

1 **ONLINE SUPPLEMENT**

2 **Ethics approval**

3 Ethics approval for the study has been obtained from the London School of Hygiene & Tropical
4 Medicine (LSHTM) ethics committee (ref: 9776) and in all five study centres. Informed
5 consent was obtained from all participants or their parents/carers before taking part.

6

7 Avon Longitudinal Study of Parents and Children (ALSPAC): Ethical approval for the UK-arm
8 of the study was obtained from the ALSPAC Ethics and Law Committee, and the Local
9 Research Ethics Committees. Consent for biological samples has been collected in accordance
10 with the Human Tissue Act (2004). REDcap was used to collect the ALSPAC data
11 (<http://projectredcap.org/resources/citations/>). Informed consent for the use of data collected
12 via questionnaires and clinics was obtained from participants following the recommendations
13 of the ALSPAC Ethics and Law Committee at the time. The ALSPAC study website contains
14 details of all the data that is available through a fully searchable data dictionary and variable
15 search tool (<http://www.bristol.ac.uk/alspac/researchers/our-data/>).

16

17 **Recruitment**

18 Participants with and without asthma were recruited from the same sources, which varied
19 between sites. For the UK, both asthmatic and non-asthmatic participants were recruited
20 through the ongoing ALSPAC study. For Brazil, Ecuador, and New Zealand, recruitment of
21 participants with and without asthma was through both existing cohort studies[1, 2] and
22 additional community recruitment (usually from surveys in schools). In Uganda, participants
23 were recruited from a larger case-control study of asthma identified through a cross-sectional
24 survey in schools [3].

25

26 **Sputum induction**

27 Sputum induction was performed when participants were stable. Aerosolised hypertonic saline
28 (4.5% w/v) was produced using an ultrasonic nebuliser (DeVilbiss Ultraneb 2000, Langen,
29 Germany) and administered orally through a mouthpiece (Hans-Rudolph Inc, Kansas City,
30 USA) for increasing intervals from 0.5-4 minutes, to a total of 15.5 minutes. Spirometry was
31 conducted between intervals, and salbutamol was administered if forced expiratory volume in
32 1 second (FEV₁) dropped to $\leq 75\%$ -predicted. Participants were subsequently encouraged to
33 produce sputum in a sterile plastic container. Sputum was separated into aliquots, with one
34 stored unprocessed at -80°C for microbiota analysis, and others processed for inflammatory
35 phenotyping.

36

37 **Sputum inflammatory phenotyping**

38 Inflammatory phenotyping was performed as described previously [4], using well-
39 characterised protocols [5]. Sputum plugs were dispersed using dithiothreitol (Sputasol, Oxoid
40 Ltd, Hampshire, England), filtered through a 60 μm filter (Millipore, County Cork, Ireland),
41 centrifuged at $400 \times g$, and resuspended in phosphate buffered saline (PBS). The resulting cell
42 suspension used to prepare cytospin slides stained using a Diff-QuikVR fixative and stain set
43 (Dade Behring, Deerfield, IL). Sputum slides were read in Wellington, New Zealand, with the
44 exception of the slides produced in Brazil (which could not be shipped overseas due to ethical
45 restrictions): these were therefore read in Brazil, with a sample of slides being remotely
46 checked (using microscopy images) by the group in Wellington.

47

48 **Sputum DNA extraction**

49 All samples underwent DNA extraction at a single facility, with extraction batching including
50 samples from at least three countries to minimise batch effects. For each sample, 100 mg of

51 unprocessed induced sputum was suspended in 300 μ L of PBS, vortexed for 10 seconds, and
52 placed on ice for 2 min. Samples were pelleted by centrifugation at 13,000 \times g for 10 min.
53 Supernatant was removed and the pellet was resuspended in 300 μ L of Tris-EDTA solution (10
54 mM Tris-HCl, 1 mM EDTA; pH 8.0; Ambion, ThermoFisher Scientific, Victoria, Australia),
55 200 mg of silica: zirconium beads (1:1 of 0.1 mm and 1.0 mm; Biospec Products, Inc., OK,
56 USA), and a single chrome bead (3.2 mm, Biospec Products, Inc., OK, USA) added. Samples
57 underwent bead-beating at 6.5 m/s for 60 sec in a FastPrep®-24 Instrument (MP Biomedicals,
58 CA, USA). Homogenised samples were heated to 95°C for 1 min, before being cooled on ice
59 for 1 min. Lysozyme (ROCHE, ThermoFisher Scientific, Victoria, Australia) and lysostaphin
60 (Sigma-Aldrich, MO, USA) were then added to a final concentration of 2 mg/mL and 0.1
61 mg/mL, respectively, and samples incubated at 37°C for 1 hr. Proteinase K (Fermentas,
62 ThermoFisher Scientific, Victoria, Australia) and sodium dodecyl sulphate (Sigma-Aldrich,
63 MO, USA) were then added to a final concentration of 1.2 mg/mL and 1.5 %, w/v, respectively.
64 Following incubation at 30 min at 56°C, 40 μ L of 5 M sodium chloride and 450 μ L of
65 phenol:chloroform:isoamyl alcohol (25:24:1; saline buffered at pH8.0; Sigma-Aldrich, MO,
66 USA) were added and samples vortexed for 30 sec. The aqueous-organic layers were separated
67 by centrifugation at 13,000 \times g for 10 min at 4°C and 400 μ L of the aqueous layer was
68 transferred to a new microfuge tube. DNA was recovered using an EZ-10 Spin column in
69 accordance with manufacturer's instructions (Bio Basic, Inc., Ontario, Canada), following
70 precipitation by the addition of 10 M ammonium acetate and 99% ethanol (Sigma Aldrich, MO,
71 USA) in a 1:10 and 1:1 ratio with sample volume, respectively. DNA was eluted in 100 μ L
72 UltraPure DNase/RNase-free distilled water (Gibco, ThermoFisher Scientific, Victoria,
73 Australia) and stored at -80°C prior to analysis.

74

75 **Quantitative PCR**

76 Abundance of all bacteria (16S), *Haemophilus influenzae* (*smpB*), *Moraxella catarrhalis*
77 (*copB*) were measured by quantitative PCR (qPCR) as described previously [6, 7], using the
78 primers and cycling conditions in **Table E1**. All assays were performed on the QuantStudio 6
79 Flex System (Thermo Fisher Scientific, Vic, Australia). Reactions were performed in triplicate
80 and averages taken. Gene copy numbers was calculated against a standard curve of a known
81 bacterial concentration and normalised per g of sputum.

82

83 **16S rRNA gene amplicon sequencing**

84 The V1-3 hypervariable region of the bacterial 16S rRNA gene was amplified from sputum
85 DNA using modified primers 27F (5'-
86 TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGAGRGTGGCTCAG-3')
87 and 519R (5'-
88 GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGTNTTACNGCGGCKGCTG-3'),
89 with Illumina adapter overhang sequences as indicated by underline. Amplicons were
90 generated, cleaned, indexed and sequenced according to the Illumina MiSeq 16S Metagenomic
91 Sequencing Library Preparation protocol with certain modifications. Briefly, an initial PCR
92 reaction contained at least 12.5 ng of DNA, 5 µL of forward primer (1 µM), 5 µL of reverse
93 primer (1 µM) and 12.5 µL of 2 × KAPA HiFi Hotstart ReadyMix (KAPA Biosystems,
94 Wilmington, MA, USA) in a total volume of 25 µL. The PCR reaction was performed on a
95 Veriti 96-well Thermal Cycler (Life Technologies) using the following program: 95 °C for 3
96 min, followed by 25 cycles of 95 °C for 30 sec, 55 °C for 30 sec and 72 °C for 30 sec and a
97 final extension step at 72 °C for 5 min. Samples were multiplexed using a dual-index approach
98 with the Nextera XT Index kit (Illumina Inc., San Diego, CA, USA) according to the
99 manufacturer's instructions. The final library was paired-end sequenced at 2 × 300 bp using a

100 MiSeq Reagent Kit v3 on the Illumina MiSeq platform. Sequencing was performed at the South
101 Australian Genomics Centre.

102

103 **Bioinformatic processing**

104 Demultiplexed sequences were processed using QIIME2 [8] (release 2021.11). Trimmed
105 paired-end reads were merged, chimeric sequences removed, denoised sequencing errors
106 corrected using the DADA2 plugin [9]. Representative sequences were aligned to the SILVA
107 database (v138) at 80% using vsearch and unassigned sequences were filtered out. Remaining
108 unique amplicon sequence variants (ASVs) were classified using the QIIME2 sklearn
109 algorithm to the SILVA database at 99% sequence similarity. ASVs that were amplified in the
110 blank extraction control were examined. Contaminant or spurious taxa, not associated with the
111 human microbiota were filtered out. The median clean and filtered read count was 4398 (IQR:
112 3086, 5861). Samples were rarefied to 1463 reads, as determined by alpha rarefaction plot. The
113 resulting ASVs were used to calculate α -diversity (Faith's phylogenetic diversity, Shannon's
114 diversity, and taxon richness), while the genus-level taxonomic assignment was converted to
115 relative abundances.

116

117 **Data analysis and visualisation**

118 Analysis of dissimilarity in square-root transformed, rarefied, genus-level microbiota
119 composition (β -diversity) was performed by permutational multivariate analysis of variance
120 (PERMANOVA) using Bray–Curtis distances, with the *adonis2* function from the 'vegan' R
121 package (v2.5-7) [10], with 9999 permutations. To assess whether unbalanced designs
122 impacted the PERMANOVA findings, assessments were repeated in Primer (v6.1.16;
123 PRIMER-E, Plymouth, United Kingdom) using a Type III model that accounts for unbalanced
124 designs. Differences in α -diversity metrics, bacterial load, and relative abundance of specific

125 taxa were calculated using ordinal logistic regression using SAS Studio (v3.81; SAS Institute
126 Inc., NC, USA). Differences in taxon relative abundance between groups of three or more
127 (countries and inflammatory phenotypes) were performed by linear discriminant analysis
128 (LDA) Effect Size (LEfSe) [11], using one-against-all multi-class analysis, and cut-offs of LDA
129 ≥ 3 and $p < 0.05$. Findings from LEfSe were confirmed by Wilcoxon signed rank test.

130

131 Microbiota dissimilarity between individuals was visualised using non-metric
132 multidimensional scaling (nMDS), calculated using the 'vegan' R package. All data were
133 visualised using GraphPad Prism (v10.0.2) except for nMDS plots and taxa bar plots, which
134 were visualised using the 'ggplot2' R package (v3.3.5).

135

136 **Supplementary Table 1:** Primers and cycling conditions for quantitative PCR

Target	Primers (5'-3')	Cycling conditions	Ref
16S	F: TCCTACGGGAGGCAGCAGT R: GGACTACCAGGGTATCTAATCCTGTT	95 °C for 15 s, 60 °C for 1 min	[12]
<i>smpB</i>	F: ATTAAATGTTGCATCAACGC R: GACTTTTGCCCACGCAC Probe: FAM- ACGRTTTTACCATAGTTGCACTTTCTC-BHQ	95 °C for 10 s, 63 °C for 1 min	[13]
<i>copB</i>	F: GTGAGTGCCGCTTTTACAACC R: TGTATCGCCTGCCAAGACAA	95 °C for 15 s, 60 °C for 1 min	[14]

137

138

139 **Supplementary Table 2:** Demographic and asthma characteristics of WASP and WASP-
 140 biome

	WASP	WASP-biome
All (n)	920	488
Female, n (%)	526 (57.2%)	256 (52.5%)
Age (years), median (IQR)	15.57 (11.76-17.93)	14.11 (11.41-17.35)
Asthma (n)	658	364
Female, n (%)	387 (58.8%)	198 (54.4%)
Age (years), median (IQR)	15.34 (11.94-17.93)	14.08 (11.17-17.15)
BMI (kg/m ²), median (IQR)	20.80 (18.12-24.22) (n=655)	20.44 (17.87-23.49) (n=361)
ACQ7 score, median (IQR)	0.50 (0.00-1.17) (n=631)	0.50 (0.0-1.17) (n=356)
ACQ7 level, n (%)	(n=632)	(n=356)
Well controlled	503 (79.6%)	290 (81.5%)
Uncontrolled	129 (20.4%)	66 (18.5%)
ICS use past 12 months, n (%)	255 (43.1%) (n=592)	118 (36.9%) (n=320)
BA use past 12 months, n (%)	402 (67.3%) (n=597)	202 (62.5%) (n=323)
FEV ₁ (% predicted), mean (STD)	94.10 (12.46) (n=628)	95.00 (12.27) (n=350)
FVC (% predicted), mean (STD)	98.48 (12.98) (n=628)	99.07 (12.30) (n=350)
Atopy, n (%)	432 (67.2) (n=643)	225 (63.0%) (n=357)
Inflammatory phenotype, n (%)		
Eosinophilic	239 (36.3%)	131 (36.0%)
Neutrophilic	68 (10.3%)	43 (11.8%)
Paucigranulocytic	328 (49.9%)	172 (47.3%)
Mixed granulocytic	23 (3.5%)	18 (4.9%)
Sputum neutrophils (%), median (IQR)	15.85 (5.50-39.86)	17.71 (6.09-47.50)
Sputum eosinophils (%), median (IQR)	1.45 (0.00-6.48)	1.50 (0.25-6.28)

141 IQR: Interquartile range, STD: Standard deviation, BMI: Body mass index, ACQ7: Asthma
 142 control questionnaire, ICS: Inhaled corticosteroids, BA: Beta-agonist, FEV₁: Forced expiratory
 143 volume in 1 second, FVC: Forced vital capacity

144

145

146 **Supplementary Table 3:** Permutational multivariate analysis of variance (PERMANOVA)
 147 output assessing the effect of country on microbiota composition in participants with asthma
 148 and participants without asthma

	R ² (%)	Pseudo-F	P(perm)	Significance
Univariate: Asthma (n=364)				
Age	2.03	7.51	<0.001	***
Sex	0.63	2.30	0.010	**
Country	7.75	7.54	<0.001	***
Multivariate: Asthma (n=364)				
Age	0.96	3.77	<0.001	***
Sex	0.43	1.67	0.069	
Country	7.75	7.61	<0.001	***
Brazil vs Ecuador	1.88	2.81	<0.001	***
Brazil vs NZ	4.55	8.90	<0.001	***
Brazil vs Uganda	4.87	6.09	<0.001	***
Brazil vs UK	4.37	3.79	<0.001	***
Ecuador vs NZ	4.75	10.80	<0.001	***
Ecuador vs Uganda	6.00	9.46	<0.001	***
Ecuador vs UK	4.52	5.30	<0.001	***
NZ vs Uganda	5.78	11.54	<0.001	***
NZ vs UK	4.19	6.65	<0.001	***
Uganda vs UK	5.56	4.94	<0.001	***
Multivariate: Non-asthma (n=124)				
Age	2.21	2.96	0.015	*
Sex	1.09	1.46	0.12	
Country	9.29	3.11	<0.001	***
Brazil vs Ecuador	5.20	1.93	0.035	*
Brazil vs NZ	3.71	2.58	0.007	**
Brazil vs Uganda	9.42	1.39	0.136	
Brazil vs UK	9.78	2.71	0.005	**
Ecuador vs NZ	6.16	5.82	<0.001	***
Ecuador vs Uganda	4.63	1.74	0.080	
Ecuador vs UK	10.37	5.61	<0.001	***
NZ vs Uganda	2.10	1.44	0.15	
NZ vs UK	2.59	2.10	0.024	*
Uganda vs UK	7.08	1.99	0.030	*

149 Significance codes: *** $P < 0.001$, ** $P < 0.01$, * $P < 0.05$.

150

151 **Supplementary Table 4:** Permutational multivariate analysis of variance (PERMANOVA)
152 output assessing the effect of asthma on microbiota composition

	R² (%)	Pseudo-F	P(perm)	Significance
Multivariate: All (n=488)				
Age	1.22	6.43	<0.001	***
Sex	0.35	1.86	0.038	*
Asthma	0.43	2.27	0.012	*
Country	6.74	8.86	<0.001	***

153

154

155 **Supplementary Table 5:** Permutational multivariate analysis of variance (PERMANOVA)
 156 output assessing the effect of inflammatory phenotypes on microbiota composition of only
 157 participants with asthma (top), as well as all participants, adjusting for asthma (bottom).

	R ² (%)	Pseudo-F	P(perm)	Significance
Multivariate: Asthma (n=364)				
Inflammatory phenotype	1.71	2.26	<0.001	***
Eosinophilic vs Paucigranulocytic	0.69	2.25	0.015	*
Eosinophilic vs Neutrophilic	1.90	3.50	<0.001	***
Eosinophilic vs Mixed	0.65	1.04	0.40	
Paucigranulocytic vs Neutrophilic	1.52	3.56	<0.001	***
Paucigranulocytic vs Mixed	0.49	1.00	0.43	
Neutrophilic vs Mixed	0.87	0.56	0.87	
Age	0.92	3.64	<0.001	***
Sex	0.45	1.80	0.042	*
Country	7.56	7.49	<0.001	***
Multivariate: All (n=488)				
Inflammatory phenotypes	1.69	2.23	<0.001	***
Eosinophilic vs No asthma	1.06	2.93	0.003	**
Paucigranulocytic vs No asthma	0.44	1.40	0.14	
Neutrophilic vs No asthma	1.97	3.59	<0.001	***
Mixed vs No asthma	0.69	1.06	0.40	
Age	1.19	6.30	<0.001	***
Sex	0.39	2.04	0.027	*
Country	6.60	8.72	<0.001	***

158 Significance codes: *** $P < 0.001$, ** $P < 0.01$, * $P < 0.05$.

159

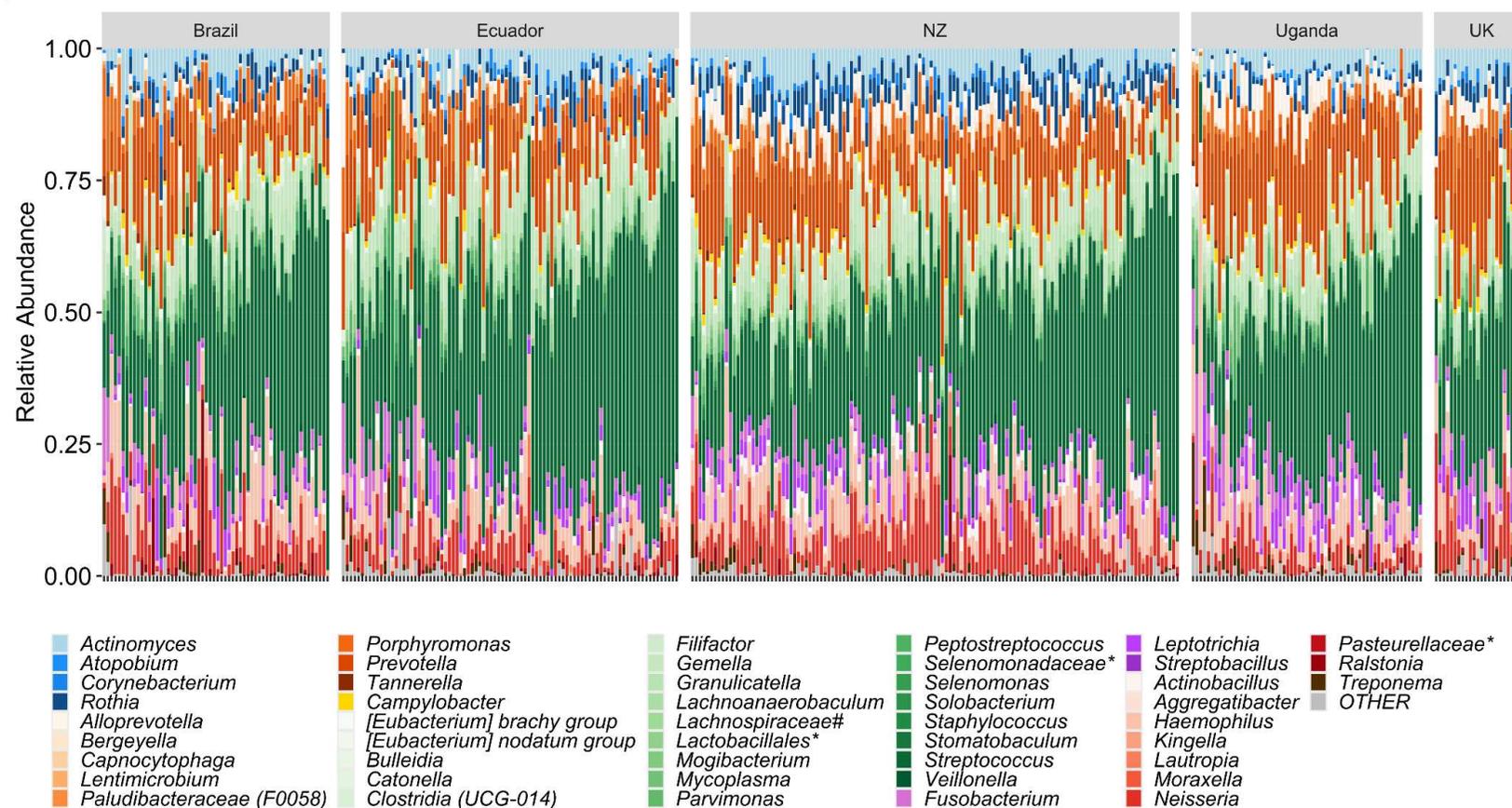
160 **Supplementary Table 6:** Permutational multivariate analysis of variance (PERMANOVA)
 161 output assessing the effect of neutrophil% and eosinophil% on microbiota composition, within
 162 country, adjusting for age and sex.#

	R ² (%)	Pseudo-F	P(perm)	Significance
Brazil				
Neutrophil%	2.96	1.81	0.057	
Eosinophil%	3.08	1.88	0.051	
Age	2.02	1.24	0.24	
Sex	2.10	1.29	0.21	
Ecuador				
Neutrophil %	1.09	0.97	0.46	
Eosinophil%	0.64	0.57	0.88	
Age	2.76	2.46	0.005	**
Sex	1.14	1.02	0.42	
New Zealand				
Neutrophil%	2.51	3.39	<0.001	***
Eosinophil%	0.10	1.36	0.17	
Age	4.04	5.46	<0.001	***
Sex	0.68	0.92	0.50	
Uganda				
Neutrophil%	2.96	1.78	0.066	
Eosinophil%	1.20	0.72	0.67	
Age	0.89	0.64	0.79	
Sex	1.50	0.92	0.48	

163 Significance codes: *** $P < 0.001$, ** $P < 0.01$, * $P < 0.05$.

164 #Analysis not performed for UK

165

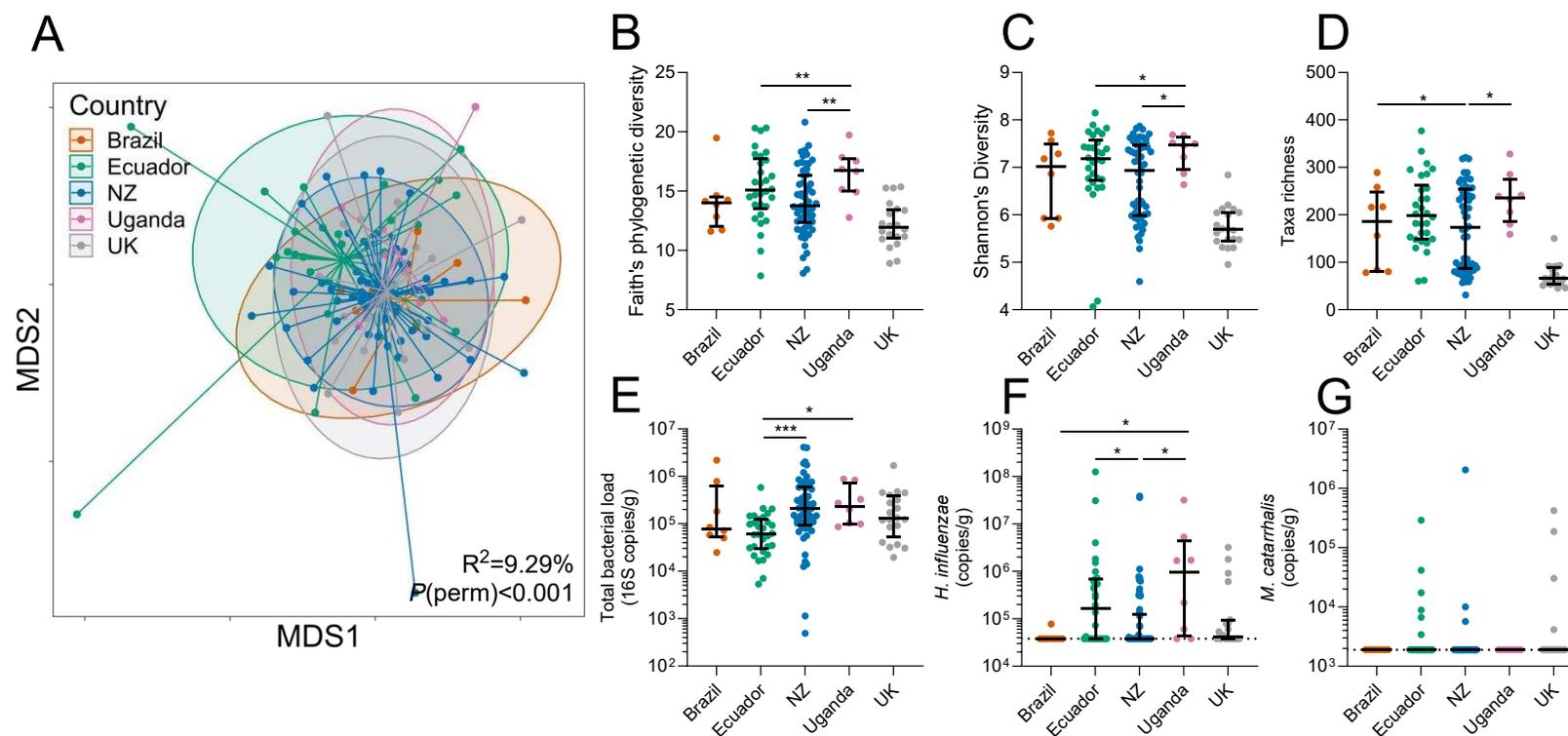


166

167 **Supplementary Figure 1: Taxa bar plot of WASP-biome.** Relative abundance distribution of taxa present in $\geq 10\%$ of individuals within a
 168 country. Taxa coloured by phylogeny where blue = Actinobacteriota (n=4), orange = Bacteroidota (n=8), gold = Campilobacterota (n=1), green =
 169 Bacillota (n=22), purple = Fusobacteriota (n=3), red = Pseudomonadota (n=9), brown = Spirochaetes (n=1). * unassigned; # uncultured.

13

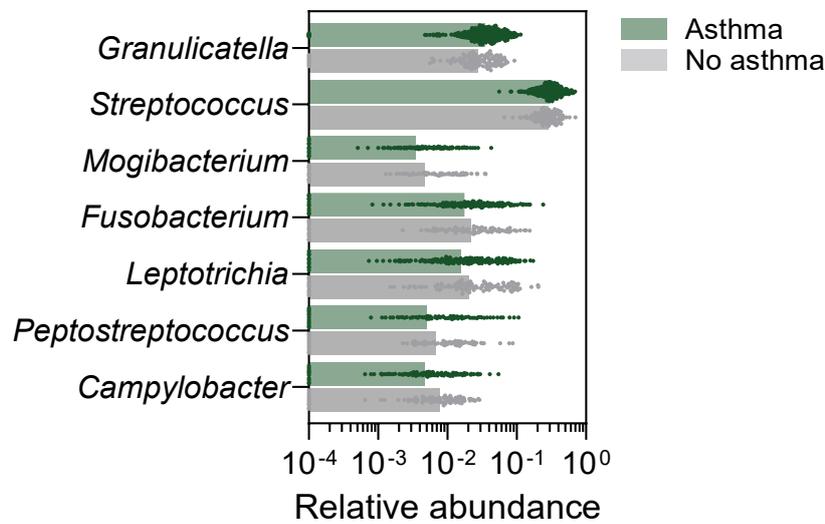
170



171

172 **Supplementary Figure 2: Sputum microbiota differs by country in non-asthmatics.** A) Nonmetric multidimensional scaling (NMDS) plot of
 173 Bray-Curtis dissimilarity. B) Shannon's diversity. C) Faith's phylogenetic diversity. D) Taxa richness. E) Total bacterial load derived from qPCR.
 174 F) *Haemophilus influenzae* abundance derived from qPCR. G) *Moraxella catarrhalis* abundance derived from qPCR. Statistics: A) Permutational
 175 multivariate analysis of variance including variables: country, age and sex; B-G) Ordinal logistic regression including variables: country, age and
 176 sex; *** adjusted $P<0.001$, ** adjusted $P<0.01$, * adjusted $P<0.05$.

14

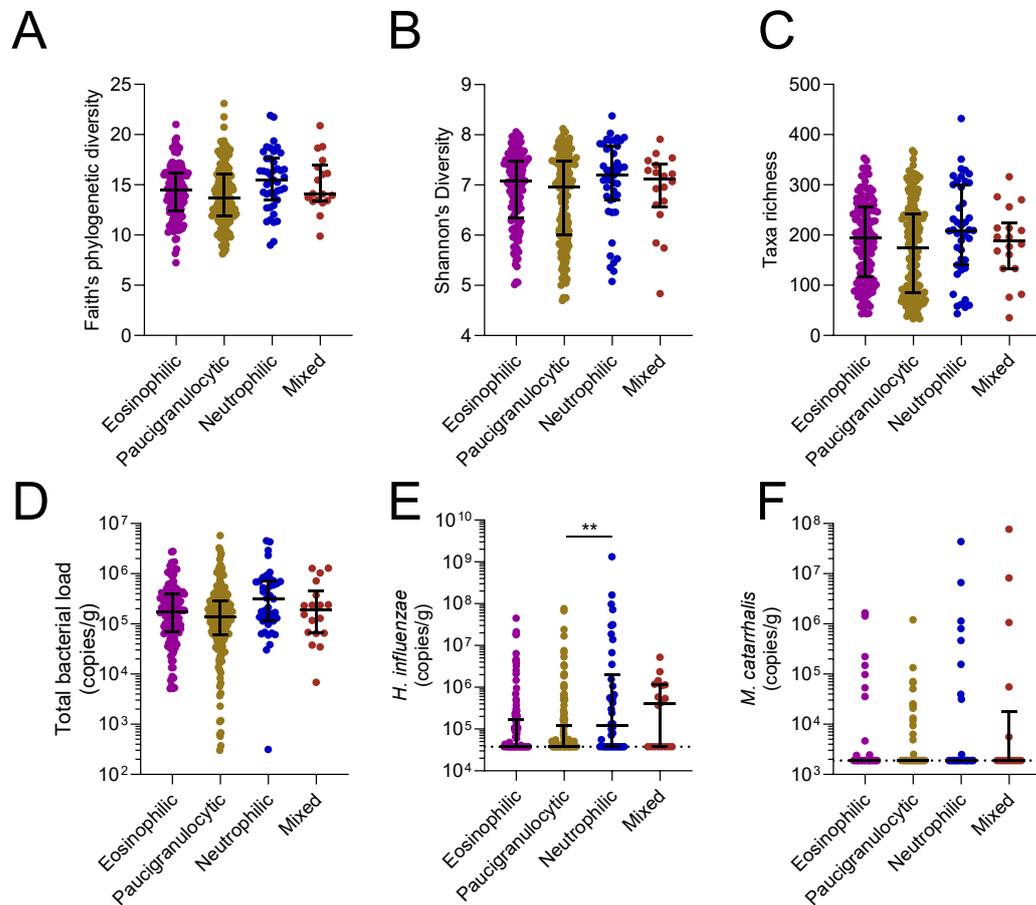


177

178 **Supplementary Figure 3: Taxa that differed significantly by asthma status.** Distribution of
179 the relative abundance of taxa either significantly higher or lower in asthmatics. Analysis
180 performed by ordinal logistic regression including variables: asthma, country, age and sex.

181

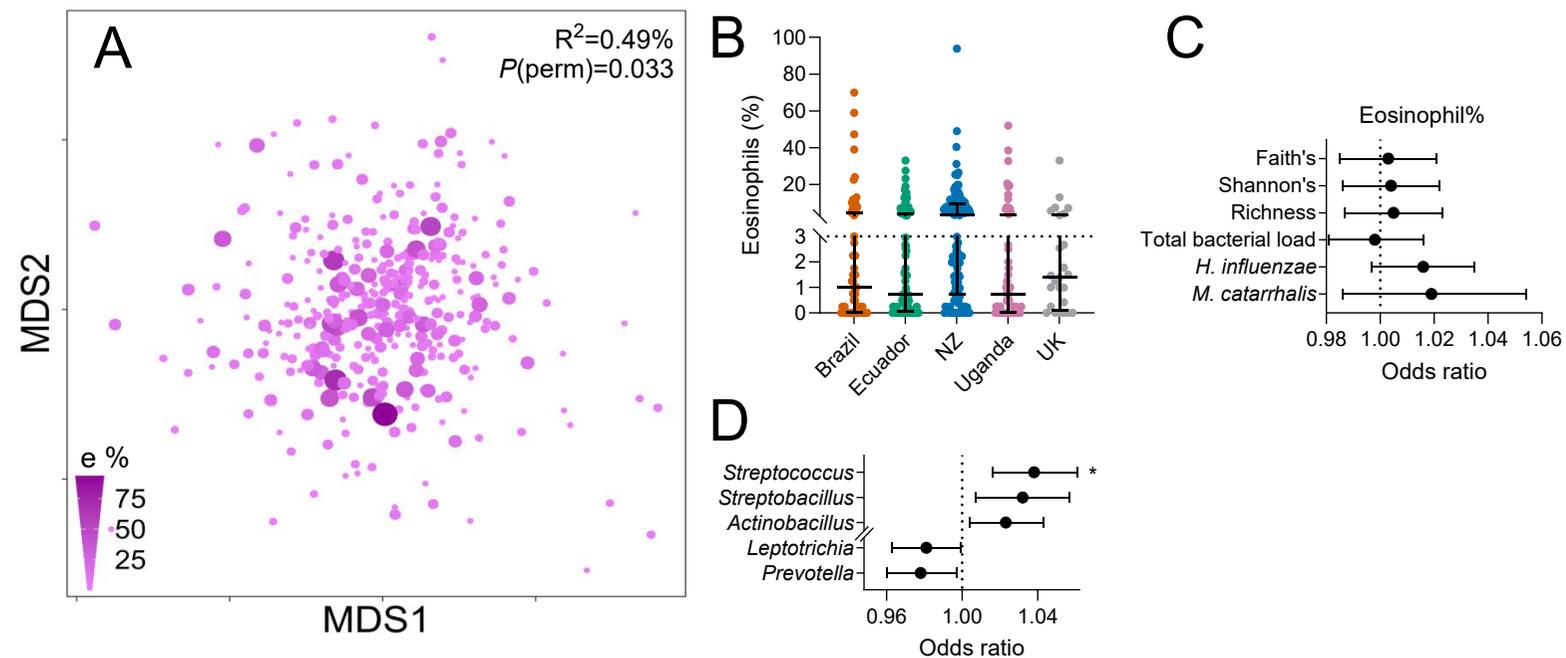
182



183

184 **Supplementary Figure 4: Sputum microbiota by inflammatory phenotype in asthmatics.**
 185 A) Faith's phylogenetic diversity. B) Shannon's diversity. C) Taxa richness. D) Total bacterial
 186 load derived from qPCR. E) *Haemophilus influenzae* abundance derived from qPCR. F)
 187 *Moraxella catarrhalis* abundance derived from qPCR. Statistics: Ordinal logistic regression
 188 including variables: inflammatory phenotype, country, age and sex; ** adjusted $P < 0.01$.

16

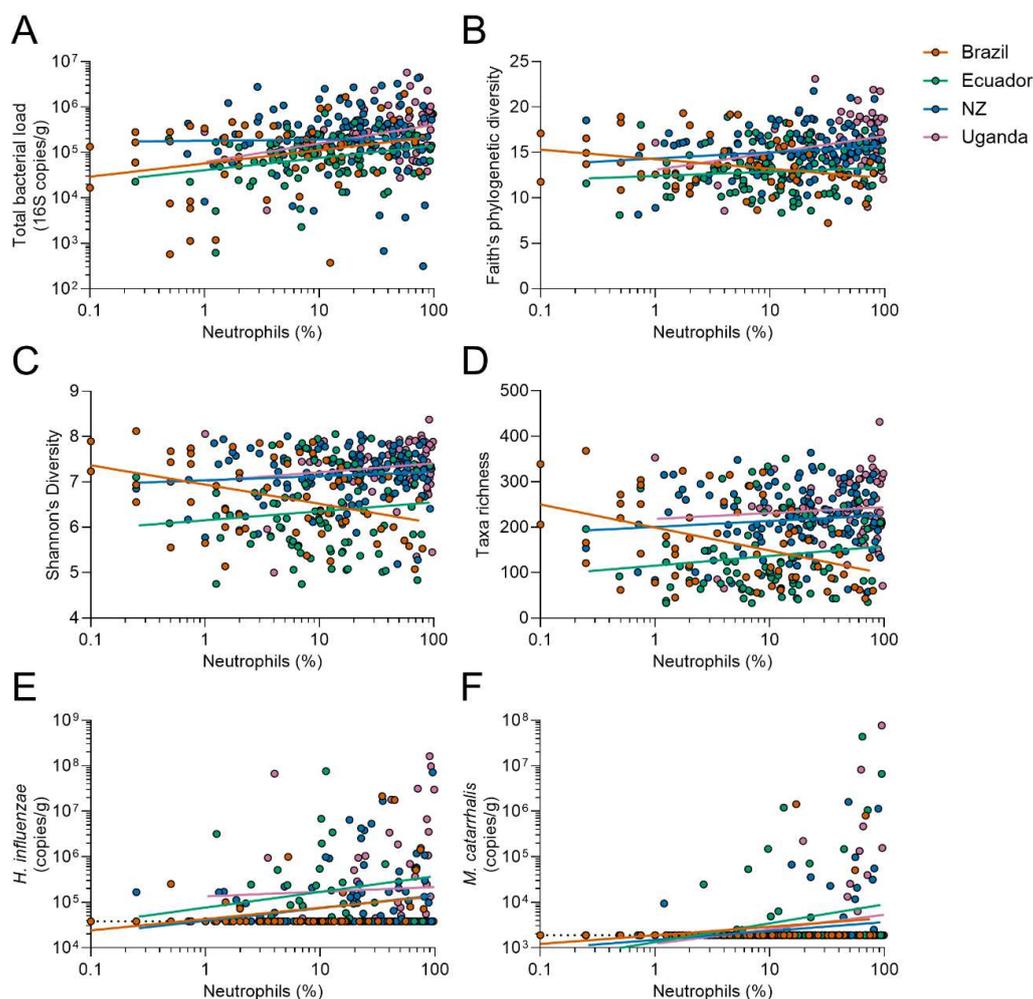


189

190 **Supplementary Figure 5: Sputum microbiota and eosinophilic phenotype in asthma.** A) Nonmetric multidimensional scaling (NMDS) plot
 191 of Bray-Curtis dissimilarity showing dispersion by eosinophil%. B) Distribution of sputum eosinophil% by country. C) Forest plot showing
 192 eosinophil% associated with α -diversity (Faith's, Shannon's, richness) and qPCR derived bacterial load (total, and *H. influenzae* and *M. catarrhalis*
 193 specific). D) Forest plot showing taxa that differed significantly by eosinophil%. Statistics: A) Permutational multivariate analysis of variance
 194 including variables: neutrophil%, eosinophil%, country, age and sex; C, D) Ordinal logistic regression including variables: neutrophil%,
 195 eosinophil%, country, age and sex. * adjusted $P < 0.05$.

17

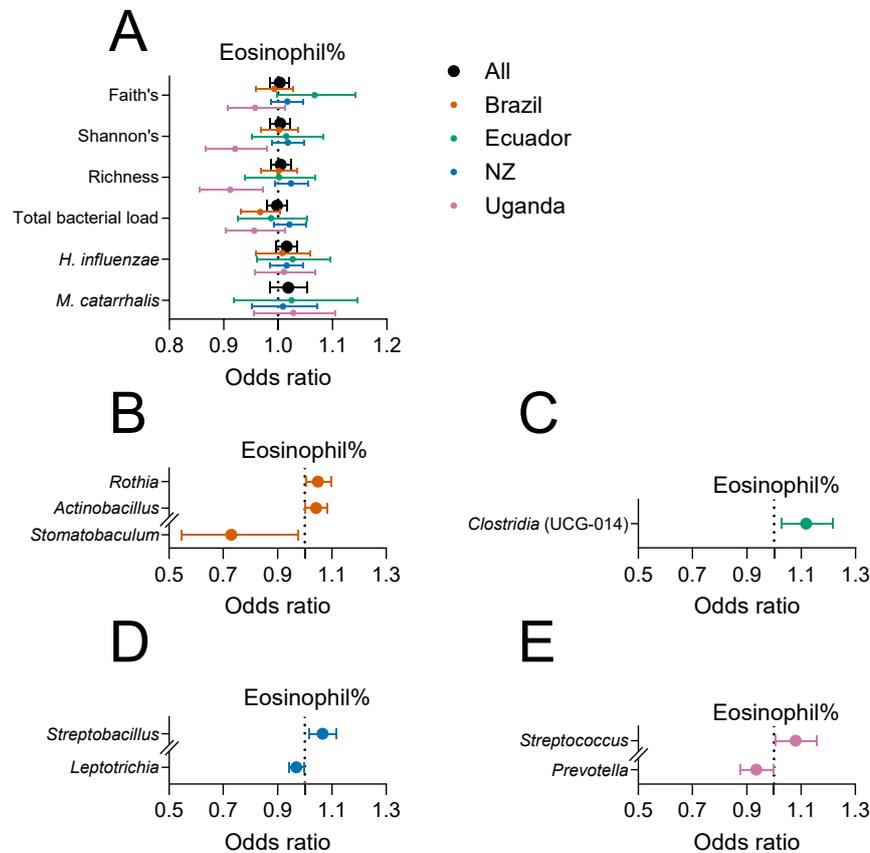
196



197

198 **Supplementary Figure 6: Within-country scatter plots between sputum microbiota**
 199 **characteristics and neutrophil%.** A) Total bacterial load derived from qPCR. B) Faith's
 200 phylogenetic diversity. C) Shannon's diversity. D) Taxa richness. E) *Haemophilus influenzae*
 201 abundance derived from qPCR. F) *Moraxella catarrhalis* abundance derived from qPCR.
 202 Linear regression lines included for visualisation. Analysis not performed for the UK due to
 203 low sample size.

18



204

205 **Supplementary Figure 7: Within-country association between sputum microbiota and**
 206 **eosinophil%.** A) Forest plot showing eosinophil% associated with α -diversity (Faith's,
 207 Shannon's, richness) and qPCR derived bacterial load (total, and *H. influenzae* and *M.*
 208 *catarrhalis* specific) in whole cohort (black), and within Brazil (orange), Ecuador (green), New
 209 Zealand (blue), and Uganda (pink). B) Forest plot showing taxa that differed by eosinophil%
 210 in Brazil. C) Forest plot showing taxa that differed by eosinophil% in Ecuador. D) Forest plot
 211 showing taxa that differed by eosinophil% in New Zealand. E) Forest plot showing taxa that
 212 differed by eosinophil% in Uganda. Statistics: Ordinal logistic regression including variables:
 213 neutrophil%, eosinophil%, age and sex; All analysis was not significant following false
 214 discovery correction. *M. catarrhalis* analysis was not performed for Brazil due to low detection
 215 frequency (3 out of 60 participants).

216

217

218 **REFERENCES**

- 219 1. Barreto ML, Cunha SS, Alcantara-Neves N, et al. Risk factors and immunological
220 pathways for asthma and other allergic diseases in children: Background and methodology of
221 a longitudinal study in a large urban center in northeastern brazil (salvador-scaala study).
222 *BMC Pulm Med* 2006; 6: 15.
- 223 2. Epton MJ, Town GI, Ingham T, et al. The new zealand asthma and allergy cohort
224 study (nza2cs): Assembly, demographics and investigations. *BMC Public Health* 2007; 7: 26.
- 225 3. Mpairwe H, Namutebi M, Nkurunungi G, et al. Risk factors for asthma among
226 schoolchildren who participated in a case-control study in urban uganda. *Elife* 2019; 8.
- 227 4. Pembrey L, Brooks C, Mpairwe H, et al. Asthma inflammatory phenotypes on four
228 continents: Most asthma is non-eosinophilic. *Int J Epidemiol* 2023; 52(2): 611-623.
- 229 5. Gibson PG. Use of induced sputum to examine airway inflammation in childhood
230 asthma. *J Allergy Clin Immunol* 1998; 102(5): S100-101.
- 231 6. Taylor SL, Leong LEX, Choo JM, et al. Inflammatory phenotypes in patients with
232 severe asthma are associated with distinct airway microbiology. *J Allergy Clin Immunol*
233 2018; 141(1): 94-103 e115.
- 234 7. Taylor SL, Leong LEX, Ivey KL, et al. Total bacterial load, inflammation, and
235 structural lung disease in paediatric cystic fibrosis. *J Cyst Fibros* 2020; 19(6): 923-930.
- 236 8. Bolyen E, Rideout JR, Dillon MR, et al. Reproducible, interactive, scalable and
237 extensible microbiome data science using qiime 2. *Nat Biotechnol* 2019; 37(8): 852-857.
- 238 9. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJ, Holmes SP. Dada2:
239 High-resolution sample inference from illumina amplicon data. *Nat Methods* 2016; 13(7):
240 581-583.
- 241 10. Dixon P. Vegan, a package of r functions for community ecology. *Journal of*
242 *Vegetation Science* 2003; 14(6): 927-930.
- 243 11. Segata N, Izard J, Waldron L, et al. Metagenomic biomarker discovery and
244 explanation. *Genome Biol* 2011; 12(6): R60.
- 245 12. Nadkarni MA, Martin FE, Jacques NA, Hunter N. Determination of bacterial load by
246 real-time pcr using a broad-range (universal) probe and primers set. *Microbiology* 2002;
247 148(1): 257-266.
- 248 13. Reddington K, Schwenk S, Tuite N, et al. Comparison of established diagnostic
249 methodologies and a novel bacterial smpb real-time pcr assay for specific detection of
250 haemophilus influenzae isolates associated with respiratory tract infections. *J Clin Microbiol*
251 2015; 53(9): 2854-2860.
- 252 14. Greiner O, Day PJR, Altwegg M, Nadal D. Quantitative detection of moraxella
253 catarrhalis in nasopharyngeal secretions by real-time pcr. *J Clin Microbiol* 2003; 41(4): 1386-
254 1390.

255