Enhanced phylogenetic insights into the microbiome of chronic rhinosinusitis through the novel application of long read 16S rRNA gene amplicon sequencing*

J.T. Connell[1,2], K. Yeo[1], G. Bouras[1], A. Bassiouni[1,2], K. Fenix[1], C. Cooksley[1], S. Vreugde[1], P.J. Wormald[1,2], A.J. Psaltis[1,2]

[1] Department of Surgery, Otolaryngology Head and Neck Surgery, University of Adelaide and the Basil Hetzel Institute for Translational Health Research, Central Adelaide Local Health Network, Adelaide, South Australia, Australia

[2] Department of Otolaryngology, Head and Neck Surgery, The Queen Elizabeth Hospital, Woodville South, South Australia, Australia

Rhinology 62: 2, 152 - 162, 2024
https://doi.org/10.4193/Rhin23.333

***Received for publication:** September 5, 2023
Accepted: December 14, 2023

Abstract

Introduction: 16S rRNA next generation sequencing (NGS) has been the de facto standard of microbiome profiling. A limitation of this technology is the inability to accurately assign taxonomy to a species order. Long read 16S sequencing platforms, including Oxford Nanopore Technologies (ONT), have the potential to overcome this limitation. The paranasal sinuses are an ideal niche to apply this technology, being a low biomass environment where bacteria are implicated in disease propagation. Characterising the microbiome to a species order may offer new pathophysiological insights.

Methodology: Cohort series comparing ONT and NGS biological conclusions. Swabs obtained endoscopically from the middle meatus of 61 CRSwNP patients underwent DNA extraction, amplification and dual sequencing (Illumina Miseq (NGS) and ONT GridION). Agreement, relative abundance, prevalence, and culture correlations were compared.

Results: Mean microbiome agreement between sequencers was 61.4%. Mean abundance correlations were strongest at a familial/genus order and declined at a species order where NGS lacked resolution. The most significant discrepancies applied to Corynebacterium and Cutibacterium, which were estimated in lower abundance by ONT. ONT accurately identified 84.2% of cultured species, which was significantly higher than NGS.

Conclusions: ONT demonstrated superior resolution and culture correlations to NGS, but underestimated core sinonasal taxa. Future application and optimisation of this technology can advance our understanding of the sinonasal microenvironment.

Key words: nanopore, Illumina, NGS, 16S, microbiome

Introduction

The 16S rRNA gene has been a reliable target for microbiome characterisation since the inception of gene sequencing technology due to its ubiquity in all bacterial species. It contains constant genetic regions common to all bacteria that are optimal for primer binding, interlaced by hypervariable regions of DNA that are unique to each species allowing taxonomic discrimination. Next generation sequencers (NGS) superseded the first generation of sequencers some two decades ago, with an innovative and efficient parallel sequencing design[(1)]. By sequencing short 2x300 base-pair(bp) segments of the 1500bp 16S rRNA gene, NGS could rapidly estimate the bacterial composition of a sampled ecosystem, leading to widespread integration[(2)]. However, the design efficiency engendered certain limitations including incorrect taxonomic assignment, taxa bias and an inability to reliably assign sequences to a species order[(3–5)]. While NGS remains the widely accepted de facto standard of 16S microbiome sequencing, a new frontier of third generation sequencing is emerging.

Full length 'long read' sequencing is an evolution of NGS, with the capacity to sequence significantly longer nucleotide segments(10,000-50,000bp)[(6)]. In this manner, the entire 1500bp 16S rRNA gene can be sequenced, offering deeper phyloge-

netic resolution[7–10]. Oxford Nanopore Technologies (ONT) is a full length sequencing platform that uses an electric gradient to drive DNA/RNA through nanopores in an artificial biosensor membrane at a sequencing rate of 400 nucleotides per second, allowing long sequences to be detected in real time[11,12]. Validation of this technology has been performed in an array of biological niches, whereby ONT was compared with NGS sequencing[13–17]. Szoboszlay compared ONT and NGS in mock and fecal samples, concluding ONT provided a more accurate representation of mock communities and superior taxonomic assignment at a species resolution[13]. Matsuo reported comparable genus relative abundance and superior species resolution for ONT when applied to fecal samples[14]. Heikema and Rozas respectively came to similar conclusions regarding genus identification in the nose and skin, while simultaneously unearthing limitations of ONT in underestimating the relative abundance of certain taxa, advocating for further methodological optimisation[15,16].

The combination of a low bacterial biomass and a heavily implicated role of bacteria in chronic rhinosinusitis (CRS) has generated a unique demand for non-culture dependent research in the paranasal sinuses, resulting in extensive application of NGS. Early NGS studies offered conflicting abundance and diversity conclusions, largely driven by heterogeneous methodology and small sample sizes[18], but progressively a consensus sinus microbiome has been established, with *Staphylococcus* and *Corynebacterium* predominating[19–24]. The largest sinus microbiome study was a multicenter international series of 410 patients, which defined a core microbiome of *Corynebacterium*, *Staphylococcus*, *Streptococcus*, *Haemophilus* and *Moraxella* in health and disease[25]. The limitation of existing studies is the inability to establish species level conclusions, which is fundamental to further advance our understanding of the role of bacteria in CRS. To date, ONT full length 16S rRNA gene amplicon sequencing has never previously been applied in the paranasal sinuses.

Herein, the objectives of this study are to compare the biological conclusions of ONT and NGS and to characterise the sinonasal microbiome to a species order. The hypothesis of this study is that ONT will provide superior taxonomic resolution and comparable taxonomic accuracy to NGS.

Materials and methods

Study design

Ethics approval was obtained from the Central Adelaide Local Health Network HREC(HREC/14/TQEHLMH/222 and LNR13604). This study was a paired design cohort series comparing ONT (GridION) and NGS (Illumina Miseq) sequencing platforms on extracted DNA from middle meatal swabs. We utilised swabs from 61 patients aged over 18 who had no documented history of immune suppression and had a clinical diagnosis of CRSwNP based on the EPOS diagnostic criteria[26]. This included multiple CRS subtypes (idiopathic CRSwNP, eosinophilic CRS and allergic fungal rhinosinusitis). Clinical subgroup analyses are not presented in the current series, as this study forms part of a broader cohort series that will later examine unique clinical parameters. Written consent was obtained from participants at the time of swab collection.

Specimen collection and DNA extraction

Samples were obtained from the middle meatus under endoscopic guidance at the commencement of endoscopic sinus surgery, using sterile, guarded, flocked swabs. Swabs were stored in -80°Celsius freezers prior to DNA extraction. Extraction was performed by the Australian Genome Research Facility (AGRF, Melbourne, Australia) utilizing the Qiagen PowerSoil Pro DNA extraction kit (QIAGEN, Hilden, Germany). Quality control was performed with the Qubit Fluorometric Assays (Thermofisher, Waltham, MA, USA) Nanodrop Microdrop Spectrophotometers (Thermofisher), and gel electrophoresis. For NGS, the V3-V4(341F-806R) region was amplified utilizing a two-stage PCR protocol (Supplementary file 1) and sequenced using the Illumina MiSeq (Illumina Inc, San Diego, CA, USA) with 300bp chemistry. ONT sequencing was performed in-house at the Basil Hetzel Institute (University of Adelaide, Australia) applying an adaptation of the default ONT PCR protocol. In summary, initial denaturation 60 seconds (95°C); denaturation 25 x 20 seconds (95°C); annealing 25 x 30 seconds (62°C); extension 25 x 120 seconds (65°C); final extension 300 seconds (65°C), using the default ONT 16S Barcoding Kit (SQK16S024, ONT, Oxford, UK). Amplicons were primed using the ONT Flow Cell Priming Kit (EXP-FLP004, ONT) and added to the flow cell of the ONT GridION sequencer.

Bioinformatics pipeline

A quality threshold of >600 reads/sample was established, yielding 47 paired samples for downstream analysis. Demultiplexed fastq files were generated from Illumina MiSeq platform v2.6.2.1 (Illumina Inc) and Real Time Analysis (RTA) v1.18.54. QIIME2 V2020.2 were used for the NGS bioinformatics pipeline, applying SILVA database for taxonomic assignment. An in-house open-source pipeline, 'Coatofarms'[27] that implements EMU v3.4.5 was used for ONT bioinformatics, applying the default EMU database for taxonomic assignment[28]. ONT taxonomic assignment was additionally performed against SILVA as a comparator.

Statistical analysis

Statistics were performed on GraphPad Prism 10(GraphPad Software, Boston, MA, USA) and 'R' (R Foundation for Statistical Computing, Vienna, Austria). Mean relative abundance (MRA) was calculated for all taxa. Wilcoxon paired signed-rank abundance

Table 1. Sequencing data - all samples.

	ONT	NGS
Samples Sequenced	61	61
Total Reads	843,203	1,764,669
Mean Reads / Sample	13,823	28,929
Number Samples <600 reads (Excluded)	13	2
Final Samples Analysed (Final Cohort)	47	47
Total Genera Identified	136	109
Mean Genera / Sample	8.0	8.3
Genera > 1% Relative Abundance	11	9
Total Species Identified	264	61
Mean Species / Sample	11.6	2.6
Species > 1% Relative Abundance	17	3

comparisons were applied to taxa with MRA>1%. Prevalence comparisons were performed with Chi-square and Fisher's exact test. Aldex2, an estimation of technical variation within each sample per taxon, was applied utilizing Dirichlet distribution and a closely related log-ratio (CLR) transformation[29]. Wilcoxon tests were used for significance. Microbiome agreement was presented as descriptive data based on a calculation of the number of taxa in agreement for each sample; and the sum of the percentage of MRA in agreement per sample, presented as a mean, median and standard deviations[15]. An example of the agreement calculation agreement is provided in Supplementary file 2. Sequencing data was compared to bacterial culture results cultivated from the same site to assess the reliability of sequencing in identifying cultured organisms. The bacterial culture results presented were performed through hospital laboratories as part of the patients' standard of care. Alpha Diversity (Shannon's diversity, Simpson's diversity) and richness were calculated for each sample/group and Wilcoxon paired tests were performed to determine significance. Beta Diversity was calculated using a Bray Curtis dissimilarity model to calculate centroid ecological distances.

Results

Demographics

Participant age ranged from 27-85 (mean=52.1; SD=13.6). Male to female ratio was 41:20. All patients had a diagnosis of CRSwNP (including eCRS, AFRS and idiopathic CRSwNP). Mean Lund-McKay score was 31.9 (SD=10.6) and mean SNOT22 was 25.9 (SD=14.5).

Sequencing output

ONT generated 843,203 total reads (mean=13,823/sample) and NGS generated 1,764,669 reads (mean=28,929/sample). 14 samples with <600 reads were excluded (final cohort n=47). ONT assigned 136 genera(mean=8/sample) including 11 with >1% MRA (Table 1). NGS assigned 109 genera (mean=8.3/sample) including 9 with >1% MRA. ONT assigned 264 species (mean=11.6/sample) including 17 with >1% MRA. NGS assigned 61 species (mean=2.6/sample) including 3 with >1% MRA. 88.13% of NGS reads could not be assigned to a species order.

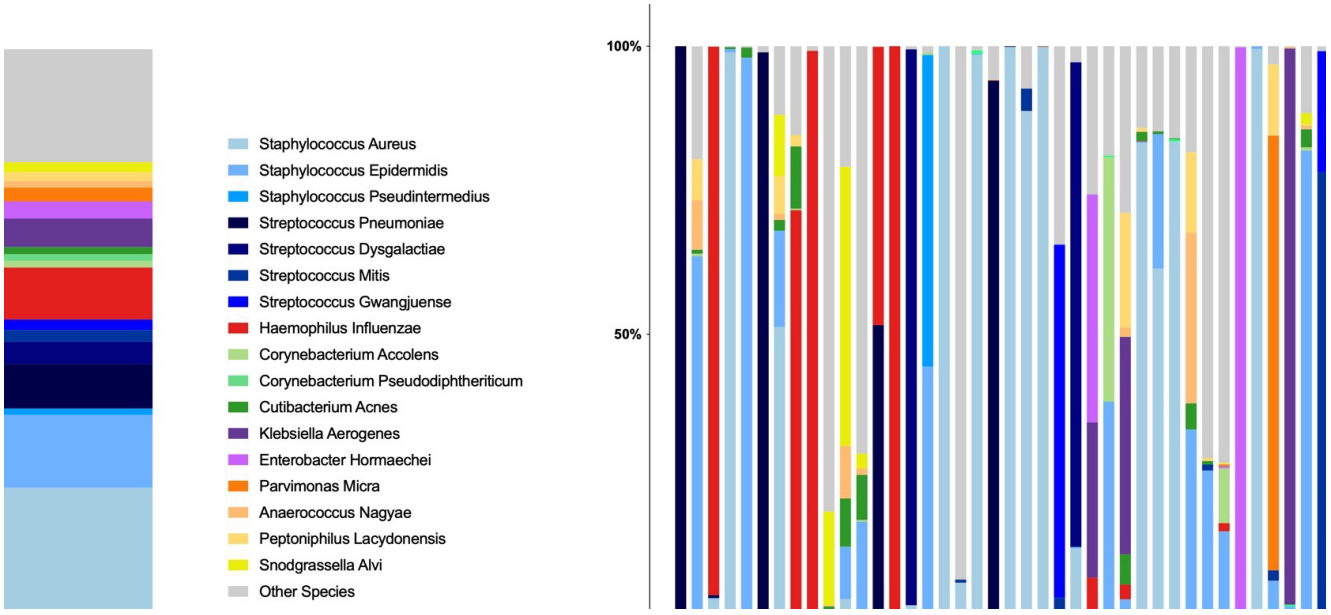


Figure 1. CRSwNP species microbiome (ONT). ONT GridION sequencing data for 47 CRSwNP patients. All species with a mean relative abundance >1% are presented. Species of low abundance are aggregated and represented as 'Other Species.' Combined mean relative abundance for all samples (left) individual sample relative abundance (right) is presented.

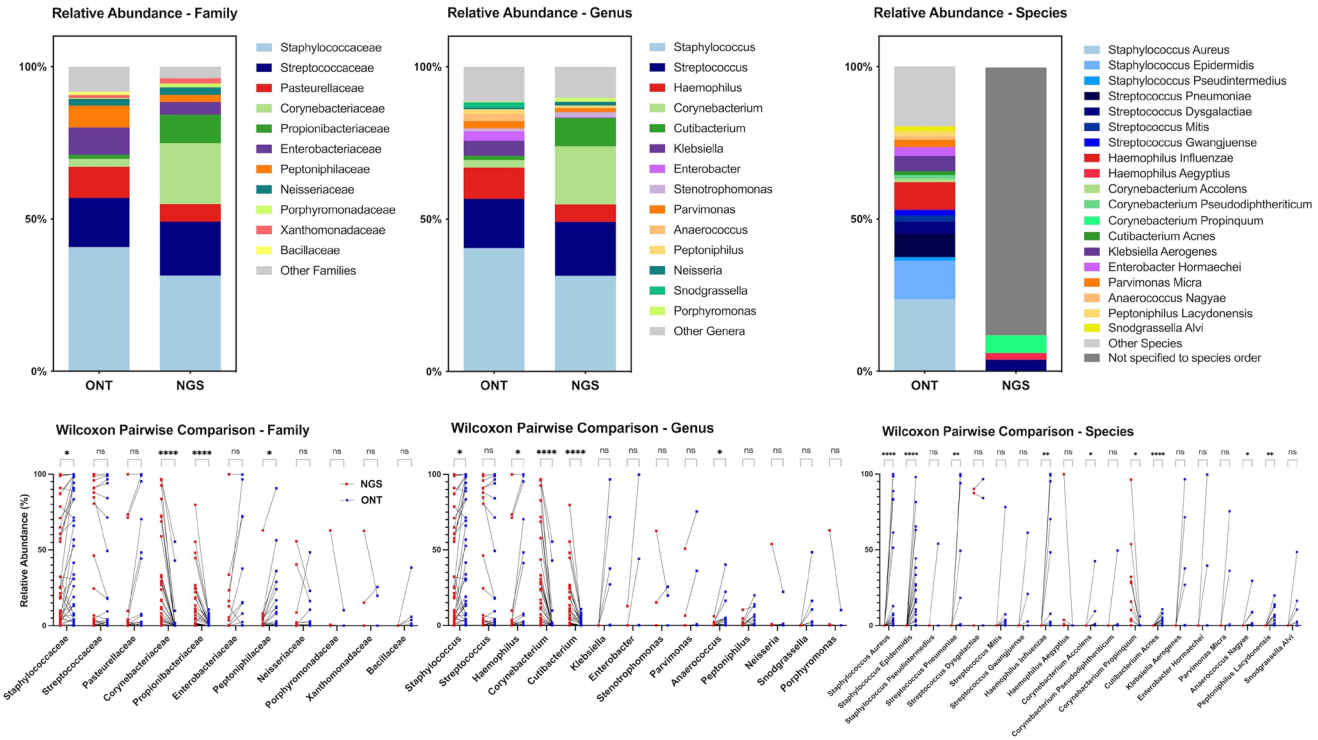


Figure 2. Paired Relative Abundance. Wilcoxon signed-rank tests.Relative Abundance Bar Graphs (Above): Mean Relative Abundance at a family (Left), Genus (Middle) and Species (Right) order, identifying all microbes with mean relative abundance >1%. The Species graph highlights the low resolution of NGS at this order with only 3 species identified in an abundance >1% and 88.13% of amplicons 'not specified to species order (dark grey). Wilcoxon Pairwise Comparison Graphs (Below): Wilcoxon paired signed-rank tests were performed for each organism >1% abundance at Family, Genus and species order, highlighting organisms with a significant difference in relative abundance based on the sequencing platform.

Species order microbiome

The sinonasal microbiome for the cohort of 47 patients presented to a species order using ONT sequencing is presented in Figure 1. 264 species were identified, including 16 discrete species of *Staphylococcus* (40.6% MRA, 85.1% prevalence). *Staphylococcus aureus* (23.7%) and *Staphylococcus epidermidis*(12.65%) had the highest MRA and prevalence, observed in 59.6%(28/47) and 61.7%(29/47) of samples respectively. *Streptococcus* (16 species, 16.1% MRA, 46.8% prevalence) and *Haemophilus* (5 species, 10.2% MRA, 40.4% prevalence) were also highly abundant, including *Streptococcus pneumoniae*, *Streptococcus dysgalactiae* and *Haemophilus influenzae*. *Corynebacterium* (9 species, 44.7% prevalence, 2.53% MRA) and *Cutibacterium* (3 different species, 42.6% prevalence, 1.34% MRA) were prevalent but low in abundance with *Corynebacterium accolens* (19.1% prevalence) and *Cutibacterium acnes* (40.4% prevalence) particularly prevalent.

Mean relative abundance comparison

ONT and NGS were strongly correlated at a familial order (Figure 2) with *Staphylococcaceae* the most abundant family for ONT and NGS. No significant difference in MRA was observed for 7 of 11 taxa with MRA >1% (Figure 2, Table 2). A statistically significant difference was observed for *Staphylococcaceae* (ONT=40.81%, NGS=31.38%, p=0.013), *Corynebacteriaceae* (ONT=2.53%; NGS=19.96% NGS, p<0.001), *Propionibacteriaceae* (ONT=1.35%; NGS=9.35%, p<0.001) and *Peptoniphilaceae* (ONT=7.26%; NGS=2.38% NGS, p=0.013). Correlations remained strong at a genus order with no significant difference in MRA for 9 of 14 genera with MRA>1% (Figure 2, Table 2). Significant differences were observed for *Haemophilus* (ONT=10.21%; NGS=5.82%, p=0.048), as well as *Staphylococcus*, *Corynebacterium*, *Cutibacterium* and *Anaerococcus* (consistent with described familial discrepancies). 3.85% of NGS MRA was familial *Enterobacteriaceae* unable to be assigned to a genus order. This likely corresponded to *Enterobacteriaceae* genera observed with ONT, including *Klebsiella* (ONT=4.98%, NGS=0%, p=0.21) and *Enterobacter* (ONT=3.07%, NGS=0.27%, p=0.57).

Correlation between ONT and NGS declined significantly at a species order due to low resolution of NGS at this phylogeny, with only one species with MRA>1% (*Streptococcus dysgalactiae*) common to both datasets. 88.13% of NGS reads were unable to be assigned to a species. NGS assigned just 61 species and only 3 had >1% MRA (*Corynebacterium propinquum* 5.81%, *Streptococcus dysgalactiae* 3.78% and *Haemophilus aegyptius* 2.24%). The remaining 58/61 assigned species had a cumulative

Table 2. ONT and NGS relative abundance and prevalence comparisons.

	Mean Relative Abundance				Prevalence (cohort=47)			
Family	**ONT (%)**	**NGS (%)**	**Wilcoxon p-value**	**Corrected p-value**	**ONT (n)**	**NGS (n)**	**Fisher's Exact**	**Chi-Square**
Staphylococcaceae	40.81	31.38	0.001	0.01	41	41	>0.99	>0.99
Streptococcaceae	16.09	17.68	0.2	0.51	22	25	0.68	0.54
Pasteurellaceae	10.28	5.84	0.01	0.057	20	16	0.52	0.4
Corynebacteriaceae	2.53	19.96	<0.001	<0.001	21	36	0.003	0.002
Propionibacteriaceae	1.34	9.35	<0.001	<0.001	20	35	0.003	0.002
Enterobacteriaceae	8.94	4.19	0.07	0.32	13	16	0.66	0.50
Peptoniphilaceae	7.26	2.38	0.002	0.01	23	25	0.84	0.68
Neisseriaceae	2.25	2.39	0.9	0.94	13	16	0.66	0.5
Porphyromonadaceae	0.22	1.37	0.13	0.41	1	4	0.36	0.17
Xanthomonadaceae	0.96	1.66	0.44	0.68	2	5	0.43	0.24
Bacillaceae	1.07	0	0.05	0.21	6	2	0.27	0.14
Other	8.25	3.8						
Genus								
Staphylococcus	40.57	31.38	0.002	0.02	40	41	>0.99	0.77
Streptococcus	16.09	17.68	0.24	0.57	22	24	0.84	0.68
Haemophilus	10.21	5.82	0.005	0.048	19	16	0.67	0.52
Corynebacterium	2.53	19.04	<0.001	<0.001	21	36	0.003	0.002
Cutibacterium	1.34	9.33	<0.001	<0.001	20	35	0.003	0.002
Klebsiella	4.98	0	0.03	0.21	6	0	0.03	0.01
Enterobacter	3.07	0.27	0.25	0.57	3	1	0.62	0.31
Stenotrophomonas	0.96	1.66	0.75	0.75	2	3	>0.99	0.65
Parvimonas	2.4	1.22	0.16	0.57	5	4	>0.99	0.73
Anaerococcus	2.36	0.37	0.002	0.02	18	16	0.83	0.67
Peptoniphilus	1.57	0.49	0.03	0.21	14	14	>0.99	>0.99
Neisseria	0.52	1.2	0.16	0.57	6	9	0.57	0.4
Snodgrassella	1.71	0	0.02	0.13	7	0	0.01	0.01
Porphyromonas	0.22	1.37	0.13	0.55	1	4	0.36	0.17
Other	11.47	10.17						
Species								
Staphylococcus aureus	23.7	0	<0.001	<0.001	28	0	<0.001	<0.001
Staphylococcus epidermidis	12.65	0	<0.001	<0.001	29	0	<0.001	<0.001
Staphylococcus pseudintermedius	1.15	0	>0.99	>0.99	1	0	>0.99	0.31
Streptococcus pneumoniae	7.73	0	<0.001	0.007	12	0	<0.001	<0.001
Streptococcus dysgalactiae	3.85	3.78	0.75	0.94	3	2	>0.99	0.65
Streptococcus mitis	2.08	0	0.008	0.08	8	0	0.006	0.003
Streptococcus gwangjuense	1.81	0	0.25	0.68	3	0	0.24	0.08
Haemophilus influenzae	9.05	0	<0.001	0.007	12	0	<0.001	<0.001
Haemophilus aegyptius	0	2.24	0.13	0.49	0	4	0.12	0.04
Corynebacterium accolens	1.19	0	0.004	0.04	9	0	0.003	0.002
Corynebacterium pseudodiphth.	1.13	0	0.008	0.08	8	0	0.006	0.003
Corynebacterium propinquum	0.13	5.81	0.002	0.02	2	12	0.007	0.004
Cutibacterium acnes	1.26	0	<0.001	<0.001	19	0	<0.001	<0.001
Klebsiella aerogenes	4.96	0	0.06	0.36	5	0	0.06	0.02

	Mean Relative Abundance				Prevalence (cohort=47)			
	ONT (%)	NGS (%)	Wilcoxon p-value	Corrected p-value	ONT (n)	NGS (n)	Fisher's Exact	Chi-Square
Enterobacter hormaechei	2.97	0	0.25	0.68	3	0	0.24	0.08
Parvimonas micra	2.4	0	0.06	0.36	5	0	0.06	0.02
Anaerococcus nagyae	1.14	0.01	0.001	0.01	11	0	<0.001	<0.001
Peptoniphilus lacydonensis	1.56	0	<0.001	0.004	13	0	<0.001	<0.001
Snodgrassella alvi	1.71	0	0.02	0.12	7	0	0.01	0.01
Other species	19.53	0.03						
Not specified to species order	0	88.13						

relative abundance of 0.03%. Only one *Staphylococcus* species was identified by NGS: *Staphylococcus equorum*, an equine-host organism identified in a single sample (MRA=0.007%). In contrast, ONT demonstrated high resolution at a species order, identifying 264 different bacterial species including 16 different *Staphylococcus*, 16 *Streptococcus*, 12 *Neisseria* and 9 *Corynebacterium* species. 17 species had an abundance of >1% with *Staphylococcus aureus* (23.7%) and *Staphylococcus epidermidis* (12.65%) the most abundant (Table 2).

Aldex2 abundance analysis

Aldex2 was utilised as an alternate differential abundance analysis, applying Wilcoxon pairwise comparisons for each genus. 196 genera were compared, with only 4 demonstrating a statistically significant difference (Supplementary file 3). These included *Corynebacterium* (p<0.001), *Laycella* (p<0.001), *Cutibacterium* (p<0.001) and *Lawsonella* (p=0.04).

Prevalence

Prevalence was compared applying Fisher's Exact and Chi-Square tests (Table 2). 10 of the 14 genera with MRA>1% showed no significant difference in prevalence. *Staphylococcus* (ONT=85.1%; NGS=87.2%; p>0.99), *Streptococcus* (ONT=46.8%; NGS=51.1%; p=0.84) and *Haemophilus* (ONT=40.4%; NGS=34.0%; p=0.67) were prevalent and comparable. *Corynebacterium* (ONT=44.7%; NGS=76.6%, p=0.003) and *Cutibacterium* (ONT=42.6%, NGS=74.5% p=0.003) were significantly more prevalent with NGS, while *Klebsiella* (6/47 ONT=12.8%, 0/47 NGS=0%, p=0.03) and *Snodgrassella* (7/47 ONT=14.9%, NGS=0%, p=0.01) were significantly more prevalent with ONT. Species order prevalence comparisons were limited due to the low resolution of NGS at this phylogeny. *Staphylococcus epidermidis* (61.7%), *Staphylococcus aureus* (59.6%) and *Cutibacterium acnes* (40.4%) had the highest prevalence by ONT (not detected with NGS).

Microbiome agreement

The sum of the percentage of MRA in agreement for each sample and the number of genera in agreement per sample was calculated as described in Supplementary file 2. The mean percentage of relative abundance agreement per sample was 61.4% (SD=32.4; median=65.35%) (Figure 3). The mean number of genera in agreement 3.7 genera/sample (SD=2.2, median=4). 9 paired samples had MRA agreement >99%, while 6 samples had <20% MRA agreement.

Bacterial culture agreement

Sequencing results were compared with bacterial cultures obtained from the same middle meatal site at the time of surgery. A total of 38 species were cultivated from the patient cohort (Figure 4). ONT accurately identified the genus of the bacteria cultured in the corresponding patient in 34/38 (89.5%) samples, which was superior to NGS 30/38 (78.9%), without statistical significance (p=0.103). At a species order, ONT accurately identified the species cultured in 32/38 (84.2%) samples which was significantly more accurate than NGS, where just 2/38 (5.3%) species were correctly identified (p<0.001). ONT demonstrated a higher identification rate of cultured bacteria to a species depth (84.2%) relative to what NGS could identify at either a genus (78.9%) or species (5.3%) depth.

Alpha and Beta diversity

Alpha diversity and richness were calculated at a genus depth using Shannon's diversity index, Simpson's index and Chao1 (Figure 5). There was no significant difference in Shannon's diversity between ONT (mean=0.85) and NGS (mean=0.73, p=0.29). There was no significant difference in richness between ONT (mean=8.2) and NGS (mean=7.9, p=0.80). Suggesting ONT produces overall highly comparable conclusions to NGS using multiple commonly applied microbiome methods to assess alpha diversity. Bray-Curtis dissimilarity calculations were performed to establish beta diversity, calculating centroid ecological distance for each sample (Figure 6). On Wilcoxon testing, the mean centroid for ONT (0.55) and NGS (0.56) were highly comparable (p=0.66) suggesting inter-sample ecological dissimilarity in the cohort produced comparable conclusions in both sequencing modalities.

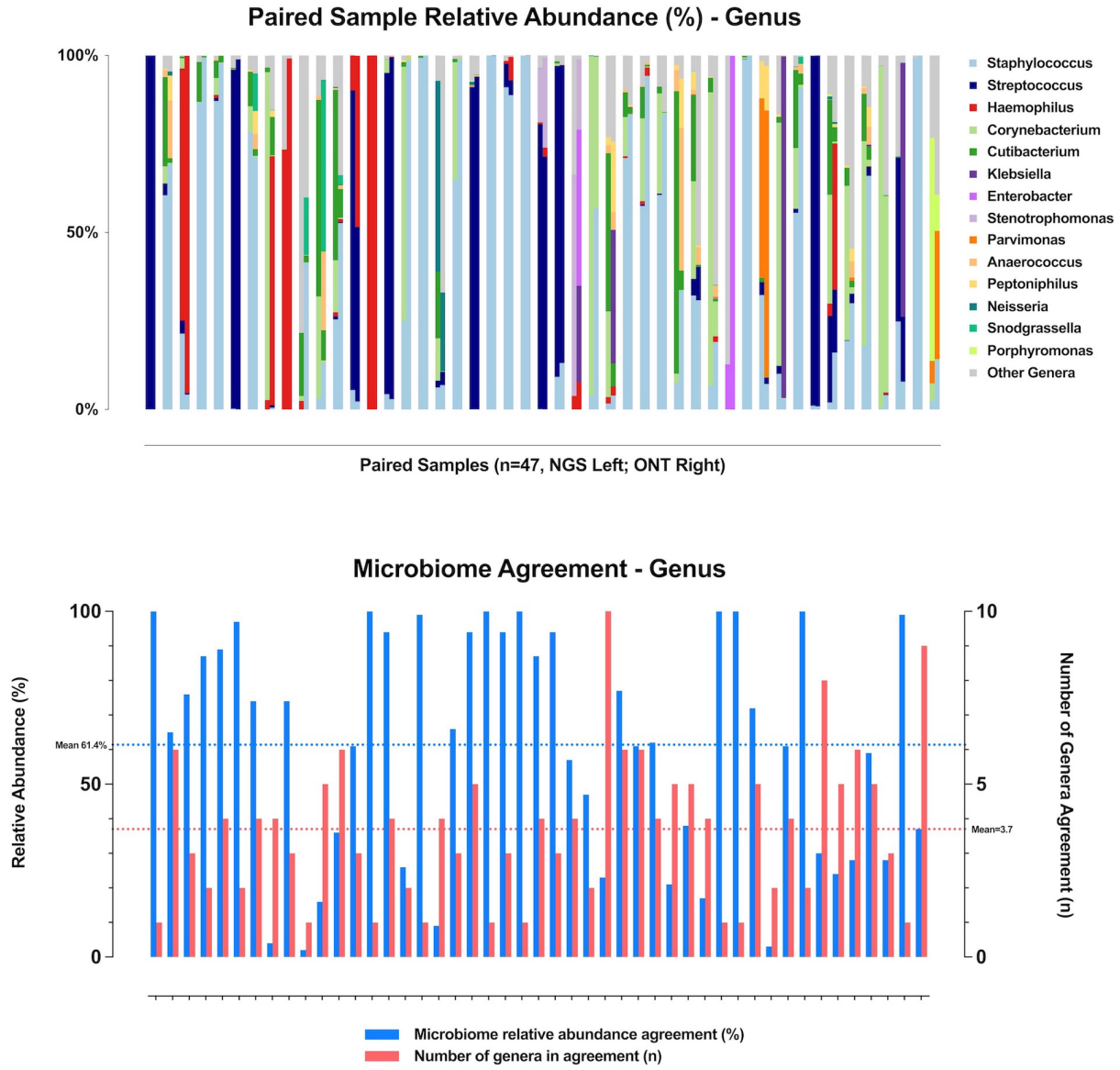


Table 2. ONT and NGS relative abundance and prevalence comparisons.

EMU versus SILVA database comparison

To confirm that the observed results related to differences between ONT and NGS and not differences in databases, we applied the SILVA database to the ONT dataset. The results were almost identical between ONT SILVA and ONT EMU dataset (Supplementary file 4). The mean difference in MRA for the 10 most abundant genera was 0.55% (SD=0.72%; range 0.01%-1.97%). The mean difference between the top 10 most abundant species was 0.85% (SD=0.66%; range 0.1%-1.68%). These results indicate that database had a negligible influence on the presented results.

Discussion

The paranasal sinuses possess a unique apposition of being both a low biomass environment and one where bacteria is heavily implicated in disease propagation. Accordingly, there is a substantial appetite for effective non-culture dependent techniques to enhance our understanding of the sinus microbiome. To our knowledge, this is the first time that ONT 16S rRNA sequencing has been applied in the paranasal sinuses, which has unearthed novel insights into the bacterial composition of CRS. By applying direct comparisons to NGS technology, we have elucidated strengths and weaknesses of each platform specific to sinonasal flora.

In this series comparable conclusions were established at genus order for abundance, alpha diversity and beta diversity. Correlations declined significantly at a species order. This mirrored the conclusions of other studies of a similar methodology, where ONT has consistently demonstrated superior phyloge-

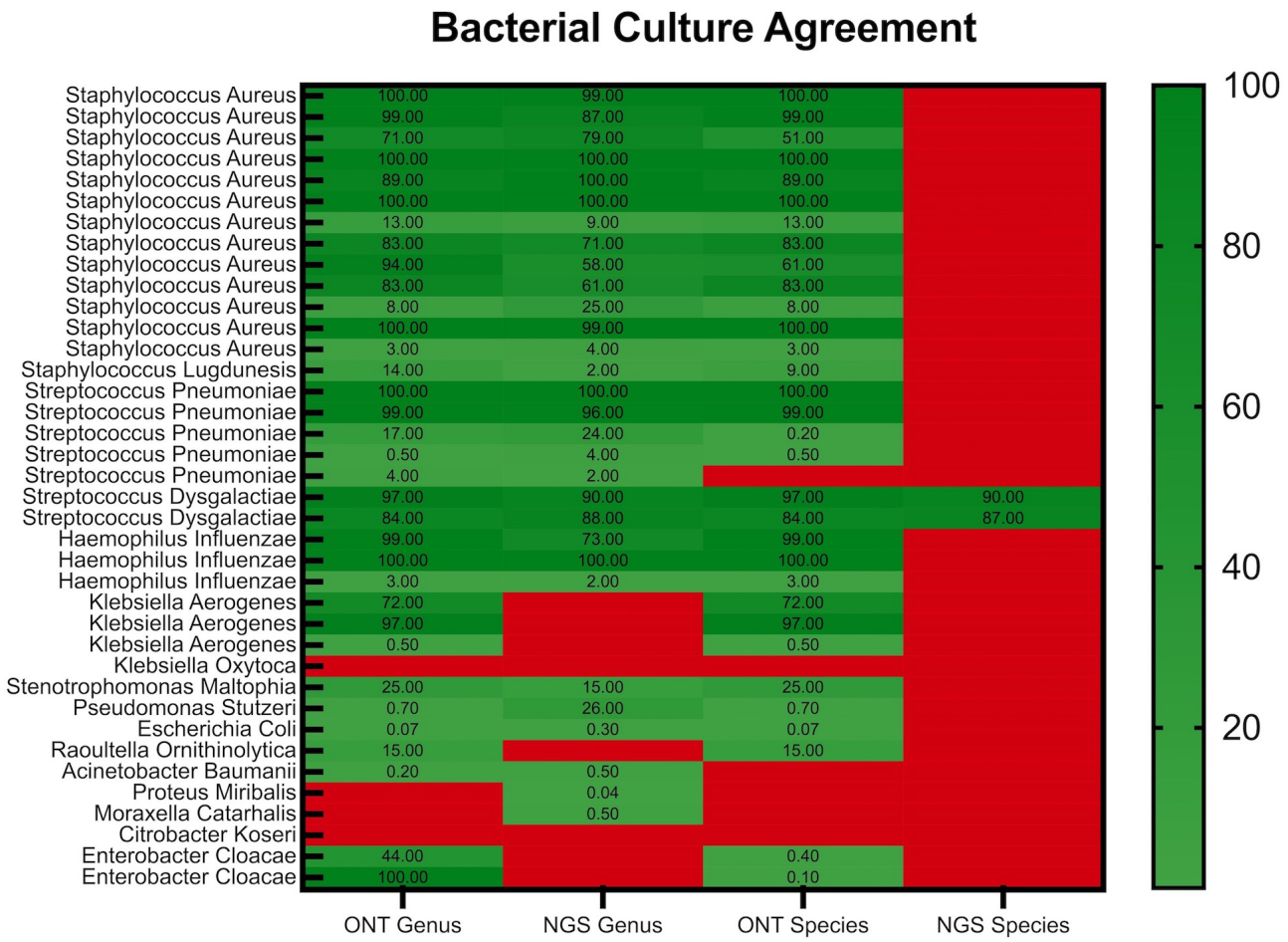


Figure 4. Bacterial culture agreement. 29/47 middle meatal swabs (from the same site that sequencing swabs were obtained) cultured 38 bacterial species, with *Staphylococcus aureus* (13) and *Streptococcus pneumoniae* (5) the most prevalent. Heatmaps assess the reliability of ONT and NGS to identify the cultured organism in the corresponding patient swab.

netic resolution[9,10,13–16]. Although select sinonasal NGS studies have reported taxa to a species order, this should be interpreted with trepidation, particularly in the absence of adjunct methods like amplicon sequence variance clustering or quantitative PCR. Yarza concluded that a minimum of 1300 nucleotides are necessary to accurately assign sequences to a species resolution[30], which significantly exceeds the limitations of NGS. Earl reported highly comparable genetic structures between *Staphylococcus aureus* and *Staphylococcus epidermidis* with 1.4% divergence and differing by as few as 23 nucleotides[31]. In our series, ONT could discriminate virulent *Staphylococcus aureus* and passive *Staphylococcus epidermidis*, two species with entirely divergent pathogenic potential. When contextualized with NGS studies that aggregate these species as a common genus, the potential to establishing clinically important inferences is limited. In this manner, ONT offers significant advantages in better understanding the microbiology of CRS.

It was critical to assess if the species conclusion established by ONT reflect the true bacteria in each sample. To assess this, we compared sequenced data to bacterial culture results obtained from the same site at the time of surgery. Correlations with the culture results were comparable at a genus order, with ONT reliably identifying 89.5% of cultured genera compared with 78.9% for NGS (p=0.103). NGS was most limited in identifying *Enterobacteriaceae* genera, (*Klebsiella*, *Enterobacter*, *Citrobacter*, *Escherichia* and *Raoultella*) which is likely due to NGS limitation of primer bias against this specific lineage. Critically, ONT correlated strongly at a species order reliably identifying 84.2% of cultured species in the patient matched sequences. This was significantly superior to NGS at a species order (5.3%; p<0.01) and notably exceeded the correlations of NGS at a genus order (78.9%; p=0.55). ONT accurately identified commonly cultured sinus species (*S. aureus*, *S. lugdunesis*, *S. dysgalactiae*, *S. pneumoniae*), difficult to culture species (*H. influenzae*) and species considered less endemic to the sinuses (*Stenotrophomonas maltophia, Pseudomonas stutzeri, Escherichia coli, Raoultella ornithinolytica, Enterobacter cloacae*). These results endorse the accuracy of ONT against a broad spectrum of common and atypical sinonasal organisms.

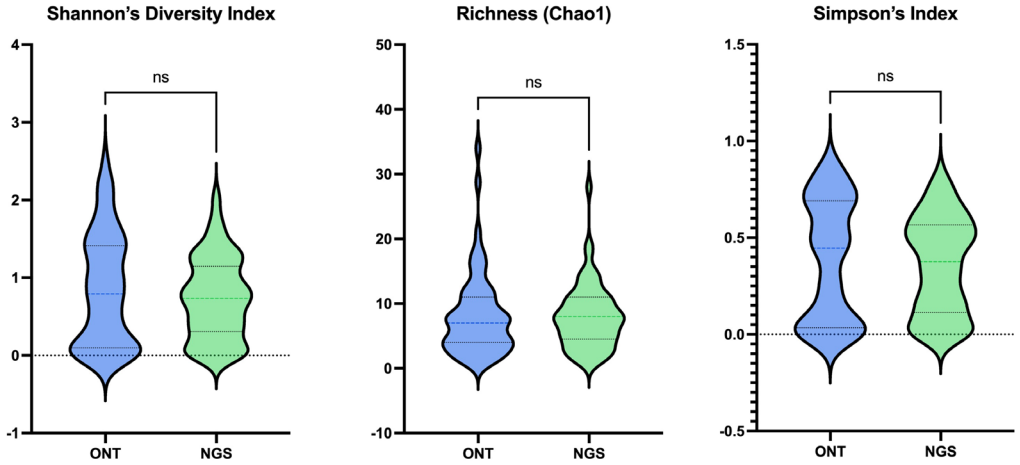


Figure 5. Alpha diversity calculations comparing ONT and NGS diversity at a genus level. No significant difference was observed on Wilcoxon paired analysis for Shannon's diversity index (mean ONT=0.85; NGS=0.74;p=0.29); Richness (Mean ONT=8.2;NGS=7.9;p=0.53) or Simpson's Index (mean ONT=0.39;NGS=0.36;p=0.52).

The most significant limitation of ONT highlighted by this series was the discrepancies of MRA for two core sinonasal genera, *Corynebacterium* and *Cutibacterium*. Both are considered sinus commensals with theoretical probiotic, antibiosis and in some cases pathogenic potential[32–35]. *Corynebacterium* is highly abundant in numerous NGS studies and was the most abundant taxa in the 'International Sinus Microbiome' study[25], yet it was significantly underrepresented in our ONT MRA dataset (ONT=2.53%, NGS=19.04%, p<0. 001). Despite its low abundance, it was the third most prevalent genus in the ONT dataset (44.7%), Similarly, *Cutibacterium* was the fourth most prevalent genus in the ONT sequences (42.6%) despite it also having underrepresented MRA (1.33% ONT, 9.33% NGS, p<0.001). The under estimation of these genera is not unique to our series. Heikema compared NGS and ONT in mock communities, identifying significant under estimation of *Corynebacterium* relative to the predicted mock community abundance[15]. Heikema deduced that the default ONT 16S rRNA primer (used in the current study) had a low affinity for the primer binding sites in multiple *Corynebacterium* species, resulting in impaired binding and low amplification. Rozas compared NGS and ONT on skin and mock cultures, identifying a similar PCR-bias with under representation of *Corynebacterium*, *Cutibacterium* and *Micrococcus*[16]. Rozas hypothesised that species with a high genetic GC-content were susceptible to under estimations of abundance which was partially validated when the PCR protocol was amended to mitigate this (changing the default polymerase from LongAmp Taq to KAPA). Like Heikema, Rozas also concluded that a lack of affinity to the ONT 16S reverse primer (1492R) was contributing to poor amplification, which was tested with the application of an alternate primer with a downstream primer target region (on the 23S gene), resulting in closer abundance approximations for the relevant species.

The current study applies the recommended ONT primers/PCR protocols, so was susceptible to the abundance discrepancies reported in other studies. Encouragingly, the described limitations appear to be amenable to protocol optimisation.

Cost and accessibility are critical additional considerations when comparing sequencing platforms. ONT has portable and benchtop products that can be readily integrated into independent laboratories. This promotes a streamlined workflow, optimised protocols, access to high-volume real-time sequencing and relinquishes the financial and time obligations of outsourcing to third parties[13]. ONT upfront costs are comparatively low, with the portable MinION and MinION Mk1C sequencers retailing at $1000-5000, and the benchtop GridION (used in this study) retailing at $50,000[36]. Consumables constitute an ongoing expense including flow cells and library preparation kits. While Illumina does not currently have a portable device, it also offers the benefit of a benchtop model. The upfront cost for Illumina Miseq is approximately $100,000[37] which may be a deterrent for independent laboratories, resulting in outsourcing of sequencing to third parties. Ongoing costs comparisons are variable , depending on the sequencing being performed, with some estimating higher per sample cost for ONT and others reporting them to be increasingly comparable[38–40].
Limitations of this study include the absence of a mock community to accurately interpret discrepancies in mean relative abundance. A commercially available mock community specific to sinonasal taxa does not exist, however our laboratory is working on a sinus-specific mock community to further optimise sequencing protocols in the future. Consistency at all stages of the wet and dry lab is critical, where feasible, for accurate methodological comparison. We have maximally strived to achieve this but

acknowledge unavoidable divergences. Sample collection and extraction were identical with DNA extracted from a common swab. Previous series have indicated that DNA from a common swab provide reproducible conclusions, even if primer targets or PCR protocols are amended[4,41], offering some assurance for consistency in our methodology. Different amplification protocols optimised for ONT and NGS were used to maximize read yield. There was necessary divergence in bioinformatics, as ONT and NGS output data are compositionally different and require different clustering/denoising algorithms and different databases. EMU (ONT) uses a custom EMU database while QIIME2 (ONT) is compatible with SILVA or other NCBI derived databases. We used compatible databases for each pipeline, but acknowledge this has potential for database bias[42]. To address this, we repeated ONT taxonomic assignment with SILVA and compared it to the original EMU results. Results were almost identical, affirming that database bias had a negligible impact. Interestingly, ONT SILVA had limitations assigning a small number of reads to a species, suggesting it may not be optimised for ONT. Therefore, EMU should be retained as the recommended ONT database. This study set out to accurately define the sinonasal microbiome to a species order. In this respect, ONT was superior to NGS and is endorsed as the preferred modality for future microbiome studies, which should include a concurrent focus on protocol optimisation. As we enter an era of full-length sequencing and metagenomics, embarking on new frontiers of microenvironmental discernment, we must simultaneously embrace the potential to reach new clinical and biological pinnacles whilst remaining humble in our understanding of the limitations and imperfections of this technology.

Conclusion

This is the first study to apply Oxford Nanopore full length 16S rRNA gene amplicon sequencing in the paranasal sinuses, which identified comparable biological conclusions to NGS at a genus order and significantly superior accuracy and resolution at a species order. Despite observed limitations in abundance estimations of sinonasal commensals, the species level characterization of sinus microbiome has provided novel insights into the microenvironment, not previously attained. Future application of full-length sequencing can advance our understanding of the role bacteria play in sinus health and disease.

Acknowledgements

The University of Adelaide 'Research Training Program Stipend' provided financial support to the primary investigator during the period this research was undertaken.

Authorship contribution

Study conception and design: AJP, SV, PJW, JC. Data collection: JC, KY, AJP, PW. Analysis and interpretation of results, draft manuscript preparation: All authors. All authors reviewed the results and approved the final version of the manuscript.

Conflict of interest

AJP is a consultant for Medtronic and Neurent and receives speaking honorariums from Storz, GSK, Sanofi and Sequiris. He is a shareholder for Chitogel. PJW is a consultant for Neilmed, Stryker, Neurent, receives royalties from Integra, Fusetec and is a shareholder for Chitogel.

Funding

No specific funding disclosures.

References

1. van Dijk EL, Jaszczyszyn Y, Naquin D, Thermes C. The third revolution in sequencing technology. Trends Genet. 2018 Sep;34(9):666–81.
2. Jervis Bardy J, Psaltis AJ. Next generation sequencing and the microbiome of chronic rhinosinusitis: a primer for clinicians and review of current research, Its Limitations, and Future Directions. Ann Otol Rhinol Laryngol. 2016 Aug;125(8):613–21.
3. Sadowsky M, Staley C, Heiner C, et al. Analysis of gut microbiota—An ever changing landscape. Gut Microbes. 2017;8:268–75.
4. Connell J, Bassiouni A, Wormald PJ, Vreugde S, Psaltis A. Optimal primer selection for sinus microbiome profiling: A comparative analysis of the V1-V3 and V3-4 16S target regions. Int Forum Allergy Rhinol. 2021 Dec;11(12):1698–702.
5. Bassiouni A, Paramasivan S, Shiffer A, et al. Microbiotyping the Sinonasal Microbiome. Front Cell Infect Microbiol. 2020 [cited 2020 Nov 8];10.
6. Dhawan D. Chapter 2 - Clinical next-generation sequencing: enabling precision medicine. In: Verma M, Barh D, editors. Progress and Challenges in Precision Medicine. Academic Press; 2017. p. 35–54.
7. Johnson JS, Spakowicz DJ, Hong BY, et al. Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. Nat Commun. 2019 Nov 6;10(1):5029.
8. Benítez-Páez A, Sanz Y. Multi-locus and long amplicon sequencing approach to study microbial diversity at species level using the MinIONTM portable nanopore sequencer. Gigascience. 2017 Jun 12;6(7):1–12.
9. Nygaard AB, Tunsjø HS, Meisal R, Charnock C. A preliminary study on the potential of Nanopore MinION and Illumina MiSeq 16S rRNA gene sequencing to characterize building-dust microbiomes. Sci Rep. 2020 Feb 21;10(1):3209.
10. Shin J, Lee S, Go MJ, et al. Analysis of the mouse gut microbiome using full-length 16S rRNA amplicon sequencing. Sci Rep. 2016 Jul 14;6(1):29681.
11. Ni Y, Liu X, Simeneh ZM, Yang M, Li R. Benchmarking of Nanopore R10.4 and R9.4.1 flow cells in single-cell whole-genome amplification and whole-genome shotgun sequencing. Comput Struct Biotechnol J. 2023 Mar 24;21:2352–64.
12. Lu H, Giordano F, Ning Z. Oxford Nanopore MinION sequencing and genome assembly. Genomics Proteomics Bioinformatics. 2016 Oct;14(5):265–79.
13. Szoboszlay M, Schramm L, Pinzauti D, Scerri J, Sandionigi A, Biazzo M. Nanopore is preferable over Illumina for 16S amplicon sequencing of the gut microbiota when species-level taxonomic classification, accurate estimation of richness, or focus on rare taxa is required. Microorganisms. 2023 Mar;11(3):804.
14. Matsuo Y, Komiya S, Yasumizu Y, et al. Full-

length 16S rRNA gene amplicon analysis of human gut microbiota using MinION™ nanopore sequencing confers species-level resolution. BMC Microbiol. 2021 Jan 26;21:35.
15. Heikema AP, Horst-Kreft D, Boers SA, et al. Comparison of Illumina versus Nanopore 16S rRNA gene sequencing of the human nasal microbiota. Genes. 2020 Sep;11(9):1105.
16. Rozas M, Brillet F, Callewaert C, Paetzold B. MinION™ nanopore sequencing of skin microbiome 16S and 16S-23S rRNA gene amplicons. Front Cell Infect Microbiol. 2021;11:806476.
17. Klair D, Dobhal S, Ahmad A, et al. Exploring taxonomic and functional microbiome of Hawaiian stream and spring irrigation water systems using Illumina and Oxford Nanopore sequencing platforms. Front Microbiol. 2023
18. Lee JT, Frank DN, Ramakrishnan V. Microbiome of the paranasal sinuses: Update and literature review. Am J Rhinol Allergy. 2016;30(1):3–16.
19. Lux CA, Wagner Mackenzie B, Johnston J, et al. Antibiotic treatment for chronic rhinosinusitis: prescription patterns and associations with patient outcome and the sinus microbiota. Front Microbiol. 2020;11:595555.
20. Siu J, Mackenzie BW, Klingler L, et al. Sinonasal and gastrointestinal bacterial composition and abundance are stable after 1 week of once-daily oral antibiotic treatment for chronic rhinosinusitis. Int Forum Allergy Rhinol. 2021 Sep;11(9):1355–66.
21. Park SC, Park IH, Lee JS, et al. Microbiome of unilateral chronic rhinosinusitis: a controlled paired analysis. Int J Environ Res Public Health. 2021 Sep 19;18(18):9878.
22. Gan W, Zhang H, Yang F, Liu S, Liu F, Meng J. The influence of nasal bacterial microbiome diversity on the pathogenesis and prognosis of chronic rhinosinusitis patients with polyps. Eur Arch Otorhinolaryngol. 2021 Apr;278(4):1075–88.
23. Zhu Z, Lan J, Wei R, et al. Microbiome and Th cytokines association in chronic rhinosinusitis with or without nasal polyp. Laryngoscope Investig Otolaryngol. 2023 Apr;8(2):335–45.
24. Kim JH, Kim SH, Lim JY, et al. Association between the sinus microbiota with eosinophilic inflammation and prognosis in chronic rhinosinusitis with nasal polyps. Exp Mol Med. 2020 Jun 29;52(6):978–87.
25. Paramasivan S, Bassiouni A, Shiffer A, et al. The international sinonasal microbiome study: a multicentre, multinational characterization of sinonasal bacterial ecology. Allergy. 2020;75(8):2037–49.
26. Fokkens WJ, Lund VJ, Hopkins C, et al. European Position Paper on Rhinosinusitis and Nasal Polyps 2020. Rhinology. 2020 Feb 20;58(Suppl S29):1–464.
27. Bouras, G. Coatofarms. GitHub Inc.; Available from: https://github.com/gbouras13/coatofarms
28. Curry KD, Wang Q, Nute MG, et al. Emu: species-level microbial community profiling of full-length 16S rRNA Oxford Nanopore sequencing data. Nat Methods. 2022 Jul;19(7):845–53.
29. Fernandes AD, Reid JN, Macklaim JM, McMurrough TA, Edgell DR, Gloor GB. Unifying the analysis of high-throughput sequencing datasets: characterizing RNA-seq, 16S rRNA gene sequencing and selective growth experiments by compositional data analysis. Microbiome. 2014 May 5;2:15.
30. Yarza P, Yilmaz P, Pruesse E, et al. Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. Nat Rev Microbiol. 2014 Sep;12(9):635–45.
31. Earl JP, Adappa ND, Krol J, et al. Species-level bacterial community profiling of the healthy sinonasal microbiome using Pacific Biosciences sequencing of full-length 16S rRNA genes. Microbiome. 2018 Oct 23;6.
32. Copeland E, Leonard K, Carney R, et al. Chronic rhinosinusitis: potential role of microbial dysbiosis and recommendations for sampling sites. Front Cell Infect Microbiol. 2018 Feb 28;8.
33. Boase S, Foreman A, Cleland E, et al. The microbiome of chronic rhinosinusitis: culture, molecular diagnostics and biofilm detection. BMC Infect Dis. 2013 Dec;13(1):210.
34. Huang S, Hon K, Bennett C, et al. Corynebacterium accolens inhibits Staphylococcus aureus induced mucosal barrier disruption. Front Microbiol. 2022;13.
35. Abreu NA, Nagalingam NA, Song Y, et al. Sinus microbiome diversity depletion and Corynebacterium tuberculostearicum enrichment mediates rhinosinusitis. Sci Translat Med. 2012 Sep 12;4(151):151ra124-151ra124.
36. Nanopore store: devices. [cited 2023 Nov 20]. Available from: https://store.nanoporetech.com/
37. Sequencing Platforms, Illumina NGS platforms. [cited 2023 Nov 20]. Available from: https://sapac.illumina.com/systems/sequencing-platforms.html
38. Cook J. Bridge Informatics. 2022 [cited 2023 Nov 20]. Battle of the Sequencers: Illumina vs Nanopore Sequencing. Available from: https://bridgeinformatics.com/battle-of-the-sequencers-illumina-vs-nanopore-sequencing/
39. Stevens BM, Creed TB, Reardon CL, Manter DK. Comparison of Oxford Nanopore Technologies and Illumina MiSeq sequencing with mock communities and agricultural soil. Sci Rep. 2023 Jun 8;13:9323.
40. Lao HY, Ng TTL, Wong RYL, et al. The clinical utility of two high-throughput 16S rRNA gene sequencing workflows for taxonomic assignment of unidentifiable bacterial pathogens in matrix-assisted laser desorption ionization-time of flight mass spectrometry. J Clin Microbiol. 2022 Jan 19;60(1):e0176921.
41. Raju SC, Lagström S, Ellonen P, et al. Reproducibility and repeatability of six high-throughput 16S rDNA sequencing protocols for microbiota profiling. J Microbiol Meth. 2018 Apr 1;147:76–86.
42. Nearing JT, Comeau AM, Langille MGI. Identifying biases and their potential solutions in human microbiome studies. Microbiome. 2021 May 18;9(1):113.

J.T. Connell
Department of Surgery
Otolaryngology Head and Neck Surgery
University of Adelaide
and the Basil Hetzel Institute for Translational Health Research
Central Adelaide Local Health Network
Adelaide
South Australia
Australia

Tel: +61 421 018 611
E-mail: james.connell@sa.gov.au

This manuscript contains online supplementary material

SUPPLEMENTARY MATERIAL

Supplementary File 1. Australian Genome Research Facility (AGRF) PCR protocol applied to the NGS dataset in our cohort. Document provided by AGRF.

Materials and methods

PCR amplification and sequencing was performed by the Australian Genome Research Facility. PCR amplicons were generated using the primers and conditions outlined in the attached table. Thermocycling was completed with an Applied Biosystem 384 Veriti and using Platinum SuperFi II mastermix (Life Technologies, Australia) for the primary PCR. The first stage PCR was cleaned using magnetic beads, and samples were visualised on 2% Sybr Egel (Thermo-Fisher). A secondary PCR to index the amplicons was performed with Platinum SuperFi II mastermix (Life Technologies, Australia). The resulting amplicons were cleaned again using magnetic beads, quantified by fluorometry (Promega Quantifluor) and normalised. The eqimolar pool was cleaned a final time using magnetic beads to concentrate the pool and then measured using a High-Sensitivity D1000 Tape on an Agilent 2200 TapeStation. The pool was diluted to 5nM and molarity was confirmed again using a Qubit High Sensitivity dsDNA assay (ThermoFisher). This was followed by sequencing on an Illumina MiSeq (San Diego, CA, USA) with a V3, 600 cycle kit (2 x 300 base pairs paired-end).

Target	Cycle	Initial	Disassociate	Anneal	Extension	Finish
16S: V3 - V4	30	98°C for 30 sec	98°C for 10 sec	60°C for 10 sec	72°C for 30 sec	72°C for 5 min

16S: V3 - V4 (341F-806R)	
Forward Primer (341F)	CCTAYGGGRBGCASCAG
Reverse Primer (806R)	GGACTACNNGGGTATCTAAT

Supplementary File 2. Microbiome agreement calculation.

Heikema et al. [1] utilised this method to calculate the cumulate relative abundance (RA) agreement between NGS and ONT paired samples and the number of genera in agreement per sample. In this example (sample 7), *Streptococcus* was identified in 95.82% RA in NGS and 98.93%RA in ONT, with agreement of 95.82% for *Streptococcus*. *Moraxella* was identified in 3.48% RA in NGS and 1.07% RA in ONT with agreement of 1.07% for *Moraxella*. No other genera were identified in both samples. So the final agreement was 95.82% + 1.07%= 96.89% relative abundance agreement and a total of 2 genera in agreement.

Sample 7	Relative Abundance (%)		Agreement	
GENUS	NGS	ONT	Relative Abundance (%)	Genera (n)
16S: V3 - V4	30	98°C for 30 sec	98°C for 10 sec	60°C for 10 sec

References

1. Heikema AP, Horst-Kreft D, Boers SA, Jansen R, Hiltemann SD, de Koning W, et al. Comparison of Illumina versus Nanopore 16S rRNA Gene Sequencing of the Human Nasal Microbiota. Genes. 2020 Sep;11(9):1105.

All samples agreement table.

Sample	Agreement	
	Relative Abundance (%)	Genera (n)
1	100	1
2	65.35	6
3	75.93	3
4	87.26	2
6	89.19	4
7	96.89	2
8	73.99	4
9	4.45	4
10	74.27	3
12	1.9	1
14	16.33	5
15	35.55	6
16	61.43	3
17	99.98	1
18	93.81	4
19	26.29	2
20	99.33	1
21	8.53	4
22	65.83	3
23	94.33	5
24	99.9	1
25	93.83	3
26	99.89	1
27	87.33	4
28	93.58	3

Sample	Agreement	
	Relative Abundance (%)	Genera (n)
29	57.25	4
30	46.71	2
32	23.24	10
33	77.21	6
34	60.64	6
35	61.58	4
37	21.15	5
38	37.87	5
40	17.49	4
41	99.81	1
42	99.94	1
43	71.84	5
44	3.26	2
45	60.59	4
47	99.75	2
48	30.35	8
54	23.91	5
55	28.24	6
56	58.58	5
58	27.69	3
59	99.35	1
61	36.9	9
Mean	**61.46**	**3.70**
SD	**32.86**	**2.15**
Median	**65.35**	**4**

Supplementary File 3.

Genus	significance (p_)
g__Corynebacterium	1.03E-06
g__Laceyella	1.57805612412796e-06
g__Cutibacterium	0.000505205
g__Lawsonella	0.043206314
g__Klebsiella	0.174471631
g__Streptococcus	0.179348012
g__Bacillus	0.263256025
g__Snodgrassella	0.275703742
g__Staphylococcus	0.280137092
g__Neobacillus	0.293867664

Genus	significance (p_)
g__Escherichia-Shigella	0.355934586
g__Crinalium	0.358337754
g__Mitochondria	0.36877446
g__Escherichia	0.412804926
g__Ralstonia	0.432531345
g__Delftia	0.450585337
g__Anabaena	0.452618158
g__Aliterella	0.48875586
g__Ochrobactrum	0.491712684
g__Cyanothece	0.495889356
g__Gloeocapsa	0.497103958

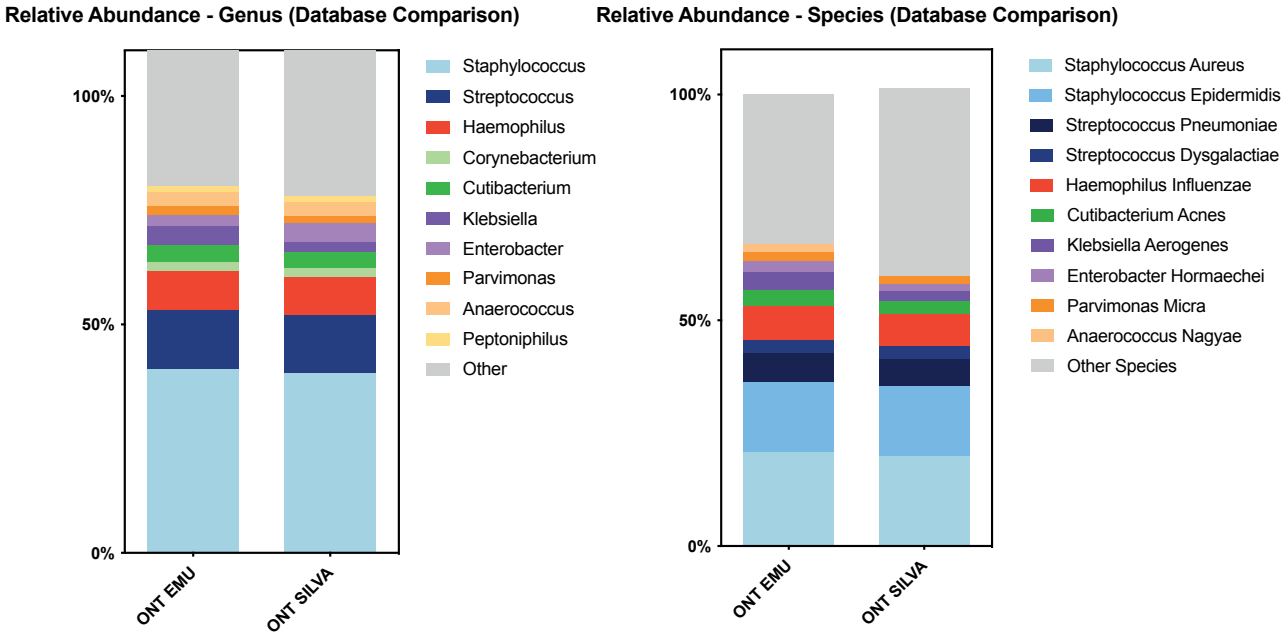
Genus	significance (p_)
g__Enterobacter	0.507617622
g__Prevotella	0.511556682
g__Pseudomonas	0.519776586
g__Veillonella	0.519861341
g__Romboutsia	0.523636563
g__Actinomyces	0.530267965
g__Bergeriella	0.531832504
g__Paracoccus	0.534042442
g__Oscillatoria	0.536983946
g__Rothia	0.537453523
g__Haemophilus	0.544526247
g__Sphingomonas	0.553864444
g__Sphingopyxis	0.554171208
g__Chryseobacterium	0.565623497
g__Calothrix	0.568926052
g__Citrobacter	0.57222826
g__Raoultella	0.573119921
g__Neisseria	0.577745528
g__Aggregatibacter	0.585500758
g__Acidihalobacter	0.586768343
g__Cylindrospermum	0.587350282
g__Prolinoborus	0.588182441
g__Porphyromonas	0.590909741
g__Pseudanabaena	0.594176747
g__Prosthecomicrobium	0.594350379
g__Methylomagnum	0.596379052
g__Baekduia	0.598417875
g__Moorea	0.602560079
g__Rubellimicrobium	0.604775457
g__Parvimonas	0.605683886
g__Ectothiorhodospira	0.605739947
g__Microcystis	0.60592725
g__Finegoldia	0.607329432
g__Turicella	0.607834368
g__Tychonema	0.610656622
g__Geobacillus	0.611129691
g__Anaerococcus	0.613417255
g__Lactobacillus	0.614169348
g__Kingella	0.617720741
g__Burkholderia-Caballeronia-Paraburkholderia	0.618423435
g__Ancalomicrobium	0.622089362
g__Dietzia	0.62263959
g__Methylococcus	0.624580333
g__Actibacterium	0.625235409
g__Geitlerinema	0.625698819
g__Rouxiella	0.626007265

Genus	significance (p_)
g__Metallibacterium	0.627295085
g__Brachybacterium	0.628348498
g__Gemella	0.628351867
g__Anaerobium	0.629516228
g__Litoreibacter	0.631044092
g__Loriellopsis	0.631672188
g__Thermicanus	0.632189385
g__Methylocaldum	0.633762839
g__Allochromatium	0.634952576
g__Steroidobacter	0.635243457
g__Comamonas	0.636556117
g__Friedmanniella	0.638104241
g__Rahnella	0.641259043
g__Alloprevotella	0.641948577
g__Phycisphaera	0.642375072
g__Salmonella	0.642521745
g__Zoogloea	0.642787128
g__Sporobacterium	0.644128447
g__Lactococcus	0.647822221
g__Herbiconiux	0.651886865
g__Actinobacillus	0.654876425
g__Acinetobacter	0.654879468
g__Trichodesmium	0.655016365
g__Algisphaera	0.655778815
g__Synechococcus	0.656475004
g__Capnocytophaga	0.658071014
g__Roseovarius	0.659010427
g__Carnobacterium	0.659759472
g__Algiphilus	0.662331206
g__Flavobacterium	0.663649759
g__Variovorax	0.663752215
g__Cnuella	0.66391496
g__Mizugakiibacter	0.666811189
g__Nakamurella	0.668825874
g__Peptostreptococcus	0.671543088
g__Negativicoccus	0.671701864
g__Acidicaldus	0.671795657
g__Brevundimonas	0.671873102
g__Aminipila	0.67265089
g__Inmirania	0.673218884
g__Methylobacterium	0.67347519
g__Nitrosococcus	0.674110389
g__Yangia	0.674404767
g__Saccharimonadaceae	0.675539927
g__Kocuria	0.675849598
g__Caedimonas	0.676523109
g__Hydrogenophaga	0.676571338

Genus	significance (p_)
g__Nitrosospira	0.677085058
g__Iphinoe	0.67743896
g__Thermus	0.678346618
g__Hafnia	0.680759403
g__Paenibacillus	0.680841011
g__Chroococcidiopsis	0.683381247
g__Abiotrophia	0.683491274
g__Liberibacter	0.68459596
g__Plasticicumulans	0.684712401
g__Leptotrichia	0.685641621
g__Granulicatella	0.686660342
g__Pseudonocardia	0.689586889
g__Lelliottia	0.689982784
g__Planifilum	0.691704481
g__Roseomonas	0.692183055
g__Peptoniphilus	0.694726346
g__Rhodococcus	0.695091889
g__Leclercia	0.695215016
g__Halothiobacillus	0.696253018
g__Blastococcus	0.699096897
g__Eikenella	0.700678119
g__Rathayibacter	0.701650756
g__Aureimonas	0.70232019
g__Rikenellaceae_RC9_gut_group	0.703549376
g__Stenotrophomonas	0.704460385
g__Thioalkalivibrio	0.704638575
g__Morganella	0.706651744
g__Mycoplasma	0.706680415
g__Reyranella	0.70937626
g__Lentimicrobium	0.709751063
g__Dialister	0.711919409
g__Facklamia	0.713662509
g__Catonella	0.713996587
g__Campylobacter	0.714434706
g__Craurococcus-Caldovatus	0.715618446
g__Longilinea	0.716472138
g__Bradyrhizobium	0.71672831
g__Clostridia	0.71766549
g__Brochothrix	0.718174767
g__Arthrobacter	0.718632703
g__Serratia	0.719038696

Genus	significance (p_)
g__Microlunatus	0.719727932
g__Anoxybacillus	0.720296124
g__Halomonas	0.720394138
g__Gloeothece	0.720513781
g__Bifidobacterium	0.725640562
g__Pelomonas	0.725642398
g__Methylocella	0.730478193
g__Treponema	0.732576869
g__Empedobacter	0.734072829
g__Corynebacteriaceae	0.734427125
g__Methylobacterium-Methylorubrum	0.737515941
g__Ahniella	0.737568165
g__Conexibacter	0.738460865
g__Proteus	0.738733637
g__Howardella	0.738854769
g__Tannerella	0.739919427
g__Candidatus	0.740859172
g__Dolosigranulum	0.740992597
g__Phreatobacter	0.742203572
g__Brevibacterium	0.743097574
g__Selenomonas	0.744125318
g__Lautropia	0.744205659
g__Micrococcus	0.744276267
g__Salinicoccus	0.744328224
g__Sporosarcina	0.745352143
g__Peptococcus	0.7457046
g__Vulcaniibacterium	0.74618523
g__Skermanella	0.746201266
g__Nocardioides	0.746863125
g__Yaniella	0.747222752
g__Achromobacter	0.747380391
g__Alloiococcus	0.750005059
g__Amphibacillus	0.75069953
g__Lentimonas	0.750914816
g__Nocardiopsis	0.751192207
g__Filifactor	0.752887712
g__Georgfuchsia	0.758495644
g__Bulleidia	0.759573958
g__Enhydrobacter	0.759803094
g__Fusobacterium	0.769515369
g__Moraxella	0.799196985

Supplementary file 4. ONT data comparisons utilising the default recommended EMU database compared with SILVA on the same raw dataset. Relative abundance conclusions are highly comparable between the two database datasets.



Genus	ONT EMU (%)	ONT SILVA (%)	Difference (%)	Species	ONT EMU (%)	ONT SILVA (%)	Difference (%)
Staphylococcus	40.21	39.29	0.92	*Staphylococcus Aureus*	20.68	19.74	0.94
Streptococcus	13	12.85	0.15	*Staphylococcus Epidermidis*	15.53	15.53	0
Haemophilus	8.48	8.12	0.36	*Streptococcus Pneumoniae*	6.45	6.06	0.39
Corynebacterium	1.95	1.97	0.02	*Streptococcus Dysgalactiae*	2.96	2.86	0.1
Cutibacterium	3.76	3.61	0.15	*Haemophilus Influenzae*	7.51	7.1	0.41
Klebsiella	4.03	2.05	1.98	*Cutibacterium Acnes*	3.52	3.03	0.49
Enterobacter	2.52	4.18	1.66	*Klebsiella Aerogenes*	4.02	2.02	2
Parvimonas	1.96	1.74	0.22	*Enterobacter Hormaechei*	2.44	1.64	0.8
Anaerococcus	3.04	3.03	0.01	*Parvimonas Micra*	1.96	1.74	0.22
Peptoniphilus	1.3	1.27	0.03	*Anaerococcus Nagyae*	1.68	0	1.68
Other	80.24	78.12	2.12	Other Species	33.25	41.8	8.55