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 Butcher1*, Oliver Sokana2, Kelvin Jack2, Eric Kalae4, Leslie Sui4, Charles Russell5, Joanna Houghton1, Christine Palmer1, Martin J Holland1, Richard T Le Mesurier6, Anthony W Solomon1,3, David C W Mabey1,3, Chrissy h. Roberts1.

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*, *Propionibacterium*, *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 (*Surgery, Antibiotics, Facial cleanliness and Environmental 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.
**Table 1.** Oligonucleotides used in this study, with assay concentrations derived from *in vitro* optimisation.

<table>
<thead>
<tr>
<th>Organism / target</th>
<th>Oligo</th>
<th>Sequence (5′-3′)</th>
<th>Concentration in final assay (nM)</th>
<th>Amplicon size</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>C. trachomatis</em> / Plasmid open reading frame 2 (pORF2)</td>
<td>F</td>
<td>CAG CTG ATA GTC CTG CTG GAG AGA</td>
<td>900</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>CAA GAG TAC ATC GGT CAA CGA AGA</td>
<td>900</td>
<td>109</td>
<td>(3, 23, 24)</td>
</tr>
<tr>
<td><em>S. aureus</em> / Staphylococcal protein A (SpA)</td>
<td>F</td>
<td>CAG CAA ACC ATG CAG ATG CTA</td>
<td>900</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>CGC TAA TGA TAA TCC ACC AAA TAC A</td>
<td>900</td>
<td>101</td>
<td>(25, 26)</td>
</tr>
<tr>
<td></td>
<td>Probe</td>
<td>[VIC] TCA AGC ATT ACC AGA AAC [MGBNFQ]</td>
<td>250</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coagulase-negative Staphylococcus / translation elongation factor (tuf)</td>
<td>F</td>
<td>TAT CCA CGA AAC TTC TAA AAC AAG TGT TAC T</td>
<td>450</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>TCT TTA GAT AAT ACG TAT ACT TCA GCT TTA AAT TT</td>
<td>450</td>
<td>204</td>
<td>(25)</td>
</tr>
<tr>
<td></td>
<td>Probe</td>
<td>[FAM] TAT TAG ACT ACG CTG AAG CTG GTG ACA ACA T [BHQ1]</td>
<td>125</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>S. pneumoniae</em> / N-acetylmuramoyl-L-alanine amidase (lytA)</td>
<td>F</td>
<td>ACG CAA TCT AGC AGA TGA AGC A</td>
<td>500</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>TCG TGC GTT TTA ATT CCA GCT</td>
<td>500</td>
<td>75</td>
<td>(27)</td>
</tr>
<tr>
<td></td>
<td>Probe</td>
<td>[FAM] GCC GAA AAC GCT TGA TAC AGG GAG [BHQ1]</td>
<td>150</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>H. influenzae</em> / L-fucokinase (fucK)†</td>
<td>F</td>
<td>ATG GCG GGA ACA TCA ATG A</td>
<td>900</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>ACG CAT AGG AGG GAA ATG GTT</td>
<td>900</td>
<td>102</td>
<td>(28)</td>
</tr>
<tr>
<td></td>
<td>Probe*</td>
<td>[FAM] CGg TAA TTg GGg TAc AT [BHQ1]</td>
<td>125</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adenoviridae / capsid assembly protein (hexon)</td>
<td>F</td>
<td>GCC ACG GTG GGG TTT TTA ACG TT</td>
<td>500</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>GCC CCA GTG GTC TTA CAT GCA CAT C</td>
<td>500</td>
<td>132</td>
<td>(29)</td>
</tr>
<tr>
<td></td>
<td>Probe</td>
<td>[HEX] TGC ACC AGA CCC GGG CTC AGG TAC TCC GA [BHQ1]</td>
<td>125</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>M. catarrhalis</em> / outer membrane protein (copB)</td>
<td>F</td>
<td>GTG AGT GCC GCT TTT TTA ACA ACC</td>
<td>900</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>TGT ATC GCC TGC CAA GAC AA</td>
<td>900</td>
<td>72</td>
<td>(30, 31)</td>
</tr>
<tr>
<td></td>
<td>Probe</td>
<td>[HEX] TGC TTT TGC AGC TGT TAG CCA GCC TAA [BHQ1]</td>
<td>250</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Lower case symbols denote locked nucleic acid bases.
These primers do not differentiate between encapsulated and noncapsulated subtype.
Study population

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

Droplet digital PCR

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

Duplex ddPCR assays were developed for (1) Adenoviridae (29) and S. pneumoniae (27), (2) H. influenzae (28) and M. catarrhalis (30, 31), and (3) S. aureus and coagulase-negative 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 2x 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]-AGTCAGTCAGCCATTACCAGGGCTGCTGG-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; CCAGCAGCCGGTAATTGGCTGACTGACT) and 14 (Read 2; AGTCAGTCAGCCATTACCAGGGCTGCTGG) 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
ddPCR data were analysed using R version 3.2.2 (33). Binomial univariate regression was used to test the relationship between each individual infection and active trachoma. The most accurate final multivariate model, determined by lowest Akaike Information Criterion (AIC) value, was determined by step-wise removal of variables from a binomial multivariate regression model incorporating all tested pathogens. The chance of the Ct association occurring due to chance was assessed by counting the number of significant results obtained from randomly re-ordering the Ct data 1000 times.

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

Between-group differences in the Shannon Diversity Index (a measure of diversity which increases with increasing species abundance and evenness) and Inverse Simpson Index (another measure of diversity where 1 is infinite diversity and 0 is no diversity) were compared using a t-test. Discriminant Analysis of Principal Components (DAPC), a multivariate clustering method to analyse 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
Ct and active trachoma was highly significant ($p = 0.001$). Active disease was neither associated with infection with at least one pathogen (any pathogen, $p = 0.185$) nor the number of concurrent infections in the same eye (number of concurrent pathogens, $p = 0.207$). Infectious loads of those who tested positive are shown figure 2. Despite some numerical differences between groups, none of the differences were statistically significant.

**Figure 2.** Target copies per swab of each pathogen identified in conjunctival swabs collected from children with and without active trachoma in the Solomon Islands. Numbers show p-values for logistic regression comparison between active trachoma case and control groups for each pathogen. None of the differences observed were statistically significant.
**Table 2.** Cases of infection in specimens from children with and without active disease. The relationship of those infections to trachoma has been tested with univariate logistic regression, and then stepwise removal of variables from multivariate regression model was used to determine the final multivariate regression model that provided the best fit for the data.

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

*Data taken from Butcher et al.(3)
16S rRNA gene sequencing

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

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

Principal Components (PC) Analysis (figure 2) revealed that 47% of the total variation in the 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.
Figure 2. (A) First and second and (B) second and third principal components describing variation between the 16S sequences identified in Solomon Island children with and without active trachoma. Dark blue spots indicate controls. Light blue spots indicate cases. Red arrows show principal component loadings.
Figure 3: Discriminant analysis of the association of 20 combined principal components with active trachoma, showing a significant discrimination of phenotype groups ($p = 0.0000013$).
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...
Staphylococcus and Streptococcus. Among the biggest contributors to between-conjunctival variation were Corynebacterium and Propionibacterium, but these genera did not differentiate TF cases from controls (figure 2). Previous investigations of the role of the conjunctival microbiome in ocular disease have also not shown significant differences. People with ocular manifestations of rosacea, Sjögren’s syndrome and healthy controls did not have significant differences in diversity or relative abundance of key phyla identified across all three groups (22). In The Gambia, there was an increased abundance of Haemophilus in cases of TF, as compared to controls. The abundance of Corynebacterium and Streptococcus varied between scarred and non-scarred individuals. However, neither trachoma-related difference was significant (20).

The microbiota of these children were heterogeneous and had subtle variations that appeared to reach significant association with TF when the full community structure was considered in comparative statistical models. It remains possible (though perhaps unlikely) that a multitude of factors, operating at the level of single bacterial species, polymicrobial communities or viral species, all contribute to the presentation of the phenotype. It is perhaps more likely that as yet unidentified viral or allergic causes could explain the prevalence of TF in the Solomon Islands.

There are some limitations to our study. Although the ddPCR assays are based on validated assays and appear to be accurate (supplementary table 1), they have not been formally evaluated in this format. By not using mechanical lysis in our extraction process, the absolute prevalence of some difficult-to-lyse Gram positive genera may have been under-estimated, although this would not impair the comparison of cases to controls, between which protocols were consistent. 16S amplicon sequencing of low biomass samples is known to result in amplification of reagent and environmental contaminants (35). This study focused on gross differences in community structure between those with and without disease and these should be independent of contaminating reads. We also employed stringent quality control measures to ensure data were not confused by artefacts of the sequencing process.
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


22. de Paiva CS, Jones DB, Stern ME, Bian F, Moore QL, Corbiere S, Streckfus CF, Hutchinson DS, Ajami NJ, Petrosino JF, Pflugfelder SC. 2016. Altered Mucosal Microbiome Diversity and


catarrhalis in nasopharyngeal secretions by real-time PCR. J Clin Microbiol 41:1386–90.


A, Schmidt T, Weinstock G, Lane D, Klindworth A, Pruesse E, Schweer T, Peplies J, Quast
C, Horn M, Glöckner 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, Sah 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,
White J, Nagarajan N, Pop M. 2014. Reagent and laboratory contamination can critically

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,
Reyment R, Beharav A, Nevo E, Fraley C, Raftery A, Cann H, Toma C de, Cazes L, Legrand
M, Morel V, Piouffe L, Bodmer J, Bodmer W, Bonne-Tamir B, Cambon-Thomsen A,
Ramachandran S, Deshpande O, Roseman C, Rosenberg N, Feldman M, Cavalli-Sforza L,
Gallo C, Balloux F, Rosenberg N, Mahajan S, Ramachandran S, Zhao C, Pritchard J,
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,


### Supplementary information

<table>
<thead>
<tr>
<th>Assay</th>
<th>Standard curve *</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LoD</td>
</tr>
<tr>
<td>S. aureus</td>
<td>4.0</td>
</tr>
<tr>
<td>Coagulase-negative Staphylococcus</td>
<td>1.2</td>
</tr>
<tr>
<td>S. pneumoniae</td>
<td>9.9</td>
</tr>
<tr>
<td>H. influenza</td>
<td>2.3</td>
</tr>
<tr>
<td>Adenoviridae</td>
<td>3.0</td>
</tr>
<tr>
<td>M. catarrhalis</td>
<td>1.0</td>
</tr>
</tbody>
</table>

CoV: Coefficient of variation; LoD: Limit of detection; R²: 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%).
Supplementary figure 2. Relative taxa abundance in age-, sex- and location-matched children with TF/TI (n = 54) and without (n = 53). Relative taxa abundance is expressed as percentage of total reads per specimen group. Genera with reads representing less than 1% of the total number of reads were combined into a group entitled ‘Other’.
Supplementary figure 3. First and second principal components coloured by (A) age group and (B) gender. Red spots indicate individuals. Red arrows indicate loadings.
Supplementary figure 4. Relative contributions of genera driving difference between active trachoma cases and controls.