bioRxiv preprint doi: https://doi.org/10.1101/548743; this version posted February 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

The international sinonasal microbiome study (ISMS): a multi-centre, multi-national collaboration characterising the microbial ecology of the sinonasal cavity

Sathish Paramasivan[1], Ahmed Bassiouni[1], Arron Shiffer[2], Matthew R Dillon[2], Emily K Cope[2], Clare Cooksley[1], Mohammad Javed Ali[3], Benjamin Bleier[4], Claudio Callejas[5], Marjolein E Cornet[6], Richard G Douglas[7], Daniel Dutra[8], Christos Georgalas[6], Richard J Harvey[9,10], Peter H Hwang[11], Amber U Luong[12], Rodney J Schlosser[13], Pongsakorn Tantilipikorn[14], Marc A Tewfik[15], Sarah Vreugde[1], Peter-John Wormald[1], J Gregory Caporaso[2], and Alkis J Psaltis[1]

[1] Department of Otolaryngology, Head and Neck Surgery, University of Adelaide, Adelaide, Australia

[2] Pathogen and Microbiome Institute, Northern Arizona University, Arizona, USA

[3] Dacryology Service, LV Prasad Institute, Hyderabad, India

[4] Department of Otolaryngology, Massachusetts Eye and Ear Infirmary, Harvard Medical School, Boston, USA

[5] Department of Otolaryngology, Pontificia Universidad Catolica de Chile, Santiago, Chile

[6] Department of Otorhinolaryngoloy, Amsterdam UMC, Amsterdam, The Netherlands

[7] Department of Surgery, University of Auckland, Auckland, New Zealand

[8] Department of Otorhinolaryngology, University of Sao Paulo, Sao Paulo, Brazil

[9] Department of Otolaryngology, Rhinology and Skull base, University of New South Wales, Sydney, Australia

[10] Faculty of Medicine and Health sciences, Macquarie University, Sydney, Australia

[11] Department of Otolaryngology -Head and Neck Surgery, Stanford University, Stanford, California, USA

[12] Department of Otolaryngology -Head and Neck Surgery, University of Texas, Texas, USA

[13] Department of Otolaryngology, Medical University of South Carolina, Charleston, South Carolina, USA

[14] Department of Otorhinolaryngology, Faculty of Medicine, Siriraj Hospital, Mahidol University, Bangkok, Thailand

[15] Department of Otolaryngology - Head and Neck Surgery, McGill University, Montreal, Canada

bioRxiv preprint doi: https://doi.org/10.1101/548743; this version posted February 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Corresponding author:

Associate Professor Alkis J Psaltis

3C (Department of Otolaryngology, Head and Neck Surgery)

The Queen Elizabeth Hospital

28 Woodville Rd

Woodville South, SA 5011

Australia

Email: alkis.psaltis@adelaide.edu.au

Phone: +61 08 8222 7158

Fax: +61 08 8222 7419

Funding information, Disclosures and Conflicts of Interest (COI):

Mohammad Javed Ali:

Receives royalties from Springer for his treatise “Principles and Practice of Lacrimal Surgery” and “Atlas of Lacrimal Drainage Disorders”.

No conflict of interest relevant to this study.

Ahmed Bassiouni, Clare Cooksley:

No conflict of interest to declare.

Benjamin Bleier:

Grant Funding: R01 NS108968-01 NIH/NINDS (Bleier PI) – This isn’t relevant to this study.

Consultant for: Gyrus ACMI Olympus, Canon, Karl Storz, Medtronic, and Sinopsys.

Equity: Cerebent, Inc, Arrinex.

COI: None relevant to this study.

bioRxiv preprint doi: https://doi.org/10.1101/548743; this version posted February 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Claudio Callejas:

No conflict of interest to declare.

J Gregory Caporaso, Matthew R Dillon, Arron Shiffer:

No conflicts of interest to declare. This work was funded in part by National Science Foundation Award 1565100 to JGC.

Emily K Cope:

Financial information: This work was partially funded under the State of Arizona Technology and Research Initiative Fund (TRIF), administered by the Arizona Board of Regents, through Northern Arizona University.

No relevant disclosures or COI.

Marjolein E Cornet:

No financial relationships or sponsors. No conflicts of interests.

Richard G Douglas:

Received consultancy fees from Lyra Therapeutics and is a consultant for Medtronic. These are not relevant to this study.

Daniel Dutra:

No conflict of interest to declare.

Christos Georgalas:

No conflicts of interest to declare.

Richard J Harvey:

Consultant with Medtronic, Olympus and NeilMed pharmaceuticals. He has also been on the speakers' bureau for Glaxo-Smith-Kline, Seqiris and Astra-Zeneca.

No direct conflict of interest to declare.

bioRxiv preprint doi: https://doi.org/10.1101/548743; this version posted February 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Peter H Hwang:

Financial Relationships: Consultancies with Arrinex, Bioinspire, Canon, Lyra Therapeutics, Medtronic, Tivic.

Conflicts of Interest: None.

Amber U Luong:

Serves as a consultant for Aerin Medical (Sunnyvale, CA), Arrinex (Redwood City, CA), Lyra Therapeutics (Watertown, MA), and Stryker (Kalamazoo, MI) and is on the advisory board for ENTvantage (Austin, TX).

Her department receives funding from Genetech/Roche (San Francisco, CA) and AstraZeneca (Cambridge, England).

No COI to declare related to this study.

Sathish Paramasivan:

Supported by a Garnett Passe and Rodney Williams Memorial Foundation Academic Surgeon Scientist Research Scholarship.

No conflicts of interest to declare.

Alkis J Psaltis:

Consultant for Aerin Devices and ENT technologies and is on the speakers' bureau for Smith and Nephew. Received consultancy fees from Lyra Therapeutics. These are not relevant to this study.

Rodney J Schlosser:

Grant support from OptiNose, Entellus, and IntersectENT (not relevant to this study). Consultant for Olympus, Meda, and Arrinex (not relevant to this study).

Pongsakorn Tantilipikorn:

No financial disclosures or conflict of interest.

bioRxiv preprint doi: https://doi.org/10.1101/548743; this version posted February 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

92 Marc A Tewfik:

93 Principal Investigator: Sanofi, Roche/Genentech, AstraZeneca.

94 Speaker/Consultant: Stryker, Ondine Biomedical, Novartis, MEDA, Mylan.

95 Royalties for book sales: Thieme.

96 Sarah Vreugde:

97 No conflicts of interest relevant to this study.

98 Peter-John Wormald:

99 Receives royalties from Medtronic, Integra, and Scopis, and is a consultant for NeilMed. These are not

100 relevant to this study.

101

bioRxiv preprint doi: https://doi.org/10.1101/548743; this version posted February 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

ABSTRACT

The sinonasal microbiome remains poorly defined, with our current knowledge based on a few cohort studies whose findings are inconsistent. Furthermore, the variability of the sinus microbiome across geographical divides remains unexplored. We characterise the sinonasal microbiome and its geographical variations in both health and disease using 16S rRNA gene sequencing of 410 individuals from across the world. Although the sinus microbial ecology is highly variable between individuals, we identify a core microbiome comprised of *Corynebacterium*, *Staphylococcus*, *Streptococcus*, *Haemophilus*, and *Moraxella* species in both healthy and chronic rhinosinusitis (CRS) cohorts. *Corynebacterium* (mean relative abundance = 44.02%) and *Staphylococcus* (mean relative abundance = 27.34%) appear particularly dominant in the majority of patients sampled. There was a significant variation in microbial diversity between countries ($p = 0.001$). Amongst patients suffering from CRS with nasal polyps, a significant depletion of *Corynebacterium* (40.29% vs 50.43%; $p = 0.02$) and over-representation of *Streptococcus* (7.21% vs 2.73%; $p = 0.032$) was identified. The delineation of the sinonasal microbiome and standardised methodology described within our study will enable further characterisation and translational application of the sinus microbiota.

Keywords

Chronic rhinosinusitis, microbiota, next-generation sequencing, 16S rRNA gene, sinus, microbiome, polyp

Abbreviations

NGS, Next-generation sequencing

CRS, chronic rhinosinusitis

CRSsNP, chronic rhinosinusitis sans nasal polyps

CRSwNP, chronic rhinosinusitis with nasal polyps

bioRxiv preprint doi: https://doi.org/10.1101/548743; this version posted February 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

MAIN TEXT

The important role of human microbiota in both health and disease has become increasingly recognised. Microbial communities encode millions of genes and associated functions which act in concert with those of human cells to maintain homeostasis.[1] Numerous studies have now established the microbiota as an important contributor to essential mammalian functions such as metabolism[2], biosynthesis[3], neurotransmission[4,5] and immunomodulation[6,7]. Characterizing the composition and diversity of normal, healthy microbial communities is a cornerstone to developing our understanding of dysbiosis, pathophysiology and, ultimately, directing therapy. To this end, advent of next-generation sequencing (NGS) has revolutionised our appreciation of the host microbiota and its polymicrobial nature.[8,9]

In many cases the entire microbial community – commensal, symbiotic, pathogenic bacteria, fungi, archaea, and viruses – play critical roles in both health and disease pathogenesis. The host-microbiota interface is particularly important in chronic mucosal inflammatory conditions where the microbiota interact directly with the host. These conditions are often poorly understood, multifactorial in nature, have heterogeneous clinical presentations and vary in treatment response.[10,11] Furthermore, single causative pathogens are rarely identified, and culture directed-antibiotics often fail to demonstrate efficacy.[12] It is plausible that a better understanding of the microbiome of such conditions may be key to unraveling their underlying pathogenesis.

The sinonasal mucosa is continuously exposed to external particulate matter and microbes, but it is relatively immunodeplete with no native secondary lymphoid organ systems.[13] The sinus microbiota is thought to play key roles in multiple extra-nasal conditions, such as providing a nidus of recurrent infection in cystic fibrosis patients[14], and otitis media.[15] In addition, there is evidence of microbial influences in the development, progression and severity of chronic rhinosinusitis (CRS).[16] This multifactorial condition, with an estimated world-wide prevalence of approximately 10%,[11] represents one of the most common diagnoses for inappropriate antibiotic prescription and is a source of significant

bioRxiv preprint doi: https://doi.org/10.1101/548743; this version posted February 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

morbidity and healthcare costs.[11,17–19] To date, despite a number of well-designed research efforts to define the nature of the sinonasal microbiome and its role in CRS pathogenesis, many uncertainties persist. This is in part due to the difficulty in bacterial collection from the nose itself. Unlike the gut and oral cavity where the bacterial burden is high and access to microbiota relatively easy (either via faecal samples or oral wash),[20,21] the sinonasal tract has a low microbial burden and access is difficult due to both the narrow nasal orifice and discomfort for the awake patient. Therefore, a replicable sample with appropriate bacterial abundance for 16S rRNA gene sequencing is currently attainable only during nasal surgery. To date, the majority of published studies have been small in size, with heterogeneous patient populations and inconsistency in collection methods, sample site, processing techniques and bioinformatics pipelines.[22] Ultimately, the consequence has been non-comparable results with no universal consensus on the constituents of the healthy sinonasal microbiota or the dysbiosis that occurs in disease.[23–29]

To address these limitations, we investigate the sinonasal microbiome using 16S rRNA gene-sequencing on a large, multi-centre, international cohort implementing consistent sampling, processing and bioinformatics methods. We aim to (1) characterise the normal sinonasal microbiome, (2) assess for any geographical or clinical influences and (3) identify any changes associated with CRS within and across geographical sites.

bioRxiv preprint doi: https://doi.org/10.1101/548743; this version posted February 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

RESULTS

Patient cohort

Middle meatus specimens were collected for 16S rRNA gene sequencing (V3-V4 hypervariable region) on the Illumina MiSeq platform (see Methods). Thirteen centres, across five continents, participated in patient sampling. 532 samples (194 healthy controls and 338 CRS patients) successfully went through all stages of transport and processing to be sent for sequencing. High-quality sequences were analysed using QIIME 2.[30] A total of 410 patients, aged between 20 and 75, reached the final stage of analysis. This population included 139 non-CRS healthy controls, 99 patients without nasal polyposis (CRSsNP) and 172 CRS patients with nasal polyposis (CRSwNP). Supplementary Figure S1 demonstrates sample distribution by centre.

The sinonasal microbiome in healthy sinuses is dominated by Corynebacterium and Staphylococcus

We first investigated the composition of the healthy sinonasal microbiome by intra-operatively sampling the 139 non-CRS control patients. (See Methods) Our analysis demonstrated the dominance of *Corynebacterium* (mean relative abundance = 48.7%; prevalence = 88.49%) and *Staphylococcus* species (mean relative abundance = 29.25%; prevalence = 79.86%) in the sinonasal microbiome of healthy patients. These were both the most abundant and prevalent genera amongst our population (Table 1). This finding has been observed in some but not all previously reported studies with variability in sampling and analysis techniques likely accounting for such discrepancies.[22]

Table 1: Abundance and prevalence of genera found in microbiota of healthy non-CRS patients.

Genera	Mean Relative Abundance (%)	Prevalence (%)
Corynebacterium	48.7	88.49
Staphylococcus	29.25	79.86
Moraxella	3.86	12.23

bioRxiv preprint doi: https://doi.org/10.1101/548743; this version posted February 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

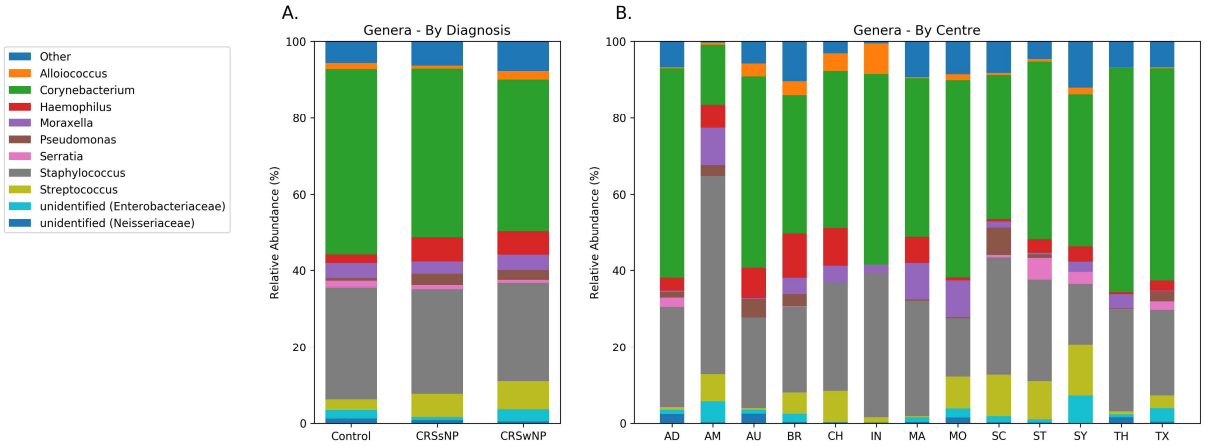
Streptococcus	2.81	20.86
Haemophilus	2.23	12.95
unidentified (Enterobacteriaceae)	2.16	13.67
Serratia	1.79	2.16
Alloiococcus	1.61	20.86
unidentified (Neisseriaceae)	1.25	12.23
Pseudomonas	0.75	2.88

Moraxella, *Streptococcus*, *Haemophilus*, *Enterobacteriaceae*, *Serratia*, *Alloiococcus*, *Neisseriaceae* and *Pseudomonas* made up the remainder of the ten most abundant genera. The lower prevalence of these remaining organisms could suggest that aside from *Corynebacterium* and *Staphylococcus*, there is a high degree of variability in the constituents of a healthy upper airway microbiome.

Microbiome composition is altered in CRSwNP and by geographical location

To explore influences of microbial composition, we examined the taxonomic profiles of our patient cohort once grouped by (1) disease cluster and (2) centre of origin (Figure 1) utilizing mixed modeling to control for the "centre" variable by assigning it as a random effect (See Methods). CRSsNP patients demonstrated no significant differences in the relative abundance of the top ten most abundant organisms when compared to healthy controls. ($p > 0.05$; mixed model analysis) The relative abundance of most organisms also remained stable between controls and CRSwNP but for two genera (Figure 1A): *Corynebacterium* was found to be significantly reduced in CRSwNP when compared to controls (40.29% vs 50.43%; mixed model analysis; $p = 0.02$) whilst *Streptococcus* was increased (7.21% vs 2.73%; mixed model analysis; $p = 0.032$). Interestingly, the most commonly cultured pathogens in CRS, *Staphylococcus* and *Pseudomonas* were similar between all cohorts.

bioRxiv preprint doi: https://doi.org/10.1101/548743; this version posted February 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



	Control	CRSsNP	CRSwNP
Alloiococcus	1.81%	1.11%	2.16%
Corynebacterium	50.43%	43.09%	40.29%
Haemophilus	2.21%	6.28%	6.02%
Moraxella	3.63%	3.53%	3.96%
Pseudomonas	0.75%	2.88%	2.59%
Serratia	1.79%	1.13%	0.74%
Staphylococcus	28.3%	27.49%	26.52%
Streptococcus	2.73%	6.41%	7.21%
unidentified Enterobacteriaceae	2.15%	0.86%	3.2%
unidentified Neisseriaceae	1.27%	0.77%	0.45%

	AD	AM	AU	BR	CH	IN	MA	MO	SC	ST	SY	TH	TX
Alloiococcus	0.3%	0.5%	3.3%	3.6%	4.6%	7.9%	0.3%	1.5%	0.6%	0.7%	1.8%	0.1%	0.3%
Corynebacterium	55.0%	15.5%	50.3%	36.5%	41.1%	50.1%	41.8%	51.8%	37.8%	46.7%	40.1%	58.8%	55.7%
Haemophilus	3.3%	6.2%	7.8%	11.4%	10.0%	0.0%	6.8%	0.8%	0.6%	3.5%	4.0%	0.5%	2.6%
Moraxella	0.3%	9.8%	0.2%	4.3%	4.3%	2.2%	9.5%	9.5%	1.6%	0.4%	2.6%	3.7%	0.2%
Pseudomonas	1.4%	2.9%	4.8%	3.1%	0.0%	0.0%	0.4%	0.3%	7.4%	0.9%	0.0%	0.2%	2.6%
Serratia	2.4%	0.0%	0.0%	0.1%	0.0%	0.0%	0.0%	0.0%	0.6%	5.6%	3.1%	0.0%	2.2%
Staphylococcus	26.3%	51.9%	23.8%	22.5%	28.4%	37.7%	30.2%	15.4%	30.6%	26.6%	15.7%	26.8%	22.5%
Streptococcus	0.7%	7.0%	0.5%	5.5%	8.2%	1.4%	0.3%	8.3%	10.8%	10.0%	13.5%	0.7%	3.3%
unidentified Enterobacteriaceae	1.1%	5.5%	1.0%	2.2%	0.0%	0.0%	1.2%	2.4%	1.6%	0.8%	7.2%	0.8%	3.5%
unidentified Neisseriaceae	2.4%	0.2%	2.5%	0.3%	0.2%	0.2%	0.3%	1.5%	0.2%	0.2%	0.0%	1.5%	0.4%

Figure 1: Microbiome taxonomic profiles by disease status and centres. *Sinonasal microbial composition of patients (n = 409) when grouped by disease and centres. Accompanying tables demonstrate the corresponding relative abundances of the top ten most abundant organisms found within our cohort. The abundances of genera in A (by disease status) have been adjusted according to the mixed model that accounted for the centre as a random variable. CRSsNP = Chronic rhinosinusitis without nasal polyps; CRSwNP = chronic rhinosinusitis with nasal polyps; Control = Healthy, non-CRS patients AD = Adelaide; AM = Amsterdam; AU = Auckland; BR = Brazil; CH = Chile; IN = India; MA = Massachusetts; MO = Montreal; SC = South Carolina; ST = Stanford; SY = Sydney; TH = Thailand; TX = Texas.*

By contrast, comparisons between centres revealed a higher degree of variability in the microbiome composition (Figure 1B). Samples from Amsterdam were significantly different from the remainder of the cohort, with increased *Staphylococcus* (51.94%) and marked depletion of *Corynebacterium* (15.51%).

bioRxiv preprint doi: https://doi.org/10.1101/548743; this version posted February 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Amongst the remaining centres, each appeared to have some variability, with individual regions displaying increased or decreased relative abundance in specific taxa. *Streptococcus*, for example, made up 13.48% of the Sydney microbiome, but was almost absent amongst the Adelaide (0.65%), Auckland (0.45%), Massachusetts (0.32%) and Thailand (0.72%) cohorts. Similar variability can be seen for almost all bacterial taxa examined (Figure 1B). While centres' samples were appeared fairly similar in microbial composition, these results would suggest that centre-specific microbiome profiles can be found. This variability may account for some of the inconsistencies observed in the literature, particularly between findings from different institutes.

Microbial diversity is influenced by geographical location

Alpha diversity amongst cohorts was performed utilising Faith's Phylogenetic Diversity (PD) index.[31] Comparison between disease states demonstrated no significant differences between controls, CRSsNP and CRSwNP (Figure 2A). Overall, alpha diversity was significantly different between centres (Kruskall-Wallis; $p < 0.001$; Figure 2B). Of particular interest was the finding of a significantly lower alpha diversity for samples from Amsterdam compared to the other centres (mean PD = 1.27, $p < 0.01$). This may be related to the compositional findings of high staphylococcal abundance in Amsterdam.

bioRxiv preprint doi: https://doi.org/10.1101/548743; this version posted February 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

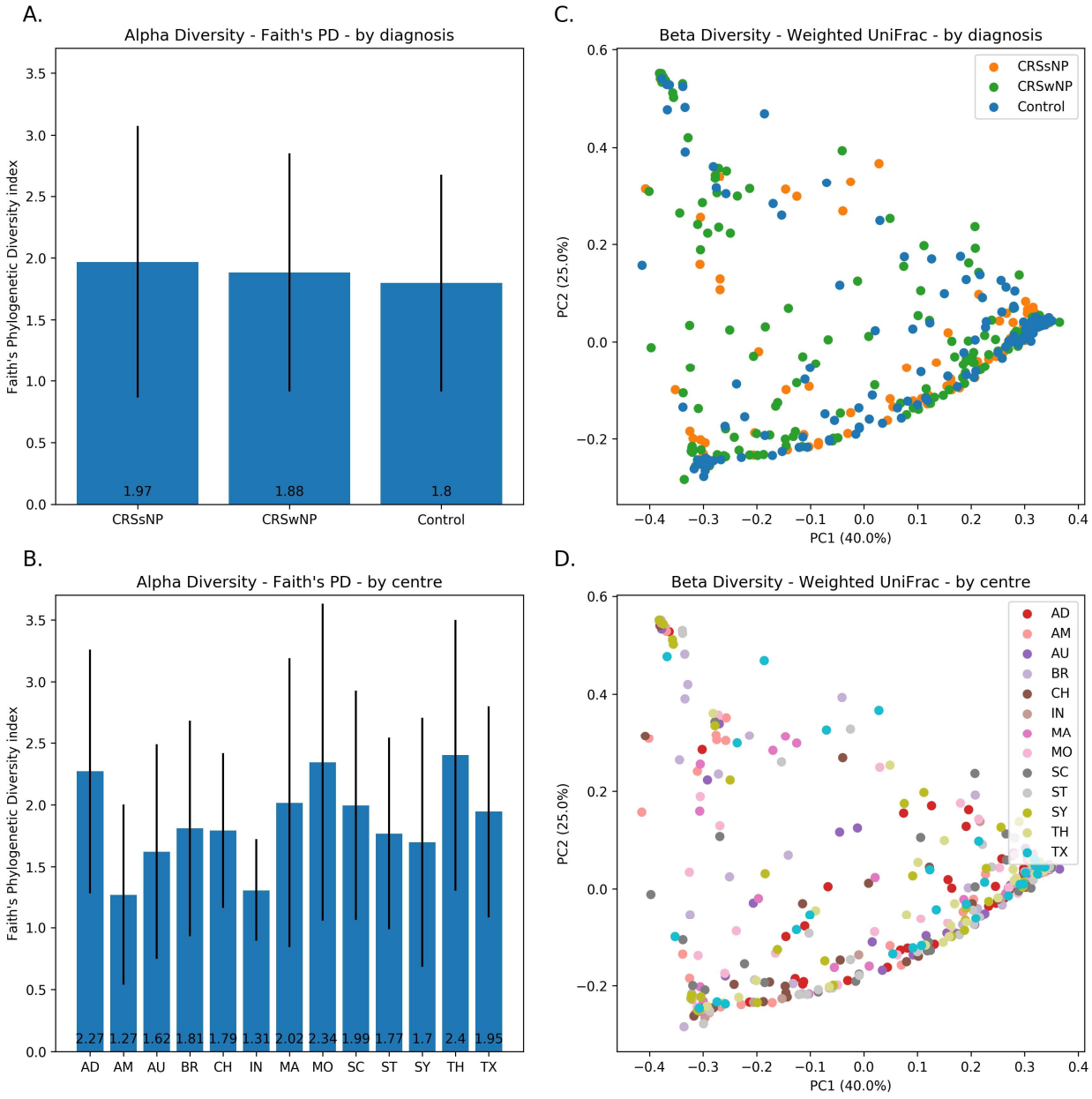


Figure 2: Alpha and beta diversity plots. *Alpha diversity, derived from Faith's Phylogenetic Diveristy Index, demonstrated for this cohort (n = 409) when grouped by disease and by collection centre. Error bars represent 95% confidence intervals. Beta diversity is demonstrated here as a Principal Coordinate Analysis (PCoA) plot of the Weighted-UniFrac distance matrix. Each dot represents a single patient. Similarities between patients are represented by their proximity to each other on the graph. Again, patients are classified by disease and centre. Component 1 (PC1) is represented on the x-axis and component 2 (PC2) on the y-axis. Patients tended towards clustering into one of three groups, as*

bioRxiv preprint doi: https://doi.org/10.1101/548743; this version posted February 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

visualised. PD = Phylogenetic diversity CRSsNP = chronic rhinosinusitis without nasal polyps; CRSwNP = chronic rhinosinusitis with nasal polyps; Control = healthy, non-CRS patients AD = Adelaide; AM = Amsterdam; AU = Auckland; BR = Brazil; CH = Chile; IN = India; MA = Massachusetts; MO = Montreal; SC = South Carolina; ST = Stanford; SY = Sydney; TH = Thailand; TX = Texas.

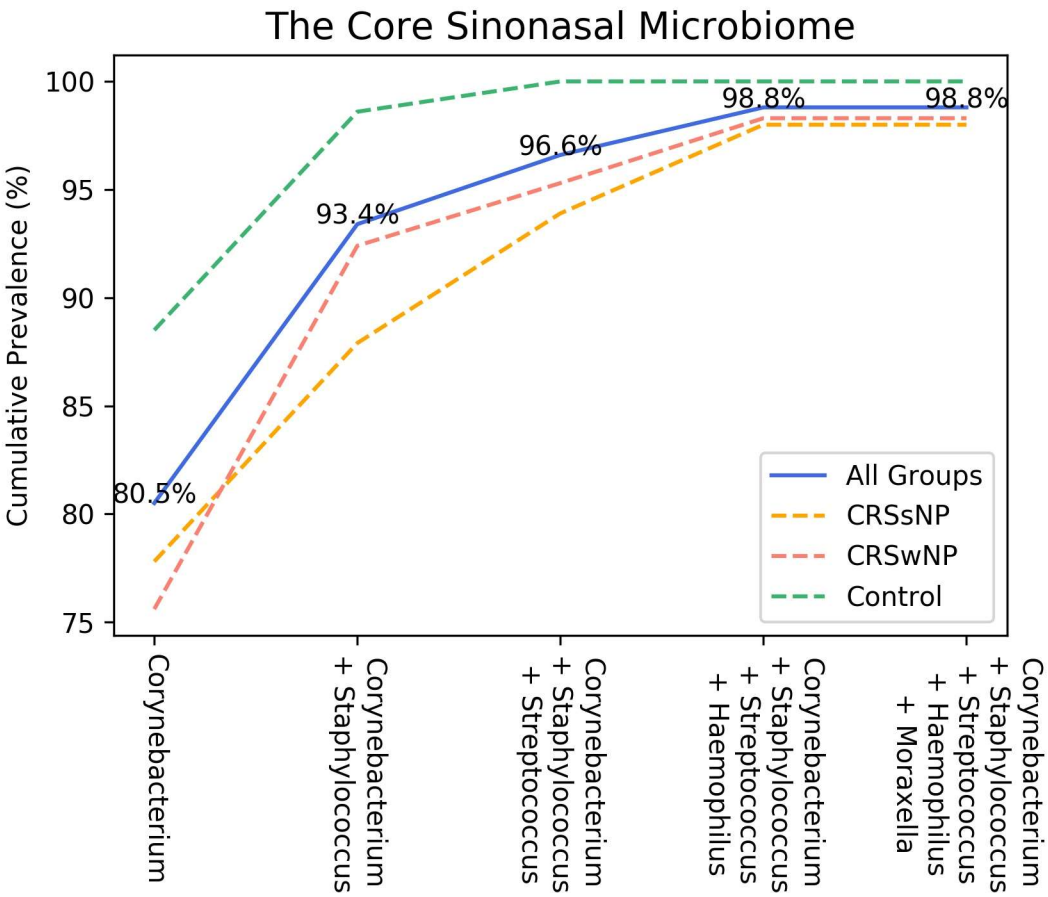
Multivariate analysis on the beta diversity distance matrix was done using PERMANOVA (with 999 permutations) to explore the significance of the diagnosis versus the centre variables in a single model. This showed a significant effect of the centre covariate (pseudo-F = 2.51; p = 0.001) on the weighted UniFrac[32] distances, while the diagnosis covariate was not significant (pseudo-F = 1.66; p = 0.1).

We also performed Principal Coordinates Analysis (PCoA) on the Weighted-UniFrac distance matrix. This did not show clustering by disease state (Figure 2C and 2D), although the plot may suggest a distribution amenable to unsupervised clustering. Patients tended to cluster into three groups on a continuum on the PCoA. These individual clusters could represent specific microbial community types, similar to what has been previously reported by Cope et al.[23] Investigation into the constitution of these groupings, and their association with underlying clinical or pathological factors in a large multi-institutional cohort remains a topic of future investigation.

The core sinonasal microbiome is composed of five genera

We defined a core sinonasal microbiome by analysing the most abundant organisms along with their cumulative prevalence. This was performed across all samples and also across the three different disease groups. The results of this investigation confirmed a high prevalence of the top five abundant genera (*Corynebacterium*, *Staphylococcus*, *Streptococcus*, *Haemophilus*, and *Moraxella*), which together reached a cumulative prevalence in 98-100% of samples in all patient groups. Presence of at least one of these taxa in nearly all patients suggests that they make up the core sinonasal microbiome, regardless of disease status (Figure 3). Amongst control patients, *Corynebacterium* and *Staphylococcus* were present in 98.6% of patients, again suggesting a likely key commensal function of these two genera in the healthy state.

bioRxiv preprint doi: https://doi.org/10.1101/548743; this version posted February 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



All Groups	80.5%	93.4%	96.6%	98.8%	98.8%
CRSsNP	77.8%	87.9%	93.9%	98.0%	98.0%
CRSwNP	75.6%	92.4%	95.3%	98.3%	98.3%
Control	88.5%	98.6%	100.0%	100.0%	100.0%

Figure 3: Cumulative microbial prevalence for the top abundant organisms. *Cumulative prevalence for the top five most abundant organisms is presented above. For all groups, we commence with the most abundant organism then add subsequent microbial taxa in a descending fashion based on relative abundance. CRSsNP = chronic rhinosinusitis without nasal polyps; CRSwNP = chronic rhinosinusitis with nasal polyps; Control = healthy, non-CRS patients.*

bioRxiv preprint doi: https://doi.org/10.1101/548743; this version posted February 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Clinical covariates do not correlate with changes to taxonomic composition

To determine whether any host factors may contribute to the stability of the sinonasal microbiome, patients were separated by known clinical variables that contribute to CRS (Figure 4). There were no significant differential abundances (of the 10 topmost abundant genera) for all six clinical covariates examined (asthma, aspirin sensitivity, diabetes, gastro-oesophageal reflux disease, smoking status and primary-versus-revision surgery). These tests were also repeated for CRS only subgroup and showed no significantly differentially abundant genera across the different covariate levels.

bioRxiv preprint doi: https://doi.org/10.1101/548743; this version posted February 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

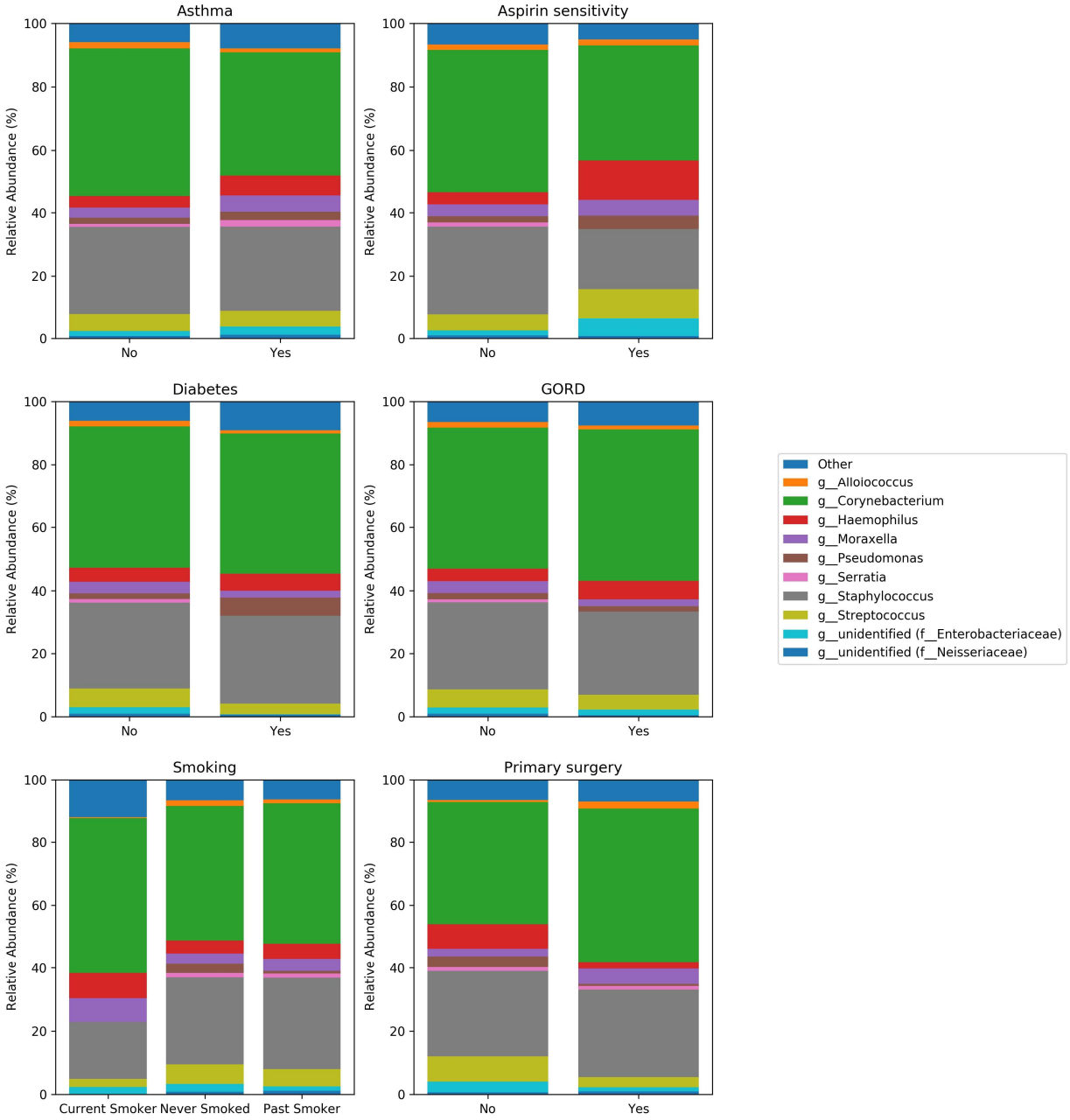


Figure 4: Relationship between clinical co-variates and microbial composition. *Relative abundances of patient cohort (n = 410) when grouped by clinical co-variates (asthma status, aspirin sensitivity, diabetes, gastro-oesophageal reflux disease, smoking and surgery status). GORD = gastro-oesophageal reflux disease.*

bioRxiv preprint doi: https://doi.org/10.1101/548743; this version posted February 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

DISCUSSION

The present study characterises the sinonasal microbiome in a large cohort of subjects from centres around the world using 16S rRNA surveillance. By adopting a unified, consistent methodology from sample acquisition to analysis we have been able to address many of the current existing limitations of currently available data.

We have identified *Corynebacterium*, *Staphylococcus*, *Moraxella*, *Streptococcus* and *Haemophilus* as the most abundant genera within the middle meatus of patients with or without CRS. This consistent finding across disease state and geography suggests that these organisms may form the core microbiome within the sinonasal tract. Interestingly, *Streptococcus*, *Moraxella* and *Haemophilus* species are traditionally respiratory tract organisms and constitute the most commonly cultured pathogens in patients with acute bacterial tonsillitis, otitis media and acute sinusitis.[13,33–35] Anatomically, the sinonasal tract connects these three distinct anatomical regions and while these organisms appear to be commensals, so it is possible that the sinuses act as a reservoir for these organisms to subsequently initiate acute infection. While relatively low bacterial burden is present within the sinonasal mucosa, a dysbiosis within the diseased state may lead to over-representation of these organisms that subsequently lead to acute infectious or inflammatory processes.

While variations across geography have been demonstrated in the gut microbiome[1] and have been suspected within the sinuses,[36] this is the first study to examine this using standardized methodology. Our results suggest that although the bacterial composition of the core microbiome was preserved across the different sites, significant differences in both alpha and beta diversity occurred according to geography. Furthermore, microbiome composition also varied between centres. Most centres demonstrated a similar consistency in relative abundances of *Corynebacterium* and Staphylococcal species. Some centres, however, appeared to exhibit a varying microbial ecology of the less abundant organisms, which were either over- or under-represented. The most distinct microbial distribution was observed in samples

bioRxiv preprint doi: https://doi.org/10.1101/548743; this version posted February 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

collected from the Amsterdam centre, which appeared to be the most distinct centre with depletion in *Corynebacterium* and over-representation in Staphylococcus. Interestingly, antibiotic use in The Netherlands is amongst the lowest in the developed world.[37] Such practices may influence the microbiota through selective microbial suppression and could partially account for the unique microbiome observed in Amsterdam. Overall, these findings suggest that there are differences in the sinonasal microbiome composition across geographical regions, but these differences would not alter the general pattern of core organisms described above. The driving factors behind geographical variations are yet to be elucidated - possible influences are diet, lifestyle, antibiotic prescribing patterns or environmental exposures.

Previous 16S rRNA gene sequence analyses of sinus microbiota have detected and ascribed importance to organisms such as *Cyanobacteria*, *Bacteroidetes*, *Propionibacterium* and *Acinetobacter*.[25,27–29,38] These organisms were either not encountered or detected only in minute abundances in our analysis and given their unexpected nature in the nasal cavity, they could represent a source of contamination (either during sample collection or from DNA extraction reagents) or artefacts from bioinformatics pipelines. *Cyanobacterium*, for example is an environmental organism and *Bacteroidetes* is typically associated with the colon.[39–41] In light of the limitations of current methodologies, reports of airway organisms that are novel, rare, and that do not form a part of the core microbiome, should therefore be interpreted cautiously. It is expected that the resolution of identification and functional characterization of bacterial species and strains in the sinuses can only improve, with advancement in sequencing technologies.

The organisms *Corynebacterium* and *Staphylococcus* are of particular interest. It was interesting to note that these organisms remained the two most abundant genera across geographical divides. Furthermore, amongst our control cohort, they were present in almost all individuals, again suggesting a key commensal function in the healthy state. Our findings of great abundance in both healthy patients and CRS patients mirrors previous studies.[22] While *Corynebacterium* has traditionally been thought of as a nasal commensal,[42] some studies suggest that certain species such as *C. tuberculostearicum* may be involved in CRS pathogenesis.[43] We could not resolve the *Corynebacterium genus* in our current analysis

bioRxiv preprint doi: https://doi.org/10.1101/548743; this version posted February 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

to the species level, given the limitations of short-read 16S rRNA gene sequencing, but it is likely that a number of different corynebacterial species reside within the nose – the majority as commensals. Recent evidence shows that certain species of *Corynebacterium* are beneficial in the nasal airways. *Corynebacterium accolens*, for example, is a common nasal colonizer, and can inhibit streptococcal growth by releasing oleaic acids through hydrolysis of host triacylglycerols.[44] In contrast staphylococcal species, and particularly *S. aureus*, by contrast have typically been viewed as potential pathogenic organisms within the nose. While being an asyptomatic coloniser in some individuals,[45] *S. aureus* contributes to severe antibiotic- and surgery-resistant CRS.[46–48] Our confirmation of high prevalence and relative abundance of *Staphylococcus* amongst controls suggests a role in maintaining a healthy sinus. Poor resolution of 16S analysis beyond the genus level results in both commensal and potentially pathogenic *Staphylococcus* species being grouped together. The outcome is that the healthy sinonasal microbiome may be composed of coagulase-negative commensal staphylococcal species, *S. aureus*, or more likely, a combination of both. If *S. aureus* is present in high abundance within the healthy sinus then perhaps this species plays a dual role within the sinuses: being essential for normal function typically, but sometimes becoming virulent at times of disease. The trigger for such a switch in roles is yet to be clarified but could be explained by the presence of *Corynebacterium*. Ramsey et al. described the ability of *C. striatum* to shift the gene profile of *S. aureus* away from virulence and towards commensalism.[49] This hypothesis would be supported by our result demonstrating reduced *Corynebacterium* amongst CRSwNP. The depletion of these organisms may allow *S. aureus* to switch on virulence genes, propagating disease. Interestingly, CRSwNP is a more severe and resistant form of disease and has previously been linked with S. aureus virulence factors and, in particular, superantigens.[50,51] Future studies and techniques which are able to identify microbiota composition to the species or strain level are required to further our understanding of these processes. Ultimately, the results from both current and previous findings highlight the complexity of the sinonasal microbiome and the myriad of functions (both human and bacterial) that act in concert to maintain homeostasis or produce disease. While we have presented data that has solidified our understanding of the upper airway microbial ecology, investigations

bioRxiv preprint doi: https://doi.org/10.1101/548743; this version posted February 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

355 into the microbiome function or metatranscriptome may provide further novel perspectives on important
356 and critical microbial and host pathways.

357 CONCLUSION

358 Understanding the characteristics of the sinonasal microbiome may provide novel insights to the normally
359 functioning upper airway. This may increase our understanding of the pathobiology of diseases such as
360 CRS. This study is the largest yet to describe the sinus microbiome and the first to examine geographical
361 variation. We demonstrate that the core microbiome is composed of *Corynebacterium*, *Staphylococcus*,
362 *Streptococcus*, *Moraxella* and *Haemophilus*. These organisms are present across disease phenotypes and
363 countries. Utilising a large cohort and standardised methodology has allowed us to better characterise the
364 sinonasal microbiome. By doing so, we have presented a foundation for future prospective studies into
365 pathological states as well as functional analysis.

366

bioRxiv preprint doi: https://doi.org/10.1101/548743; this version posted February 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

METHODS

Participating centres

A total of fourteen centres participated in the completion of this investigation. Thirteen centres provided patient samples for utilisation within the study. One centre (Northern Arizona University) provided bioinformatics expertise and consultation. The project was approved by the respective institutional human research ethics boards of all sample-collection centres (Table S1). Participating centres are listed below:

1. University of Adelaide, Adelaide, Australia (Lead Centre)
2. University of Sydney, Sydney, Australia
3. University of Auckland, Auckland, New Zealand
4. Siriraj Hospital, Mahidol University, Bangkok, Thailand
5. LV Prasad Institute, Hyderabad, India
6. University of Sao Paulo, Sao Paulo, Brazil
7. Catholic University of Chile, Santiago, Chile
8. Academic Medical Centre, Amsterdam, The Netherlands
9. McGill University, Montreal, Canada
10. Stanford University, California, USA
11. Harvard Medical School, Boston, USA
12. Medical University of South Carolina, Charleston, USA
13. University of Texas, Texas, USA
14. Northern Arizona University, Arizona, USA (Bioinformatics expertise)

Study design

This study was a multi-centre, international collaborative investigation with a prospective, cross-sectional design. In total, thirteen centres across 9 countries provided samples for analysis. Written consent to tissue and clinical data collection was procured from all participants prior to surgery. Collection was performed during either endoscopic sinonasal surgery or ancillary otolaryngological procedures, such as tonsillectomy, septoplasty or skull base tumour resections. Individuals were classified as having CRS if they fulfilled the criteria outlined in the International Consensus statement on Allergy and Rhinology:

bioRxiv preprint doi: https://doi.org/10.1101/548743; this version posted February 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Rhinosinusitis.[11] Further sub-classification according to the absence (CRSsNP) or presence (CRSwNP) of nasal polyps was performed by endoscopic assessment at the time of surgery. Patients who underwent ancillary otolaryngological procedures and had no clinical or radiological evidence of CRS formed the healthy control cohort.

Metadata collection

Clinical data was collected using standardized patient questionnaires undertaken by each participant at the time of study consent. In non-English speaking countries, the questionnaire was translated by a qualified interpreter and checked for accuracy by the lead investigator from that country. Collected data included patient demographics, medical history, ethnicity, social and environmental exposures as well as quality of life scoring via the validated SNOT-22 and Visual Analogue Scores (VAS). The standardized questionnaires used (which we termed the Open Source Sinonasal Survey "OS3") are freely available at https://github.com/adelaide-orl/os3.

Sample collection and transport

Each centre was asked to provide microbiome swabs from non-CRS controls and CRS patients for analysis. Patients were all anaesthetised at the time of sample acquisition. All samples were collected in a standardized manner prior to commencement of any surgical intervention, whilst the nasal cavity remained unaltered. Microbiome swabs were collected intra-operatively using guarded and endoscopically-guided Copan Flocked swabs (COPAN ITALIA, Brescia, Italy) to sample the middle meatus.[24,52] Swab heads were subsequently separated into sterile cryotubes (Sarstedt, Numbrecht, Germany), placed on ice immediately, and then transported to -80°C storage. Once all samples had been acquired from a centre, they were then transported to the lead centre (Adelaide) using a secure cold chain (Cryoport, Irvine, CA, USA). This was to ensure standardized down-stream processing. The standard Cryoport containers used for shipment of samples are liquid nitrogen dewars capable of keeping stored samples at a stable temperature for at least 15 days. These containers have continuous temperature-

bioRxiv preprint doi: https://doi.org/10.1101/548743; this version posted February 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

418 monitoring to ensure preservation of the cold chain throughout the shipment. Any evidence of
419 temperature disturbance, or displacement or damage to transported cryotubes resulted in those samples
420 being excluded from further processing.

421 **DNA Extraction**

422 All DNA extraction was performed at the lead centre using Qiagen DNeasy Blood and Tissue Kit
423 (Qiagen, Hilden, Germany), as per manufacturer's instructions. Total DNA was extracted from all clinical
424 samples as well as a DNA-negative control with extraction reagents only. All extractions were undertaken
425 in strict sterile conditions, utilising new equipment for each sample to exclude cross-contamination. In
426 brief, swab heads were prepared for extraction by being cut to 2-3mm pieces and placed into a
427 microcentrifuge tube. A lysozyme (Sigma, St. Louis, USA) solution at 20mg/ml in Lysis buffer (20 mM
428 Tris-Cl, pH8; 2mM sodium EDTA; 1.2% Triton X-100, filter sterilised; Sigma, St Louis, USA) was
429 added to each sample and left overnight at room temperature. Samples were then homogenised using
430 5mm steel beads and a Tissue Lyser II (Qiagen, Hilden, Germany) at 15Hz for 20 seconds. Steel beads
431 were then removed prior to further homogenisation with 50mg glass beads, again using the Tissue Lyser
432 II at 30Hz for 5 minutes. Proteinase K and Buffer AL (Qiagen, Hilden, Germany) were added to each
433 sample and left to incubate for 30 minutes at 56°C. Tubes were then centrifuged briefly to collect beads
434 and supernatant transferred to a fresh microcentrifuge tube. After addition of 100% ethanol to
435 supernatant, the new mixture was pipetted into DNeasy Mini Spin columns (Qiagen, Hilden, Germany).
436 Subsequent extraction of DNA from supernatant mixture was as per manufacturer instructions, with a
437 total of 100ul of DNA extracted per sample. Concentration was determined using a NanoDrop 1000
438 Spectrophotometer (Thermo Scientific, Massachusetts, USA). Extracted DNA was stored at -80°C until
439 sequencing. Any samples that were suspected to be improperly handled, contaminated or did not pass
440 quality control during processing were excluded. A total of 532 samples (126 CRSsNP; 212 CRSwNP;
441 194 Controls) passed all stages of transport and processing to be sent to the sequencing facility
442 (Australian Genome Research Facility; AGRF, Melbourne, Vic, Australia).

bioRxiv preprint doi: https://doi.org/10.1101/548743; this version posted February 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

443 **Polymerase Chain Reaction (PCR) Amplification of the 16S rRNA gene and sequencing**

444 PCR and sequencing were performed by AGRF. Libraries were generated by amplifying (341F–806R)
445 primers against the V3–V4 hypervariable region of the 16S rRNA gene. (CCTAYGGGRBGCASCAG
446 forward primer; GGACTACNNGGGTATCTAAT reverse primer).[53] PCR was done using AmpliTaq
447 Gold 360 master mix (Life Technologies, Mulgrave, Australia) following a two-stage PCR protocol (29
448 cycles for the first stage; and 14 cycles for the second, indexing stage). Concentrations of the resulting
449 amplified amplicons were measured using fluorometry (Invitrogen Picogreen; Thermo Fisher Scientific,
450 Waltham, MA, USA). Amplicons were normalised according to the obtained concentrations prior to
451 sequencing. Sequencing was done on the Illumina MiSeq platform (Illumina Inc., San Diego, CA, USA)
452 with the 300-base-pairs paired-end chemistry over 8 runs.

453 **Bioinformatics pipeline**

454 Demultiplexed fastq files were received from the sequencing facility. We used the new QIIME 2 (version
455 2018.11)[30] for our bioinformatics pipeline, utilizing various QIIME 2 plugins at each step. Forward and
456 reverse reads were joined using PEAR[54] through the QIIME 2 plugin q2-pear
457 (https://github.com/bassio/q2-pear). Joined sequences were then quality-filtered using the QIIME 2 plugin
458 q2-quality-filter,[55] with minimum quality 20, according to recommendations.[56] This was followed by
459 abundance-filtering applied on the reads, according to the method by Wang et al.[57], through the python
460 implementation in the QIIME 2 plugin q2-abundance-filtering (https://github.com/bassio/q2-abundance-
461 filtering). Denoising and Amplicon Sequence Variant (ASV) formation were done using deblur[56] through
462 the q2-deblur plugin using the parameters (trim-size = 435; min-size = 1; min-size = 1). Taxonomy
463 assignment was done against the Greengenes 16S reference database (the 99% clustered similarity
464 sequences),[58] version 13.8 (August 2013) using the BLAST-based classifier implemented in QIIME 2
465 (q2-feature-classifier)[59] and which implements a Lowest Common Ancestor (LCA) consensus algorithm.
466 To address limitations of de novo trees generated from short-length ASVs, we utilized the SATé-enabled
467 phylogenetic placement (SEPP) technique[60] for insertion of the ASVs into the high-quality tree generated

bioRxiv preprint doi: https://doi.org/10.1101/548743; this version posted February 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

from the 99% OTUs Greengenes reference database, and the ASVs that did not fit anywhere into the tree were filtered out of the ASV table.

A rarefaction depth cut-off was chosen at 400 before downstream diversity analysis and comparisons of relative abundances of taxa. Alpha rarefaction plots of unique number of ASVs in each sample (i.e. richness), as well as Shannon's diversity index confirmed almost all samples reaching a plateau at this depth indicating sufficient sampling. (Supplementary Figure S2) Applying this depth yielded 410 (out of 532) samples for downstream analysis. Taxa were mostly compared at the genus level. Mean relative abundance as well as prevalence of the genera were calculated for each group. Faith's phylogenetic diversity index[31] was used for alpha diversity and weighted Unifrac[32] distance matrices were calculated for beta diversity analyses. Diversity metrics were generated through sci-kit bio version 0.5.3.

Statistical analysis

Statistics were done using packages from the Python Scientific Stack[61] and R (R Foundation for Statistical Computing, Vienna, Austria) through the Jupyter notebook interface[62], utilizing the assistance of packages from the Scientific Python[61] stack (numpy, scipy, pandas, statsmodels), scikit-bio (https://github.com/biocore/scikit-bio) and omicexperiment (https://www.github.com/bassio/omicexperiment).

We investigated the relative abundances of genera in different subgroups using linear mixed model analysis (R packages "lme4" and "lmerTest"). Linear mixed effects modelling was performed to control for the "centre" variable, which was included in the model as a random effect. The mixed models were fit using the restricted maximum likelihood (REML) as implemented in the default method in the "lme4" package. Mean fixed effects of variables were extracted from the model objects in R using the R package "emmeans", (the successor to "lsmeans").[63] The p values for the covariates in the mixed models were generated using t-tests using Satterthwaite's method as implemented in the "lmerTest" package.[64]

bioRxiv preprint doi: https://doi.org/10.1101/548743; this version posted February 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Comparison of mean relative abundances of the top 10 taxa between centres, and comparison of mean alpha diversity indices between disease groups and centres, were performed using Mann-Whitney-Wilcoxon tests, with multiple comparisons correction using the Benjamini-Hochberg method.[65] For multivariate analysis of beta diversity metrics, we employed permutational multiple analysis of variance (PERMANOVA)[66] implemented in the function “adonis” from the R package “vegan”.[67]

bioRxiv preprint doi: https://doi.org/10.1101/548743; this version posted February 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Supplementary Tables

Table S1: Ethics approvals.

Centre	Approval
University of Adelaide, Adelaide, Australia	The Queen Elizabeth Hospital Human Research Ethics Committee (approval HREC/14/TQEHLMH/222)
University of Sydney, Sydney, Australia	HREC Reference number: LNR/14/SVH/373
University of Auckland, Auckland, New Zealand	HDC Northern Regional Ethics Committee approval NTX/08/12/126/AM13
Siriraj Hospital, Mahidol University, Bangkok , Thailand	Siriraj Institutional Review Board: Certificate of Approval 246/2559 (EC1)
LV Prasad Institute, Hyderabad, India	Local ethics approval number LEC-01-15-001
University of Sao Paulo, Sao Paulo, Brazil	Local ethics approval number CAAE: 46421415.8.0000.0068
Catholic University of Chile, Santiago, Chile	Scientific Ethics Committee, Faculty of Medicine, Pontifical Catholic University of Chile (reference number 14-435)
Academic Medical Center, Amsterdam, The Netherlands	METC number: W14_282 # 14.17.0338
McGill University, Montreal, Canada	McGill University Health Centre, Biomedical D Research Ethics Board, Study Number 14-433 BMD
Stanford University, California, USA	Ethics Protocol ID : 33326
Harvard Medical School, Boston, USA	Massachusetts Eye and Ear IRB Approval #651070-9
Medical University of South Carolina, Charleston, USA	Ethics/Institutional Review Board: Pro00038979
University of Texas, Texas, USA	Ethics approval HSC-MS-16-0853

bioRxiv preprint doi: https://doi.org/10.1101/548743; this version posted February 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

499 *Table S2A: Abundance and prevalence of genera in CRSsNP patients.*

Genera	Mean Relative Abundance (%)	Prevalence (%)
Corynebacterium	44.43	77.78
Staphylococcus	27.42	69.7
Haemophilus	6.22	19.19
Streptococcus	5.98	27.27
Moraxella	3.18	8.08
Pseudomonas	2.88	15.15
Serratia	1.16	2.02
unidentified (Moraxellaceae)	0.88	4.04
unidentified (Enterobacteriaceae)	0.86	11.11
unidentified (Neisseriaceae)	0.81	20.2

500

bioRxiv preprint doi: https://doi.org/10.1101/548743; this version posted February 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

501 *Table S2B: Abundance and prevalence of genera in CRSwNP patients.*

Genera	Mean Relative Abundance (%)	Prevalence (%)
Corynebacterium	40.01	75.58
Staphylococcus	25.74	73.26
Streptococcus	7.35	31.98
Haemophilus	6.12	15.12
Moraxella	3.9	9.88
unidentified (Enterobacteriaceae)	3.2	16.86
Pseudomonas	2.59	6.4
Alloiococcus	2.2	22.67
Klebsiella	0.88	2.33
Acinetobacter	0.74	5.23

502

bioRxiv preprint doi: https://doi.org/10.1101/548743; this version posted February 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Supplementary Figures

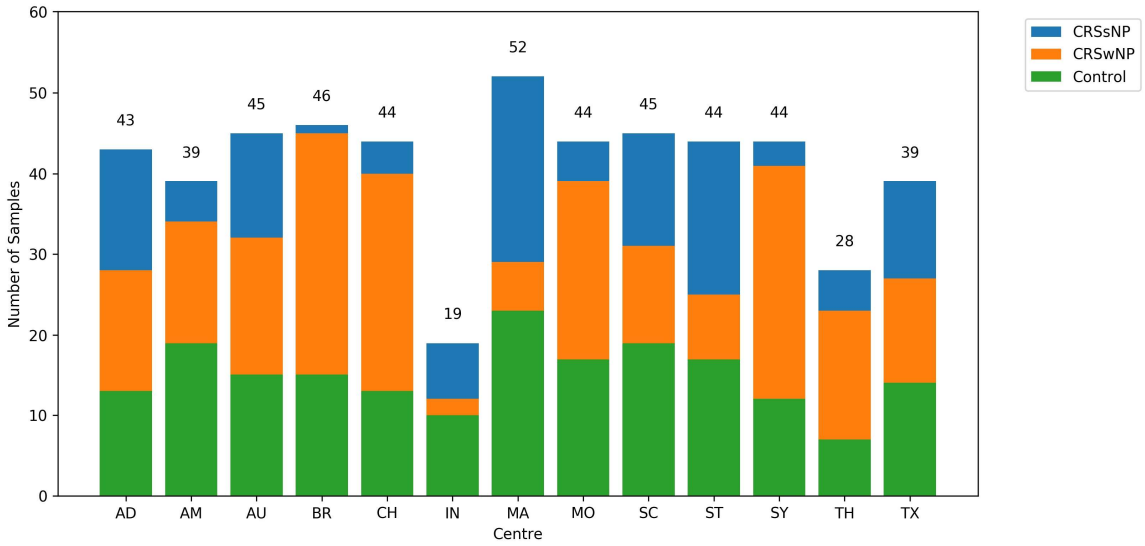


Figure S1: Collected sample disease characteristics by centre. *Figure demonstrating the disease characteristics of patients from each centre. Each bar represents cumulative numbers. Control patients are indicated by, CRSsNP by blue and CRSwNP by orange. Final contributing numbers are indicated at the peak of each bar. AD = Adelaide; AM = Amsterdam; AU = Auckland; BR = Brazil; CH = Chile; IN = India; MA = Massachusetts; MO = Montreal; SC = South Carolina; ST = Stanford; SY = Sydney; TH = Thailand; TX = Texas.*

bioRxiv preprint doi: https://doi.org/10.1101/548743; this version posted February 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

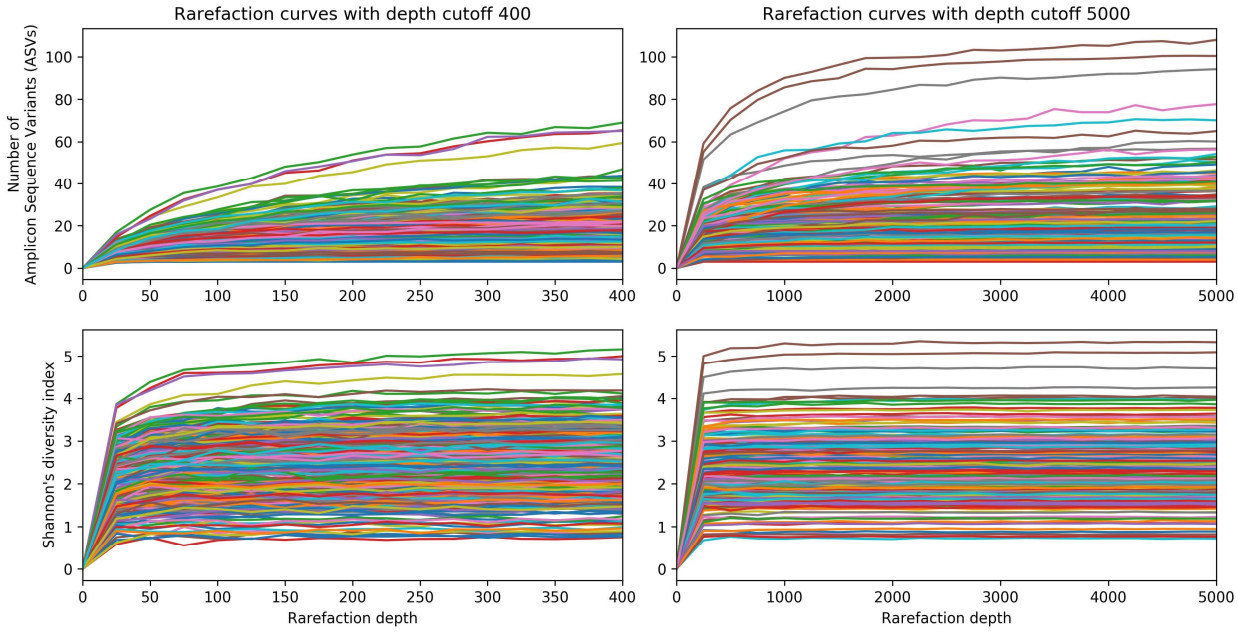


Figure S2: Rarefaction analysis: Rarefaction curves plot. *A rarefaction depth of 400 was chosen before downstream diversity analysis and comparisons of relative abundances of taxa. Rarefaction plots (at depth cutoffs 400 and 5000) of richness (total number of ASVs per sample) and Shannon's alpha diversity index confirmed almost all samples reaching a plateau at this depth indicating sufficient sampling. Each line in the plot represents a single sample. ASVs = Amplicon Sequence Variants.*

bioRxiv preprint doi: https://doi.org/10.1101/548743; this version posted February 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

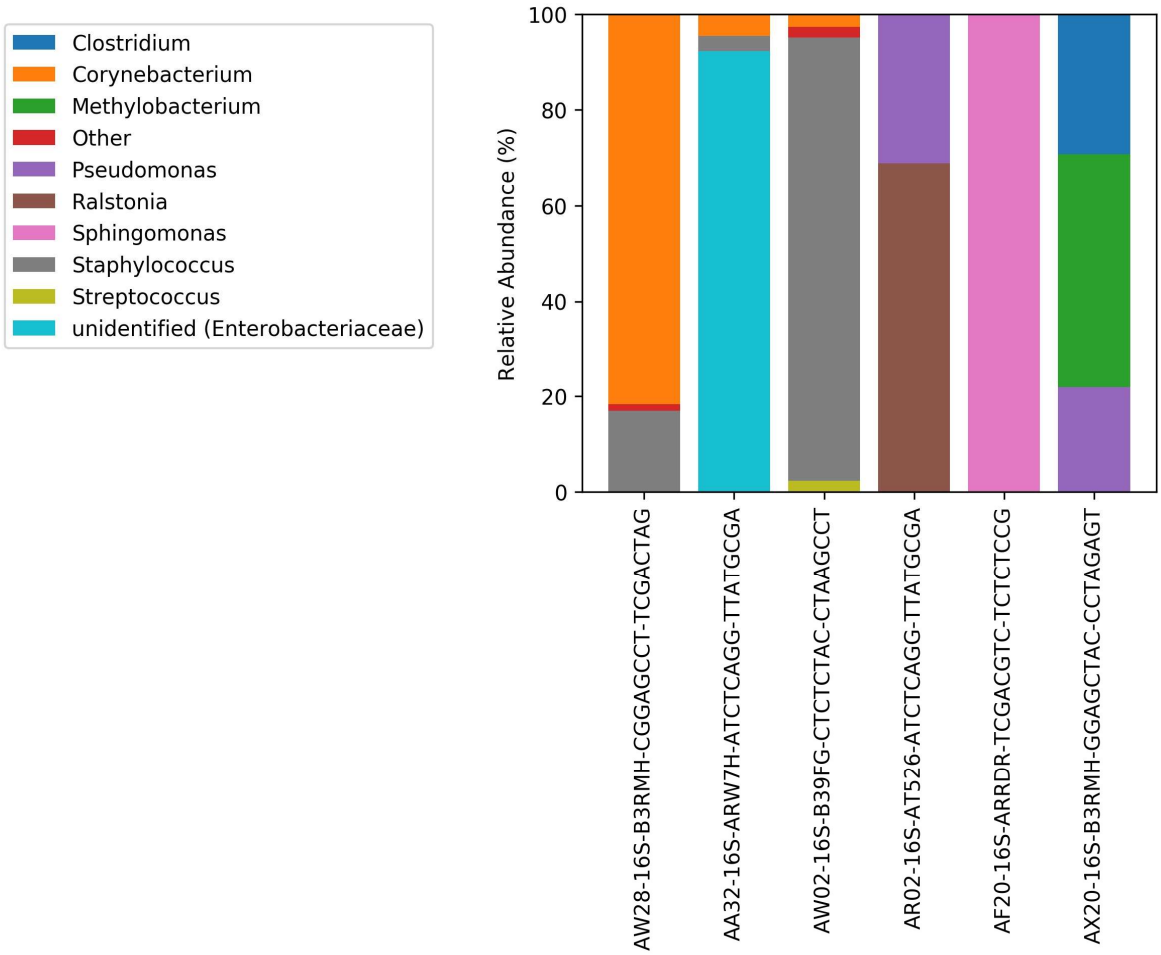


Figure S3: Taxa bar plots of the negative template control samples (i.e. DNA-negative samples with extraction reagents only). *The rightmost three (of the six) samples were below the rarefaction threshold depth of 400 reads.*

bioRxiv preprint doi: https://doi.org/10.1101/548743; this version posted February 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

References

1. Yatsunenko, T. *et al.* Human gut microbiome viewed across age and geography. *Nature* **486,** 222–227 (2012).

2. Trompette, A. *et al.* Gut microbiota metabolism of dietary fiber influences allergic airway disease and hematopoiesis. *Nature Medicine* **20,** 159–166 (2014).

3. Lynch, S. V. & Pedersen, O. The Human Intestinal Microbiome in Health and Disease. *The New England Journal of Medicine* **375,** 2369–2379 (2016).

4. Cryan, J. F. & O'Mahony, S. M. The microbiome-gut-brain axis: From bowel to behavior. *Neurogastroenterology and Motility: The Official Journal of the European Gastrointestinal Motility Society* **23,** 187–192 (2011).

5. Bhattacharjee, S. & Lukiw, W. J. Alzheimer's disease and the microbiome. *Frontiers in Cellular Neuroscience* **7,** 153 (2013).

6. Fulde, M. & Hornef, M. W. Maturation of the enteric mucosal innate immune system during the postnatal period. *Immunological Reviews* **260,** 21–34 (2014).

7. Lloyd, C. M. & Marsland, B. J. Lung Homeostasis: Influence of Age, Microbes, and the Immune System. *Immunity* **46,** 549–561 (2017).

8. Arumugam, M. *et al.* Enterotypes of the human gut microbiome. *Nature* **473,** 174–180 (2011).

9. Eckburg, P. B. *et al.* Diversity of the Human Intestinal Microbial Flora. *Science* **308,** 1635–1638 (2005).

10. Xavier, R. J. & Podolsky, D. K. Unravelling the pathogenesis of inflammatory bowel disease. *Nature* **448,** 427–434 (2007).

11. Orlandi, R. R. *et al.* International Consensus Statement on Allergy and Rhinology: Rhinosinusitis. *International Forum of Allergy & Rhinology* **6 Suppl 1,** S22–209 (2016).

12. Head, K. *et al.* Systemic and topical antibiotics for chronic rhinosinusitis. *The Cochrane Database of Systematic Reviews* **4,** CD011994 (2016).

13. *Cummings Otolaryngology - Head and Neck Surgery*. (Elsevier Health Sciences, 2014).

14. Boutin, S. *et al.* Comparison of microbiomes from different niches of upper and lower airways in children and adolescents with cystic fibrosis. *PloS One* **10,** e0116029 (2015).

15. van Dongen, T. M. A. *et al.* Evaluation of concordance between the microorganisms detected in the nasopharynx and middle ear of children with otitis media. *The Pediatric Infectious Disease Journal* **32,** 549–552 (2013).

16. Psaltis, A. J. & Wormald, P.-J. Therapy of Sinonasal Microbiome in CRS: A Critical Approach. *Current Allergy and Asthma Reports* **17,** 59 (2017).

17. Benninger, M. S. *et al.* Adult chronic rhinosinusitis: Definitions, diagnosis, epidemiology, and pathophysiology. *Otolaryngology–Head and Neck Surgery: Official Journal of American Academy of Otolaryngology-Head and Neck Surgery* **129,** S1–32 (2003).

bioRxiv preprint doi: https://doi.org/10.1101/548743; this version posted February 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

18. Khalid, A. N., Quraishi, S. A. & Kennedy, D. W. Long-term quality of life measures after functional endoscopic sinus surgery. *American Journal of Rhinology* **18,** 131–136

19. Gliklich, R. E. & Metson, R. The health impact of chronic sinusitis in patients seeking otolaryngologic care. *Otolaryngology–Head and Neck Surgery: Official Journal of American Academy of Otolaryngology-Head and Neck Surgery* **113,** 104–109 (1995).

20. Ahn, J. *et al.* Oral microbiome profiles: 16S rRNA pyrosequencing and microarray assay comparison. *PloS One* **6,** e22788 (2011).

21. Franzosa, E. A. *et al.* Relating the metatranscriptome and metagenome of the human gut. *Proceedings of the National Academy of Sciences of the United States of America* **111,** E2329–2338 (2014).

22. Anderson, M., Stokken, J., Sanford, T., Aurora, R. & Sindwani, R. A systematic review of the sinonasal microbiome in chronic rhinosinusitis. *American Journal of Rhinology & Allergy* **30,** 161–166 (2016).

23. Cope, E. K., Goldberg, A. N., Pletcher, S. D. & Lynch, S. V. Compositionally and functionally distinct sinus microbiota in chronic rhinosinusitis patients have immunological and clinically divergent consequences. *Microbiome* **5,** 53 (2017).

24. Bassiouni, A., Cleland, E. J., Psaltis, A. J., Vreugde, S. & Wormald, P.-J. Sinonasal microbiome sampling: A comparison of techniques. *PloS One* **10,** e0123216 (2015).

25. Cleland, E. J., Bassiouni, A., Vreugde, S. & Wormald, P.-J. The bacterial microbiome in chronic rhinosinusitis: Richness, diversity, postoperative changes, and patient outcomes. *American Journal of Rhinology & Allergy* **30,** 37–43 (2016).

26. Frank, D. N. *et al.* The Human Nasal Microbiota and Staphylococcus aureus Carriage. *PLoS ONE* **5,** e10598 (2010).

27. Aurora, R. *et al.* Contrasting the microbiomes from healthy volunteers and patients with chronic rhinosinusitis. *JAMA otolaryngology– head & neck surgery* **139,** 1328–1338 (2013).

28. Stephenson, M.-F. *et al.* Molecular characterization of the polymicrobial flora in chronic rhinosinusitis. *Journal of otolaryngology - head & neck surgery = Le Journal d'oto-rhino-laryngologie et de chirurgie cervico-faciale* **39,** 182–187 (2010).

29. Choi, E.-B. *et al.* Decreased diversity of nasal microbiota and their secreted extracellular vesicles in patients with chronic rhinosinusitis based on a metagenomic analysis. *Allergy* **69,** 517–526 (2014).

30. Bolyen, E. *et al. QIIME 2: Reproducible, interactive, scalable, and extensible microbiome data science*. (PeerJ Inc., 2018). doi:10.7287/peerj.preprints.27295v1

31. Faith, D. P. Conservation evaluation and phylogenetic diversity. *Biological Conservation* **61,** 1–10 (1992).

32. Lozupone, C. A., Hamady, M., Kelley, S. T. & Knight, R. Quantitative and qualitative beta diversity measures lead to different insights into factors that structure microbial communities. *Applied and environmental microbiology* **73,** 1576–1585 (2007).

33. Massa, H. M., Cripps, A. W. & Lehmann, D. Otitis media: Viruses, bacteria, biofilms and vaccines. *The Medical Journal of Australia* **191,** S44–49 (2009).

bioRxiv preprint doi: https://doi.org/10.1101/548743; this version posted February 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

598 34. Leibovitz, E., Broides, A., Greenberg, D. & Newman, N. Current management of pediatric acute otitis
599 media. *Expert Review of Anti-Infective Therapy* **8,** 151–161 (2010).

600 35. Gul, M. *et al.* The comparison of tonsillar surface and core cultures in recurrent tonsillitis. *American*
601 *Journal of Otolaryngology* **28,** 173–176

602 36. Wagner Mackenzie, B. *et al.* Bacterial community collapse: A meta-analysis of the sinonasal
603 microbiota in chronic rhinosinusitis. *Environmental Microbiology* **19,** 381–392 (2017).

604 37. Goossens, H., Ferech, M., Coenen, S., Stephens, P. & European Surveillance of Antimicrobial
605 Consumption Project Group. Comparison of outpatient systemic antibacterial use in 2004 in the United
606 States and 27 European countries. *Clinical Infectious Diseases: An Official Publication of the Infectious*
607 *Diseases Society of America* **44,** 1091–1095 (2007).

608 38. Hauser, L. J. *et al.* Sinus culture poorly predicts resident microbiota. *International Forum of Allergy*
609 *& Rhinology* (2014). doi:10.1002/alr.21428

610 39. Moisander, P. H. *et al.* Unicellular cyanobacterial distributions broaden the oceanic N2 fixation
611 domain. *Science (New York, N.Y.)* **327,** 1512–1514 (2010).

612 40. Van de Waal, D. B. *et al.* Reversal in competitive dominance of a toxic versus non-toxic
613 cyanobacterium in response to rising CO2. *The ISME journal* **5,** 1438–1450 (2011).

614 41. Krieg, N. R., Ludwig, W., Euzéby, J. & Whitman, W. B. Phylum XIV. Bacteroidetes phyl. Nov. in
615 *Bergey's manual of systematic bacteriology* 25–469 (Springer, 2010).

616 42. Kaspar, U. *et al.* The culturome of the human nose habitats reveals individual bacterial fingerprint
617 patterns. *Environmental Microbiology* **18,** 2130–2142 (2016).

618 43. Abreu, N. A. *et al.* Sinus microbiome diversity depletion and Corynebacterium tuberculostearicum
619 enrichment mediates rhinosinusitis. *Science translational medicine* **4,** 151ra124 (2012).

620 44. Bomar, L., Brugger, S. D., Yost, B. H., Davies, S. S. & Lemon, K. P. Corynebacterium accolens
621 Releases Antipneumococcal Free Fatty Acids from Human Nostril and Skin Surface Triacylglycerols.
622 *mBio* **7,** e01725–01715 (2016).

623 45. Gorwitz, R. J. *et al.* Changes in the prevalence of nasal colonization with Staphylococcus aureus in
624 the United States, 2001-2004. *The Journal of Infectious Diseases* **197,** 1226–1234 (2008).

625 46. Archer, N. K. *et al.* Staphylococcus aureus biofilms: Properties, regulation, and roles in human
626 disease. *Virulence* **2,** 445–459

627 47. Cleland, E. J. *et al.* Probiotic manipulation of the chronic rhinosinusitis microbiome. *International*
628 *Forum of Allergy & Rhinology* **4,** 309–314 (2014).

629 48. Singhal, D., Foreman, A., Bardy, J.-J. & Wormald, P.-J. Staphylococcus aureus biofilms: Nemesis of
630 endoscopic sinus surgery. *The Laryngoscope* **121,** 1578–1583 (2011).

631 49. Ramsey, M. M., Freire, M. O., Gabrilska, R. A., Rumbaugh, K. P. & Lemon, K. P. Staphylococcus
632 aureus Shifts toward Commensalism in Response to Corynebacterium Species. *Frontiers in Microbiology*
633 **7,** 1230 (2016).

bioRxiv preprint doi: https://doi.org/10.1101/548743; this version posted February 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

50. Clark, D. W., Wenaas, A., Citardi, M. J., Luong, A. & Fakhri, S. Chronic rhinosinusitis with nasal polyps: Elevated serum immunoglobulin E is associated with Staphylococcus aureus on culture. *International Forum of Allergy & Rhinology* **1,** 445–450

51. Ou, J. *et al.* Staphylococcus aureus superantigens are associated with chronic rhinosinusitis with nasal polyps: A meta-analysis. *European archives of oto-rhino-laryngology: official journal of the European Federation of Oto-Rhino-Laryngological Societies (EUFOS): affiliated with the German Society for Oto-Rhino-Laryngology - Head and Neck Surgery* **271,** 2729–2736 (2014).

52. Warnke, P., Frickmann, H., Ottl, P. & Podbielski, A. Nasal Screening for MRSA: Different Swabs ! *PLoS ONE* **9,** e111627 (2014).

53. Yu, Y., Lee, C., Kim, J. & Hwang, S. Group-specific primer and probe sets to detect methanogenic communities using quantitative real-time polymerase chain reaction. *Biotechnology and Bioengineering* **89,** 670–679 (2005).

54. Zhang, J., Kobert, K., Flouri, T. & Stamatakis, A. PEAR: A fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics* **30,** 614–620 (2014).

55. Bokulich, N. A. *et al.* Quality-filtering vastly improves diversity estimates from Illumina amplicon sequencing. *Nature Methods* **10,** 57–59 (2013).

56. Amir, A. *et al.* Deblur Rapidly Resolves Single-Nucleotide Community Sequence Patterns. *mSystems* **2,**

57. Wang, J. *et al. Minimizing spurious features in 16S rRNA gene amplicon sequencing*. (PeerJ Inc., 2018). doi:10.7287/peerj.preprints.26872v1

58. DeSantis, T. Z. *et al.* Greengenes, a Chimera-Checked 16S rRNA Gene Database and Workbench Compatible with ARB. *Applied and Environmental Microbiology* **72,** 5069–5072 (2006).

59. Bokulich, N. A. *et al.* Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. *Microbiome* **6,** 90 (2018).

60. Janssen, S. *et al.* Phylogenetic Placement of Exact Amplicon Sequences Improves Associations with Clinical Information. *mSystems* **3,**

61. Oliphant, T. E. Python for Scientific Computing. *Computing in Science & Engineering* **9,** 10–20 (2007).

62. Kluyver, T. *et al.* Jupyter Notebooks a publishing format for reproducible computational workflows. in *Positioning and Power in Academic Publishing: Players, Agents and Agendas* (eds. Loizides, F. & Scmidt, B.) 87–90 (IOS Press, 2016). doi:10.3233/978-1-61499-649-1-87

63. Lenth, R. V. Least-Squares Means: The R Package lsmeans. *Journal of Statistical Software* **69,** 1–33 (2016).

64. Kuznetsova, A., Brockhoff, P. B. & Christensen, R. H. B. lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software* **82,** (2017).

65. Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)* **57,** 289–300 (1995).

bioRxiv preprint doi: https://doi.org/10.1101/548743; this version posted February 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

672 66. Anderson, M. J. A new method for non-parametric multivariate analysis of variance. *Austral Ecology*
673 **26,** 32–46 (2001).

674 67. Oksanen, J. *et al. Vegan: Community Ecology Package*. (2018).