, *Klebsiella* and *Escherichia/Shigella* ($p_{\text{adj}} \leq 0.009$; **Figure 3a** and **Supplementary Figure 5**). Taken together, rural living individuals lacked BloSSUM taxa and had

higher abundance of VANISH taxa such as *Prevotella* and *Succinivibrio*, whereas urban individuals were enriched with BloSSUM taxa like *Bacteroides*.

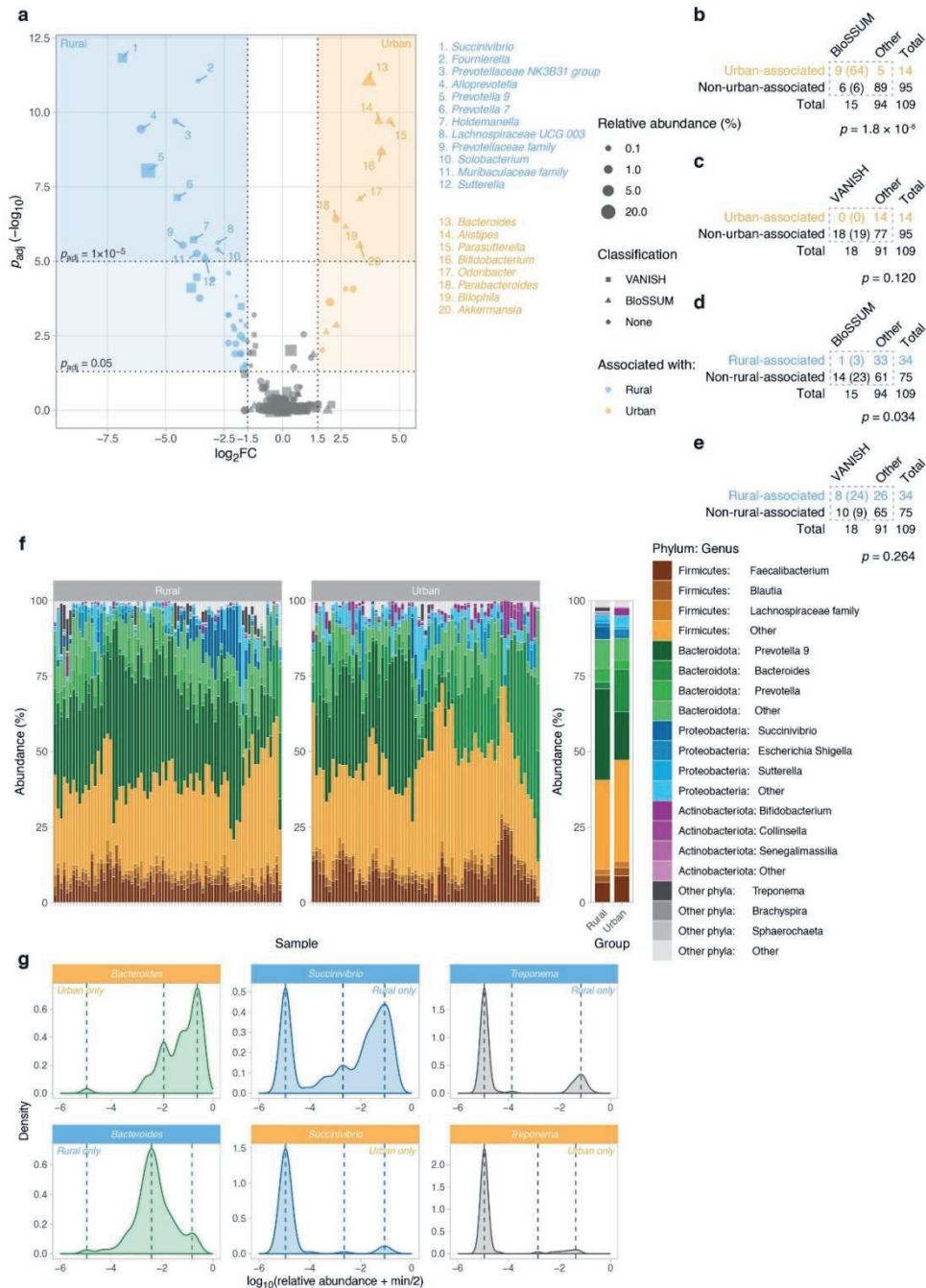


Figure 3 | Differential abundance analyses indicate specific taxa are related to urban/rural living.

a) Volcano plot visualizing genera enriched in rural or urban living individuals. Statistical significance between groups (rural/urban living) was assessed using linear models with genus-level abundance as outcome, adjusting for vaccination status, age and sex. Genera with $p_{adj} < 0.05$ and $\log_2$ -fold change (FC) < -1.5 or > 1.5 are colored and genera with a $p_{adj} < 1 \times 10^{-5}$ and $\log_2$ FC < -1.5 or > 1.5 were annotated. The shape of the data points indicates whether these genera were previously considered to belong to VANISH (volatile and/or associated negatively with industrialized societies of humans) or BloSSUM (bloom or selected in societies of urbanization/modernization) taxa. b-e 2×2 tables showing the number of BloSSUM (b and d) or VANISH taxa (c and e) across urban- and rural-associated taxa. The number between brackets indicates the percentage of VANISH/BloSSUM/other taxa out of the total number of (non-)-rural-/urban-associated taxa. Statistical significance was assessed using Fisher's Exact tests. f) Stacked bar chart showing the genus-level taxonomic composition of each sample as relative abundances. Bar colors are based on a hierarchical color palette with hues specified by Phylum and shades specified by Genus. Samples are arranged based on Bray-Curtis dissimilarities and the plot is divided by living location. Right of the per-sample bar chart, a bar chart showing mean relative abundance for each taxon per group is shown. g) Density plots showing the relative abundance distribution of *Bacteroides*, *Succinivibrio* and *Treponema* within rural or urban living individuals. Modes are indicated by dashed lines.

Rural and urban-associated differences in diet are linked to microbiota profiles

Given the multimodal distribution of the relative abundance of specific genera not only between rural and urban living individuals, but also within individuals from the same location (**Figure 3g**), we next assessed the impact of dietary history on microbiota composition. To study this, we clustered samples into two Community State Types (CSTs; **Supplementary Figure 6a** and **Supplementary Figure 6b**) using Dirichlet-multinomial modelling (DMM). CST1 was characterized by a (non-significant) enrichment of VANISH bacteria (25% vs 12%, Fisher's Exact test $p = 0.106$), including *Prevotella*, *Alloprevotella*, *Holdemanella* and *Succinivibrio*, whereas CST2 was strongly enriched for BloSSUM taxa (67% vs 7%, Fisher's Exact test $p = 6.5 \times 10^{-6}$), like *Bacteroides*, *Alistipes*, *Parasutterella* and *Bifidobacterium* (**Figure 4a** and **Supplementary Figure 6c-f**).

Although the majority of samples from rural living individuals clustered in CST1 (93.2%), urban living individuals were split across CST1 (30.4%) and CST2 (69.6%). Baseline characteristics between urban living individuals with CST1 (referred to as 'rural-like urbanites') and urban living individuals with CST2 did not differ, except that

a higher proportion of rural-like urbanites was male (71.5% vs 37.5%, $p = 0.009$; **Supplementary Table 3**). Comparing rural-like urbanites to urban living individuals revealed highly divergent profiles, with differences similar to those observed when assessing CST1 compared to CST2. Conversely, rural-like urbanites showed much more subtle differences compared to rural living individuals, with enrichment of *Haemophilus* and *Akkermansia*, but lack of rural-associated *Succinivibrio* and *Fournierella* (109 genera tested; $p_{adj} < 0.05$ and $\log_2$ -transformed fold change (FC) > 1.5; **Supplementary Figure 7**).

We did not observe differences in either lifestyle or food score between rural-like urban or urban individuals (**Figure 4b**). However, individual food variables did show differences, indicating that rural-like urbanites less frequently consume rice (logistic regression analyses adjusted for all other food variables and sex, $\beta = -1.95$, $p = 0.032$) and more beans and/or peas ($\beta = 0.85$, $p = 0.078$) compared to urbanites belonging to CST2 (i.e. non-rural-like urbanites). Regardless, rural-like urbanites still more frequently consumed rice compared to rural living individuals ($\beta = 2.20$, $p = 0.033$). Furthermore, rural-like urbanites ate significantly less ugali (maize stiff porridge) compared to rural living individuals ($\beta = -3.38$, $p = 0.023$), yet ugali consumption was no different from other urban living individuals (CST2) ($\beta = -0.73$, $p = 0.212$). Last, we found that rural-like urban individuals less frequently eat potatoes compared to both urban living individuals (CST2; $\beta = -0.815$, $p = 0.087$) and rural living individuals ($\beta = -2.02$, $p = 0.100$; **Figure 4c**). Together, these findings suggest urban living individuals less frequently consuming starch/carbohydrate-rich dietary products (potatoes and rice) and more fiber-rich products (beans and/or peas) harbor a microbiota reminiscent of that of rural living individuals, despite slight differences in dietary habits compared to this latter group.

Genus-level microbiota association tests (linear models including all food variables, vaccination, age and sex) were stratified by living location, given the strong collinearity between living location and dietary habits. Among rural living adults, we detected seven significant associations, suggesting enrichment of *Prevotella*, *Prevotellaceae* and *Eubacterium ventriosum* and lower abundance of *Odoribacter* with increased

consumption of locally brewed beer. *Odoribacter* abundance was also negatively associated with the consumption of bananas, as was *Alistipes* abundance (Figure 4d). Within urban living individuals, we detected five significant associations, indicating frequent ugali consumption is linked to the depletion of *Oscillibacter*, *Christensenellaceae*, *Clostridia*, *Eubacterium eligens* and *Eubacterium siraeum* (Figure 4e).

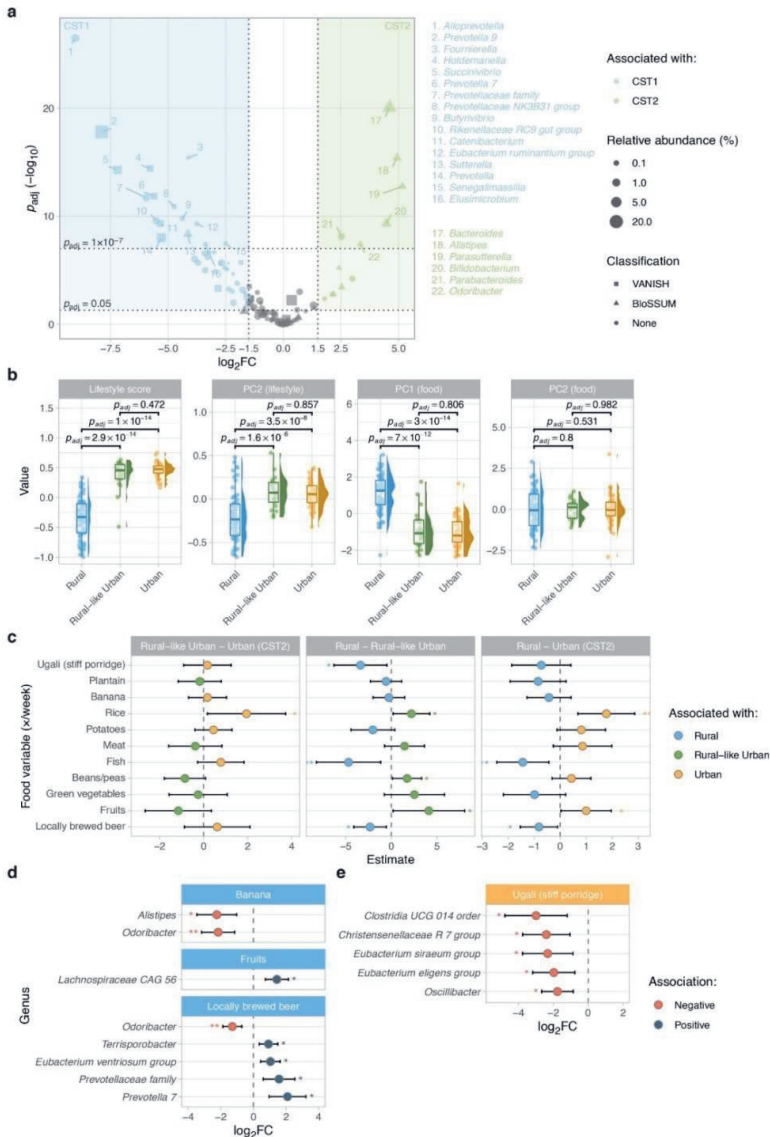


Figure 4 | Frequent consumption of carbohydrate-rich dietary products is associated with urbanization.

a) Volcano plot visualizing genera enriched in CST1 and CST2. Statistical significance between CSTs was assessed using linear models with genus-level abundance as outcome (no adjustment for covariates). Genera with $p_{adj} < 0.05$ and $\log_2$ -fold change (FC) < -1.5 or > 1.5 are colored and genera with a $p_{adj} < 1 \times 10^{-7}$ and $\log_2$ FC < -1.5 or > 1.5 were annotated. The shape of the data points indicates whether these genera were previously considered to belong to VANISH (volatile and/or associated negatively with industrialized societies of humans) or BloSSUM (bloom or selected in societies of urbanization/modernization) taxa. b) Differences in lifestyle scores (PC1/PC2; Supplementary Figure 3) and food scores (PC1/PC2; Supplementary Figure 4) between rural living individuals, rural-like urbanites (urban living individuals belonging to community state type [CST]1; Supplementary Figure 6) and urban living individuals (urban living individuals with CST2). See legend Figure 2a-b. Statistical significance between groups ($n = 3$) was assessed using linear models with lifestyle/food score as outcome, adjusting for vaccination status, age and sex. Pairwise contrasts were extracted and adjusted using Tukey's post-hoc test. c) Association between food variables and group (rural living individuals, rural-like urbanites (urban living + CST1) or urban living individuals with CST2). For each comparison (panel), a separate logistic regression model was fitted including all food variables and sex. Model estimates are depicted along the x-axis (colored points). Colors indicate the group with which a given food variable is positively associated. Whiskers denote 95% confidence intervals (CIs; Wald-method). d-e) Association between food variables and genera (present in $\geq 10\%$ of samples at $> 0.1\%$ abundance across day 56 samples; $n = 109$) within rural (d) and urban living individuals (e). Results were stratified by food variable and only food variables with any significant ($p < 0.05$) associations are shown. $\log_2$ fold change (FC) is shown along the x-axis. Whiskers denote 95% confidence intervals (CIs; Wald-method). Asterisks denote statistical significance (*, $p \leq 0.05$; **, $p \leq 0.01$; ***, $p \leq 0.001$).

Rural-like microbiota profile is linked to yellow fever neutralizing antibody titers

Next, we investigated whether vaccine immunogenicity differed between rural and urban living individuals. Across all vaccinated individuals ($N = 155$), no significant differences in yellow fever Plaque Reduction Neutralization Test (PRNT) (PRNT₅₀ and PRNT₉₀) were found between rural and urban living individuals at 4 weeks post-vaccination (generalized linear mixed effects regression [GLMER], adjusted for age and sex, $p = 0.161$ and $p = 0.226$, respectively, **Supplementary Figure 8a-b**). However, for the microbiota sub-cohort (i.e. with stool sample available, $N = 121$), yellow fever neutralizing antibodies (PRNT₅₀) at 4 weeks post-vaccination were higher in rural (geometric mean titer [GMT], 954 [95% CI, 742 – 1,226]) compared to urban living individuals (656 [95% CI, 490 – 880], $p = 0.042$). This was similar for PRNT₉₀ values

($p = 0.032$, **Figure 5b** and **Supplementary Figure 8d**). Antibody titers (PRNT₅₀) were lower at 6 months compared to 4 weeks post-vaccination for both urban (PRNT₅₀; $\beta = -1.16$, $p = 0.082$) and rural living individuals (PRNT₅₀; $\beta = -1.48$, $p = 0.040$), with urban living individuals showing a slightly stronger and statistically significant drop (**Figure 5b**). Similar results were found when considering all individuals ($N = 155$; **Supplementary Figure 8c**).

To assess whether microbiota profiles impact subsequent YF-17D-induced vaccine responses, we first ascertained that vaccination at day 0 had no impact on microbiota profiles collected at day 56. No significant difference in overall microbiota composition between day 56 samples from vaccinated compared to non-vaccinated subjects were found (**PERMANOVA-test, adjusted for living location, age and sex, $R^2 = 0.5\%$, $p = 0.787$; Figure 2d and Supplementary Figure 9a**). Also, paired day 0 and day 56 samples of vaccinated individuals did not show a consistent direction of movement (**Supplementary Figure 9b-c, PERMANOVA-test, $R^2 = 2.3\%$, $p = 0.422$**). Within- and between-subject distances between sample pairs and permutation tests across between-subject distances similarly indicated no statistically significant effect of vaccination (**Supplementary Figure 9d-e**). Therefore, day 56 samples were considered representative of day 0 (i.e. before vaccination) and were used as such for downstream analyses.

Next, yellow fever neutralizing antibodies (PRNT₅₀ at 4 weeks post-vaccination), were linked to Shannon diversity estimates, adjusting for living location, the interaction between living location and Shannon diversity, age and sex. Shannon diversity was negatively associated with antibody titers in rural-living individuals ($\beta = -0.345$, $p = 0.034$), but positively in urban-living individuals (**interaction term; $\beta = 0.496$, $p = 0.025$; Supplementary Figure 8e**). Similar results were found when considering PRNT₅₀ at 6 months post-vaccination and PRNT₉₀ titers.

Following, yellow fever neutralizing antibodies (PRNT₅₀ or PRNT₉₀) were related to CSTs, showing that CST1 (enriched in rural living individuals) is linked to higher antibody titers compared to CST2 (PRNT₅₀; $p = 0.005$, **Figure 5c and Supplementary Figure 8f [PRNT₉₀]**). Of note, this association showed large model estimates and was

highly significant, as opposed to initial comparisons between rural and urban living individuals, indicating microbial signatures better explain antibody variation than living location. Though antibody titers were higher 4 weeks post-vaccination, individuals with CST1 showed a much stronger and faster drop over time (**PRNT₅₀; $\beta = -1.96$, $p = 0.002$**) compared to CST2 (**$\beta = -0.02$, $p = 0.805$; Figure 5d**).

We found that within urban living individuals, there was a striking variation in antibody titers (particularly PRNT₅₀) linked to CSTs, with rural-like urbanites showing higher PRNT₅₀ compared to urban living individuals with a CST2-profile (**GMT 95%CI, 1,064 [723 – 1,565] vs 548 [381 – 787], $p = 0.022$, Figure 5e and Supplementary Figure 8g [PRNT₉₀]**). Furthermore, in rural-like urbanites, stronger waning was seen (**PRNT₅₀ at 6 months compared to 4 weeks post-vaccination; $\beta = -0.502$, $p = 2.5 \times 10^{-4}$**) compared to rural living individuals ($\beta = -0.116$, $p = 0.065$) and urbanites with CST2-profile (**$\beta = -0.015$, $p = 0.856$, Figure 5f**). Findings for PRNT₉₀ were similar but less outspoken (**Supplementary Figure 10**). Together, these findings indicate that microbiota profiles might be an important driver of yellow fever neutralizing antibodies in this population.

Differential abundance analyses at genus level ($n = 109$) revealed no significant associations between specific genera and yellow fever neutralizing antibodies (**PRNT₅₀ or PRNT₉₀**) at 4 weeks or 6 months post-vaccination (linear models, adjusting for either 1) living location, age and sex or 2) age and sex). This may indicate that the bacterial community in its totality, rather than single genera, relates to yellow fever vaccine immunogenicity.

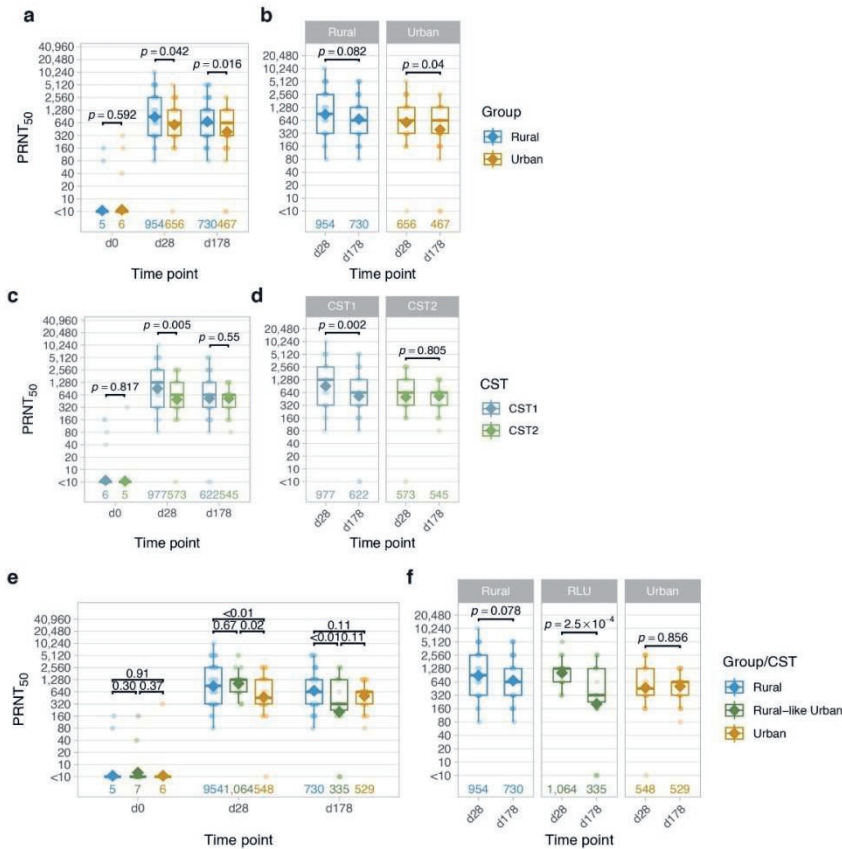


Figure 5 | Yellow fever neutralizing antibody titer variation across living location and with microbiota profiles ($N = 121$). a-f) Boxplots (see legend Figure 2a-b) showing yellow neutralizing antibody titers (PRNT₅₀) in vaccinated individuals from whom a stool sample was collected ($N = 121$) at baseline (day 0, d0), 4 weeks post-vaccination (day 28, d28) and 6 months post-vaccination (day 178, d178). a, c, e) Participants were compared according to group (rural/urban living location; a) CST (c) and group/CST (i.e. rural living individuals, rural-like urbanites [CST1] and urban living individuals [CST2]) (e). b, d, f) Boxplots showing yellow neutralizing antibody titers (PRNT₅₀) at 4 weeks post-vaccination (day 28, d28) compared to 6 months post-vaccination (day 178, d178) for group (rural/urban living location; b), CST (d) and group/CST (f). Diamond-shaped data points and color values at the bottom of the plot denote log₁₀-base geometric mean titers. Statistical analyses were performed using a generalized linear mixed effects model with log₁₀-transformed PRNT₅₀-values as outcome, group, CST or group/CST, time point, the interaction between group, CST or group/CST and time point, age and sex as fixed effects and participant ID as random effects. Pairwise comparisons of estimated marginal means between groups were computed at each time point (a, c and d) or between day 28 and day 178 for each group (b, d and f). Values under the limit of detection (<10) were assumed to have a value of 5.

Discussion

Here, we assessed the associations between rural or urban living location, gut microbiota profiles and yellow fever vaccine responses. We identified striking differences in microbiota community composition between rural and urban living individuals, which appeared in part related to dietary habits. Gut microbiota profile variation was linked to yellow fever vaccine immunogenicity and waning, at least at a microbiota community, but not a genus-level.

We show that rural living individuals harbor a gut microbiota enriched for *Succinivibrio*, *Treponema* and *Prevotella*, which is consistent with literature[20,26,28,39], and likely explained by variation in dietary habits. Higher abundance of these microbiota members has been associated with the digestion of plant-rich diets and the production of short-chain fatty acids[24,40,41]. Indeed, we found that Tanzanians living in rural settings consume more grain-based food products such as ugali (maize stiff porridge), vegetables and local beer made by a mix of fermented bananas and finger millet[26]. Strikingly, within those living in rural areas, local beer consumption in particular was associated with a higher abundance of *Prevotella*, suggesting it may be a particularly important driver of the rural microbiota signature.

Urban living individuals exhibited a higher abundance of *Bacteroides*, *Parabacteroides*, *Enterobacteriaceae* and *Bifidobacterium*, which may be related to the consumption of a more carbohydrate-rich diet, again largely aligning with literature[27,31,42]. Interestingly, approximately half of the urban living individuals showed a high abundance of *Prevotella*, which we identified as characteristic of the rural microbial signature. The microbiota composition of this subgroup of urban living individuals co-clustered with rural living individuals based on their microbial make-up. Despite that, these rural-like urbanites showed a different dietary consumption pattern from both rural living individuals and their urban counterparts, with high consumption of beans and peas (rich in proteins, carbohydrates and dietary fibers) and lower rates of potato and rice consumption. Within the urban population we found that frequent

consumption of ugali (maize stiff porridge) is related to a reduced abundance of *Eubacterium siraeum*. *Eubacterium siraeum* has previously been linked to the consumption of a Western diet and systemic inflammation[43], possibly indicating that ugali consumption limits the establishment of microbes related to industrialization.

Genera associated with rural or urban living location were classified as either BloSSUM or VANISH taxa. The distinction between these taxa has been based on previous work comparing the gut microbiome of a Tanzanian population of Hadza hunter-gatherers to that of populations living industrialized lifestyles[22]. Although rural living individuals in our study showed a lack of industrialized lifestyle-associated BloSSUM taxa, we did not observe a significant enrichment for VANISH taxa, suggesting that indeed the population of rural living individuals recruited in our study is possibly in transition between traditional and industrialized lifestyles. Conversely, a clear signature of Westernized microbiota profiles was detected in urban living individuals, indicated by significant enrichment of BloSSUM taxa. Apart from diet, there may be other host and environmental factors contributing to the observed differences in microbiota composition between rural and urban living individuals. Among others, agricultural activities, such as livestock farming[44,45], and environmental exposures, including indoor cooking, wood stove cooking or exposure to pollutants, may impact microbiota composition[46]. Helminth infections have previously been linked to microbiota changes[47-50], but since helminth infection prevalence was low in our rural cohort (4.3%), we presume this was not a main driver of rural urban differences in our study. We detected slightly higher yellow fever neutralizing antibody responses 4 weeks post-vaccination in participants living in rural compared to urban settings. The difference was statistically significant in the microbiota sub-cohort, but not in the total vaccinated cohort. At 6 months, titers remained higher in rural settings. This is not in line with the general hypothesis that vaccine immunogenicity in rural populations is lower than urban urban-living individuals, especially within LMICs, which has been observed for a wide range of vaccines[6,51]. Yellow fever vaccine, which is highly effective, has previously been used to model human immune responses to vaccines in general. As such, several studies assessed yellow fever vaccine immunogenicity, showing lower antibody titers and rapid

waning in Ugandan compared to Swiss individuals and lower seroprevalence rates in rural Ghanaian compared to urban Malian infants[52,53]. In both studies, the observed differences in immunogenicity are relatively small compared to other vaccines, such as for malaria[54,55], and may be obscured by between-country differences. Moreover, discordant results have been reported in some studies assessing factors associated with yellow fever vaccine immunogenicity[56,57]. Regardless, a recent study in Uganda (POPVAC)[58], did show higher yellow fever vaccine antibody titers in urban living individuals compared to individuals living in rural settings, which is in contrast with our findings and at least in part could be explained by the high helminth and malaria infection prevalence in the Ugandan study. Also, the rural sites sampled in our study are relatively developed compared to other rural sites in Tanzania and some other East African countries, with better access to health care and nutrition[59,60]. Last, rural individuals in our study live at high altitude, which induces a hypoxic state, as reflected by their high hemoglobin levels. We speculate this causes an increase of transcriptional factor hypoxia-inducible factor (HIF), which induces metabolic and phenotypic changes in B cells and boosts B cell differentiation and antigen switching, thus resulting in higher antibody titers in these rural individuals[61-64]. We found that especially within urban living individuals, a rural-like microbiota profile was related to relatively higher antibody titers, but stronger waning, as opposed to a more *Prevotella*-depleted, industrialized microbial community composition. Despite this, no significant associations were detected between specific genera and yellow fever vaccine-induced neutralizing antibody titers. This lack of genus-level associations could be explained by insufficient statistical power or limited sequencing resolution from 16S-rRNA-sequencing. Alternatively, it may suggest that the overall bacterial community, rather than individual genera, is linked to yellow fever vaccine immunogenicity. The microbiota has been suggested to potentiate vaccine responses by providing adjuvating signals, although this seems most relevant to vaccines other than yellow fever vaccine, which has endogenous adjuvants and is therefore able to engage with toll-like receptors (TLRs) and RIG-1 receptors itself³². Alternatively, we speculate higher yellow fever antibody titers observed in rural living individuals and rural-like urbanites could be related to increased consumption of fiber-rich diets which enriches for short-chain fatty acid (SCFA)-producing bacteria, such as *Prevotella*[40]. In B cells, SCFAs increase

acetyl-CoA and metabolic activity, resulting in increased antibody production[65]. Indeed, a study showed that mice fed with a diet low in dietary fibers had lower production of SCFAs and a reduced specific antibody response towards pathogens[65]. It is possible that the microbiota-associated increase in antibody titers observed in rural individuals in our study has been obscured in other rural populations by factors such as a high prevalence of helminth infections and other infections such as malaria.

Although the gut microbiota of rural and urban living individuals in LMICs has been described previously, there remains a significant gap in the literature regarding gut microbiota profiles from LMIC populations currently transitioning from traditional to industrialized lifestyles. Our work expands on this topic, but also furthers our understanding on microbiome-host associations in the context of yellow fever vaccination. Our study also has limitations, such as the use of 16S-rRNA sequencing rather than shotgun sequencing. Additionally, as this was primarily an observational study aimed at identifying associations between the gut microbiome and the yellow fever vaccine responses, we could not assess causal effects. Also, we cannot exclude the possibility that (unmeasured) microbiota-independent factors (e.g. altitude, genetic variation and historic microbial exposures) drive vaccine immunogenicity, although the observed variation within the urban subgroup advocates against this. Fungal and viral microbiota may covary with the bacterial microbiota and explain the residual variation in vaccine responses we identified. We lacked longitudinal stool samples for most participants, as most samples were collected post-vaccination, which may have influenced the findings. However, our extensive analyses, also leveraging the paired samples we had available, indicated no major impact of the vaccine on the gut microbiota. This warrants our assumption that the microbiota measured post-vaccination is reflective of the baseline microbiota in these individuals.

To further explore microbe-host interactions in the context of yellow fever vaccination, future studies could include measurements of the cellular immune responses, which have previously been shown to differ at least between HIC and LMIC-populations[52]. Taken together, we show strong differences in gut microbiota profiles between rural and urban living individuals, with part of urban living individuals currently

transitioning towards a more industrialized microbiota profile. Microbiota variation within urban living individuals was associated with moderate variation in yellow fever vaccine-induced neutralizing antibodies and antibody waning. These findings suggest that gut microbiota profiles may impact vaccine responsiveness in a vaccine- and context-dependent manner. Identifying adult subpopulations where microbiota influences vaccine responsiveness creates opportunities for research on microbiota-based interventions.

Methods

Study design

This is a prospective longitudinal cohort study (CapTan) in a healthy, 18-35-year-old Tanzanian population recruited in a rural and urban Moshi, Northern Tanzania. Volunteers were enrolled and randomized into a vaccine group receiving a single 0.5 mL intramuscular dose of the yellow fever vaccine (YF-17D, Sanofi-Pasteur) and a control group. Randomization occurred by allocating every sixth individual to the control group. No placebo was administered to the control group. Neither the volunteers nor study personnel were blinded to group allocation. Plasma samples were collected at baseline (pre-vaccination), day 28 (4 weeks post-vaccination), 56 (8 weeks post-vaccination) and 178 (~6 months post-vaccination). Paired stool samples were collected at baseline and on day 56 for 11 individuals. For 143 individuals, stools samples were collected on day 56 only. All questionnaires and clinical samples were collected by a well-trained study team consisting of medical doctors, nurses and laboratory scientists. All samples were processed according to established standard operational procedures and good clinical and laboratory practice principles.

The study protocol was approved by the Ethical Board of the Kilimanjaro Christian Medical University College (No. 2588) and by the Tanzania National Ethical Committee Board (NIMR/HQ/R.8a/Vol.IX/4089). The study was registered under The Pan African Clinical Trial Registry (PACTR) with trial number PACTR202405738173023 on 03 May 2024. Data were collected in REDCap, a cloud-

based electronic data collection system, with a server hosted at the Kilimanjaro Clinical Research Institute (KCRI) in Tanzania.

Description of study areas

The study was conducted in rural and urban Moshi located in the Kilimanjaro region (total population of 1.9 million (**Figure 1a**). The rural study area is at an elevation of 2,000-2,100 meters above sea level, while urban Moshi is at 700-950 meters above sea level. The district of Moshi City (urban Moshi) is the administrative, commercial and educational center of the Kilimanjaro region, having 331,733 inhabitants. Most people practice a Western lifestyle with good general sanitation and access to clean water. The main ethnicities are Chagga and Pare. Formal business is the main activity, followed by government and public employment, while a small proportion of people are involved in agricultural and entrepreneurial activities. Rural Moshi has about 535,803 inhabitants who are mainly involved in farming activities. Most people have access to clean water, but a small proportion use borehole water sources. People live in large family units, and their main economic activities are subsistence farming and animal husbandry. The primary ethnicity is Chagga, and people follow Chagga traditions, such as drinking local brews made from banana or plantain.

Participant screening and enrollment

In rural Moshi, study information was given through community leaders and announcements during mass gatherings in mosques, churches and village meetings. In urban Moshi, study information was distributed using leaflets and through community leaders, office announcements and university gatherings. Eligible participants (aged 18-35 years and with permanent residency of a given location) were asked to enroll in the study. Following informed consent, 233 participants were voluntarily screened for in- and exclusion criteria. Exclusion criteria were having an acute or chronic disease (including HIV, tuberculosis, cancer, cardiovascular disease, gastrointestinal diseases, recurrent infections, liver, renal, endocrine or neurological disorders), ≥ 2 hospital admissions times/year in the last year, chronic use of antibiotics or corticosteroids, use

of other immunosuppressive drugs, recent receipt of another vaccine, lactation, positive pregnancy test, history of blood product transfusion, hemoglobin level ≤ 8.5 g/dL, testing positive HIV or malaria, low or high blood pressure ($\leq 90/60$ mmHg and $\geq 140/90$ mmHg, respectively) or high blood glucose (≥ 7.1 mmol/L fasting or ≥ 11.1 mmol/L random glucose).

Participants were screened for HIV infection (SDBIOLINE HIV-1/2 3.0kit, LOT:03ADG020A), malaria (Malaria Ag p.f/Pan, Ref: 05FK60, LOT:05EDG018A) and soil-transmitted helminth such as hookworms (*Ancylostoma duodenale* and *Necator americanus*), *Trichuris trichiura*, *Ascaris lumbricoides*, *Strongyloides stercoralis* and *Schistosoma mansoni* using Kato-Katz or POC-CCA (*Schistosoma mansoni* and *Schistosoma haematobium*; butch no:220701075). Furthermore, hemoglobin levels were measured (HemoCue Hb 301(CE:1450820055) and random blood glucose was assessed (ACCU-CHECK glucose test strips, Roche Diabetic care,06993761001). Weight and height were measured using a well-calibrated machine (RGZ-160, made in China), and blood pressure was measured using OMRON(SN:202111007949V). All individuals with abnormal laboratory or clinical findings except those with parasitic infections received nurse counselling, referral, or treatment before being excluded. Based on exclusion criteria, 48 of 233 participants were excluded.

Lifestyle questionnaire

Questionnaires were adapted from the Tanzania Demographic and Health Survey and Malaria Indicator Survey (TDHS-MIS) and a food history questionnaire previously applied to a similar population[54,60]. We collected information on the water source, toilet type, available cooking facilities and the materials used to construct the floor, roof and walls of the house. We assessed exposure to livestock by collecting information on the number of milk cows, cattle, goats, sheep, horses and poultry owned. Inquiries were made on land ownership and possession of non-productive assets, including radios, televisions, computers, refrigerators, ironing tools (whether powered by charcoal or electricity), watches, motorcycles, trucks, animal-drawn carts,

generators and motorboats. Using food questionnaires, data was collected on the frequency of dietary products participants consume weekly, including ugali (maize stiff porridge), plantain, banana, rice, potatoes, meat, fish, beans/peas, green vegetables, fruits and locally brewed beer.

Microbiota profiling

Stool sample collection

Stool samples were collected at baseline and/or at day 56. For rural individuals, stool samples were collected at the outpatient clinic or at home. For urban individuals, stool was collected at the KCMC or at home. Feces were initially stored in a dry stool container, before it was transferred into DNA/RNA Shield Fecal Collection (Zymo Research, Irvine, California, USA) for transportation and storage (median [IQR] time between stool production and storage in medium, 58.5 [15.0-113.0] minutes). Samples were stored in a -80°C freezer at the KCMC (median [IQR] time between storage in medium and storage in freezer, 120.5 [93.0-168.5 minutes]). Samples were transported to The Netherlands on dry ice and stored at -80°C before further processing. Samples of individuals who took antibiotics between study enrollment and stool sample collection at day 56 were excluded ($n = 9$).

Bacterial DNA isolation

DNA was extracted from 250 μ L diluted feces (in DNA/RNA Shield) by Repeated Bead Beating (RBB) combined with purification using the chemagic DNA Stool 200 Kit H96 (Revvity, Waltham, Massachusetts, USA). Briefly, 250 μ L diluted feces was homogenized with 1.0 mL Lysis Buffer 1 of the chemagic DNA Stool 200 Kit H96 by vortexing in a 2 mL tube (with screw cap) containing 0.5 g of sterile zirconia beads $\varnothing$ 0.1mm (BioSpec, Cat. No. 11079101z). Cells were mechanically lysed on a FastPrep-24™ 5G Instrument (MP Bio, Irvine, California, USA) at 5.5 m/s for 3 cycles of 1 minute. Samples were subsequently centrifuged (16,000 $\times$ g 4°C, for 5 min) and the supernatant was transferred to a new tube to which 30 μ L Proteinase K was added, mixed and incubated for 10 minutes at 70°C. Thereafter, samples were incubated for another 5 minutes at 95°C followed by centrifugation for 5 minutes at high speed

(13,000 rpm). Deep-well plates were filled with 800 μ L of the supernatant (lysates) and further isolation was done on a chemagic™ 360 instrument (Revvity, Waltham, Massachusetts, USA) according to the manufacturer's instructions. Finally, purified DNA was quantified on a Qubit Fluorometer (ThermoFischer Scientific, Waltham, Massachusetts, USA) using the dsDNA HS Assay Kit (Invitrogen, Waltham, Massachusetts, USA).

16S-rRNA gene amplicon sequencing

The variable regions of the V3-V4 of the bacterial 16S rRNA gene were PCR amplified from each DNA sample in a single reaction workflow with simultaneous indexing and target amplification using the EasySeq™ 16S Microbiome Library Prep Kit (NimaGen, Nijmegen, the Netherlands) according to the manufacturer's instructions. Amplicon libraries were sequenced on an Illumina MiSeq instrument (Illumina, Eindhoven, The Netherlands) (MiSeq Reagent Kit v3, 2 $\times$ 300 cycles, 10% PhiX) to generate paired-end reads of 300 bases in length in both directions.

16S rRNA gene amplicon sequence data processing

Forward and reverse primers were removed using cutadapt v4.7⁶⁶. Following, using DADA2 (v1.28.0) paired-end sequences were filtered and trimmed (maxEE = 2, truncLen = 240/210bp), denoised, merged (minOverlap = 10, maxMismatch = 0) and a sequence table was constructed. Chimeras were identified and removed (method = 'consensus'). After denoising and merging, any ASVs with a length of <350 or >500 bases were discarded. ASVs were annotated up to genus-level using the DADA2 implementation of the naïve Bayesian classifier based on the SILVA v138.1 reference database. Species-level annotations were added using the *addSpecies()*-function⁶⁷. ASVs not assigned to the kingdom Archaea/Bacteria or annotated as Mitochondria (family) or Chloroplast (class) were removed.

Yellow fever neutralizing antibody titer measurements

Blood sample collection and processing

Blood samples were collected in 5mL EDTA tubes on day 0 (pre-vaccination), day 28 (4 weeks post-vaccination; presumed peak in antibody titers) and

day 178 (~6 months post-vaccination; detect delayed response/assess weaning status). Blood samples were transported to a clinical laboratory for plasma separation through centrifugation. Plasma samples were stored at -80°C within two hours. Samples were shipped between institutions on dry ice.

Yellow fever plaque reduction neutralization tests

Plasma neutralizing antibodies against the yellow fever virus from all individuals were quantified using plaque reduction neutralization tests (PRNTs). Results are presented as PRNT₅₀ and PRNT₉₀ titers, which correspond to the reciprocal of the plasma dilution, achieving a 50% and 90% reduction in virus plaque-forming units per 0.1 mL of the reference 17D virus preparation, respectively. Assay quality control included determining the 50% and 90% neutralization cut-off values for each assay via back titration of the virus inoculum. Further details are provided elsewhere⁶⁸.

Data analysis

All data preprocessing and statistics were performed in R v4.3.3 within R Studio v2024.04.1+748 (Boston, MA). Microbiota-specific analyses and visualization, including alpha- and beta-diversity analyses and microbiota association tests were performed using the microViz v0.12.1 R package⁶⁹. All statistical tests were two-sided and *p*-values were corrected for multiple testing using the Benjamini-Hochberg procedure (referred to as *p_{adj}*-values). *P*-/*p_{adj}*-values < 0.05 were considered statistically significant.

Baseline characteristics

Descriptives for baseline and lifestyle questionnaire data were generated using the gtsummary v1.7.2 R package.

Lifestyle and food score

Lifestyle questionnaire data (*N* = 185) were mode imputed (missing values) and subjected to either Multiple Correspondence Analysis (MCA) or Principal Component Analysis (PCA) to calculate lifestyle and food score, respectively. Lifestyle score was calculated by applying MCA to categorical questionnaire data (*n* = 38 manually curated

lifestyle-related questions; 21 on assets, 11 on food and 6 on housing) (FactoMineR v2.7 R package, *MCA()*-function; PC1 defined as 'lifestyle score'). (Cumulative) variable category contributions were extracted and shown. To calculate a food score, variables capturing the frequency of consumption of dietary products ($n = 11$; 0, 1, 2-4 or $\geq 5 \times / \text{week}$) were converted into ordinal variables and used as input variables for a PCA ordination (*PCA()*-function). Coordinates of samples and variable categories/variables were visualized in biplots. For lifestyle score analyses, (cumulative) variable category contributions were extracted and shown.

Alpha-diversity analysis

Diversity analyses were performed on unfiltered and non-rarefied raw read counts (ASVs present in ≥ 2 samples). Shannon diversity was primarily used as a measure for within-sample/alpha-diversity, leveraging both species richness and evenness (ASV- and genus-level). In addition, we compared the number of observed ASVs or genera. Statistical significance between groups was assessed using linear models with observed richness or Shannon diversity as outcome, adjusting for vaccination status, age, sex and sequencing depth.

Beta-diversity analysis

Beta-diversity analyses were performed on centered-log-ratio [CLR]-transformed genus-level abundance matrices. CLR-transformed abundances were used in principal components analysis (PCA) to visualize major patterns of microbiota variation. Associations between rural/urban living location, host and environmental variables (age, sex, helminth infection status, highest level of education, body mass index (BMI), lifestyle and food scores (PC1 and PC2), vaccination status and technical variables (sequencing depth) were explored by multivariable Permutational Multivariate Analysis of Variance (Aitchison-distance matrix; PERMANOVA), adjusting for living location, vaccination, age and sex as appropriate. For variables for which collinearity with living location was expected (i.e. lifestyle/food scores, helminth infection status and sex), models were adjusted for vaccination, age and sex (dropping living location). Paired day 0 and day 56 samples from vaccinated individuals were tested for

differences between time points (indicating a vaccination effect) using a PERMANOVA with permutations restricted within participant.

Microbiota clustering

Microbiota profiles ($n = 154$ samples) were clustered into Community State Types (CSTs) by fitting a Dirichlet Multinomial model (DirichletMultinomial v1.46.0 R-package; *dmn()*-function). Genus-level count data were taken as input and models were fitted with 1 to 10 components. The optimal number of Dirichlet components was determined by inspecting measures of fit (Laplace, AIC and BIC). Species associated with either CST1 or CST2 were identified using differential abundance analyses (MaAsLin2; no adjustment for covariates, see below).

Food-variable association tests

To assess the association between food variables (frequency of consumption of dietary products per week) and living location and living location/CST (including rural-like urban individuals), we used logistic regression analysis, adjusting for all other food variables and sex.

Differential abundance analysis

Differentially abundant taxa between groups were identified using linear models, like the implementation in the MaAsLin2 framework (default parameters; log₂-normalisation)⁶⁸. For each comparison genera present at >0.1% relative abundance in $\geq 10\%$ of samples were selected. We explored the associations between 1) rural/urban living location (adjusted for vaccination status, age and sex), 2) CST1 and CST2 (no adjustment for covariates), 3) urban (urban + CST2)/rural living individuals and rural-like urbanites (urban + CST1) (no adjustment for covariates), 4) food variables (stratified by living location; adjusted for other food variables and sex), and 5) log₁₀-transformed yellow fever neutralizing antibodies (adjusted for living location, age and sex or age and sex) and relative abundances of individual taxa. When comparing >2 groups (e.g. urban/rural/rural-like urbanites), estimates and p -values for each pairwise comparison between groups (e.g. urban vs rural) were extracted using the *emmeans* v1.8.5 R-package. p -values were adjusted per rank, term and contrast, as appropriate,

and resulting BH-adjusted p -values of below 0.05 were considered statistically significant.

BloSSUM and VANISH taxa

Classification into BloSSUM (bloom or selected in societies of urbanization/modernization) and VANISH (volatile and/or associated negatively with industrialized societies of humans) taxa was based on a publication describing gut microbiota profiles in a non-industrialized population of Hadza hunter-gatherers. Lists of BloSSUM and VANISH taxa were extracted from the publication and collapsed at genus-level. Genera classified as both BloSSUM- and VANISH-associated were dropped. For differentially abundant genera associated with rural/urban living or CST1/CST2, we tested for significant enrichment of either BloSSUM or VANISH taxa using Fisher's Exact tests.

Multimodal analyses

Based on visual inspection of stacked bar plots, we manually selected several genera for which the (multimodal) distribution of relative abundance was assessed in a post-hoc analysis using the multimode R-package. The number of modes was determined using the *nmodes()*-function, based on inspection of a stairs plot across a range of density bandwidths. Mode location was determined by the *locmodes()*-function.

Yellow fever neutralizing antibody titers

Log₁₀-base geometric mean yellow fever neutralizing antibody titers (PRNT₅₀ and PRNT₉₀) at baseline (pre-vaccination), day 28 (4 weeks post-vaccination) and at day 178 (~6 months post-vaccination) were calculated. Values under the limit of detection (<10) were assumed to have a value of 5. Statistical analyses were performed using a generalized linear mixed effects model with log₁₀-transformed PRNT₅₀/PRNT₉₀-values as outcome, group, CST or group/CST, time point, the interaction between group, CST or group/CST and time point, age and sex as fixed effects and participant ID as random effects. Pairwise comparisons of estimated marginal means between groups were

computed at each time point or between day 28 and day 178 for each group (emmeans v1.10.0 R package).

Data availability

16S-rRNA sequencing data from this study are available from NCBI under BioProject accession number PRJNA1141956. Other data are available from the corresponding author upon reasonable request.

Acknowledgements

This study is part of the EDCTP2 program supported by the European Union and by grants from the Dutch Research Organization (NWO) through the Spinoza prize awarded to Maria Yazdanbakhsh, the European Research Council (ERC) via the ERC Advanced Grant ‘REVERSE’ awarded to Maria Yazdanbakhsh (Grant No: 101055179), the LUMC Excellent Student Fellowship awarded to Marloes M.A.R. van Dorst and the LUMC Global PhD Fellowship awarded to Jeremia J. Pyuza. We want to acknowledge all clinical and research staff at KCRI and KCMC in Tanzania who helped to make this study possible. We would also like to acknowledge the work of John Kayiwa and the team at the Uganda Virus Research Institute (UVRI) in Uganda and members of the Penders laboratory at Maastricht UMC+ for performing 16S-rRNA sequencing. Finally, we would like to thank all volunteers who participated in this study.

Authors contributions

Conceptualization, J.J.P., M.M.A.R.v.D., M.Y.; Methodology, J.J.P., M.M.A.R.v.D., S.E.M., S.P.J., P.C.W.H., L.W., M.Y.; Data Curation, J.J.P., D.B., W.A.A.d.S.P.; Formal Analysis, J.P., K.S., J.J.P., D.B., M.M., M.Y., W.A.A.d.S.P.; Investigation, J.J.P., M.M.A.R.v.D., N.A., A.N., Y.K., M.K.; Writing - Original Draft, J.J.P., M.Y., W.A.A.d.S.P.; Review & Editing, W.A.A.d.S.P., M.M., J.P., K.S., V.I.K., E.R.S., A.M., S.E.M., S.P.J., P.C.W.H., L.W., M.Y.; Supervision, J.P., S.P.J., S.E.M., P.C.W.H., W.A.A.d.S.P., M.Y.; Funding Acquisition, M.Y. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interest.

References

- Montero DA, Vidal RM, Velasco J, et al. Two centuries of vaccination: historical and conceptual approach and future perspectives. *Front Public Health*. 2024;11. doi:10.3389/fpubh.2023.1326154
- Alghamdi S. The role of vaccines in combating antimicrobial resistance (AMR) bacteria. *Saudi J Biol Sci*. 2021;28(12):7505-7510. doi:10.1016/j.sjbs.2021.08.054
- Sinumvayo JP, Munezero PC, Tope AT, et al. Vaccination and vaccine-preventable diseases in Africa. *Sci Afr*. 2024;24:e02199. doi:10.1016/j.sciaf.2024.e02199
- Micoli F, Bagnoli F, Rappuoli R, Serruto D. The role of vaccines in combatting antimicrobial resistance. *Nat Rev Microbiol*. 2021;19(5):287-302. doi:10.1038/s41579-020-00506-3
- Greenwood B. The contribution of vaccination to global health: past, present and future. *Philosophical Transactions of the Royal Society B: Biological Sciences*. 2014;369(1645):20130433. doi:10.1098/rstb.2013.0433
- Lynn DJ, Benson SC, Lynn MA, Pulendran B. Modulation of immune responses to vaccination by the microbiota: implications and potential mechanisms. *Nat Rev Immunol*. 2022;22(1):33-46. doi:10.1038/s41577-021-00554-7
- Nkurunungi G, Nassuuna J, Natukunda A, et al. The effect of intensive praziquantel administration on vaccine-specific responses among schoolchildren in Ugandan schistosomiasis-endemic islands (POPVAC A): an open-label, randomised controlled trial. *Lancet Glob Health*. 2024;12(11):e1826-e1837. doi:10.1016/S2214-109X(24)00280-8
- van Riet E, Retra K, Adegnika AA, et al. Cellular and humoral responses to tetanus vaccination in gabonese children. *Vaccine*. 2008;26(29-30):3690-3695. doi:10.1016/j.vaccine.2008.04.067
- Riet E van, Adegnika AA, Retra K, et al. Cellular and Humoral Responses to Influenza in Gabonese Children Living in Rural and Semi-Urban Areas. *J Infect Dis*. 2007;196(11):1671-1678. doi:10.1086/522010
- Lynn DJ, Pulendran B. The potential of the microbiota to influence vaccine responses. *J Leukoc Biol*. 2018;103(2):225-231. doi:10.1189/jlb.5MR0617-216R
- Zimmermann P, Curtis N. Factors That Influence the Immune Response to Vaccination. *Clin Microbiol Rev*. 2019;32(2). doi:10.1128/CMR.00084-18
- Falahi S, Kenarkoobi A. Host factors and vaccine efficacy: Implications for COVID-19 vaccines. *J Med Virol*. 2022;94(4):1330-1335. doi:10.1002/jmv.27485
- Ng SC, Peng Y, Zhang L, et al. Gut microbiota composition is associated with SARS-CoV-2 vaccine immunogenicity and adverse events. *Gut*. 2022;71(6):1106-1116. doi:10.1136/gutjnl-2021-326563
- Huda MN, Lewis Z, Kalanetra KM, et al. Stool Microbiota and Vaccine Responses of Infants. *Pediatrics*. 2014;134(2):e362-e372. doi:10.1542/peds.2013-3937
- Huda MN, Ahmad SM, Alam MJ, et al. Bifidobacterium Abundance in Early Infancy and Vaccine Response at 2 Years of Age .

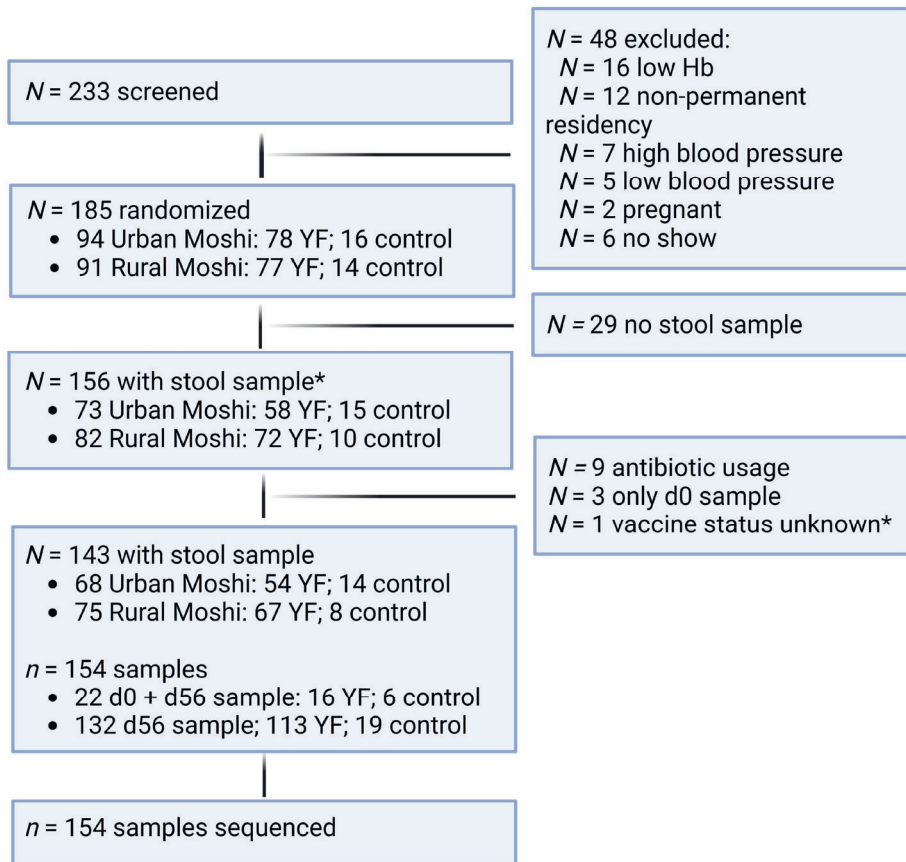
- Pediatrics*. 2019;143(2). doi:10.1542/peds.2018-1489
16. Harris V, Ali A, Fuentes S, et al. Rotavirus vaccine response correlates with the infant gut microbiota composition in Pakistan. *Gut Microbes*. 2018;9(2):93-101. doi:10.1080/19490976.2017.1376162
 17. Fix J, Chandrashekhar K, Perez J, et al. Association between Gut Microbiome Composition and Rotavirus Vaccine Response among Nicaraguan Infants. *Am J Trop Med Hyg*. 2020;102(1):213-219. doi:10.4269/ajtmh.19-0355
 18. Harris VC, Armah G, Fuentes S, et al. Significant Correlation Between the Infant Gut Microbiome and Rotavirus Vaccine Response in Rural Ghana. *J Infect Dis*. 2017;215(1):34-41. doi:10.1093/infdis/jiw518
 19. Ray S, Narayanan A, Vesterbacka J, et al. Impact of the gut microbiome on immunological responses to COVID-19 vaccination in healthy controls and people living with HIV. *NPJ Biofilms Microbiomes*. 2023;9(1):104. doi:10.1038/s41522-023-00461-w
 20. Schnorr SL, Candela M, Rampelli S, et al. Gut microbiome of the Hadza hunter-gatherers. *Nat Commun*. 2014;5(1):3654. doi:10.1038/ncomms4654
 21. Smits SA, Leach J, Sonnenburg ED, et al. Seasonal cycling in the gut microbiome of the Hadza hunter-gatherers of Tanzania. *Science (1979)*. 2017;357(6353):802-806. doi:10.1126/science.aan4834
 22. Carter MM, Olm MR, Merrill BD, et al. Ultra-deep sequencing of Hadza hunter-gatherers recovers vanishing gut microbes. *Cell*. 2023;186(14):3111-3124.e13. doi:10.1016/j.cell.2023.05.046
 23. Yatsunenko T, Rey FE, Manary MJ, et al. Human gut microbiome viewed across age and geography. *Nature*. 2012;486(7402):222-227. doi:10.1038/nature11053
 24. Ayeni FA, Biagi E, Rampelli S, et al. Infant and Adult Gut Microbiome and Metabolome in Rural 25. Bassa and Urban Settlers from Nigeria. *Cell Rep*. 2018;23(10):3056-3067. doi:10.1016/j.celrep.2018.05.018
 25. Das B, Ghosh TS, Kedia S, et al. Analysis of the Gut Microbiome of Rural and Urban Healthy Indians Living in Sea Level and High Altitude Areas. *Sci Rep*. 2018;8(1):10104. doi:10.1038/s41598-018-28550-3
 26. Stražar M, Temba GS, Vlamakis H, et al. Gut microbiome-mediated metabolism effects on immunity in rural and urban African populations. *Nat Commun*. 2021;12(1):4845. doi:10.1038/s41467-021-25213-2
 27. van der Vossen EWJ, Davids M, Bresser LRF, et al. Gut microbiome transitions across generations in different ethnicities in an urban setting—the HELIUS study. *Microbiome*. 2023;11(1):99. doi:10.1186/s40168-023-01488-z.
 28. Rampelli S, Schnorr SL, Consolandi C, et al. Metagenome Sequencing of the Hadza Hunter-Gatherer Gut Microbiota. *Current Biology*. 2015;25(13):1682-1693. doi:10.1016/j.cub.2015.04.055

29. 30. Maghini DG, Oduaran OH, Wirbel J, et al. Expanding the human gut microbiome atlas of Africa. Published online March 14, 2024. doi:10.1101/2024.03.13.584859
30. Tamburini FB, Maghini D, Oduaran OH, et al. Short- and long-read metagenomics of urban and rural South African gut microbiomes reveal a transitional composition and undescribed taxa. *Nat Commun.* 2022;13(1):926. doi:10.1038/s41467-021-27917-x
31. Ramaboli MC, Ocvirk S, Khan Mirzaei M, et al. Diet changes due to urbanization in South Africa are linked to microbiome and metabolome signatures of Westernization and colorectal cancer. *Nat Commun.* 2024;15(1):3379. doi:10.1038/s41467-024-46265-0
32. Oh JZ, Ravindran R, Chassaing B, et al. TLR5-Mediated Sensing of Gut Microbiota Is Necessary for Antibody Responses to Seasonal Influenza Vaccination. *Immunity.* 2014;41(3):478-492. doi:10.1016/j.immuni.2014.08.009
33. Tang B, Tang L, He W, et al. Correlation of gut microbiota and metabolic functions with the antibody response to the BBIBP-CorV vaccine. *Cell Rep Med.* 2022;3(10):100752. doi:10.1016/j.xcrm.2022.100752
34. Pabst O, Hornef M. Gut Microbiota: A Natural Adjuvant for Vaccination. *Immunity.* 2014;41(3):349-351. doi:10.1016/j.immuni.2014.09.002
35. Lynn MA, Tumes DJ, Choo JM, et al. Early-Life Antibiotic-Driven Dysbiosis Leads to Dysregulated Vaccine Immune Responses in Mice. *Cell Host Microbe.* 2018;23(5):653-660.e5. doi:10.1016/j.chom.2018.04.009
36. Nyangahu DD, Happel AU, Wendoh J, et al. Bifidobacterium infantis associates with T cell immunity in human infants and is sufficient to enhance antigen-specific T cells in mice. *Sci Adv.* 2023;9(49). doi:10.1126/sciadv.ade1370
37. Nadeem S, Maurya SK, Das DK, Khan N, Agrewala JN. Gut Dysbiosis Thwarts the Efficacy of Vaccine Against Mycobacterium tuberculosis. *Front Immunol.* 2020;11. doi:10.3389/fimmu.2020.00726
38. De Filippo C, Cavalieri D, Di Paola M, et al. Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. *Proceedings of the National Academy of Sciences.* 2010;107(33):14691-14696. doi:10.1073/pnas.1005963107
39. Gomez A, Petrzalkova KJ, Burns MB, et al. Gut Microbiome of Coexisting BaAka Pygmies and Bantu Reflects Gradients of Traditional Subsistence Patterns. *Cell Rep.* 2016;14(9):2142-2153. doi:10.1016/j.celrep.2016.02.013
40. Precup G, Vodnar DC. Gut Prevotella as a possible biomarker of diet and its eubiotic versus dysbiotic roles: a comprehensive literature review. *British Journal of Nutrition.* 2019;122(2):131-140. doi:10.1017/S0007114519000680
41. Vacca M, Celano G, Calabrese FM, Portincasa P, Gobetti M, De Angelis M. The Controversial Role of Human Gut Lachnospiraceae. *Microorganisms.* 2020;8(4):573. doi:10.3390/microorganisms8040573
42. Ou J, Carbonero F, Zoetendal EG, et al. Diet, microbiota, and microbial metabolites in colon cancer risk in rural Africans and African

- Americans. *Am J Clin Nutr.* 2013;98(1):111-120. doi:10.3945/ajcn.112.056689
43. Newman TM, Shively CA, Register TC, et al. Diet, obesity, and the gut microbiome as determinants modulating metabolic outcomes in a non-human primate model. *Microbiome.* 2021;9(1):100. doi:10.1186/s40168-021-01069-y
 44. Sun J, Liao XP, D'Souza AW, et al. Environmental remodeling of human gut microbiota and antibiotic resistome in livestock farms. *Nat Commun.* 2020;11(1):1427. doi:10.1038/s41467-020-15222-y
 45. Sudatip D, Mostacci N, Thamlikitkul V, Oppliger A, Hilty M. Influence of occupational exposure to pigs or chickens on human gut microbiota composition in Thailand. *One Health.* 2022;15:100463. doi:10.1016/j.onehlt.2022.100463
 46. Nel Van Zyl K, Whitelaw AC, Hesselning AC, Seddon JA, Demers AM, Newton-Foot M. Association between clinical and environmental factors and the gut microbiota profiles in young South African children. *Sci Rep.* 2021;11(1):15895. doi:10.1038/s41598-021-95409-5
 47. Rosa BA, Supali T, Gankpala L, et al. Differential human gut microbiome assemblages during soil-transmitted helminth infections in Indonesia and Liberia. *Microbiome.* 2018;6(1):33. doi:10.1186/s40168-018-0416-5
 48. Mondot S, Poirier P, Abou-Bacar A, et al. Parasites and diet as main drivers of the Malagasy gut microbiome richness and function. *Sci Rep.* 2021;11(1):17630. doi:10.1038/s41598-021-96967-4
 49. Walusimbi B, Lawson MAE, Nassuuna J, et al. The effects of helminth infections on the human gut microbiome: a systematic review and meta-analysis. *Frontiers in Microbiomes.* 2023;2. doi:10.3389/frmbi.2023.1174034
 50. Schneeberger PHH, Gueuning M, Welsche S, et al. Different gut microbial communities correlate with efficacy of albendazole-ivermectin against soil-transmitted helminthiasis. *Nat Commun.* 2022;13(1):1063. doi:10.1038/s41467-022-28658-1
 51. van Dorst MMAR, Pyuza JJ, Nkurunungi G, et al. Immunological factors linked to geographical variation in vaccine responses. *Nat Rev Immunol.* 2024;24(4):250-263. doi:10.1038/s41577-023-00941-2
 52. Muyanja E, Ssemaganda A, Ngauv P, et al. Immune activation alters cellular and humoral responses to yellow fever 17D vaccine. *Journal of Clinical Investigation.* 2014;124(7):3147-3158. doi:10.1172/JCI75429
 53. Domingo C, Fraissinet J, Ansah PO, et al. Long-term immunity against yellow fever in children vaccinated during infancy: a longitudinal cohort study. *Lancet Infect Dis.* 2019;19(12):1363-1370. doi:10.1016/S1473-3099(19)30323-8
 54. Jongo SA, Shekalaghe SA, Church LWP, et al. Safety, Immunogenicity, and Protective Efficacy against Controlled Human Malaria Infection of Plasmodium falciparum Sporozoite Vaccine in Tanzanian Adults. *Am J Trop Med Hyg.* 2018;99(2):338-349. doi:10.4269/ajtmh.17-1014
 55. Querec TD, Akondy RS, Lee EK, et al. Systems biology approach predicts immunogenicity of the yellow fever vaccine in humans. *Nat Immunol.* 2009;10(1):116-125. doi:10.1038/ni.1688
 56. Monath TP, Cetron MS, McCarthy K, et al. Yellow Fever 17D Vaccine Safety and

- Immunogenicity in the Elderly. *Hum Vaccin*. 2005;1(5):207-214. doi:10.4161/hv.1.5.2221
57. Monath TP, Nichols R, Archambault WT, et al. Comparative safety and immunogenicity of two yellow fever 17D vaccines (ARILVAX and YF-VAX) in a phase III multicenter, double-blind clinical trial. *Am J Trop Med Hyg*. 2002;66(5):533-541. doi:10.4269/ajtmh.2002.66.533
 58. A, Nkurunungi G, Zirimenya L, et al. Schistosome and malaria exposure and urban-rural differences in vaccine responses in Uganda: a causal mediation analysis using data from three linked randomised controlled trials. *Lancet Glob Health*. 2024;12(11):e1860-e1870. doi:10.1016/S2214-109X(24)00340-1
 59. TNBS. Tanzania population and housing census- Tanzania National Bureau of statistics(TNBS). <https://www.nbs.go.tz/index.php/en/census-surveys/population-and-housing-census>. Published online 2022.
 60. TDHS-MIS. Tanzania Demographic and Health Survey and Malaria Indicator Survey (TDHS-MIS) 2015-16. <https://dhsprogram.com/publications/publication-FR321-DHS-Final-Reports.cfm>. Published online 2016.
 61. Cho SH, Raybuck AL, Stengel K, et al. Germinal centre hypoxia and regulation of antibody qualities by a hypoxia response system. *Nature*. 2016;537(7619):234-238. doi:10.1038/nature19334
 62. Taylor CT, Scholz CC. The effect of HIF on metabolism and immunity. *Nat Rev Nephrol*. 2022;18(9):573-587. doi:10.1038/s41581-022-00587-8
 63. Abbott RK, Thayer M, Labuda J, et al. Germinal Center Hypoxia Potentiates Immunoglobulin Class Switch Recombination. *The Journal of Immunology*. 2016;197(10):4014-4020. doi:10.4049/jimmunol.1601401
 64. McGettrick AF, O'Neill LAJ. The Role of HIF in Immunity and Inflammation. *Cell Metab*. 2020;32(4):524-536. doi:10.1016/j.cmet.2020.08.002
 65. Kim M, Qie Y, Park J, Kim CH. Gut Microbial Metabolites Fuel Host Antibody Responses. *Cell Host Microbe*. 2016;20(2):202-214. doi:10.1016/j.chom.2016.07.001
 66. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J*. 2011;17(1):10. doi:10.14806/ej.17.1.200
 67. Edgar RC. Updating the 97% identity threshold for 16S ribosomal RNA OTUs. *Bioinformatics*. 2018;34(14):2371-2375. doi:10.1093/bioinformatics/bty113
 68. Mallick H, Rahnavard A, McIver LJ, et al. Multivariable association discovery in population-scale meta-omics studies. *PLoS Comput Biol*. 2021;17(11):e1009442. doi:10.1371/journal.pcbi.1009442
 69. Barnett D, Arts I, Penders J. microViz: an R package for microbiome data visualization and statistics. *J Open Source Softw*. 2021;6(63):3201. doi:10.21105/joss.03201

Supplementary Figure



Supplementary Figure 1 | Flowchart of volunteer recruitment, randomization and stool sample collection.

N = number of volunteers; *n* = number of samples. YF = Yellow Fever; d0 = day 0/baseline sample; d56 = day 56 sample. Low Hb, Hb < 8.5 g/dL; low blood pressure, blood pressure ≤90/60mmHg; high blood pressure, blood pressure ≥140/90mmHg. Non-permanent residency indicates individuals who upon further questioning had not lived at a given location for their whole life. *1 volunteer without recorded group/vaccine status at the moment of sample shipment.



Supplementary Figure 2 | Heatmap visualizing lifestyle questionnaire data. $N = 185$ participants.

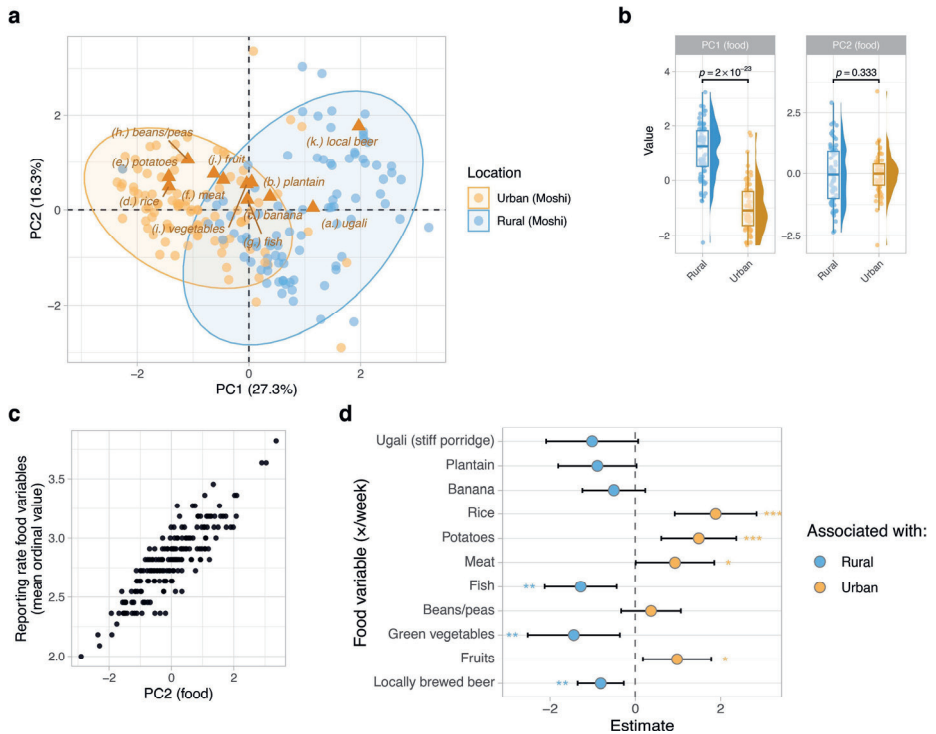
Values represent the number of participants. Colors indicate the percentage of the total. Comparisons between groups (rural/urban living) were performed using Fisher's exact or chi-squared tests. Asterisks denote statistical significance (NS, non-significant; *, $p \leq 0.05$; **, $p \leq 0.01$; ***, $p \leq 0.001$, $p \leq 0.0001$). See Supplementary Table 2.



Supplementary Figure 3 | Multiple Correspondence Analysis (MCA) based on questionnaire data to generate lifestyle score.

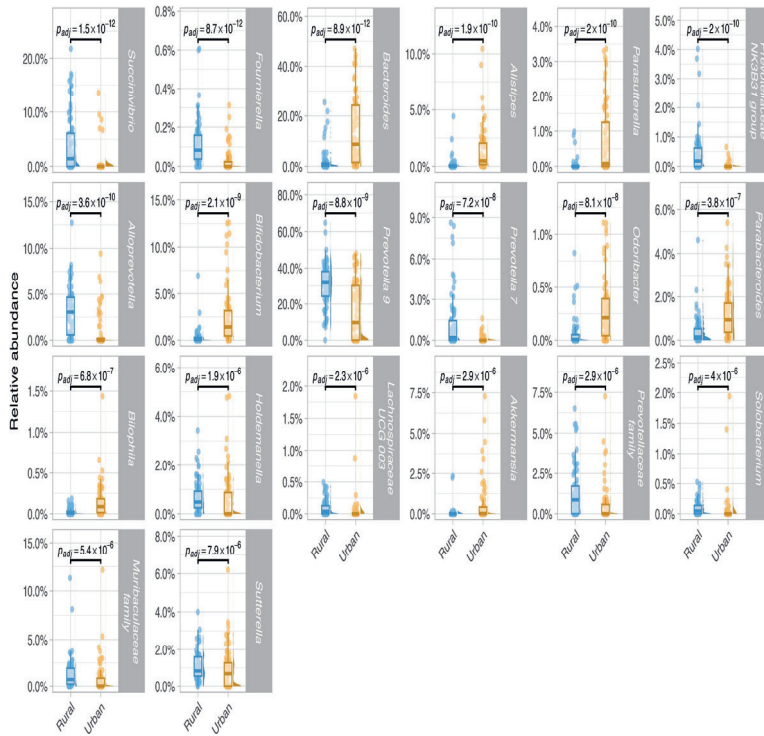
a) MCA was applied to categorical questionnaire data (38 manually curated questions; 21 on assets, 11 on food and 6 on housing) ($N = 185$ individuals). Data points are colored based on location. Ellipses reflect the data spread at a 95% confidence level. Density plots show the distribution of PC1 (lifestyle score) (x-axis) and PC2 (y-axis) score. b) Differences in lifestyle scores (PC1/PC2) between rural and urban living individuals. See legend **Figure 2a-b**. Statistical significance between groups was assessed using linear models with lifestyle score as outcome, adjusting for age and sex. c) Cumulative contributions (in percentage) of the variable categories by questionnaire data category

(i.e. housing, assets and food). d, f) Coordinates of each variable category most strongly contributing to PC1 (d) or PC2 (f) across dimensions 1 and 2. e, g) Contributions (in percentage) of variable categories to e) PC1 or lifestyle score or g) PC2 (based on lifestyle variables). Bars are colored based on PC1/PC2 scores as appropriate. h) Difference in lifestyle score across the overall cohort ($N = 185$) compared to the microbiota subcohort ($N = 143$), stratified by urban/rural living location. Statistical significance between groups was assessed using a linear model with lifestyle score as outcome.



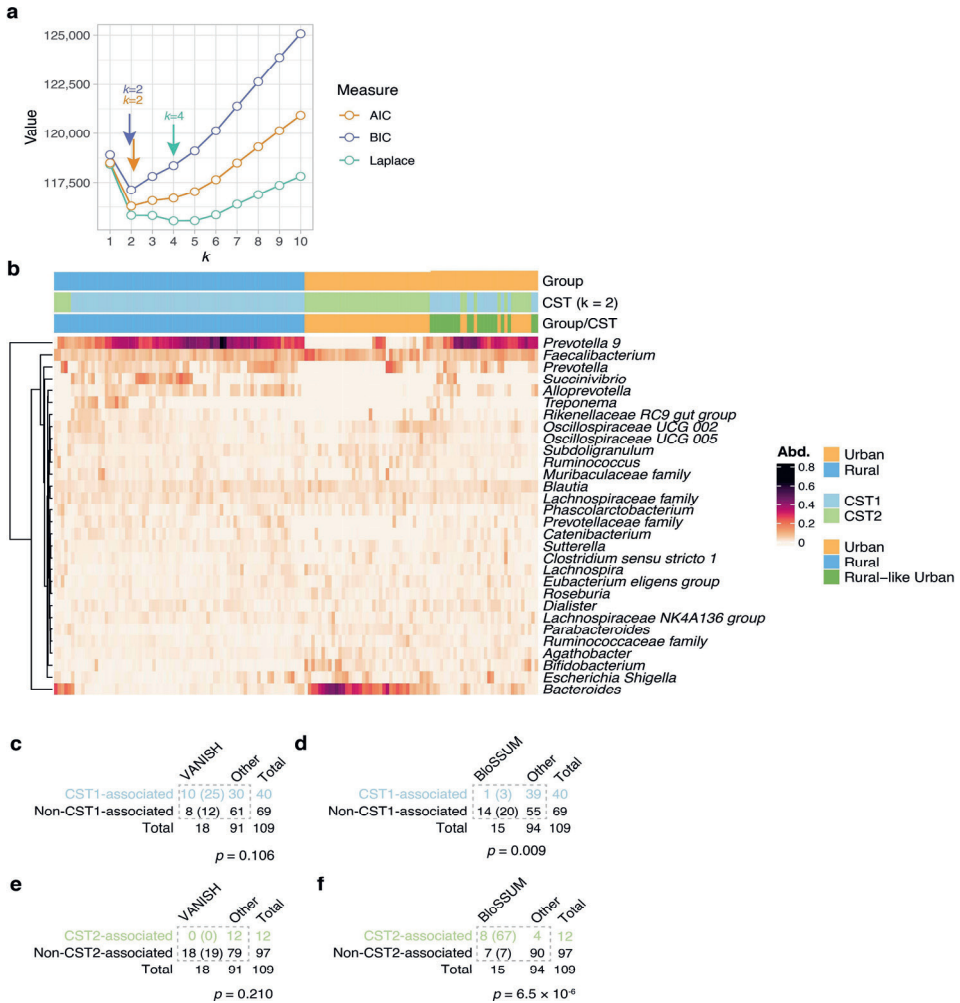
Supplementary Figure 4 | Principal component analysis (PCA) based on food variables to generate a food score.

a) PCA was applied to 11 food-related variables (encoded as ordinal variables) ($N = 185$ individuals). Data points are colored based on location. Ellipses reflect the data spread at a 95% confidence level. b) Differences in food scores (PC1/PC2) between rural and urban living individuals. See legend Figure 2a-b. Statistical significance between groups was assessed using linear models with food score as outcome, adjusting for age and sex. c) Relation between PC2 food score and reporting rate of food variables (expressed as the mean ordinal value across 11 food variables for each individual). d) Association between food variables and living location (rural/urban; $N = 185$). Statistical significance was assessed using a logistic regression analysis with group as outcome and all food variables and sex as predictors. For each food variable, model estimates are depicted along the x-axis (colored points). Colors indicate the group with which a given food variable is positively associated with. Whiskers denote 95% confidence intervals (CIs; Wald-method). Asterisks denote statistical significance (*, $p \leq 0.05$; **, $p \leq 0.01$; ***, $p \leq 0.001$).



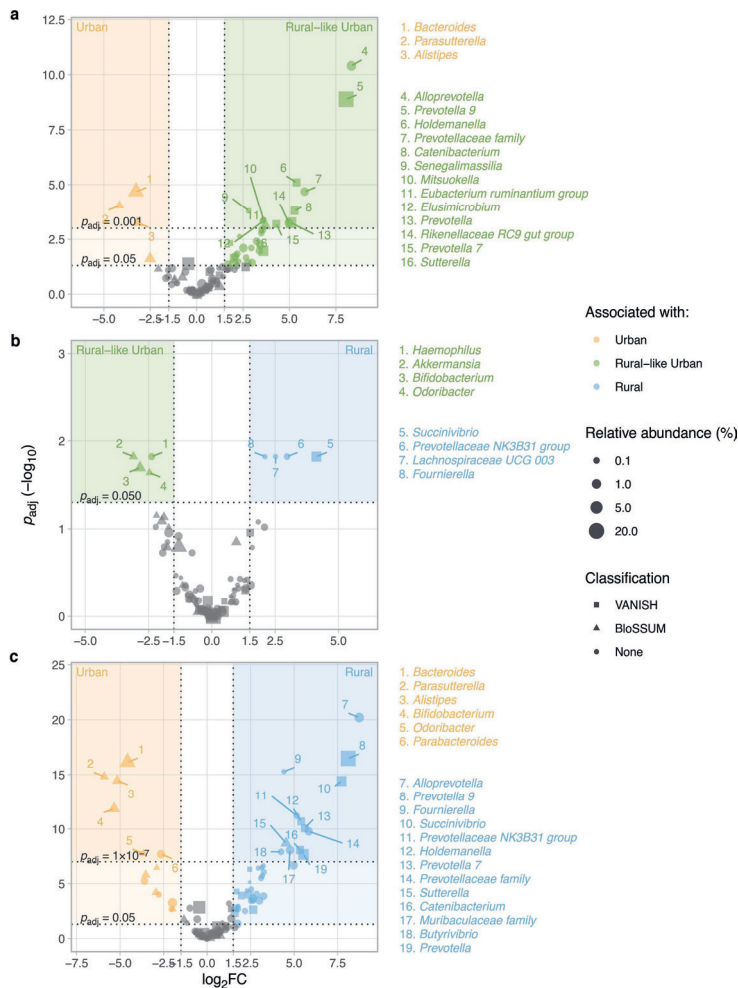
Supplementary Figure 5 | Boxplots of genera associated with rural/urban living location.

Only genera with $p_{adj} < 5 \times 10^{-5}$ and $\log_2FC < -1.5$ or > 1.5 are shown. Statistical significance between groups (rural/urban living) was assessed using linear models with genus-level abundance as outcome, adjusting for vaccination status, age and sex. See legend **Figure 2a-b**.



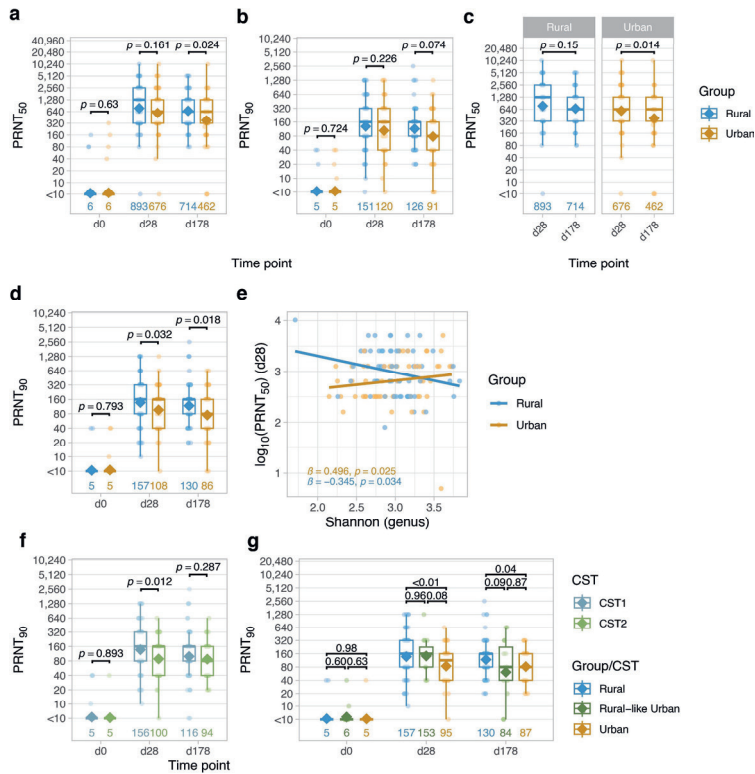
Supplementary Figure 6 | Microbiota profiles cluster in distinct Community State Types (CSTs) using Dirichlet-multinomial modelling (DMM).

a) DMM goodness-of-fit (AIC, BIC and Laplace) for $k = 1-10$ mixture components. For each goodness-of-fit measure, the optimal number of components is denoted using a colored arrow. b) Heatmap showing the relative abundance of the 30 most abundant genera for each sample. Samples are ordered based on living location; within living location samples are ordered according to Bray-Curtis dissimilarities (based on genus-level relative abundance) between samples. Taxa are ordered based on using hierarchical clustering (Euclidean distance) with optimal leaf ordering. Top horizontal bars indicate 1) living location, 2) CST allocation and 3) living location and CST membership combined. Urban living individuals belonging to CST1 were classified as rural-like urbanites since rural individuals almost exclusively clustered in CST1. For these comparisons, ‘Urban’ denotes urban individuals belonging to CST2. c-f 2×2 tables showing the number of VANISH (c, e) or BloSSUM taxa (d, f) across CST1- and CST2-associated taxa. Statistical significance was assessed using Fisher’s Exact tests.



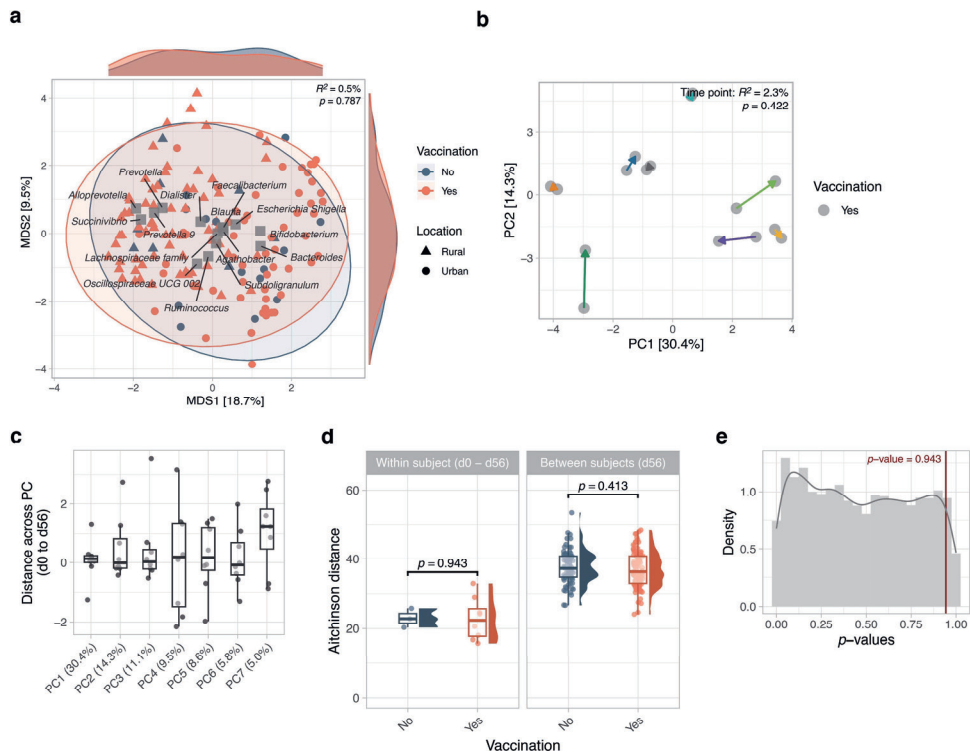
Supplementary Figure 7 | Differential abundance analyses urban/rural living individuals and rural-like urbanites.

a-c) Volcano plots visualizing genera enriched in a) urban (urban + CST2) living individuals compared to rural-like urbanites, b) rural-like urbanites compared to rural living individuals and c) urban (urban + CST2) compared to rural living individuals. Statistical significance between groups (rural/urban living/rural-like urbanites) was assessed using linear models with genus-level abundance as outcome (no adjustment for covariates). Model estimates and p -values for each contrast were extracted using the *emmeans* R-package. P -values were adjusted for multiple testing for each contrast. Genera with $p_{adj} < 0.05$ and $\log_2$ -fold change (FC) < -1.5 or > 1.5 are colored and genera with a $p_{adj} < 0.001$ (a), $p_{adj} < 0.05$ (b) or $p_{adj} < 1 \times 10^{-7}$ (c) and $\log_2$ FC < -1.5 or > 1.5 were annotated. The shape of the data points indicates whether these genera were previously considered to belong to VANISH (volatile and/or associated negatively with industrialized societies of humans) or BioSUM (bloom or selected in societies of urbanization/modernization) taxa.



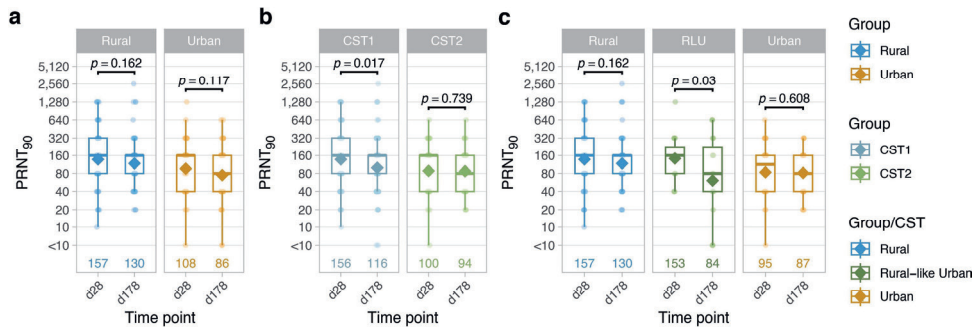
Supplementary Figure 8 | Yellow fever neutralizing antibody titer variation across living location and with microbiota profiles.

a-f) Boxplots (see legend Figure 2a-b) showing yellow neutralizing antibody titers (a; PRNT₅₀ and b-f; PRNT₉₀) in all vaccinated individuals ($N = 155$, a-c) and in those from whom a stool sample was collected ($N = 121$, d-f) at baseline (day 0, d0), 4 weeks post-vaccination (day 28, d28) and 6 months post-vaccination (day 178, d178). Boxplots show comparisons according to group (rural/urban living location; a-b), CST (e) and group/CST (i.e. rural living individuals, rural-like urbanites [CST1] and urban living individuals [CST2]) (f). c) Boxplots showing yellow neutralizing antibody titers (PRNT₅₀) at 4 weeks post-vaccination (day 28, d28) compared to 6 months post-vaccination (day 178, d178) for group (rural/urban living location). Diamond-shaped data points and color values at the bottom of the plot denote $\log_{10}$ -base geometric mean titers. Statistical analyses were performed using a generalized linear mixed effects model with $\log_{10}$ -transformed PRNT₅₀/PRNT₉₀-values as outcome, group, CST or group/CST, time point, the interaction between group, CST or group/CST and time point, age and sex as fixed effects and participant ID as random effects. Pairwise comparisons of estimated marginal means between groups were computed at each time point (a-b, d-f) or between day 28 and day 178 for each group (c). Values under the limit of detection (<10) were assumed to have a value of 5.



Supplementary Figure 9 | Association between YF-17D vaccination on gut microbiota profiles.

a) Principal component analysis (PCA) biplot using CLR-transformed genus-level microbiota features across day 56 samples. Data points, ellipses and side plots are colored based on vaccination status. See legend Figure 1C. R^2 and statistical significance of the association between vaccination status and the overall microbiota composition was assessed using PERMANOVA-test (1,000 permutations), while adjusting for living location, age and sex. b) Principal coordinate analysis of vaccinated volunteers with paired (day 0 – day 56) stool samples ($N = 8$, $n = 16$). Arrows indicate direction of ‘movement’ across PC1 and PC2 from day 0 to day 56. c) Boxplots (see legend Figure 2a-b) showing the distance from day 0 to day 56 across PC1-PC7. d) Within-subject distances (day 0 – day 56) between vaccinated ($N = 8$) and non-vaccinated individuals ($N = 3$) compared to between-subject distances (day 56). Distances between subjects were calculated for subjects with matching group (rural/urban living) and vaccination status (yes/no YF-17D vaccination). Up to 50 distances per stratum (rural – vaccine, urban – vaccine, rural – no vaccine and urban – no vaccine) were randomly selected ($n = 28$ distances were selected for rural – no vaccine). Statistical significance was assessed using a Student’s t-test with Aitchison distance as dependent variable and vaccine status as independent variable. e) Permutation test based on subsampled between-subject distances between paired vaccinated and non-vaccinated subjects ($n = 8$ and $n = 3$, respectively). Histogram and line plot indicate the density distribution of p -values across 5,000 iterations. The ‘true’ observed within-subject p -value is shown in dark red.



Supplementary Figure 10 | Change in yellow fever neutralizing antibody titers at day 28 compared to day 178.

a-c) Boxplots (see legend Figure 2a-b) showing yellow neutralizing antibody titers (PRNT₉₀) at 4 weeks post-vaccination (day 28, d28) compared to 6 months post-vaccination (day 178, d178) for rural compared to urban living individuals (a), CST1 compared to CST2 (b) and group/CST (i.e. rural living individuals, rural-like urbanites [CST1] [RLU] and urban living individuals [CST2]) (c) ($N = 121$). Diamond-shaped data points and color values at the bottom of the plot denote log₁₀-base geometric mean titers. Statistical analyses were performed using a generalized linear mixed effects model with log₁₀-transformed PRNT₉₀-values as outcome, group, CST or group/CST, time point, the interaction between group, CST or group/CST and time point, age and sex as fixed effects and participant ID as random effects. Pairwise comparisons of estimated marginal means between day 28 and day 178 were computed for each group. Values under the limit of detection (<10) were assumed to have a value of 5.

Supplementary tables

Supplementary Table 1 | Baseline characteristics of the study population (N = 185).

Variable	Overall, N = 185	Urban Moshi, N = 94	Rural Moshi, N = 91	p-value
Female sex	84 (45%)	50 (53%)	34 (37%)	0.031
Age	23.7 (21.2, 27.5)	23.6 (21.8, 26.1)	23.9 (20.3, 29.7)	0.813
Age categories				0.022
18-25	115 (62%)	66 (70%)	49 (54%)	
26-36	70 (38%)	28 (30%)	42 (46%)	
BMI	22.5 (20.0, 25.1)	23.6 (20.4, 26.4)	22.1 (19.9, 24.6)	0.076
BMI classification				0.228
<18.5	20 (11%)	9 (9.6%)	11 (12%)	
18.5-24.9	118 (64%)	55 (59%)	63 (69%)	
25.0-29.9	34 (18%)	22 (23%)	12 (13%)	
>30	13 (7.0%)	8 (8.5%)	5 (5.5%)	
Systolic blood pressure (mmHg)	116 (105, 124)	111 (102, 121)	120 (110, 129)	0.002
Diastolic blood pressure (mmHg)	71 (65, 78)	70 (64, 76)	72 (67, 79)	0.125
Hemoglobin level g/dl	15.00 (13.70, 16.20)	14.45 (12.70, 15.88)	15.40 (14.20, 16.50)	<0.001
Random blood sugar, mmol¹^^	5.20 (4.80, 5.80)	5.30 (4.90, 6.00)	4.90 (4.70, 5.45)	<0.001
Highest level of education				<0.001
Primary	52 (28%)	2 (2.1%)	50 (55%)	
Secondary	85 (46%)	47 (50%)	38 (42%)	
College	21 (11%)	20 (21%)	1 (1.1%)	
University	27 (15%)	25 (27%)	2 (2.2%)	
Helminth infection^a	8 (4.3%)	0 (0%)	8 (8.8%)	0.003
Insurance status	80 (43%)	77 (82%)	3 (3.3%)	<0.001
Occupation				<0.001
Farming	39 (21%)	13 (14%)	26 (29%)	
Elementary occupation	59 (32%)	10 (11%)	49 (54%)	
Student	3 (1.6%)	3 (3.2%)	0 (0%)	
Employed/business owner	75 (41%)	64 (68%)	11 (12%)	
Not employed	1 (0.5%)	0 (0%)	1 (1.1%)	
Other	8 (4.3%)	4 (4.3%)	4 (4.4%)	
Received yellow fever vaccine	155 (84%)	78 (83%)	77 (85%)	0.763

N = 185 participants. Values represent number of participants (percentage of total) and median (interquartile range [IQR]) for categorical and continuous variables, respectively. Comparisons between locations were performed using Fisher's exact, chi-squared and Mann-Whitney U-test for categorical and continuous variables, respectively. ^aStool was tested for helminths using the Kato-Katz method, testing for *Schistosoma mansoni*, *Ascaris lumbricoides*, hookworm and *Trichuris trichuria*. Additionally, urine was tested for *Schistosoma haematobium* and *Schistosoma mansoni* using the POC-CCA method.

Supplementary Table 2 | Descriptives of lifestyle score variables (*N* = 185).

Characteristic	Urban Moshi, <i>N</i> = 94	Rural Moshi, <i>N</i> = 91	p-value
House floor			<0.001
Hard floor (tile, cement, concrete, wood)	94 (100%)	74 (81%)	
Soil/sand	0 (0%)	17 (19%)	
House walls			<0.001
Cement, brick or stone	94 (100%)	71 (78%)	
Cane, palm, trunks, bamboo	0 (0%)	14 (15%)	
Mud (with poles)	0 (0%)	6 (6.6%)	
House roof			0.497
Roof tiles	2 (2.1%)	0 (0%)	
Metal sheets	92 (98%)	91 (100%)	
Water source			0.156
Tap water (piped through house or at yard)	92 (98%)	83 (91%)	
Public standpipe	1 (1.1%)	1 (1.1%)	
Spring	1 (1.1%)	4 (4.4%)	
Pond-water or stream	0 (0%)	2 (2.2%)	
Other	0 (0%)	1 (1.1%)	
Toilet facility			<0.001
Flush to piped sewage or septic tank	63 (67%)	18 (20%)	
Pour flush latrine	30 (32%)	43 (47%)	
Pit latrine	1 (1.1%)	30 (33%)	
Cooking place			<0.001
In a separate room used as kitchen	59 (63%)	14 (15%)	
In a separate building used as kitchen	32 (34%)	77 (85%)	
In a room used for living or sleeping	3 (3.2%)	0 (0%)	
Total number of milk cows			0.036
None	77 (82%)	68 (75%)	
1-4	11 (12%)	22 (24%)	
5-9	5 (5.3%)	1 (1.1%)	
10+	1 (1.1%)	0 (0%)	
Total number of other cattle			0.004
None	93 (99%)	81 (89%)	
1-4	1 (1.1%)	10 (11%)	
5-9	0 (0%)	0 (0%)	
10+	0 (0%)	0 (0%)	
Total number of horses			>0.999
None	94 (100%)	91 (100%)	
1-4	0 (0%)	0 (0%)	
5-9	0 (0%)	0 (0%)	

10+	0 (0%)	0 (0%)	
Total number of goats			<0.001
None	86 (91%)	37 (41%)	
1-4	3 (3.2%)	35 (38%)	
5-9	3 (3.2%)	17 (19%)	
10+	2 (2.1%)	2 (2.2%)	
Total number of sheep			0.273
None	92 (98%)	86 (95%)	
1-4	2 (2.1%)	5 (5.5%)	
5-9	0 (0%)	0 (0%)	
10+	0 (0%)	0 (0%)	
Total number of chicken/poultry			<0.001
None	39 (41%)	23 (25%)	
1-4	3 (3.2%)	13 (14%)	
5-9	5 (5.3%)	20 (22%)	
10+	47 (50%)	35 (38%)	
Agricultural land (hectares)			<0.001
None	47 (50%)	67 (74%)	
1-4	29 (31%)	23 (25%)	
5-9	7 (7.4%)	1 (1.1%)	
10+	11 (12%)	0 (0%)	
Connected to electricity	94 (100%)	60 (66%)	<0.001
Working radio	90 (96%)	74 (81%)	0.002
Working television	92 (98%)	44 (48%)	<0.001
Working computer	76 (81%)	5 (5.5%)	<0.001
Working refrigerator	84 (89%)	7 (7.7%)	<0.001
Working rechargeable battery or generator	41 (44%)	31 (34%)	0.183
An iron (charcoal/electric)	91 (97%)	51 (56%)	<0.001
Watch	92 (98%)	45 (49%)	<0.001
Mobile phone	94 (100%)	90 (99%)	0.492
Bicycle	41 (44%)	7 (7.7%)	<0.001
Motorcycle	29 (31%)	13 (14%)	0.007
Animal drawn cart	2 (2.1%)	0 (0%)	0.497
Car or truck	70 (74%)	11 (12%)	<0.001
Boat with a motor	1 (1.1%)	0 (0%)	>0.999
Ugali (maize stiff porridge) (×/week)			<0.001
0	10 (11%)	0 (0%)	
1	39 (41%)	5 (5.5%)	
2-4	40 (43%)	65 (71%)	
≥5	5 (5.3%)	21 (23%)	

Plantain (×/week)			0.134
0	17 (18%)	7 (7.7%)	
1	63 (67%)	68 (75%)	
2-4	8 (8.5%)	12 (13%)	
≥5	6 (6.4%)	4 (4.4%)	
Banana (×/week)			0.001
0	5 (5.3%)	0 (0%)	
1	52 (55%)	32 (35%)	
2-4	27 (29%)	45 (49%)	
≥5	10 (11%)	14 (15%)	
Rice (×/week)			<0.001
0	0 (0%)	0 (0%)	
1	3 (3.2%)	35 (38%)	
2-4	22 (24%)	45 (49%)	
≥5	68 (73%)	11 (12%)	
Missing	1	0	
Potatoes (×/week)			<0.001
0	2 (2.1%)	5 (5.5%)	
1	25 (27%)	67 (74%)	
2-4	31 (33%)	17 (19%)	
≥5	36 (38%)	2 (2.2%)	
Meat (×/week)			0.007
0	0 (0%)	0 (0%)	
1	8 (8.5%)	14 (15%)	
2-4	44 (47%)	56 (62%)	
≥5	42 (45%)	21 (23%)	
Fish (×/week)			<0.001
0	2 (2.1%)	0 (0%)	
1	61 (65%)	39 (43%)	
2-4	20 (21%)	45 (49%)	
≥5	11 (12%)	7 (7.7%)	
Beans/peas (×/week)			<0.001
0	5 (5.3%)	0 (0%)	
1	12 (13%)	40 (44%)	
2-4	28 (30%)	37 (41%)	
≥5	49 (52%)	13 (14%)	
Missing	0	1	
Green vegetables (×/week)			0.347
0	0 (0%)	0 (0%)	
1	6 (6.5%)	4 (4.4%)	

2-4	29 (31%)	21 (23%)	
≥5	58 (62%)	65 (72%)	
Missing	1	1	
Fruits (×/week)			<0.001
0	0 (0%)	0 (0%)	
1	3 (3.2%)	27 (30%)	
2-4	24 (26%)	18 (20%)	
≥5	67 (71%)	46 (51%)	
Locally brewed beer (×/week)			<0.001
0	68 (73%)	41 (46%)	
1	22 (24%)	13 (14%)	
2-4	0 (0%)	10 (11%)	
≥5	3 (3.2%)	26 (29%)	
Missing	1	1	

$N = 185$ participants. Values represent number of participants (percentage of total). Comparisons between locations were performed using Fisher's exact or chi-squared tests. All variables ($n = 38$ variables), after mode imputation, were used to construct the lifestyle score. See **Supplementary Figure 2**.

Supplementary Table 3 | Baseline characteristics of the rural-like urbanites compared to other urban living individuals ($N = 69$).

Variable	Overall, $N = 69$	Rural-like Urban, $N = 21$	Urban, $N = 48$	p-value
Female sex	36 (52%)	6 (29%)	30 (63%)	0.009
Age	23.0 (21.2, 24.6)	23.1 (21.8, 24.6)	22.9 (21.2, 24.5)	0.676
Age categories				>0.999
18-25	55 (80%)	17 (81%)	38 (79%)	
26-36	14 (20%)	4 (19%)	10 (21%)	
BMI	23.1 (20.4, 26.4)	23.8 (20.7, 26.3)	22.8 (19.8, 26.5)	0.825
BMI classification				0.575
<18.5	6 (8.7%)	3 (14%)	3 (6.3%)	
18.5-24.9	42 (61%)	11 (52%)	31 (65%)	
25.0-29.9	16 (23%)	6 (29%)	10 (21%)	
>30	5 (7.2%)	1 (4.8%)	4 (8.3%)	
Systolic blood pressure (mmHg)	110 (102, 120)	110 (107, 120)	111 (100, 121)	0.481
Diastolic blood pressure (mmHg)	70 (63, 76)	70 (60, 77)	70 (64, 75)	0.958
Hemoglobin level g/dl	14.30 (12.60, 15.60)	14.90 (14.20, 15.80)	13.65 (12.43, 15.28)	0.022
Random blood sugar, mmol-1^^	5.30 (4.90, 6.00)	5.30 (5.10, 5.60)	5.30 (4.90, 6.00)	0.759
Highest level of education				0.593
Primary	2 (2.9%)	1 (4.8%)	1 (2.1%)	
Secondary	38 (55%)	10 (48%)	28 (58%)	

College	14 (20%)	4 (19%)	10 (21%)	
University	15 (22%)	6 (29%)	9 (19%)	
Insurance status	60 (87%)	18 (86%)	42 (88%)	>0.999
Occupation				0.690
Farming	10 (14%)	2 (9.5%)	8 (17%)	
Elementary occupation	7 (10%)	2 (9.5%)	5 (10%)	
Employed/business owner	49 (71%)	17 (81%)	32 (67%)	
Other	3 (4.3%)	0 (0%)	3 (6.3%)	
Received yellow fever vaccine	55 (80%)	15 (71%)	40 (83%)	0.332

$N = 69$ urban living participants. Values represent number of participants (percentage of total) and median (interquartile range [IQR]) for categorical and continuous variables, respectively. Comparisons between locations were performed using Fisher's exact, chi-squared and Mann–Whitney U-test for categorical and continuous variables, respectively.