Active trachoma cases in the Solomon Islands have varied polymicrobial community structures but do not associate with individual non-chlamydial pathogens of the eye.

Robert M R Butcher[1]*, Oliver Sokana[2], Kelvin Jack[2], Eric Kalae[4], Leslie Sui[4], Charles Russell[5], Joanna Houghton[1], Christine Palmer[1], Martin J Holland[1], Richard T Le Mesurier[6], Anthony W Solomon[1,3], David C W Mabey[1,3], Chrissy h. Roberts[1].

Author contact details:

[1] Clinical Research Department, London School of Hygiene & Tropical Medicine, Keppel Street, London WC1E 7HT, UK

[2] Eye Department, Solomon Island Ministry of Health and Medical Services, Honiara, Solomon Islands

[3] Hospital for Tropical Diseases, University College London Hospitals, Mortimer Market, London

[4] Primary Care Department, Lata Hospital, Lata, Santa Cruz, Solomon Islands

[5] Bellona Rural Health Centre, Bellona, Solomon Islands

[6] The Fred Hollows Foundation, Carlton, Victoria, Australia

Abstract

Background

Several non-chlamydial microbial pathogens are associated with clinical signs of active trachoma in trachoma endemic communities with a low prevalence of ocular *Chlamydia trachomatis* (*Ct*) infection. We observed a low prevalence of *Ct* in the Solomon Islands, therefore hypothesised that the high prevalence of active trachoma could be explained by a common non-chlamydial infection or by a dominant polymicrobial community dysbiosis.

Methods

We studied DNA from conjunctival swabs collected from 257 Solomon Islanders with active trachoma and 257 matched controls. Droplet digital PCR was used to test for *Adenoviridae, Haemophilus influenzae, Moraxella catarrhalis, Staphylococcus aureus,* coagulase-negative *Staphylococcus* and *Streptococcus pneumoniae*. Polymicrobial community diversity and composition were studied by sequencing of hypervariable regions of the 16S ribosomal ribonucleic acid gene in a subset of 54 cases and 53 controls.

Results

Although *Ct* was associated with active trachoma, the number of infections was low (cases: 3.9%, controls: 0.4%). Estimated prevalence (cases, controls) of each nonchlamydial infection was as follows: *S. aureus* (1.9%, 1.9%), *Adenoviridae* (1.2%, 1.2%), coagulase-negative *Staphylococcus (5.8%, 4.3%), H. influenzae* (7.4%, 11.7%)*, M. catarrhalis* (2.3%, 4.7%) and *S. pneumoniae* (7.0%, 6.2%). There was no statistically significant association between clinical signs of trachoma and presence or load of any of the non-*Ct* infections that were assayed. Inter-individual variations in the conjunctival microbiome were characterised by differences in the levels of *Corynebacterium*, *Proprionibacterium*, *Helicobacter* and *Paracoccus*, but diversity and relative abundance of these specific genera did not differ between cases and controls.

Discussion

It is unlikely that the prevalent trachoma-like follicular conjunctivitis in the Solomon Islands has a dominant bacterial aetiology. Before implementing community-wide azithromycin distribution for trachoma, policy makers should consider that clinical signs of trachoma can occur without the involvement of any azithromycin-susceptible organism.

Introduction

Trachoma, caused by *Chlamydia trachomatis* (*Ct*), is the leading infectious cause of preventable blindness (1), and is targeted for elimination as a public health problem by 2020 through the SAFE strategy (**S**urgery, **A**ntibiotics, **F**acial cleanliness and **E**nvironmental improvement). The decision to implement community-wide trachoma control interventions, which include mass drug administration (MDA) with azithromycin, is based on population prevalence estimates of one clinical sign of active trachoma (trachomatous inflammation–follicular [TF]) in the 1–9-year-old age group (2). In the Solomon Islands, the prevalence of active trachoma is sufficient to warrant MDA but the prevalence of TT suggests trachoma may not pose a significant public health problem. A recent study of trachoma in a treatment-naïve population of the Solomon Islands estimated the TF prevalence at 26% but, surprisingly, conjunctival *Ct* infection was detected in only 1.3% of 1–9 year olds (3). Whilst *Ct* infection is not always detectable in TF cases (4, 5), infection prevalence in the Solomon Islands is far lower than is seen in other countries with similar TF prevalence (3).

Ct is not the only cause of follicular conjunctivitis. *Moraxella lacunata*, Parinaud's oculo-glandular syndrome, atopic conjunctivitis (6), adenovirus (7), and other *Chlamydia* serotypes (8) and species (9), among others (10), have been suggested as possible causes of TF in trachoma-endemic populations. A greater diversity of pathogens can be cultured from the conjunctivae of people with TF (10) and, among bacterial species isolated during a study in Tanzanian children, *Streptococcus pneumoniae*, *Haemophilus influenzae* B and *Haemophilus influenzae* non-type B associated more strongly than *Ct* with clinical signs of active trachoma (10). In communities which had received MDA in The Gambia, *S. pneumoniae* and *H. influenzae* type b infection correlated closely with signs of active trachoma, whereas *Moraxella catarrhalis* and *Staphylococcus aureus* did not (11). Common non-chlamydial pathogens are also associated with trachomatous scarring (TS) (12) and recurrence of trachomatous trichiasis (TT) after surgery (13).

Although traditionally thought to be 'sterile', several studies have now described polymicrobial communities colonising the conjunctival epithelium (14, 15). *Staphylococcus*, *Streptococcus*, *Haemophilus* and *Moraxella* genera can be readily cultured from swabs taken from the inferior fornix (10–12). Attempts to detect bacteria from the conjunctiva have shown community diversity and composition to vary significantly between individuals. It is unclear whether a 'core' microbiota persists at the conjunctiva, but profiles closely related to that of the skin have been reported (16). *Corynebacterium*, *Propionibacterium*, *Staphylococcus* and *Streptococcus* have been consistently dominant, whereas other genera such as *Acinetobacter, Brevundimonas, Pseudomonas, Bradyrhizobium, Sphingomonas, Bacillus, Simonsiella* and *Elizabethkingia* have been identified more sporadically (17–19). Conjunctival polymicrobial community composition is known to vary with age and season (20) and appears to be responsive to external stimuli such as regular contact lens wear (21), however, significant associations with disease have yet to be described (20, 22).

We hypothesised that clinical signs of active trachoma in the Solomon Islands where ocular *Ct* is uncommon could be explained by a common non-chlamydial infection or by a dominant polymicrobial community dysbiosis.

Methods

Study ethics

Ethical approval was granted by the London School of Hygiene & Tropical Medicine (6319/6360) and the Solomon Islands National Health Research (HRC13/18) Ethics Committees. A parent or guardian provided written consent for each child participant.

110 **Table 1.** Oligonucleotides used in this study, with assay concentrations derived from *in vitro* optimisation.

Organism / target	Oligo	Sequence (5'-3')	Concentration in final assay (nM)	Amplicon size	Ref
C. trachomatis / *Plasmid open reading frame 2 (pORF2)*	F	CAG CTT GTA GTC CTG CTT GAG AGA	900	109	(3, 23, 24)
	R	CAA GAG TAC ATC GGT CAA CGA AGA	900		
	Probe	[FAM] CCC CAC CAT TTT TCC GGA GCG A [BHQ1]	200		
S. aureus / Staphylococcal protein A (*SpA*)	F	CAG CAA ACC ATG CAG ATG CTA	900	101	(25, 26)
	R	CGC TAA TGA TAA TCC ACC AAA TAC A	900		
	Probe	[VIC] TCA AGC ATT ACC AGA AAC [MGBNFQ]	250		
Coagulase-negative *Staphylococcus* / translation elongation factor (*tuf*)	F	TAT CCA CGA AAC TTC TAA AAC AAC TGT TAC T	450	204	(25)
	R	TCT TTA GAT AAT ACG TAT ACT TCA GCT TTG AAT TT	450		
	Probe	[FAM] TAT TAG ACT ACG CTG AAG CTG GTG ACA ACA T [BHQ1]	125		
S. pneumoniae / N-acetylmuramoyl-L-alanine amidase (*lytA*)	F	ACG CAA TCT AGC AGA TGA AGC A	500	75	(27)
	R	TCG TGC GTT TTA ATT CCA GCT	500		
	Probe	[FAM] GCC GAA AAC GCT TGA TAC AGG GAG [BHQ1]	150		
H. influenzae / L-fuculokinase (*fucK*)†	F	ATG GCG GGA ACA TCA ATG A	900	102	(28)
	R	ACG CAT AGG AGG GAA ATG GTT	900		
	Probe*	[FAM] CGg TAa TTg GGa TCc AT [BHQ1]	125		
Adenoviridae / capsid assembly protein (hexon)	F	GCC ACG GTG GGG TTT CTA AAC TT	500	132	(29)
	R	GCC CCA GTG GTC TTA CAT GCA CAT C	500		
	Probe	[HEX] TGC ACC AGA CCC GGG CTC AGG TAC TCC GA [BHQ1]	125		
M. catarrhalis / outer membrane protein (*copB*)	F	GTG AGT GCC GCT TTT ACA ACC	900	72	(30, 31)
	R	TGT ATC GCC TGC CAA GAC AA	900		
	Probe	[HEX] TGC TTT TGC AGC TGT TAG CCA GCC TAA [BHQ1]	250		

111 * Lower case symbols denote locked nucleic acid bases

112 † These primers do not differentiate between encapsulated and noncapsulated subtype

113 Study population

114

115 The samples tested during this study were a sub-set of specimens from a population-based
116 trachoma prevalence survey of 3674 people (1135 children aged 1–9 years) in 32 clusters from
117 Temotu and Rennell & Bellona Provinces, Solomon Islands; data from that survey are presented
118 elsewhere (3). Sterile swabs were passed three times over the right conjunctiva of every child
119 before storage in 300 μL RNALater® (Life Technologies). Swabs were kept cool in the field for up
120 to 24hrs then frozen (3). Inclusion criteria for cases (n = 257) were age 1–9 years, detectable
121 human DNA in the conjunctival swab and a field grade of TF or trachomatous inflammation—
122 intense (TI) in the right eye. An equal number (n = 257) of age, gender and island matched controls
123 without TF or TI were selected using the 'e1071' R package. A random subset of cases (n = 54) and
124 controls (n = 53) were further characterized using V1-V3 16S rRNA gene sequencing. Field controls
125 (clean swabs passed within 20 cm of seated participants, without touching them, and then treated
126 identically to subjects' specimens), extraction controls (extraction carried out in the absence of a
127 specimen and PCR carried out on the eluate) and PCR controls (PCR carried out in the absence of
128 any added material) were used to determine background levels of microbial contamination in the
129 16S rRNA gene sequencing protocol.

130

131 Droplet digital PCR

132

133 All individuals in this sample have previously been tested for *Ct* and human RPP30 using DNA
134 extracted from conjunctival swabs with the Qiagen AllPrep DNA/RNA kit. We chose not to use
135 mechanical lysis on these low biomass specimens due to the reported lower yields compared to
136 chemical lysis (32).

137

138 Duplex ddPCR assays were developed for (1) *Adenoviridae* (29) and *S. pneumoniae* (27), (2) *H.*
139 *influenzae* (28) and *M. catarrhalis* (30, 31), and (3) *S. aureus* and coagulase-negative
140 *Staphylococcus* (25, 26) based on published assays. Assay performance was assessed by repeat

testing of synthetic standard dilution series' and cultured pathogen material before rolling out to clinical samples (Supplementary Table 1). Each assay contained 10 μL 2× ddPCR Supermix for probes (Bio-rad, Hemel Hempstead, UK), primers and probes at custom concentrations (Table 1) and 8 μL of template DNA. Thermal cycling was 10'00'' at 95°C then 40x (0'15'' at 95°C | 1'00'' at 60°C) then 12'00'' at 98°C.(23) A positive result for a clinical specimen was defined as >95% confidence of non-zero target load, as described previously.(23)

16S rRNA gene sequencing

An approximately 530-bp region of the 16S ribosomal RNA (rRNA) gene (variable regions 1-3) was amplified using forward (modified 27F; 5'-[adaptor]-AGAGTTTGGATCCTGGCTCAG-3') and custom barcoded reverse primers (534R; 5'-[adaptor]-[barcode]-AGTCAGTCAGCCATTACCGCGGCTGCTGG-3'). Each 15 μL reaction contained 7.5 μL 2x Phusion High Fidelity Master Mix (New England Biosciences, MA, USA), 0.45 μL DMSO, 0.1 μM primers and 5.55 μL DNA. The thermal cycling conditions were 0'30'' at 98°C then 31x (0'10'' at 98°C | 0'30'' at 62°C | 0'15'' at 72°C) then 7'00'' at 72°C. Amplicons were cleaned using 0.6 v/v AMPure XP beads (Beckman Coulter, CA, USA), quantified using Qubit (Thermo Fisher Scientific, MA, USA) then pooled at equimolar concentrations. The 4 nM sequencing library was mixed 0.75 v/v with a 4 nM Phi-X control library (Illumina, CA, USA). 3 μL of 100 μM custom read primers were mixed in to wells 12 (Read 1; CTACACTATGGTAATTGTAGAGTTTGGATCCTGGCTCAG), 13 (Index; CCAGCAGCCGCGGTAATGGCTGACTGACT) and 14 (Read 2; AGTCAGTCAGCCATTACCGCGGCTGCTGG) of the MiSeq reagents cartridge before 2 x 300 bp paired-end sequencing with the 600-cycle MiSeq v3 sequencing reagent kit on the MiSeq platform using a standard protocol (Illumina, CA, USA). 96 uniquely barcoded specimens were run in multiplex on each MiSeq run.

Data analysis

168

169 ddPCR data were analysed using R version 3.2.2 (33). Binomial univariate regression was used to
170 test the relationship between each individual infection and active trachoma. The most accurate final
171 multivariate model, determined by lowest Akaike Information Criterion (AIC) value, was determined
172 by step-wise removal of variables from a binomial multivariate regression model incorporating all
173 tested pathogens. The chance of the *Ct* association occurring due to chance was assessed by
174 counting the number of significant results obtained from randomly re-ordering the *Ct* data 1000
175 times.

176

177 Raw 16S amplicon sequences were directly assigned to genera using Illumina BaseSpace '16S
178 Metagenomics' app version 1.0.1.0, which uses algorithms from Wang and colleagues (34). Two
179 clinical specimens yielded fewer than 1000 reads in total and were removed from the analysis.
180 Amplicons were sequenced from six no-template control (NTC) specimens (two 'field' controls, two
181 'extraction' controls, two 'PCR' controls, defined above) to identify any contaminants endogenous to
182 the collection, PCR or sequencing process. Any genera represented by more than 600 reads in the
183 six combined NTCs (average 100 reads per NTC) were eliminated from the clinical specimen
184 analysis. This resulted in the removal of 21 genera, which accounted for 95.1% of the reads from
185 NTCs and 88.3% of the reads from clinical specimens. The genera removed included common
186 contaminants of reagent kits (35) such as *Pelomonas*, as well as some previously described
187 conjunctival microbiota constituents such as *Brevundimonas*, *Staphylococcus* and *Streptococcus*
188 (17, 20).

189

190 Between-group differences in the Shannon Diversity Index (a measure of diversity which increases
191 with increasing species abundance and evenness) and Inverse Simpson Index (another measure of
192 diversity where 1 is infinite diversity and 0 is no diversity) were compared using a t-test.
193 Discriminant Analysis of Principal Components (DAPC), a multivariate clustering method to analyse
194 complex datasets, was performed using the 'adegenet' package in R (36) and cross-validation was

used to determine the optimal number of principal components to include in a discriminant function aimed at separating cases from controls on the basis of their polymicrobial community structures.

Results

Specimen set demographics

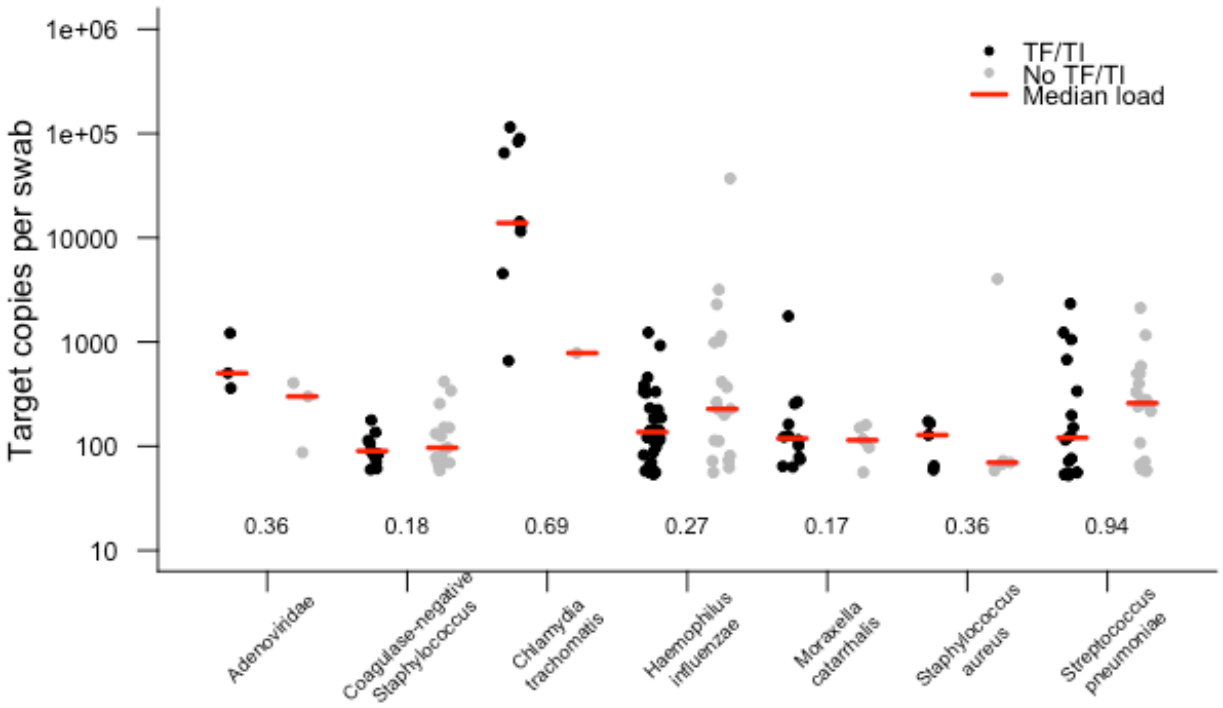
A total of 257 cases and 257 controls (n = 514) were tested. All 257 cases had TF, and one also had TI. Case status was defined by clinical signs in the right eye but respectively 8/257 (3.1%) and 236/257 (91.8%) of the controls and cases had TF in the left eye. Both case and control groups were 38% female. Mean age was 5.6 years (cases) and 5.5 years (controls, student's t test p = 0.76). There was no significant difference between cases and controls in terms of the clusters represented (Kolmogorov-Smirnov test p = 0.97) and all 32 clusters were represented. The case specimens had higher loads of human DNA than the controls (18030 versus 9354 copies/swab; p = 0.00003). There were no significant differences in age, gender or location within the subset selected for 16S-amplicon sequencing.

Quantitative PCR tests for ocular pathogens

In this study, 19.8% of children had evidence of infection with at least one of the targeted organisms (table 2). We considered the prevalence of *Ct* in this sample set to be too low (96.1% of TF cases were infection-negative) to account for the level of active disease. The prevalence of *Adenoviridae* (1.2%) and *S. aureus* (1.9%) were both very low. There was no association between active trachoma and infection with *H. influenzae, S. pneumoniae, S. aureus,* coagulase-negative *Staphylococcus* or *M. catarrhalis. Ct* infection was associated with active trachoma (logistic regression p = 0.026; odds ratio: 10.4). The association between *Ct* and active trachoma was still significant in a multivariate analysis (p = 0.025). The permuted p-value for the association between

223 *Ct* and active trachoma was highly significant (p = 0.001). Active disease was neither associated
224 with infection with at least one pathogen (any pathogen, p = 0.185) nor the number of concurrent
225 infections in the same eye (number of concurrent pathogens, p = 0.207). Infectious loads of those
226 who tested positive are shown figure 2. Despite some numerical differences between groups, none
227 of the differences were statistically significant.

228



229

230 **Figure 2.** Target copies per swab of each pathogen identified in conjunctival swabs collected from children with and
231 without active trachoma in the Solomon Islands. Numbers show p-values for logistic regression comparison between
232 active trachoma case and control groups for each pathogen. None of the differences observed were statistically
233 significant.

234

235 **Table 2.** Cases of infection in specimens from children with and without active disease. The relationship of those
236 infections to trachoma has been tested with univariate logistic regression, and then stepwise removal of variables from
237 multivariate regression model was used to determine the final multivariate regression model that provided the best fit for
238 the data.

Pathogen	ddPCR result	No TF/TI (%; n=257)	TF/TI (%; n=257)	Total (%; n=514)	Univariate odds ratio (95% CI)	Univariate p-value	Multivariate odds ratio (95% CI)	Multivariate p-value
Adenoviridae	Positive	3 (1.2)	3 (1.2)	6 (1.2)	1.00 (0.67 – 1.50)	1.000	-	-
	Negative	254 (98.8)	254 (98.8)	508 (98.8)				
*Chlamydia trachomatis**	Positive	1 (0.4)	10 (3.9)	11 (2.1)	10.36 (1.96 – 190.92)	**0.026**	10.64 (2.15 – 196.00)	**0.025**
	Negative	256 (99.6)	247 (96.1)	503 (97.9)				
Coagulase-negative *Staphylococcus*	Positive	15 (5.8)	11 (4.3)	26 (5.1)	0.72 (0.32 – 1.59)	0.423	-	-
	Negative	242 (94.2)	246 (95.7)	488 (94.9)				
Haemophilus influenzae	Positive	19 (7.4)	30 (11.7)	49 (9.5)	1.66 (0.91 – 3.07)	0.101	-	-
	Negative	238 (92.6)	227 (88.3)	465 (90.4)				
Moraxella catarrhalis	Positive	6 (2.3)	12 (4.7)	18 (3.5)	2.05 (0.78 – 5.96)	0.158	2.13 (0.81 – 6.20)	0.137
	Negative	251 (97.7)	245 (95.3)	496 (96.5)				
Staphylococcus aureus	Positive	5 (1.9)	5 (1.9)	10 (1.9)	1.00 (0.28 – 3.64)	1.000	-	-
	Negative	252 (98.1)	252 (98.1)	504 (98.1)				
Streptococcus pneumoniae	Positive	18 (7.0)	16 (6.2)	34 (6.6)	0.88 (0.43 – 1.77)	0.723	-	-
	Negative	239 (93.0)	241 (93.8)	480 (93.4)				
Any pathogen	Positive	45 (17.5)	57 (22.2)	102 (19.8)	0.75 (0.48 – 1.15)	0.185	Not included	
	Negative	212 (82.4)	200 (77.8)	412 (80.2)				
Different pathogen species in same eye	0	212 (82.4)	200 (77.8)	412 (80.2)	1.21 (0.90 – 1.63)	0.207	Not included	
	1	28 (10.9)	35 (13.6)	63 (12.3)				
	>1	17 (6.6)	22 (8.6)	39 (7.6)				

239 *Data taken from Butcher *et al.*(3)

240

241 16S rRNA gene sequencing

242

243 The total number of genera identified across all clinical specimens and NTCs was 659, however,
244 125/659 (19%) had a cumulative total of ten reads or fewer. The median number of reads per
245 clinical specimen was 55,104 (IQR: 40,705 – 89,541) and the median percentage of reads mapped
246 to genus level was 51.9%. The median number of genera identified was 151 (IQR: 117 – 212) per
247 clinical specimen. Following removal of presumed contaminants, approximately 40% of cleaned
248 reads were assigned to genera that each constituted less than 1% of the total polymicrobial
249 community. 40 genera were represented by at least 1% of remaining reads in clinical samples.

250

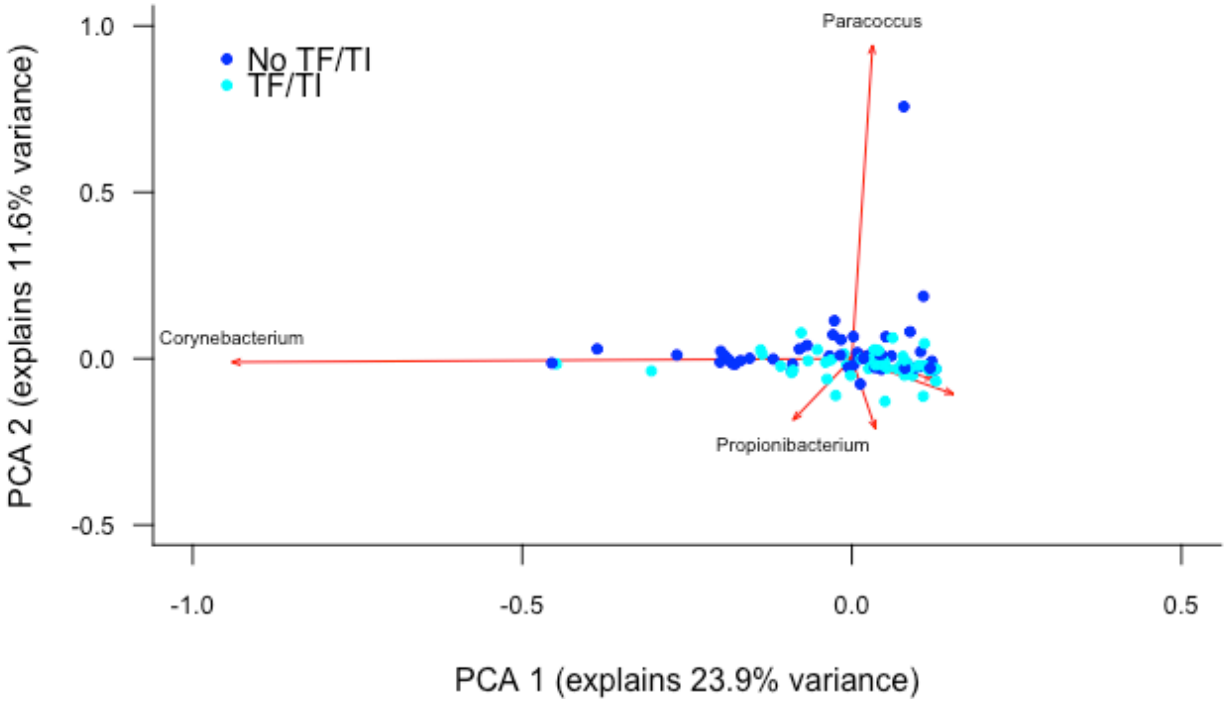
251 The diversity of bacterial communities was in general low; the mean Inverse Simpsons Diversity
252 index over all specimens was 0.061 (IQR: 0.046 – 0.092). There was no significant difference in
253 alpha-diversity between cases and controls by either measure of diversity. The median Inverse
254 Simpson's Diversity index in cases was 0.061 [IQR: 0.048 – 0.095] versus controls 0.060 [IQR:
255 0.044 – 0.078] (Student's t-test p = 0.56). The median Shannon Diversity Index was 3.38 in cases
256 and 3.37 in controls (Student's t-test p = 0.96). In cases of active trachoma, the most dominant
257 bacterial genera were *Corynebacterium* (12.0% of total reads), *Proprionibacterium* (6.2%) and
258 *Helicobacter* (4.8%). In controls, the most dominant genera were *Corynebacterium* (13.9%),
259 *Paracoccus* (5.2%), *Proprionibacterium* (4.7%), and *Neisseria* (4.1%) (supplementary figure 2). The
260 genus level membership of the bacterial community varied between cases and controls. Of 21
261 genera found in specimens from those with active trachoma, 10 were not found in controls,
262 including *Helicobacter*, *Mesoplasma, Brachybacterium* and *Haemophilus*. Of 23 genera found in
263 controls, 12 were not found in cases, including *Neisseria, Prevotella, Rhodococcus* and
264 *Porphyromonas*.

265

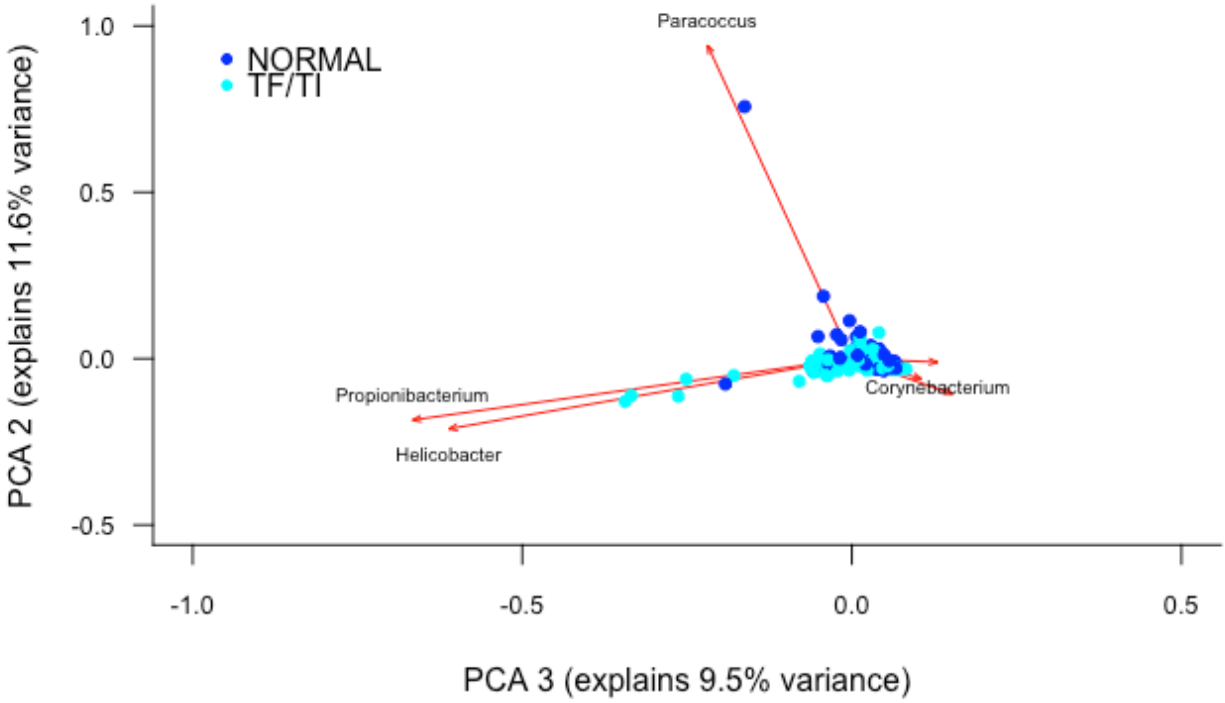
266 Principal Components (PC) Analysis (figure 2) revealed that 47% of the total variation in the
267 microbiomes of children in the Solomon Islands is dominated by *Corynebacterium* (PC1),

Paracoccus (PC2), *Propionibacterium* (PC2,PC3) and *Helicobacter* (PC3). While individual cases or controls have distinctive profiles dominated by these genera, the majority of cases and controls are indistinguishable (figure 2) using the first three PCs (explaining 45% of total variation). We condensed PCs 1-20 into a single discriminant function, which explained 92% of variation in the data. In this analysis, the genus-level polymicrobial community structure is significantly different between cases and controls (logistic regression $p = 0.0000013$) (figure 3). That discriminant function was dominated by a number of genera such as *Curvibacterium* and *Mesoplasma* (Supplementary Figure 4). Cross validation to predict group membership using discriminant functions of between 10 and 90 PCs had a low success rate (median 53.3% success, range: 50.1 – 62.3%). This suggests that while differences do exist between the polymicrobial communities in cases and controls, they are too subtle and varied to be predictive of phenotype, at least when working with this number of specimens.

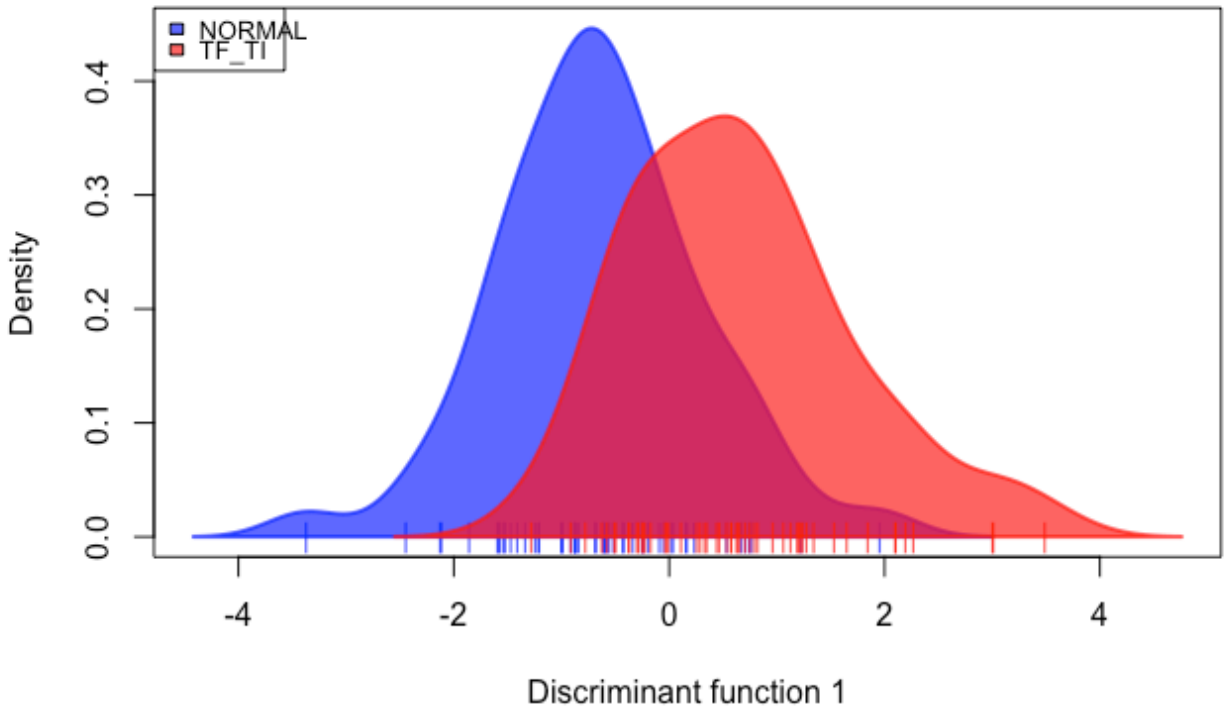
281 **A**



282 **B**



283 **Figure 2. (A)** First and second and **(B)** second and third principal components describing variation between the 16S
284 sequences identified in Solomon Island children with and without active trachoma. Dark blue spots indicate controls. Light
285 blue spots indicate cases. Red arrows show principal component loadings.



286

287 **Figure 3:** Discriminant analysis of the association of 20 combined principal components with active trachoma, showing a

288 significant discrimination of phenotype groups (p = 0.0000013).

289

290

Discussion

Follicular conjunctivitis meeting WHO criteria for the trachoma phenotype "TF" is highly prevalent in the Solomon Islands, but the prevalence of *Ct* infection is curiously low. Although *Ct* was the only pathogen to associate with TF in this study, the prevalence was much lower than may be expected of a population with this level of TF (3). By ddPCR testing for a number of bacteria and viruses that have previously been linked to TF, we can discount the possibility that any of them can account either singly or *en masse* for the high TF prevalence. Both *S. pneumoniae* and *H. influenzae* were detected at moderate prevalence in our population, whilst we found very few cases of *Adenoviridae* and *M. catarrhalis* infection. Comparator data from other populations are scarce, but it is clear that staphylococcal species were detected in the conjunctivae in the Solomon Islands at a substantially lower prevalence (7%) than in children in either Tanzania (14.8% *S. epidermidis*) (10), Sierra Leone (20% *S. aureus,* 29% coagulase-negative *Staphylococcus*) (37) or The Gambia (post MDA, 14.7% *S. aureus*) (11). Through 16S amplicon sequencing and community profiling, we have shown that there is apparently no dominant bacterial genus associated with this disease. The dominant features of variation in conjunctival bacterial communities in Solomon Islands children (*Corynebacterium, Propionibacterium)* have consistently been identified in other studies. Other genera (e.g., *Paracoccus*, *Helicobacter, Haemophilus)* have not previously been identified in 16S studies, whilst some reported in other studies (e.g., *Simonsella, Pseudomonas*) were not found in these specimens.(17–20) Many studies have sequenced the V3-4 region of the 16S gene whereas we targeted V1-3. V region choice has been shown to have an impact on the genera identified at other mucosal sites (38) and, while no data have yet emerged on how this affects profiles from the eye, it is possible this could account for some of the differences. Consistent with previous studies (35), we found a background of genera that amplified in NTCs that also appeared in clinical specimens. The microbiological biomass at the conjunctiva is known to be very low, even compared with other nearby sites such as skin or oral mucosa (18). When true resident bacteria are scarce, 16S rRNA gene PCR readily amplifies reagent contaminants. We took stringent measures to exclude reagent contaminants, but this resulted in the removal of some important genera such as

319 *Staphylococcus* and *Streptococcus.* Among the biggest contributors to between-conjunctiva
320 variation were *Corynebacterium* and *Propionibacterium*, but these genera did not differentiate TF
321 cases from controls (figure 2). Previous investigations of the role of the conjunctival microbiome in
322 ocular disease have also not shown significant differences. People with ocular manifestations of
323 rosacea, Sjögren's syndrome and healthy controls did not have significant differences in diversity or
324 relative abundance of key phyla identified across all three groups (22). In The Gambia, there was an
325 increased abundance of *Haemophilus* in cases of TF, as compared to controls. The abundance of
326 *Corynebacterium* and *Streptococcus* varied between scarred and non-scarred individuals. However,
327 neither trachoma-related difference was significant (20).
328
329 The microbiota of these children were heterogeneous and had subtle variations that appeared to
330 reach significant association with TF when the full community structure was considered in
331 comparative statistical models. It remains possible (though perhaps unlikely) that a multitude of
332 factors, operating at the level of single bacterial species, polymicrobial communities or viral species,
333 all contribute to the presentation of the phenotype. It is perhaps more likely that as yet unidentified
334 viral or allergic causes could explain the prevalence of TF in the Solomon Islands.
335
336 There are some limitations to our study. Although the ddPCR assays are based on validated assays
337 and appear to be accurate (supplementary table 1), they have not been formally evaluated in this
338 format. By not using mechanical lysis in our extraction process, the absolute prevalence of some
339 difficult-to-lyse Gram positive genera may have been under-estimated, although this would not
340 impair the comparison of cases to controls, between which protocols were consistent. 16S amplicon
341 sequencing of low biomass samples is known to result in amplification of reagent and environmental
342 contaminants (35). This study focused on gross differences in community structure between those
343 with and without disease and these should be independent of contaminating reads. We also
344 employed stringent quality control measures to ensure data were not confused by artefacts of the
345 sequencing process.
346

Public health scenarios such as the one in the Solomon Islands will become increasingly common as trachoma elimination programmes reduce the global prevalence of *Ct,* and the positive predictive value of TF for ocular Ct infection declines as a result (39). MDA might be inappropriately delivered when clinical signs of trachoma are the only indicator used for programmatic decision making. While MDA is effective for the treatment of trachoma (40), mass antibiotic exposure may increase macrolide resistance in nonchlamydial bacteria (41) and theoretically could make the population more susceptible to later infections by preventing accumulation of acquired immunity to *Ct* (42) and decisions to undertake such a program should therefore be carefully considered. The case for using tests for *Ct* infection during trachoma surveys is strengthened by our data. The value of nucleic acid amplification tests for detecting non-chlamydial infection remains questionable, as multiple infectious agents might be important, and these are likely to differ between populations. Key questions now include whether non-chlamydial follicular conjunctivitis such as this responds to MDA and whether it is linked to incident or progressive scarring which may lead to blindness.

Acknowledgements

We thank the residents of the communities who took part in this study, Victoria Miari and Emma Cobb (Pathogen Molecular Biology Department, London School of Hygiene & Tropical Medicine, UK) for providing bacterial culture isolates for PCR assay validation, and David Nelson (Indiana University, Indianapolis, US) for the provision of 16S sequencing primers and protocols.

Financial statement

Fieldwork was jointly funded by the United Kingdom's Department for International Development Global Trachoma Mapping Project grant (ARIES: 203145) to Sightsavers, and by the Fred Hollows Foundation, Australia (1041). Laboratory costs were funded by the Fred Hollows Foundation, Australia (1041).

RMRB and AWS were funded by a Wellcome Trust Intermediate Fellowship to AWS (098521).

OS, KJ, EK, LS and CR were employed by the Solomon Islands Ministry of Health and Medical Services for the duration of the survey.

MJH is supported by the Wellcome Trust (093368/Z/10/Z).

ChR is supported by the Wellcome Trust Institutional Strategic Support Fund (105609/Z/14/Z).

Author Contributions

Conceived and designed the study: RMRB, OS, RTLM, AWS, DCWM, ChR.

Performed the fieldwork: RMRB, OS, EK, KJ, LS, CR.

Performed the experiments: RMRB, MJH, JH, CP, ChR.

Analysed the data: RMRB, ChR.

Wrote the manuscript: RMRB, ChR.

Revised and approved the manuscript: RMRB, OS, EK, KJ, LS, CR, JH, CP, MJH, RTLM, AWS, DCWM, ChR.

References

1. Bourne RA, Stevens GA, White RA, Smith JL, Flaxman SR, Price H, Jonas JB, Keeffe J, Leasher J, Naidoo K, Pesudovs K, Resnikoff S, Taylor HR, Vision Loss Expert G. 2013. Causes of vision loss worldwide, 1990-2010: a systematic analysis. Lancet Glob Heal 1:e339-49.
2. World Health Organization. 2003. Report of the 2nd Global Scientific Meeting on Trachoma. 25-27 August. Geneva, Switzerland.

397 3. Butcher RMR, Sokana O, Jack K, Macleod CK, Marks ME, Kalae E, Sui L, Russell C, Tutill
398 HJ, Williams RJ, Breuer J, Willis R, Le Mesurier RT, Mabey DCW, Solomon AW, Roberts CH.
399 2016. Low Prevalence of Conjunctival Infection with Chlamydia trachomatis in a Treatment-
400 Naïve Trachoma-Endemic Region of the Solomon Islands. PLoS Negl Trop Dis 10:e0004863.
401 4. Thein J, Zhao P, Liu H, Xu J, Jha H, Miao Y, Pizzarello L, Tapert L, Schachter J, Mabon M,
402 Osaki-Holm S, Lietman T, Paxton A. 2002. Does clinical diagnosis indicate ocular chlamydial
403 infection in areas with a low prevalence of trachoma? Ophthalmic Epidemiol 9:263–9.
404 5. Grassly NC, Ward ME, Ferris S, Mabey DC, Bailey RL. 2008. The natural history of trachoma
405 infection and disease in a Gambian cohort with frequent follow-up. PLoS Negl Trop Dis
406 2:e341.
407 6. Dawson CR, Jones BR, Tarizzo ML, World Health Organization. 1981. Guide to trachoma
408 control in programmes for the prevention of blindness. Geneva, Switzerland.
409 7. Jhanji V, Chan TCYY, Li EYMM, Agarwal K, Vajpayee RB. 2015. Adenoviral
410 keratoconjunctivitis. Surv Ophthalmol 60:435–443.
411 8. Andersson P, Harris SR, Smith HMBS, Hadfield J, O'Neill C, Cutcliffe LT, Douglas FP, Asche
412 LV, Mathews JD, Hutton SI, Sarovich DS, Tong SYC, Clarke IN, Thomson NR, Giffard PM.
413 2016. Chlamydia trachomatis from Australian Aboriginal people with trachoma are
414 polyphyletic composed of multiple distinctive lineages. Nat Commun 7:10688.
415 9. Lietman T, Brooks D, Moncada J, Schachter J, Dawson C, Dean D. 1998. Chronic follicular
416 conjunctivitis associated with Chlamydia psittaci or Chlamydia pneumoniae. Clin Infect Dis
417 26:1335–40.
418 10. Burton MJ, Hu VH, Massae P, Burr SE, Chevallier C, Afwamba IA, Courtright P, Weiss HA,
419 Mabey DCW, Bailey RL. 2011. What is causing active trachoma? The role of nonchlamydial
420 bacterial pathogens in a low prevalence setting. Invest Ophthalmol Vis Sci 52:6012–6017.
421 11. Burr SE, Hart JD, Edwards T, Baldeh I, Bojang E, Harding-Esch EM, Holland MJ, Lietman
422 TM, West SK, Mabey DCW, Sillah A, Bailey RL. 2013. Association between ocular bacterial
423 carriage and follicular trachoma following mass azithromycin distribution in The Gambia.
424 PLoS Negl Trop Dis 7:e2347.

425 12. Hu VH, Massae P, Weiss HA, Chevallier C, Onyango JJ, Afwamba IA, Mabey DCW, Bailey
426 RL, Burton MJ. 2011. Bacterial infection in scarring trachoma. Invest Ophthalmol Vis Sci
427 52:2181–6.
428 13. Burton MJ, Adegbola RA, Kinteh F, Ikumapayi UN, Foster A, Mabey DCW, Bailey RL. 2007.
429 Bacterial infection and trachoma in the gambia: a case control study. Invest Ophthalmol Vis
430 Sci 48:4440–4.
431 14. Kugadas A, Gadjeva M. 2016. Impact of Microbiome on Ocular Health. Ocul Surf 14:342–9.
432 15. Lu LJ, Liu J. 2016. Focus: Microbiome: Human Microbiota and Ophthalmic Disease. Yale J
433 Biol Med 89:325.
434 16. Zhou Y, Gao H, Mihindukulasuriya KA, La Rosa PS, Wylie KM, Vishnivetskaya T, Podar M,
435 Warner B, Tarr PI, Nelson DE, Fortenberry JD, Holland MJ, Burr SE, Shannon WD,
436 Sodergren E, Weinstock GM. 2013. Biogeography of the ecosystems of the healthy human
437 body. Genome Biol 14:R1.
438 17. Dong Q, Brulc JM, Iovieno A, Bates B, Garoutte A, Miller D, Revanna K V, Gao X,
439 Antonopoulos DA, Slepak VZ, Shestopalov VI. 2011. Diversity of bacteria at healthy human
440 conjunctiva. Invest Ophthalmol Vis Sci 52:5408–13.
441 18. Doan T, Akileswaran L, Andersen D, Johnson B, Ko N, Shrestha A, Shestopalov V, Lee CS,
442 Lee AY, Van Gelder RN. 2016. Paucibacterial Microbiome and Resident DNA Virome of the
443 Healthy Conjunctiva. Investig Opthalmology Vis Sci 57:5116.
444 19. Huang Y, Yang B, Li W. 2016. Defining the normal core microbiome of conjunctival microbial
445 communities. Clin Microbiol Infect 22:643.e7-643.e12.
446 20. Zhou Y, Holland MJ, Makalo P, Joof H, Roberts CH, Mabey D, Bailey RL, Burton MJ,
447 Weinstock GM, Burr SE. 2014. The conjunctival microbiome in health and trachomatous
448 disease: a case control study. Genome Med 6:99.
449 21. Shin H, Price K, Albert L, Dodick J, Park L, Dominguez-Belloa MG. 2016. Changes in the eye
450 microbiota associated with contact lens wearing. MBio 7:1–6.
451 22. de Paiva CS, Jones DB, Stern ME, Bian F, Moore QL, Corbiere S, Streckfus CF, Hutchinson
452 DS, Ajami NJ, Petrosino JF, Pflugfelder SC. 2016. Altered Mucosal Microbiome Diversity and

453 Disease Severity in Sjögren Syndrome. Sci Rep 6:23561.
454 23. Roberts CH, Last A, Molina-Gonzalez S, Cassama E, Butcher R, Nabicassa M, McCarthy E,
455 Burr SE, Mabey DC, Bailey RL, Holland MJ. 2013. Development and Evaluation of a Next-
456 Generation Digital PCR Diagnostic Assay for Ocular Chlamydia trachomatis Infections. J Clin
457 Microbiol 51:2195–203.
458 24. Last AR, Roberts CH, Cassama E, Nabicassa M, Molina-Gonzalez S, Burr SE, Mabey DCW,
459 Bailey RL, Holland MJ. 2013. Plasmid copy number and disease severity in naturally
460 occurring ocular Chlamydia trachomatis infection. J Clin Microbiol 52:324.
461 25. Okolie CE, Wooldridge KG, Turner DP, Cockayne A, James R. 2015. Development of a new
462 pentaplex real-time PCR assay for the identification of poly-microbial specimens containing
463 Staphylococcus aureus and other staphylococci, with simultaneous detection of
464 staphylococcal virulence and methicillin resistance markers. Mol Cell Probes 29:144–50.
465 26. Nakagawa S, Taneike I, Mimura D, Iwakura N, Nakayama T, Emura T, Kitatsuji M, Fujimoto
466 A, Yamamoto T. 2005. Gene sequences and specific detection for Panton-Valentine
467 leukocidin. Biochem Biophys Res Commun 328:995–1002.
468 27. Carvalho M da GS, Tondella ML, McCaustland K, Weidlich L, McGee L, Mayer LW,
469 Steigerwalt A, Whaley M, Facklam RR, Fields B, Carlone G, Ades EW, Dagan R, Sampson
470 JS. 2007. Evaluation and improvement of real-time PCR assays targeting lytA, ply, and psaA
471 genes for detection of pneumococcal DNA. J Clin Microbiol 45:2460–6.
472 28. Meyler KL, Meehan M, Bennett D, Cunney R, Cafferkey M. 2012. Development of a
473 diagnostic real-time polymerase chain reaction assay for the detection of invasive
474 Haemophilus influenzae in clinical samples. Diagn Microbiol Infect Dis 74:356–62.
475 29. Heim A, Ebnet C, Harste G, Pring-Akerblom P. 2003. Rapid and quantitative detection of
476 human adenovirus DNA by real-time PCR. J Med Virol 70:228–39.
477 30. Kais M, Spindler C, Kalin M, Ortqvist A, Giske CG. 2006. Quantitative detection of
478 Streptococcus pneumoniae, Haemophilus influenzae, and Moraxella catarrhalis in lower
479 respiratory tract samples by real-time PCR. Diagn Microbiol Infect Dis 55:169–78.
480 31. Greiner O, Day PJR, Altwegg M, Nadal D. 2003. Quantitative detection of Moraxella

481 catarrhalis in nasopharyngeal secretions by real-time PCR. J Clin Microbiol 41:1386–90.
482 32. Abusleme L, Hong B-Y, Dupuy AK, Strausbaugh LD, Diaz PI. 2014. Influence of DNA
483 extraction on oral microbial profiles obtained via 16S rRNA gene sequencing. J Oral Microbiol
484 6.
485 33. R Core Team. 2014. R: A Language and Environment for Statistical Computing. R Found
486 Stat Comput.
487 34. Wang Q, Garrity GM, Tiedje JM, Cole JR. 2007. Naive Bayesian Classifier for Rapid
488 Assignment of rRNA Sequences into the New Bacterial Taxonomy. Appl Environ Microbiol
489 73:5261–5267.
490 35. Salter SJ, Cox MJ, Turek EM, Calus ST, Cookson WO, Moffatt MF, Turner P, Parkhill J,
491 Loman NJ, Walker AW, Kunin V, Engelbrektson A, Ochman H, Hugenholtz P, Fv W, Göbel U,
492 Stackebrandt E, Kulakov L, McAlister M, Ogden K, Larkin M, O'Hanlon J, McAlister M,
493 Kulakov L, O'Hanlon J, Larkin M, Ogden K, Kéki Z, Grébner K, Bohus V, Márialigeti K, Tóth
494 E, Bohus V, Kéki Z, Márialigeti K, Baranyi K, Patek G, Schunk J, Tóth E, McFeters G,
495 Broadaway S, Pyle B, Egozy Y, Nogami T, Ohto T, Kawaguchi O, Zaitsu Y, Sasaki S, Shen
496 H, Rogelj S, Kieft T, Rand K, Houck H, Maiwald M, Ditton H, Sonntag H, Doeberitz M von K,
497 Tanner M, Goebel B, Dojka M, Pace N, Corless C, Guiver M, Borrow R, Edwards-Jones V,
498 Kaczmarski E, Fox A, Grahn N, Olofsson M, Ellnebo-Svedlund K, Monstein H, Jonasson J,
499 Newsome T, Li B, Zou N, Lo S, Mohammadi T, Reesink H, Vandenbroucke-Grauls C,
500 Savelkoul P, Barton H, Taylor N, Lubbers B, Pemberton A, Laurence M, Hatzis C, Brash D,
501 Oberauner L, Zachow C, Lackner S, Högenauer C, Smolle K, Berg G, Duc M La, Kern R,
502 Venkateswaran K, Ling Z, Liu X, Luo Y, Yuan L, Nelson K, Wang Y, Xiang C, Li L, Benítez-
503 Páez A, Álvarez M, Belda-Ferre P, Rubido S, Mira A, Tom I, Amar J, Serino M, Lange C,
504 Chabo C, Iacovoni J, Mondot S, Lepage P, Klopp C, Mariette J, Bouchez O, Perez L,
505 Courtney M, Marre M, Klopp P, Lantieri O, Doré J, Charles M, Balkau B, Burcelin R, Branton
506 W, Ellestad K, Maingat F, Wheatley B, Rud E, Warren R, Holt R, Surette M, Power C,
507 Borewicz K, Pragman A, Kim H, Hertz M, Wendt C, Isaacson R, Dong Q, Brulc J, Iovieno A,
508 Bates B, Garoutte A, Miller D, Revanna K, Gao X, Antonopoulos D, Slepak V, Shestopalov V,

509 Xuan C, Shamonki J, Chung A, DiNome M, Chung M, Sieling P, Lee D, Kuehn J, Gorden P,
510 Munro D, Rong R, Dong Q, Plummer P, Wang C, Phillips G, Srinivas G, Möller S, Wang J,
511 Künzel S, Zillikens D, Baines J, Ibrahim S, Boissière A, Tchioffo M, Bachar D, Abate L, Marie
512 A, Nsango S, Shahbazkia H, Awono-Ambene P, Levashina E, Christen R, Morlais I,
513 McKenzie V, Bowers R, Fierer N, Knight R, Lauber C, Carlos C, Torres T, Ottoboni L, Cheng
514 X, Tian X, Wang Y, Lin R, Mao Z, Chen N, Xie B, Davidson S, Powell R, James S, Knowlton
515 C, Veerapaneni R, D'Elia T, Rogers S, Shtarkman Y, Koçer Z, Edgar R, Veerapaneni R,
516 D'Elia T, Morris P, Rogers S, DeLeon-Rodriguez N, Lathem T, Rodriguez-R L, Barazesh J,
517 Anderson B, Beyersdorf A, Ziemba L, Bergin M, Nenes A, Konstantinidis K, Turner P, Turner
518 C, Jankhot A, Helen N, Lee S, Day N, White N, Nosten F, Goldblatt D, Willerslev E, Hansen
519 A, Poinar H, Kearney M, Spindler J, Wiegand A, Shao W, Anderson E, Maldarelli F, Ruscetti
520 F, Mellors J, Hughes S, Grice S Le, Coffin J, Cooper A, Poinar H, Roberts C, Ingham S,
521 Segal L, Alekseyenko A, Clemente J, Kulkarni R, Wu B, Chen H, Berger K, Goldring R, Rom
522 W, Blaser M, Weiden M, Bhatt A, Freeman S, Herrera A, Pedamallu C, Gevers D, Duke F,
523 Jung J, Michaud M, Walker B, Young S, Earl A, Kostic A, Ojesina A, Hasserjian R, Ballen K,
524 Chen Y, Hobbs G, Antin J, Soiffer R, Baden L, Garrett W, Hornick J, Marty F, Meyerson M,
525 Deragon J, Sinnett D, Mitchell G, Potier M, Labuda D, Sarkar G, Sommer S, Klaschik S,
526 Lehmann L, Raadts A, Hoeft A, Stuber F, Tamariz J, Voynarovska K, Prinz M, Caragine T,
527 Hilali F, Saulnier P, Chachaty E, Andremont A, Heininger A, Binder M, Ellinger A, Botzenhart
528 K, Unertl K, Döring G, Silkie S, Tolcher M, Nelson K, Carroll N, Adamson P, Okhravi N,
529 Mohammadi T, Reesink H, Vandenbroucke-Grauls C, Savelkoul P, Hughes M, Beck L, Skuce
530 R, Vaishampayan P, Probst A, Duc M La, Bargoma E, Benardini J, Andersen G,
531 Venkateswaran K, Rueckert A, Morgan H, Patel P, Garson J, Tettmar K, Ancliff S, McDonald
532 C, Pitt T, Coelho J, Tedder R, Champlot S, Berthelot C, Pruvost M, Bennett E, Grange T,
533 Geigl E, Chang S, Hsu H, Cheng J, Tseng C, Bonfert T, Csaba G, Zimmer R, Friedel C,
534 Knights D, Kuczynski J, Charlson E, Zaneveld J, Mozer M, Collman R, Bushman F, Knight R,
535 Kelley S, Eren A, Maignen L, Sul W, Murphy L, Grim S, Morrison H, Sogin M, Morris A, Beck
536 J, Schloss P, Campbell T, Crothers K, Curtis J, Flores S, Fontenot A, Ghedin E, Huang L,

Jablonski K, Kleerup E, Lynch S, Sodergren E, Twigg H, Young V, Bassis C, Venkataraman A, Schmidt T, Weinstock G, Lane D, Klindworth A, Pruesse E, Schweer T, Peplies J, Quast C, Horn M, Gl°Ckner F, Cooper P, Walker A, Reyes J, Chico M, Salter S, Vaca M, Parkhill J, Schloss P, Westcott S, Ryabin T, Hall J, Hartmann M, Hollister E, Lesniewski R, Oakley B, Parks D, Robinson C, Sahl J, Stres B, Thallinger G, Horn D Van, Weber C, Kozich J, Westcott S, Baxter N, Highlander S, Schloss P, Quince C, Lanzen A, Davenport R, Turnbaugh P, Bolger A, Lohse M, Usadel B, Kiełbasa S, Wan R, Sato K, Horton P, Frith M, Huson D, Mitra S, Ruscheweyh H, Weber N, Schuster S, Schloss P, Gevers D, Westcott S, White J, Nagarajan N, Pop M. 2014. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. BMC Biol 12:87.

36. Jombart T, Devillard S, Balloux F, Falush D, Stephens M, Pritchard J, Pritchard J, Stephens M, Donnelly P, Corander J, Waldmann P, Sillanpaa M, Tang J, Hanage W, Fraser C, Corander J, Lee C, Abdool A, Huang C, Jombart T, Jombart T, Devillard S, Dufour A, Pontier D, Jombart T, Pontier D, Dufour A, McVean G, Novembre J, Stephens M, Patterson N, Price A, Reich D, Price A, Patterson N, Plenge R, Weinblatt M, Shadick N, Reich D, Hotelling H, Hotelling H, Pearson K, Liu N, Zhao H, Fisher R, Lachenbruch P, Goldstein M, Aitchison J, Reyment R, Beharav A, Nevo E, Fraley C, Raftery A, Cann H, Toma C de, Cazes L, Legrand M, Morel V, Piouffre L, Bodmer J, Bodmer W, Bonne-Tamir B, Cambon-Thomsen A, Ramachandran S, Deshpande O, Roseman C, Rosenberg N, Feldman M, Cavalli-Sforza L, Rosenberg N, Pritchard J, Weber J, Cann H, Kidd K, Zhivotovsky L, Feldman M, Wang S, Lewis C, Jakobsson M, Ramachandran S, Ray N, Bedoya G, Rojas W, Parra M, Molina J, Gallo C, Balloux F, Rosenberg N, Mahajan S, Ramachandran S, Zhao C, Pritchard J, Feldman M, Rambaut A, Pybus O, Nelson M, Viboud C, Taubenberger J, Holmes E, Russell C, Jones T, Barr I, Cox N, Garten R, Gregory V, Gust I, Hampson A, Hay A, Hurt A, Smith D, Lapedes A, Jong J de, Bestebroer T, Rimmelzwaan G, Osterhaus A, Fouchier R, Holmes E, Ghedin E, Miller N, Taylor J, Bao Y, George KS, Grenfell B, Salzberg S, Fraser C, Lipman D, Young J, Palese P, Benson D, Karsch-Mizrachi A, Lipman D, Ostell J, Wheeler D, Larkin M, Blackshields G, Brown N, Chenna R, McGettigan P, McWilliam H, Valentin F, Wallace I, Wilm

565 A, Lopez R, Waterhouse A, Procter J, Martin D, Clamp M, Barton G, Paradis E, Claude J,
566 Strimmer K, Handley L, Manica A, Goudet J, Balloux F, Serre D, Paabo S, Corander J,
567 Marttinen P, Siren J, Tang J, Francois O, Ancelet S, Guillot G, Hunley K, Healy M, Long J,
568 Kittles R, Weiss K, Manica A, Prugnolle F, Balloux F, Prugnolle F, Manica A, Balloux F,
569 Romero I, Manica A, Handley L, Balloux F, Amos W, Hoffman J, Fraley C, Raftery A, Peres-
570 Neto P, Jackson D, Somers K, Saporta G, Paradis E, Dray S, Dufour A, Schwarz G, Evanno
571 G, Regnaut S, Goudet J, Jakobsson M, Rosenberg N, Chessel D, Dufour A, Thioulouse J,
572 Dray S, Dufour A, Chessel D, Venables W, Ripley B, Nei M. 2010. Discriminant analysis of
573 principal components: a new method for the analysis of genetically structured populations.
574 BMC Genet 11:94.

575 37. Capriotti JA, Pelletier JS, Shah M, Caivano DM, Ritterband DC. 2009. Normal ocular flora in
576 healthy eyes from a rural population in Sierra Leone. Int Ophthalmol 29:81–4.

577 38. Walker AW, Martin JC, Scott P, Parkhill J, Flint HJ, Scott KP. 2015. 16S rRNA gene-based
578 profiling of the human infant gut microbiota is strongly influenced by sample processing and
579 PCR primer choice. Microbiome 3:26.

580 39. Ramadhani AM, Derrick T, Macleod D, Holland MJ, Burton MJ. 2016. The Relationship
581 between Active Trachoma and Ocular Chlamydia trachomatis Infection before and after Mass
582 Antibiotic Treatment. PLoS Negl Trop Dis 10:e0005080.

583 40. Solomon AW, Holland MJ, Alexander NDE, Massae PA, Aguirre A, Natividad-Sancho A,
584 Molina S, Safari S, Shao JF, Courtright P, Peeling RW, West SK, Bailey RL, Foster A, Mabey
585 DCW. 2004. Mass treatment with single-dose azithromycin for trachoma. N Engl J Med
586 351:1962–71.

587 41. Seidman JC, Coles CL, Silbergeld EK, Levens J, Mkocha H, Johnson LB, Muñoz B, West
588 SK. 2014. Increased carriage of macrolide-resistant fecal E. coli following mass distribution of
589 azithromycin for trachoma control. Int J Epidemiol 43:1105–13.

590 42. Brunham RC, Rekart ML. 2008. The Arrested Immunity Hypothesis and the Epidemiology of
591 Chlamydia Control. Sex Transm Dis 35:53–54.

592

593

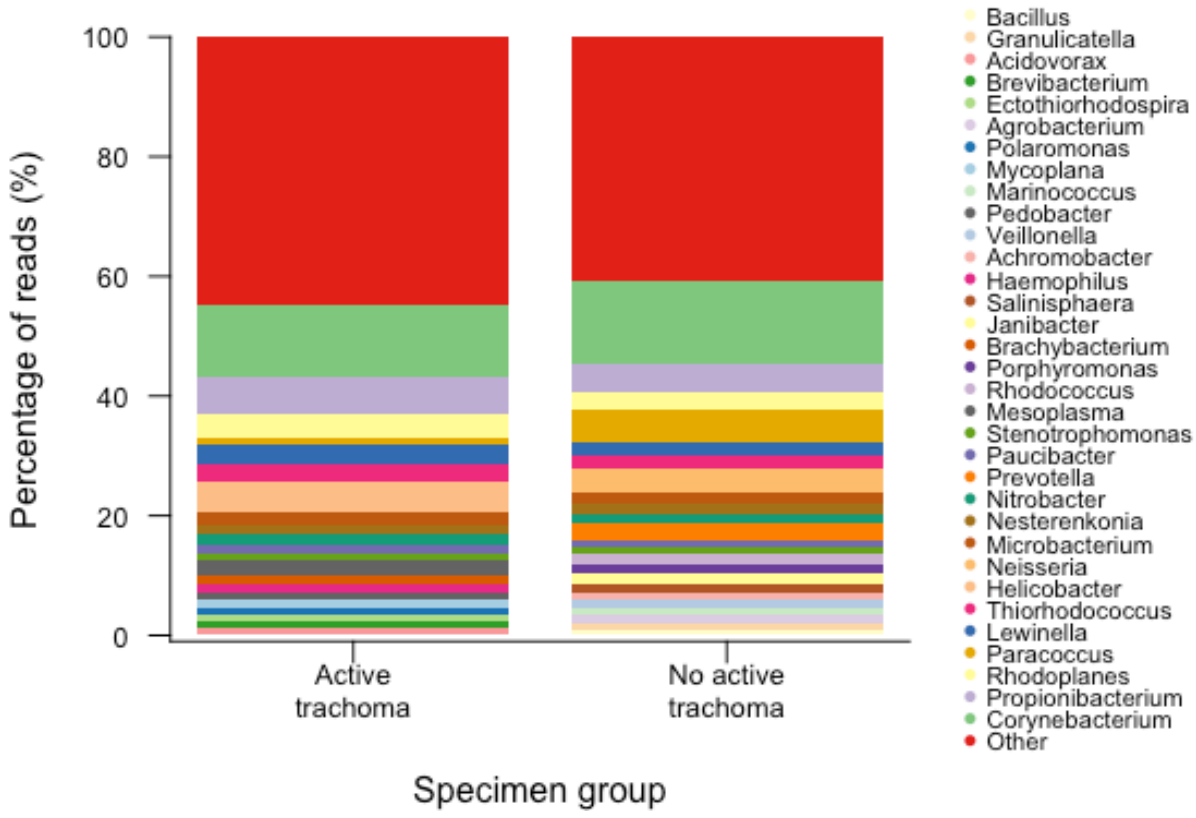
Supplementary information

Assay	Standard curve *		
	LoD	R^2	CoV (%)
S. aureus	4.0	0.991	14.4
Coagulase-negative *Staphylococcus*	1.2	0.994	14.7
S. pneumoniae	9.9	0.998	11.9
H. influenza	2.3	0.997	14.7
Adenoviridae	3.0	0.997	9.1
M. catarrhalis	1.0	0.993	7.8

CoV: Coefficient of variation; LoD: Limit of detection: R^2: Coefficient of determination

* Values calculated from 10-fold dilution series of PCR product between approximately 10^5 and 10^1 copies per target per test with 5 technical replicates at each dilution point.

Supplementary table 1: Each PCR product for dilution into standards was prepared using TaqMan Universal II PCR mix (Life Technologies, Paisley, UK), cleaned with Qiagen MinElute PCR product kit (Qiagen, Manchester, UK) and serially diluted from 1:10^6 to 1:10^{12} in ten-fold dilution steps. Each series was tested in five technical replicates with duplex ddPCR assays to determine the reproducibility of the assay. Target concentrations between one and 10 copies/µL were reproducibly detected by all six assays. The coefficient of determination for all assays was in excess of 0.99 when fitted to a linear regression model. All six assays were highly reproducible, and the mean coefficient of variance (CoV) was 12.1% (range: 7.8 – 14.7%).



609

610

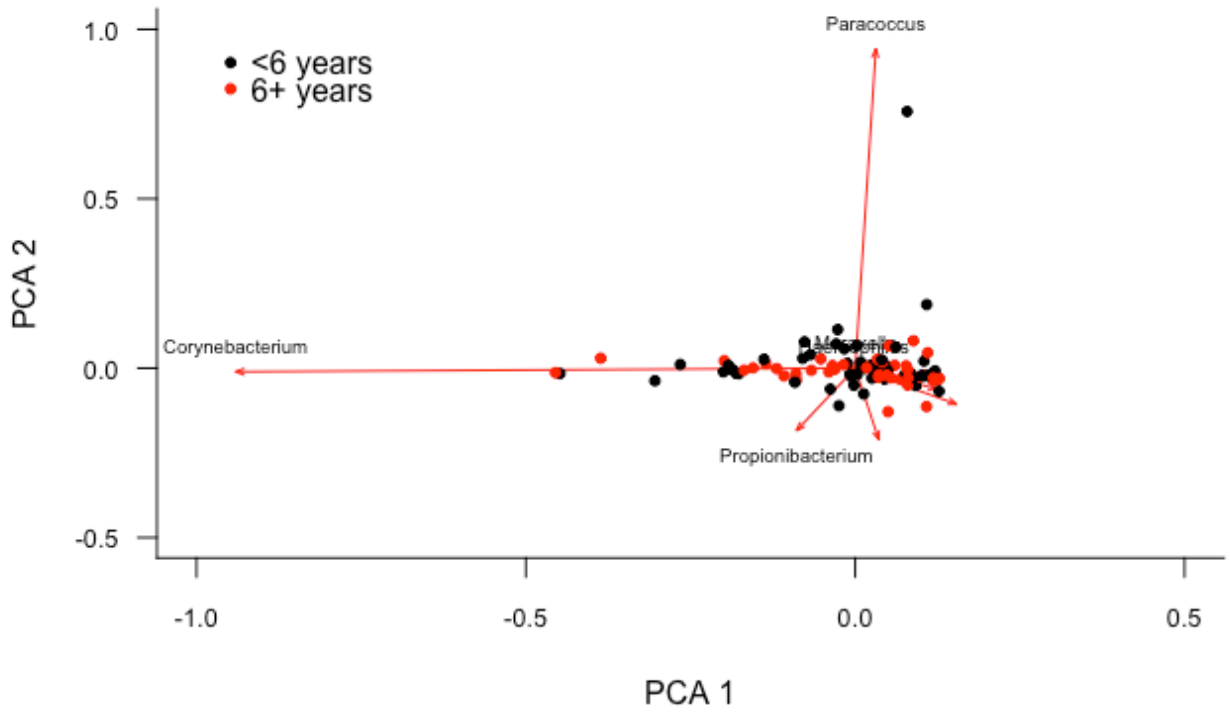
611 **Supplementary figure 2.** Relative taxa abundance in age-, sex- and location-matched children with TF/TI (n = 54) and

612 without (n = 53). Relative taxa abundance is expressed as percentage of total reads per specimen group. Genera with

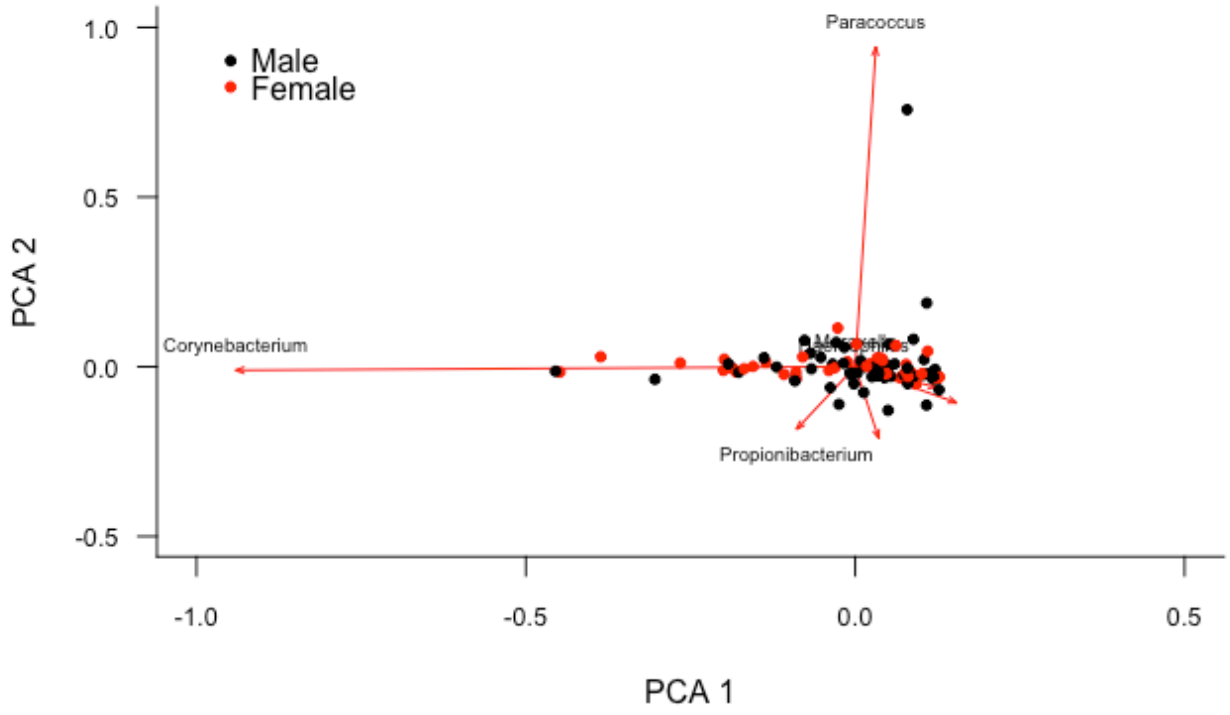
613 reads representing less than 1% of the total number of reads were combined into a group entitled 'Other'.

614

615 **A**



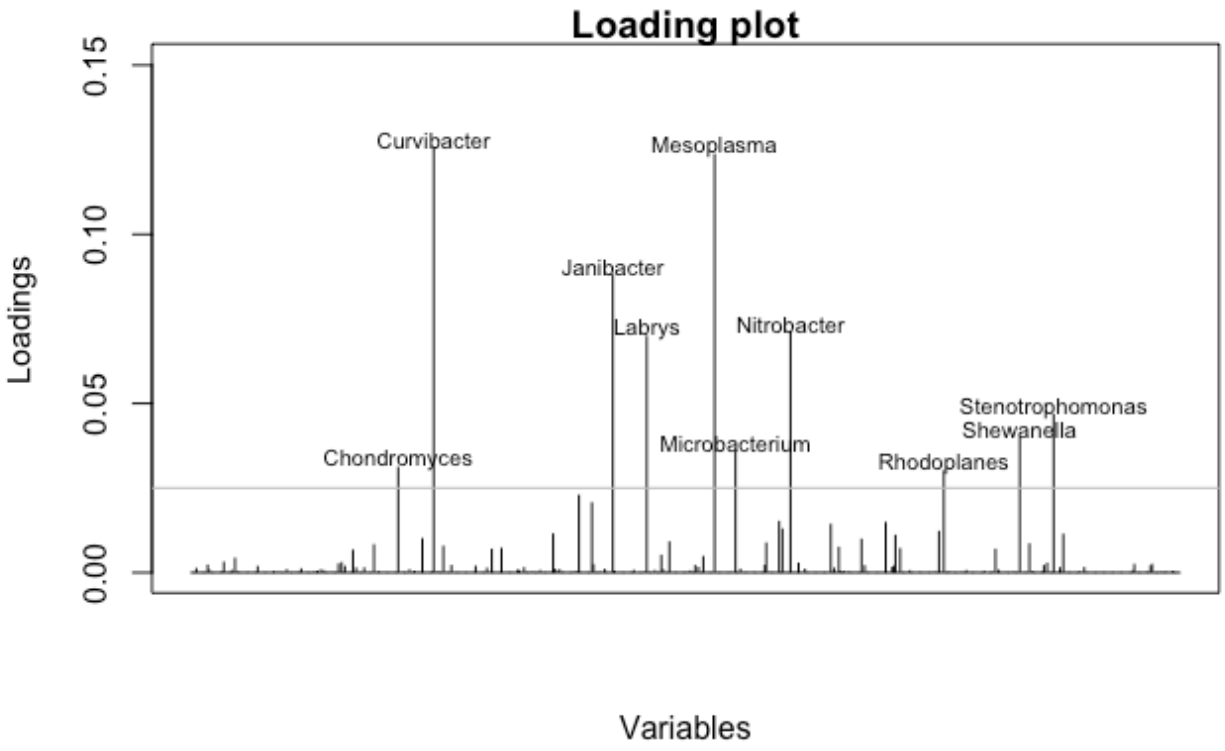
616 **B**



617

618 **Supplementary figure 3.** First and second principal components coloured by **(A)** age group and **(B)** gender. Red spots

619 indicate individuals. Red arrows indicate loadings.



620

621 **Supplementary figure 4.** Relative contributions of genera driving difference between active trachoma cases and controls