UC Irvine
UC Irvine Previously Published Works

Title

Host DNA depletion on frozen human respiratory samples enables successful metagenomic sequencing for microbiome studies.

Permalink

Journal

Authors

Kim, Minsik

Parrish, Raymond

Tisza, Michael

et al.

Publication Date

DOI

Host DNA depletion on frozen human respiratory samples enables successful metagenomic sequencing for microbiome studies

Check for updates

Minsik Kim [1,2,3], Raymond C. Parrish II[1], Michael J. Tisza[4], Viral S. Shah[1], Thi Tran[1], Matthew Ross[4], Juwan Cormier[4], Aribah Baig[1,5], Ching-Ying Huang [1], Laura Brenner[1,2], Isabel Neuringer[1], Katrine Whiteson [6], J. Kirk Harris[7], Amy D. Willis [8] & Peggy S. Lai [1,2] ✉

Most respiratory microbiome studies use amplicon sequencing due to high host DNA. Metagenomics sequencing offers finer taxonomic resolution, phage assessment, and functional characterization. We evaluated five host DNA depletion methods on frozen nasal swabs from healthy adults, sputum from people with cystic fibrosis (pwCF), and bronchoalveolar lavage (BAL) from critically ill patients. Median sequencing depth was 76.4 million reads per sample. Untreated nasal, sputum, and BAL had 94.1%, 99.2%, and 99.7% host reads, respectively. Host depletion effects varied by sample type, generally increasing microbial reads, species and functional richness; this was mediated by higher effective sequencing depth. Rarefaction curves showed species richness saturation at 0.5–2 million microbial reads. Most methods did not change Morisita-Horn dissimilarity for BAL and nasal samples although the proportion of gram-negative bacteria decreased for sputum from pwCF. Freezing did not affect the viability of *Staphylococcus aureus* but reduced the viability of *Pseudomonas aeruginosa* and *Enterobacter spp*.; this was mitigated by adding a cryoprotectant. QIAamp-based host depletion minimally impacted gram-negative viability even in non-cryoprotected frozen isolates. While some host depletion methods may shift microbial composition, metagenomics sequencing without host depletion severely underestimates microbial diversity of respiratory samples due to shallow effective sequencing depth and is not recommended.

The respiratory microbiome has been associated with the development or exacerbation of a broad range of lung diseases ranging from respiratory infections, chronic lung diseases such as asthma, chronic obstructive pulmonary disease, and lung cancer[1]. However, a major barrier to progress is the high host DNA content of many respiratory samples, leading the respiratory microbiome field to rely on amplicon sequencing targeting the small subunit ribosomal RNA (SSU-rRNA) gene regions (typically 16S rRNA for bacteria and ITS2 for fungi) to describe respiratory microbial communities. While SSU-rRNA profiling is less costly and not limited by host DNA content, it has shortcomings compared to metagenomic next-generation sequencing (mNGS). Each kingdom requires different regions (bacterial or archaeal 16S rRNA genes, eukaryotic 18S rRNA or inter-spacer (ITS) region) to be amplified and sequenced while there is no conserved marker gene region for viruses. Common targets of 16S rRNA short reads such as the V3-V4 regions can only reliably identify taxonomy at the genus level[2,3]. Untargeted mNGS addresses some of these limitations of amplicon sequencing including cross-kingdom characterization of microbial communities and the ability to identify microbes at the species or strain level[4]

[1]Division of Pulmonary and Critical Care Medicine, Massachusetts General Hospital, Boston, MA, USA. [2]Department of Medicine, Harvard Medical School, Boston, MA, USA. [3]Department of Biological Engineering, Inha University, Incheon, Republic of Korea. [4]Alkek Center for Metagenomics and Microbiome Research, Baylor College of Medicine, Houston, TX, USA. [5]College of Science, Northeastern University, Boston, MA, USA. [6]Department of Molecular Biology & Biochemistry, University of California, Irvine, Irvine, CA, USA. [7]Department of Pediatrics, University of Colorado Anschutz Medical Campus, Aurora, CO, USA. [8]Department of Biostatistics, University of Washington School of Public Health, Seattle, WA, USA. ✉e-mail: plai@mgh.harvard.edu

(a degree of taxonomic and functional resolution critical for the design of future microbiome-targeted interventions[5]). mNGS additionally can assess DNA viruses, most notably phage communities targeting bacterial and archaeal hosts[6]. mNGS can also assess functional profiles[7,8], which is important given the ability of multiple microbes to perform the same community function[9].

During mNGS both mammalian host and microbial DNA is sequenced. While this is not a problem in certain sample types such as stool, which typically has less than 0.5% host DNA content[10], other sample types such as vacuumed dust and skin swabs have on average 50% host DNA[10,11]. One of the most challenging biospecimens are respiratory samples; with saliva and nasopharyngeal swabs averaging 90% and >99.9% host DNA[11], respectively. A proposed solution for mNGS of samples with high host content has been deeper sequencing, which may be tractable for samples with less than 90% host content. However, for biospecimens with >99% host DNA content, even ultra-deep sequencing is unlikely to overcome the challenges of undersampling due to inadequate effective sequencing depth after host read removal[12]. Other proposed solutions include culture enrichment to increase microbial load prior to mNGS[13] though abundance estimates no longer reflect in vivo conditions after culture. Some media, e.g., artificial sputum medium recipes also contain salmon sperm DNA[14], which will also be sequenced during mNGS and may overwhelm microbial-derived reads.

An alternate strategy to address the challenges of mNGS for low-biomass and high-host content samples is selective degradation or binding of human DNA prior to sequencing[15,16]. For example, osmotic lysis followed by propidium monoazide treatment to cross-link free DNA (lyPMA) has been employed for saliva frozen with cryoprotectants[11]. A benzonase-based approach has been tailored for sputum[17] and later for skin swabs and saliva[18]. Commercial kits have also been developed and tested in tissue specimens[19,20] and nasopharyngeal aspirate[21]. These studies focused on either treatment of never-frozen samples that required immediate processing at time of sample collection, or samples frozen with cryoprotectants[11] and did not assess differences between different respiratory samples. Host depletion at time of sample collection is resource-intensive and the requirement for cryoprotectants before freezing limits the generalizability of these tested methods as most longstanding cohort studies with biorepositories have not added cryoprotectants to respiratory specimens. Some respiratory specimens, such as human sputum, have natural cryoprotectant properties[22] suggesting that optimal host DNA depletion approaches may differ based on the underlying sample matrix, while other studies have shown that the degree of inflammation for chronic disease vary across proximal vs. lower airways which influences the amount of extracellular microbial DNA[23] (this is removed during host DNA depletion). A head-to-head comparison of host DNA depletion approaches for metagenomics across diverse samples along the respiratory tract continuum has not been performed, particularly for samples frozen without added cryoprotectants.

To address these challenges to the respiratory microbiome field, we evaluated the efficacy of 5 different commonly used methods for host depletion before mNGS using whole bronchoalveolar lavage fluid (BAL), nasal swabs, and spontaneously expectorated sputum frozen without cryoprotectants collected from ongoing human observational studies. Host depletion efficiency was evaluated based on sequencing failure rate, host DNA proportion, final non-human reads, non-viral microbial species richness, viral species richness, predicted functional richness, potential bias compared to the untreated community, and presence of contamination. Potential bias from host depletion was further assessed with viability studies.

Results
Host depletion efficiency
The methods compared in this study are as follows: lyPMA, developed for saliva by Marotz et al.[11]; Benzonase, a treatment tailored for sputum by Neslon et al.[17]; and HostZERO, MolYsis, and QIAamp, which are commercial kits developed by Zymo, Molzym, and Qiagen respectively (Fig. 1).

Fig. 1 | Overview of study design. Samples collected from the same participant were aliquoted so that paired comparisons could be made between treated and untreated samples. For nasal samples, it was only feasible to collect 4 swabs from a participant at the same time, thus a total of 10 swabs for the untreated condition was required to allow for paired treated and untreated comparisons.
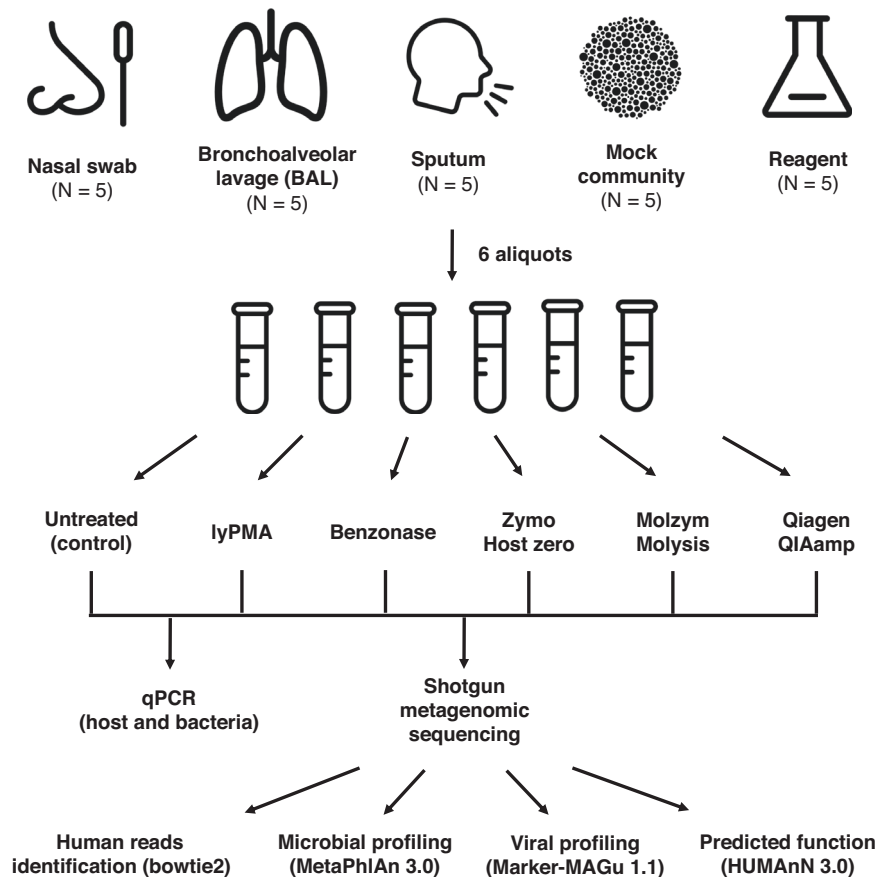
Summary statistics describing the effect of host depletion on human and bacterial DNA quantified by qPCR, library preparation and sequencing failure rates, proportion of host mapped reads, effective sequencing depth (final reads after human read removal), non-viral microbial, viral, and predicted functional richness are summarized in Table 1. Results of statistical models to quantify the effects of host depletion are summarized in Table 2. Based on qPCR, most host depletion methods decreased both total host and bacterial DNA for all sample types (Supplementary Fig. 1), although the degree of human DNA reduction far exceeded that for bacterial DNA. Thirteen samples out of 157 (including negative reagent-only controls) failed library prep based on fragment analysis but were nevertheless still sequenced for further analyses. Four lyPMA, two HostZERO, and four MolYsis treated nasal samples failed library prep. lyPMA, HostZERO, and MolYsis each failed library prep for one BAL sample (Table 1).

The median sequencing depth of all respiratory samples was 76.4 [interquartile range 46–138.8] million reads. After removal of host reads, untreated samples had a median of 0.33, 4.82, and 0.60 million reads for BAL, nasal, and sputum, respectively. Two BAL samples, one untreated and one lyPMA treated, had no microbial mapped reads (Supplementary Fig. 2) and were considered to have failed microbial sequencing.

Host DNA content was 99.7%, 94.1%, and 99.2% for BAL, nasal, and sputum samples, respectively, based on mNGS. Overall, the proportion of host DNA decreased after host depletion treatment (Supplementary Fig. 3B, D) though treatment was more effective for nasal and sputum compared to BAL samples. The proportion of host DNA estimated by sequencing and by qPCR were highly correlated ($R^2 = 0.92$, Supplementary Fig. 4A) and had a high agreement (Bland-Altman plot Supplementary Fig. 4B), indicating that qPCR using the primers tested can be used to reliably estimate host DNA content prior to mNGS. Change in % host DNA differed by sample type and host depletion treatment (Table S1). For BAL, the most effective treatment was HostZERO which decreased host DNA proportion by 18.3 [5.6–30.9]%, followed by MolYsis (17.7 [5.1–30.3]%). For nasal, all treatments besides Benzonase led to significant differences in % host content, with the most effective methods being QIAamp (75.4 [54.0–96.9]% decrease) and HostZERO (73.6 [52.1–94.9]% decrease). For sputum, the most efficient methods were MolYsis (69.6 [58.0–81.3]% decrease) and HostZERO (45.5 [33.8–57.1]% decrease).

Most host depletion treatments significantly increased final reads after host read removal though efficacy differed by sample type (Table S2). Untreated BAL had 0.3 million final reads, which significantly increased after all treatments except lyPMA. Specifically, all commercial kits resulted in a 10-fold increase (effect size of 1.0 after $\log_{10}$ transformation) in final reads. For nasal swabs, QIAamp increased final reads by 13-fold and HostZERO by 8-fold. For sputum, all treatments increased final reads; MolYsis, HostZERO, and QIAamp had the largest effect sizes, increasing final reads by 100-fold, 50-fold, and 25-fold, respectively.

Host depletion increases observed microbial (both non-viral and viral) species and predicted functional richness by increasing effective sequencing depth

Species-level non-viral and predicted viral profiles for untreated and treated samples are depicted in Fig. 2. Host depletion leads to a higher effective sequencing depth (final non-human reads), and thus we evaluated the effect of host depletion on observed species richness. Overall, species richness increased after host depletion (Table 1), although the magnitude of increase differed based on sample type and treatment (Table S3, Figs. 3A, B, S5 and S6). For BAL, only MolYsis showed significantly increased non-viral microbial species richness compared to the untreated samples (by 19 [7–31] species). For nasal swabs, HostZERO, QIAamp, and MolYsis increased non-viral microbial species richness by 10, 8, and 6, respectively. All host depletion treatments significantly increased the non-viral microbial species richness of sputum, with the largest effect sizes being from MolYsis, HostZERO, and QIAamp, with an increase of 113, 103, and 85 respectively. Changes in viral species richness were attenuated for nasal and BAL

samples. For BAL, most of the samples failed to identify any viral community members (Fig. 2). For sputum, HostZERO, MolYsis, and QIAamp increased viral species richness by 92, 118, and 47.

To determine whether higher final non-human reads explain the increase in species richness after host depletion, we performed a causal mediation analysis with host depletion method, final non-human reads, and species richness as the exposure, mediator, and outcome, respectively (Table S4). Besides lyPMA, all the treatments showed a significant indirect effect on non-viral microbial species richness. The proportion mediated by HostZERO, MolYsis, and QIAamp was over 50% of the total effect, indicating that the increase in species richness after host depletion was largely explained by the increase in final reads. Similar results were seen when evaluating predicted microbial functional richness.

Potential bias in microbial community composition due to host depletion treatment

Most host depletion methods rely on the observation that host cells are more vulnerable to lysis than microbial cells and thus require microbial cells to be intact. However, gram-negative bacteria are potentially more vulnerable to the effects of freezing and host depletion compared to gram-positive bacteria or fungi, and we show that host depletion methods decreased both human and bacterial DNA concentrations. Thus we evaluated the effect of host depletion on the relative abundance of gram-negative bacteria present in a mock community preserved in DNA/RNA Shield (Zymo) (Supplementary Fig. 7) as well as in respiratory samples (Supplementary Fig. 8, Table S5). In analyses stratified by sample type, the effect was the strongest in the mock community compared to respiratory samples, which was expected as the mock community we used is stored in DNA/RNA Shield, a common nucleic acid stabilizing agent that contains a mild detergent to inactivate infectious agents and prevent further microbial growth. Host depletion treatment did not decrease the relative abundance of gram-negative bacteria in BAL, only lyPMA changed the relative abundance of gram-negative bacteria in nasal samples (increased by 19.4%), while all host depletion treatments decreased the relative abundance of gram-negative bacteria in sputum; note that all sputum samples were obtained from patients with cystic fibrosis. Key members of the cystic fibrosis airway community, such as *Pseudomonas aeruginosa*, are known to produce large amounts of extracellular DNA[24], and most host depletion protocols rely on removal of extracellular DNA after lysis of host cells. Microbes could have been ingested by neutrophil extracellular traps (NETs) increasing extracellular microbial DNA as NETosis is induced in response to microbial cues[25].

Changes in overall microbial community structure were analyzed using Morisita-Horn dissimilarity. PCoA plots stratified by sample type suggested the presence of strong sample-type-specific treatment effects (Supplementary Fig. 9). PERMANOVA analysis was conducted (Table S6), and all treatments showed sample-type-specific effects. To better quantify potential bias from host depletion treatment, we calculated Morisita-Horn dissimilarities between paired samples (untreated and host-depleted) from the same subject and sample type (Fig. 3C, Table S7) and used this continuous measure as an outcome in linear mixed effects models. For non-viral microbial communities, only lyPMA changed the paired Morisita-Horn dissimilarities for BAL. lyPMA, Benzonase, and MolYsis changed the paired Morisita-Horn dissimilarities of nasal samples. All the treatments changed the paired Morisita-Horn dissimilarities of sputum. The change in viral community structure (Fig. 3D, Table S7) also indicated that similar alterations were introduced for sputum samples after host depletion, while different methods were associated with significant bias for nasal samples, and most BAL samples lacked identified viral species making assessment of bias challenging.

Effect of host depletion treatment on differential abundance of non-viral and viral microbes

To determine whether there are species-specific effects of host depletion treatment for non-viral microbes, we conducted differential abundance analysis using linear mixed-effect models after centered log-ratio

Table 1 | Summary statistics of host depletion treatment effect stratified by sample type and host depletion treatment

Sample	Treatment	N	Human DNA pg/µL	Bacterial DNA pg/µL	Sequencing fail[a] N (%)	QC'd reads reads x 10^6	Host reads %	Final reads[b] reads x 10^6	Microbial species richness[c]	Viral species richness[c]	Predicted functional richness[c]
BAL	Untreated	5	1512.8 (1237.9, 9836.4)	12.6 (10.5, 37.8)	1[d] (20%)	129.5 (52.5, 129.9)	99.7 (99.6, 99.7)	0.3 (0.3, 0.4)	3 (3, 5)	1 (1, 1)	8 (7, 56)
	lyPMA	5	2139.7 (60.4, 6255.6)	8.4 (0.3, 17.4)	2[d] (40%)	46.7 (28.6, 110.0)	99.1 (97.8, 99.5)	0.6 (0.4, 1.0)	7 (6, 8)	2 (1, 2)	86 (54, 141)
	Benzonase	5	59.6 (47.8, 70.1)	0.9 (0.7, 2.3)	0 (0%)	149.3 (129.7, 183.7)	98.8 (98.7, 98.9)	1.7 (1.6, 2.2)	6 (5, 7)	14 (9, 20)	162 (116, 179)
	HostZERO	5	6.8 (2.3, 7.5)	0.4 (0.3, 1.1)	1 (20%)	31.9 (18.4, 35.1)	83.7 (76.8, 87.2)	2.4 (1.3, 8.2)	8 (7, 11)	17 (10, 24)	226 (132, 237)
	MolYsis	5	7.6 (6.6, 25.2)	2.0 (0.3, 4.6)	1 (20%)	39.0 (29.0, 39.3)	92.5 (92.5, 93.6)	2.9 (1.3, 15.6)	17 (7, 45)	23 (13, 37)	232 (226, 264)
	QIAamp	5	33.1 (32.0, 79.5)	0.5 (0.3, 1.9)	0 (0%)	132.4 (119.6, 137.5)	98.3 (92.3, 98.6)	2.6 (1.0, 10.2)	8 (6, 14)	27 (15, 39)	231 (32, 234)
Nasal	Untreated	10[e]	340.2 (202.3, 685.8)	22.9 (16.9, 26.6)	0 (0%)	106.2 (63.7, 138.7)	94.1 (92.8, 97.9)	4.8 (1.0, 8.7)	12 (9, 13)	10 (5, 13)	153 (125, 167)
	lyPMA	5	2.6 (0.8, 9.2)	0.3 (0.3, 0.3)	4 (80%)	7.9 (6.9, 9.7)	91.2 (35.6, 91.6)	0.7 (0.6, 0.8)	5 (5, 7)	3 (2, 3)	146 (133, 181)
	Benzonase	5	12.8 (1.9, 78.8)	6.1 (5.4, 10.2)	0 (0%)	47.1 (41.7, 53.2)	78.7 (77.8, 94.8)	2.8 (2.6, 10.4)	8 (7, 14)	4 (3, 5)	200 (136, 203)
	HostZERO	5	0.5 (0.1, 0.7)	7.6 (3.3, 15.8)	2 (40%)	24.5 (11.7, 55.2)	8.9 (2.7, 30.4)	24.3 (9.7, 50.3)	20 (19, 20)	15 (11, 17)	224 (211, 226)
	MolYsis	5	0.4 (0.0, 0.8)	1.6 (1.1, 5.8)	4 (80%)	8.1 (5.0, 34.9)	49.9 (5.0, 78.4)	3.2 (1.7, 25.3)	12 (11, 21)	6 (4, 18)	210 (198, 216)
	QIAamp	5	2.1 (0.9, 7.1)	28.8 (24.7, 30.7)	0 (0%)	56.2 (54.9, 58.5)	20.1 (15.7, 23.2)	46.3 (45.0, 46.7)	18 (17, 20)	17 (8, 21)	223 (204, 227)
Sputum	Untreated	5	39,231.5 (19,448.0, 59,430.9)	245.3 (220.3, 311.0)	0 (0%)	69.2 (68.0, 75.6)	99.2 (98.9, 99.2)	0.6 (0.6, 0.9)	13 (10, 15)	2 (2, 2)	159 (157, 160)
	lyPMA	5	9779.5 (994.2, 11,437.5)	97.8 (25.9, 100.9)	0 (0%)	89.7 (42.0, 105.2)	96.4 (92.5, 98.3)	2.5 (1.5, 4.4)	50 (43, 65)	11 (7, 16)	266 (259, 276)
	Benzonase	5	154.3 (141.3, 349.0)	33.2 (21.3, 53.0)	0 (0%)	84.0 (82.0, 87.1)	94.2 (92.9, 94.5)	4.7 (4.5, 5.9)	94 (84, 95)	33 (29, 36)	288 (282, 322)
	HostZERO	5	49.4 (11.7, 57.7)	38.9 (33.6, 39.0)	0 (0%)	106.2 (61.6, 114.8)	61.7 (37.5, 68.0)	29.1 (23.6, 36.7)	131 (112, 133)	89 (67, 162)	328 (313, 340)
	MolYsis	5	13.6 (8.0, 26.3)	28.4 (24.1, 30.4)	0 (0%)	105.6 (90.8, 115.7)	32.8 (17.0, 33.8)	61.1 (55.6, 83.7)	126 (123, 135)	140 (93, 153)	328 (326, 333)
	QIAamp	5	241.6 (196.3, 273.5)	64.3 (34.6, 71.0)	0 (0%)	102.4 (100.9, 106.0)	88.2 (68.9, 88.6)	11.6 (11.3, 38.9)	103 (90, 114)	45 (42, 47)	293 (291, 296)

Number of samples in each experimental group, human and bacterial DNA quantified by qPCR, number of samples did not pass the library QC or failed sequencing (defined as no microbial mapped reads), QC'd reads, % host mapped reads, final reads, average observed non-viral microbial species richness, viral species richness, and predicted functional richness. Values depicted as N (%) or median (interquartile range).
[a]Samples that failed library prep were still sequenced. Sequenced samples with 0 microbial mapped reads after QC and host read removal were considered to have failed sequencing.
[b]Final reads were calculated after removing samples that failed sequencing.
[c]Samples resulted in no identified viruses were considered NAs and omitted from calculations but included in statistical calculations as zeros.
[d]Includes 1 sample that showed zero microbial species richness.
[e]Only 4 nasal could be collected from one individual at each time point thus with 6 treatment conditions, 4 nasal were collected from 10 participants and randomized to each host depletion method.

Table 2 | Overall effect sizes (95% CIs) of host depletion method based on statistical models

Sample	Treatment	% Host	log₁₀(Final reads)	Microbial species richness	Viral species richness	Predicted functional richness	Microbial bias (Morisita-Horn)	Viral bias (Morisita-Horn)	% Gram-negative
BAL	lyPMA	−3.1 (−15.7, 9.5)	0.4 (−0.2, 0.9)	1.8 (−10.7, 14.2)	0.4 (−10.7, 14.2)	63.0 (−48.7, 174.7)	0.3 (0.1, 0.6)*	−ᵃ	6.4 (−21.2, 34.1)
	Benzonase	−1.1 (−13.8, 11.5)	0.8 (0.3, 1.3)*	5.7 (−6.2, 17.6)	5.6 (−6.2, 17.6)	137.2 (25.5, 248.9)*	0.1 (−0.1, 0.3)	-	15.3 (−12.3, 43.0)
	HostZERO	−18.3 (−30.9, −5.6)*	1.0 (0.4, 1.5)**	8.9 (−3.0, 20.8)	6.6 (−3.0, 20.8)	177.8 (66.1, 289.5)**	0.3 (0.0, 0.5)	-	8.6 (−19.0, 36.3)
	MolYsis	−17.7 (−30.3, −5.1)*	1.0 (0.5, 1.6)**	18.9 (7.0, 30.8)*	15.0 (7.0, 30.8)*	203.2 (91.5, 314.9)***	0.2 (0.0, 0.4)	-	3.4 (−24.3, 31.0)
	QIAamp	−6.3 (−18.9, 6.3)	1.0 (0.5, 1.6)**	9.5 (−2.4, 21.4)	10.6 (−2.4, 21.4)	139.4 (27.7, 251.1)*	0.3 (0.0, 0.5)	-	7.3 (−20.4, 35.0)
Nasal	lyPMA	−27.7 (−49.0, −6.3)*	−0.5 (−1.0, −0.1)*	−4.8 (−9.2, −0.4)	−6.4 (−9.2, −0.4)*	7.8 (−26.7, 42.3)	0.2 (0.1, 0.3)*	0.5 (0.1, 0.3)***	19.4 (14.1, 24.6)***
	Benzonase	−20.0 (−41.4, 1.5)	0.1 (−0.3, 0.6)	−0.4 (−4.8, 4.1)	−3.9 (−4.8, 4.1)	23.4 (−11.1, 57.9)	0.2 (0.0, 0.3)*	0.1 (0.0, 0.3)	1.9 (−3.4, 7.5)
	HostZERO	−73.6 (−94.9, −52.1)***	0.9 (0.4, 1.3)**	10.0 (5.6, 14.5)***	4.1 (5.6, 14.5)	69.4 (34.9, 103.9)**	0.0 (−0.1, 0.2)	0.1 (−0.1, 0.2)	0.0 (−5.3, 5.6)
	MolYsis	−50.6 (−72.0, −29.3)***	0.2 (−0.2, 0.7)	6.2 (1.8, 10.6)*	−0.2 (1.8, 10.6)	56.2 (21.7, 90.7)**	0.2 (0.1, 0.3)**	0.2 (0.1, 0.3)	2.3 (−3.0, 7.6)
	QIAamp	−75.4 (−96.9, −54.0)***	1.1 (0.6, 1.5)***	7.8 (3.3, 12.2)**	4.1 (3.3, 12.2)	69.4 (34.9, 103.9)***	0.1 (0.0, 0.3)	0.4 (0.0, 0.3)**	0.1 (−5.4, 5.5)
Sputum	lyPMA	−3.8 (−15.4, 7.8)	0.5 (0.3, 0.8)**	37.6 (19.5, 55.7)**	4.2 (19.5, 55.7)	86.0 (19.6, 152.4)*	0.3 (0.1, 0.6)**	0.2 (0.1, 0.6)	−40.9 (−52.6, −29.1)***
	Benzonase	−6.3 (−17.9, 5.4)	0.8 (0.6, 1.1)***	66.6 (48.5, 84.7)***	19.2 (48.5, 84.7)	91.6 (25.2, 158.0)**	0.5 (0.3, 0.7)***	0.3 (0.3, 0.7)	−52.5 (−64.2, −40.7)***
	HostZERO	−45.5 (−57.1, −33.8)***	1.7 (1.4, 1.9)***	103.0 (84.9, 121.1)***	91.8 (84.9, 121.1)***	145.8 (79.4, 212.2)***	0.6 (0.4, 0.8)***	0.7 (0.4, 0.8)**	−59.9 (−71.6, −48.1)***
	MolYsis	−69.6 (−81.3, −58.0)***	2.0 (1.7, 2.2)***	112.8 (94.7, 130.9)***	118.4 (94.7, 130.9)***	150.4 (84.0, 216.8)***	0.6 (0.4, 0.8)***	0.8 (0.4, 0.8)***	−59.9 (−71.6, −48.1)***
	QIAamp	−18.7 (−30.3, −7.1)*	1.4 (1.2, 1.7)***	85.2 (67.1, 103.3)***	46.6 (67.1, 103.3)*	123.6 (57.2, 190.0)***	0.6 (0.4, 0.8)***	0.7 (0.4, 0.8)***	−60.6 (−72.3, −48.8)***

Changes were assessed using linear mixed-effect models (features ~ host depletion treatment + (1|subject)) stratified by sample type; untreated samples modeled as the comparator group. Sample sizes were $n = 30$, 35 and 30 for BAL, nasal, and sputum samples, respectively. For microbial bias, samples with no reads were omitted from the analysis, and sample sizes were $n = 28$, 35, and 30 for BAL, nasal, and sputum samples, respectively. For viral bias, samples with no reads were omitted from the analysis, and sample sizes were $n = 35$ and 29 for nasal and sputum samples, respectively. *p-value < 0.05, **p-value < 0.01 and ***p-value < 0.001. Values depicted effect size (95% confidence intervals) derived from statistical models. With BAL, MolYsis was the optimal method. lyPMA, Benzonase, and QIAamp showed low depletion efficiency, HostZERO did not increase richness. With nasal swabs, QIAamp was the optimal method. lyPMA and Benzonase showed no microbial richness increase, HostZERO showed high library prep failure rate, MolYsis showed high bias. With sputum HostZERO was the optimal method. lyPMA and Benzonase showed low depletion efficiency, MolYsis and QIAamp showed higher bias.
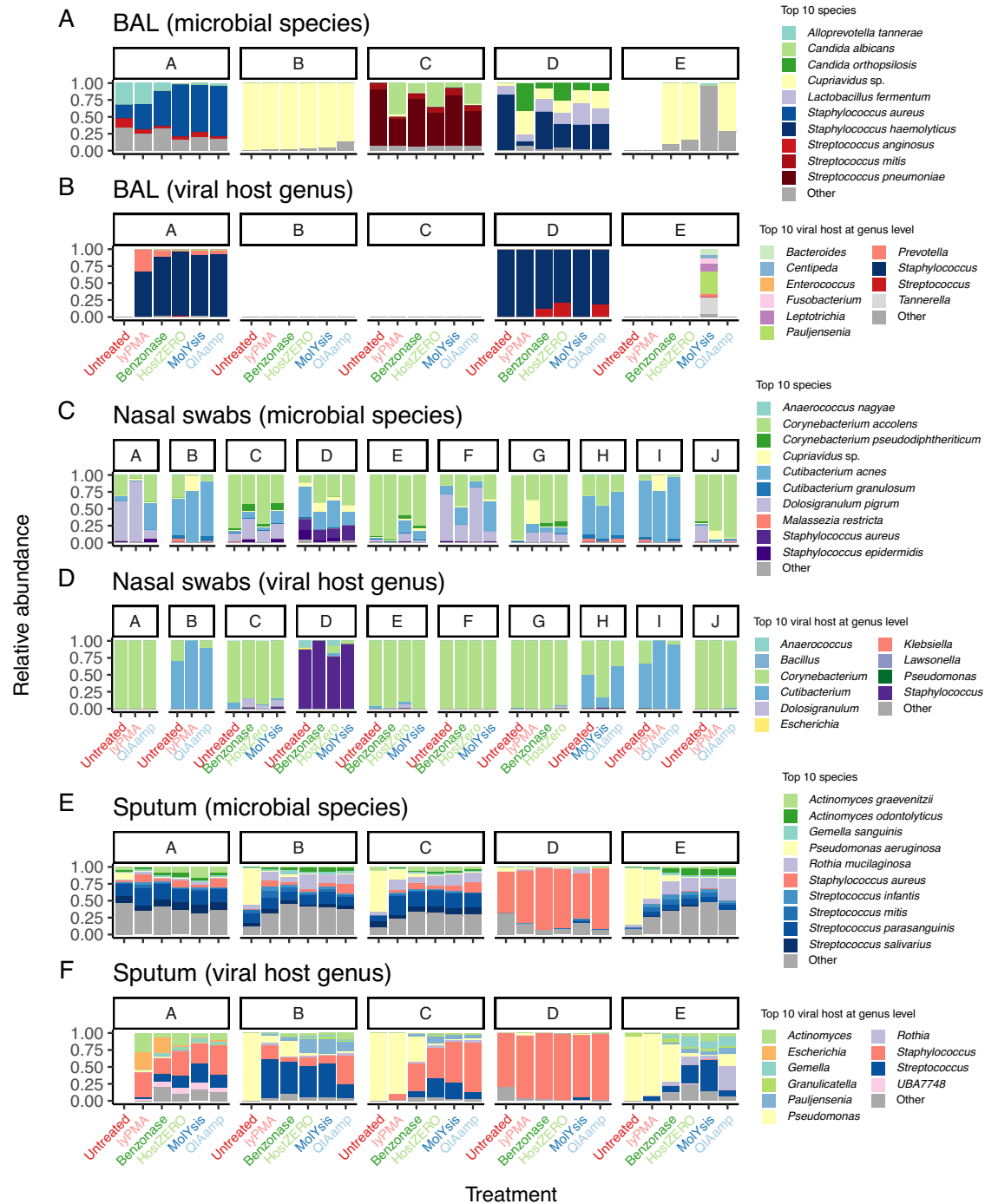ᵃNote the majority of BAL samples had no viral species identified thus unable to perform statistical modeling.

Fig. 2 | Sample relative read abundances of microbial non-viral and viral clades stratified by sample type and subject. Note viral clades depicted using predicted microbial (largely bacterial) host to facilitate interpretability as most DNA viruses identified by metagenomics sequencing are bacteriophages. Prediction algorithms for phage viral hosts reliable only at the genus level, thus viral profiles depicted at this level. Bronchoalveolar lavage (BAL) from critically ill patients **A** Microbes (non-viral) and **B** Viruses. Nasal swab samples from healthy adults **C** Microbes (non-viral) and **D** Viruses. Spontaneously expectorated sputum from people living with cystic fibrosis **E** Microbes (non-viral) and **F** Viruses. Empty space indicates samples that failed sequencing (no microbial reads identified). Nasal swab samples were collected from 10 different subjects as it is not feasible to collect more than 4 nasal swabs per participant at any given time, thus the experimental design was modified to ensure an equal number of replicates for each host depletion group, resulting in a larger number of control samples for nasal swabs.

transformation, accounting for the fixed effects of sample type and treatment and considering the random effect of each subject (Table S8). Given that most of the significant associations were due to sample type (Fig. 4A), we then performed differential abundance analyses stratified by sample type (Fig. 5A, C, E, top 20 non-viral microbial species by minimum q-value). For BAL, host depletion treatment did not lead to differentially abundant taxa. For nasal samples, 19 taxa were differentially abundant at a

significant level of $q < 0.1$. For sputum samples, 111, 102, 101, 86, and 82 taxa were differentially abundant compared to untreated sputum samples for QIAamp, MolYsis, HostZero, Benzonase, and lyPMA treatments, respectively.

The same analysis was conducted for viruses (evaluated at the level of predicted host genus to facilitate comparison with non-viral microbes), and the result is illustrated in Fig. 4B. When evaluating all sample types, 2, 15, 17,

Fig. 3 | Alpha and beta diversity stratified by sample type and treatment method. Species richness in mean values ± SD for microbial non-viral (**A**) and viral (**B**) communities. Boxplot of potential bias measured by Morisita-Horn dissimilarity (1 – Morisita-Horn similarity index) between each host depletion method and corresponding untreated sample for non-viral microbial (**C**) and viral (**D**) communities. Note most BAL samples had no detected viral communities. Statistical significance was tested with linear mixed-effect model adjusting for repeated measures in a participant as a random effect variable. *p-value < 0.05, **p-value < 0.01 and ***p-value < 0.001.
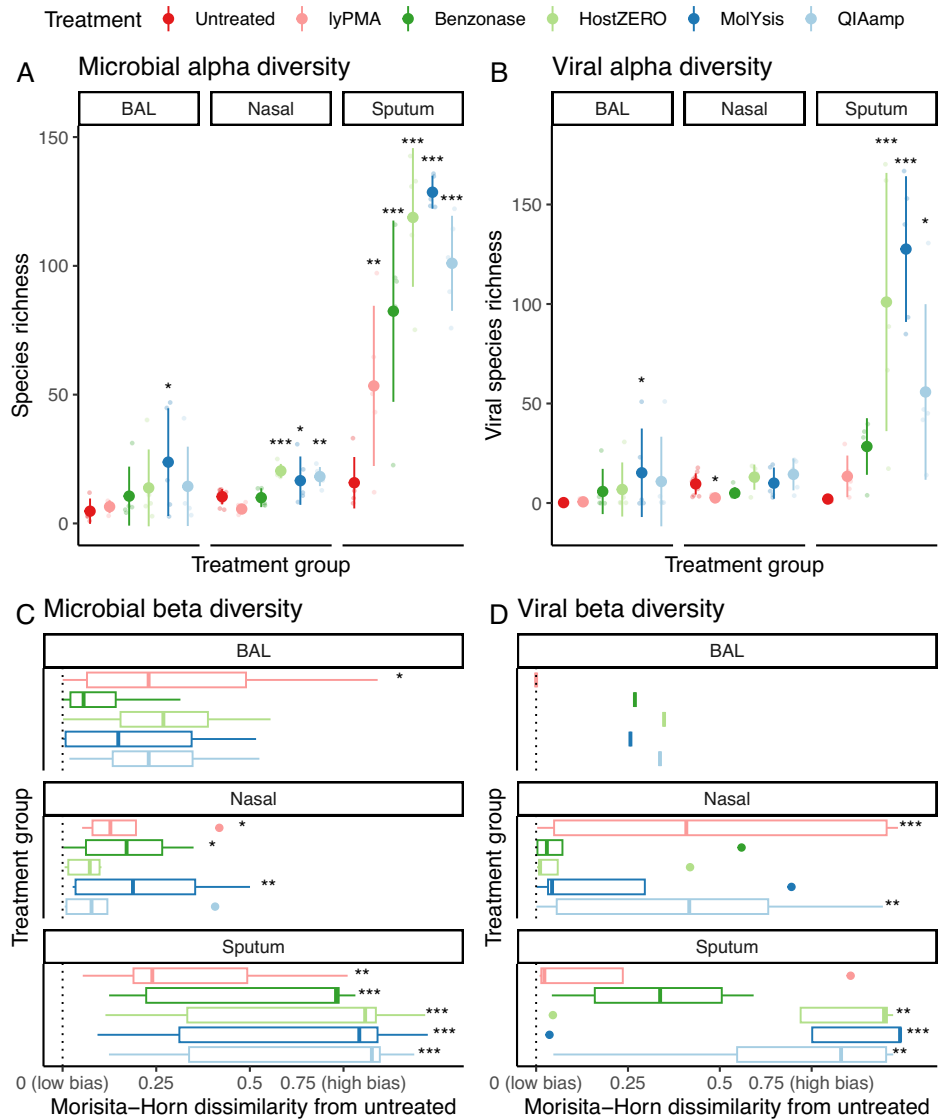
and 8 viral clades were differentially abundant after Benzonase, HostZero, MolYsis, and QIAamp treatments respectively, while 37 were differentially abundant by sample type. The predicted microbial host of the differentially abundant viruses largely did not correlate with the differentially abundant bacteria with only a few genera having the same pattern from both the viral and non-viral microbial clades (*Actinomyces, Escherichia, Pseudomonas,* and *Streptococcus*); Fig. 5B, D, F, top 20 host genus by minimum q-value and Supplementary Fig. 10, Spearman's correlation between differentially abundant microbial species and viral host genus.

Effect of host depletion treatment on predicted microbial community function
Similar to species richness, host depletion increased the richness of predicted microbial community functions (Table 1). Most of the treatments significantly increased predicted functional richness (Supplementary Fig. 11A and Table S9). For BAL, MolYsis, HostZERO, QIAamp, and Benzonase treatment increased functional richness by 203, 178, 139, and 137 pathways, respectively. For nasal, HostZERO, QIAamp, and MolYsis increased functional richness by 70, 70, and 56 pathways, respectively. For sputum, all treatments increased functional richness, with the largest effect size seen with MolYsis (150), HostZERO (146), and QIAamp (124). Compared to taxonomic profiles, Morisita-Horn dissimilarities in functional profiles showed smaller changes after treatment for nasal and sputum but higher for BAL (Supplementary Fig. 11B). Larger numbers of predicted functions were

differently abundant in CPM (copies per million) with differences based on sample type (Supplementary Fig. 12). For BAL, pathways unable to be identified in untreated BAL were detected after most host depletion treatments (Supplementary Fig. 13).

Sensitivity analysis for potential effect of contamination
Given that increasing effective sequencing depth was associated with increased species and predicted microbial richness (Supplementary Fig. 5), we performed a sensitivity analysis to ensure that the increased species richness was not due to the introduction of contaminants, given the low-biomass nature of most respiratory samples. Although extracted nucleic acids and sequencing libraries prepared from reagent-only negative controls had undetectable DNA concentrations (Table S10), we pooled larger volumes of the libraries from reagent-only negative controls for sequencing in order to assess the possible effect of contamination. We identified potential contaminants using two approaches: the approach implemented in the 'decontam' R package[26] and the approach implemented in the 'tinyvamp' R package[27] (Supplementary Fig. 14). Even after removing the 7 species identified as potential contaminants by decontam (Table S11), and analyzing corrected relative abundances estimated by tinyvamp, host depletion treatment increased species richness. Using mixed effects models stratified by sample type on these decontaminated datasets, the increase in species richness from host depletion treatment remained significant (Table S12).
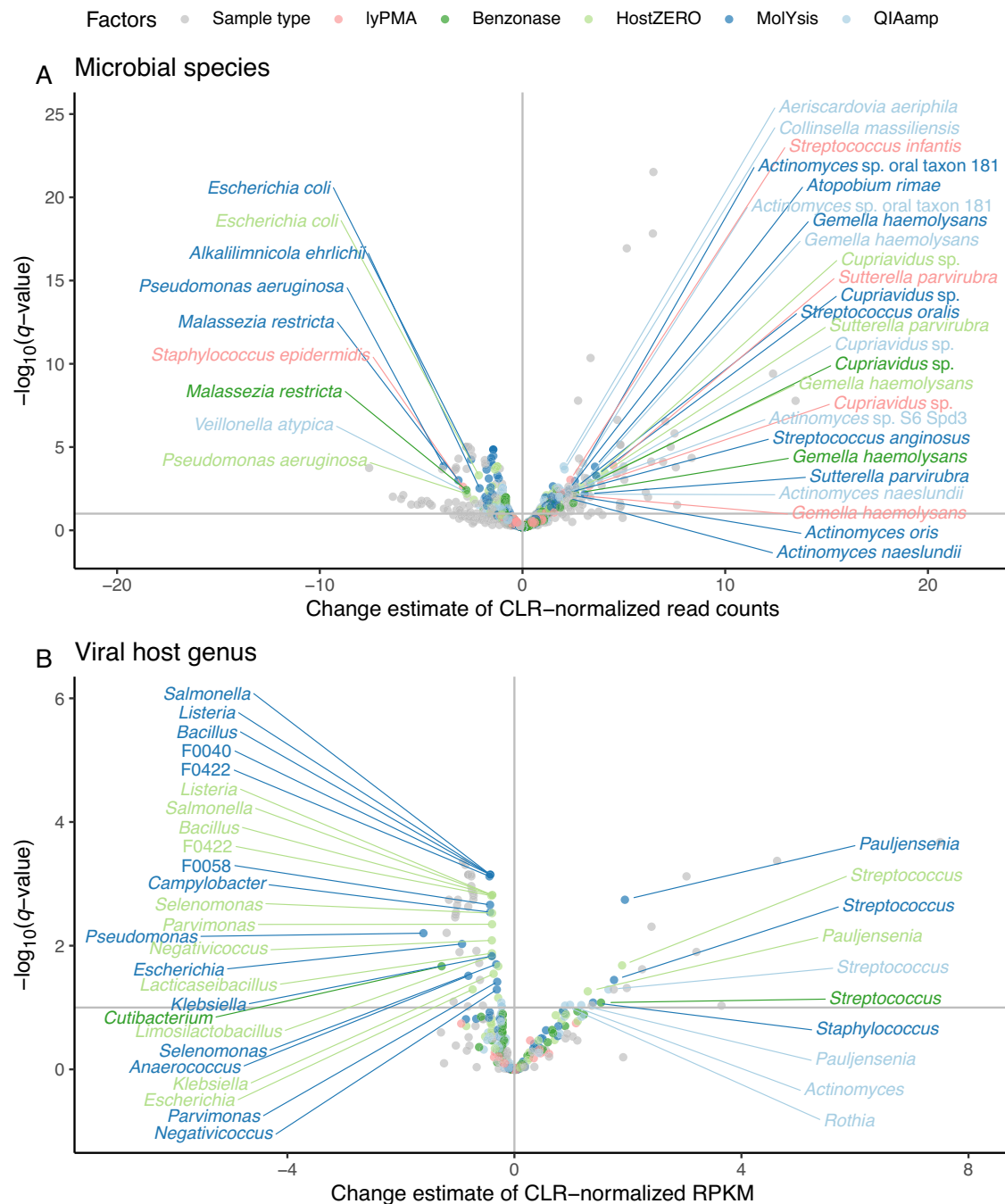
Fig. 4 | Volcano plot of differential abundance. A Non-viral microbes and **B** viral host genus by sample type and host depletion treatment. Each dot represents association analyzed by differential abundance analysis. Microbial taxa that showed strong significant changes (q-val < 0.01, |effect size| > 2) and viral host genus with significant changes (q-val < 0.1, |effect size| > 0.3) were labeled with their names. The analysis was conducted with linear mixed-effect model (feature ~ sample type + lyPMA + Benzonase + HostZERO + MolYsis + QIAamp, random effect = subject id) after centered-log ratio normalization.

Effect of freezing, DNA/RNA shield, and host depletion on microbial cell viability

To validate the effect of freezing without and with a cryoprotectant (glycerol), addition of DNA/RNA shield, and each host depletion method on the viability of typical gram-negative and gram-positive bacteria found in the respiratory tract, CFU testing was conducted for isolates from sputum samples (Fig. 6). Except for DNA/RNA shield, which universally rendered all isolates non-viable, experimental conditions exhibited species-specific effects on bacterial viability. In analyses stratified by species, freezing did not impact the viability of the gram-positive isolate (*Staphylococcus aureus*), while it reduced the viability of gram-negative

isolates (*Pseudomonas aeruginosa* and *Achromobacter* spp.); this effect was largely mitigated when a cryoprotectant (glycerol) was added (Table S13). When comparing the effect of host depletion on viability, MolYsis had the largest negative effect size (−4.7 (−5.8, −3.6), while the intercept was 9.0 (8.3, 9.7)) on viability for *Staphylococcus aureus* and thus testing for this method was not performed for the other isolates. Of the host depletion methods tested, QIAamp had the lowest impact on viability (both on samples without and with glycerol as a cryoprotectant); in multivariate analyses, it was the only host depletion treatment that did not reduce the viability of the tested gram-negative species even in non-cryopreserved frozen samples (Table S14).
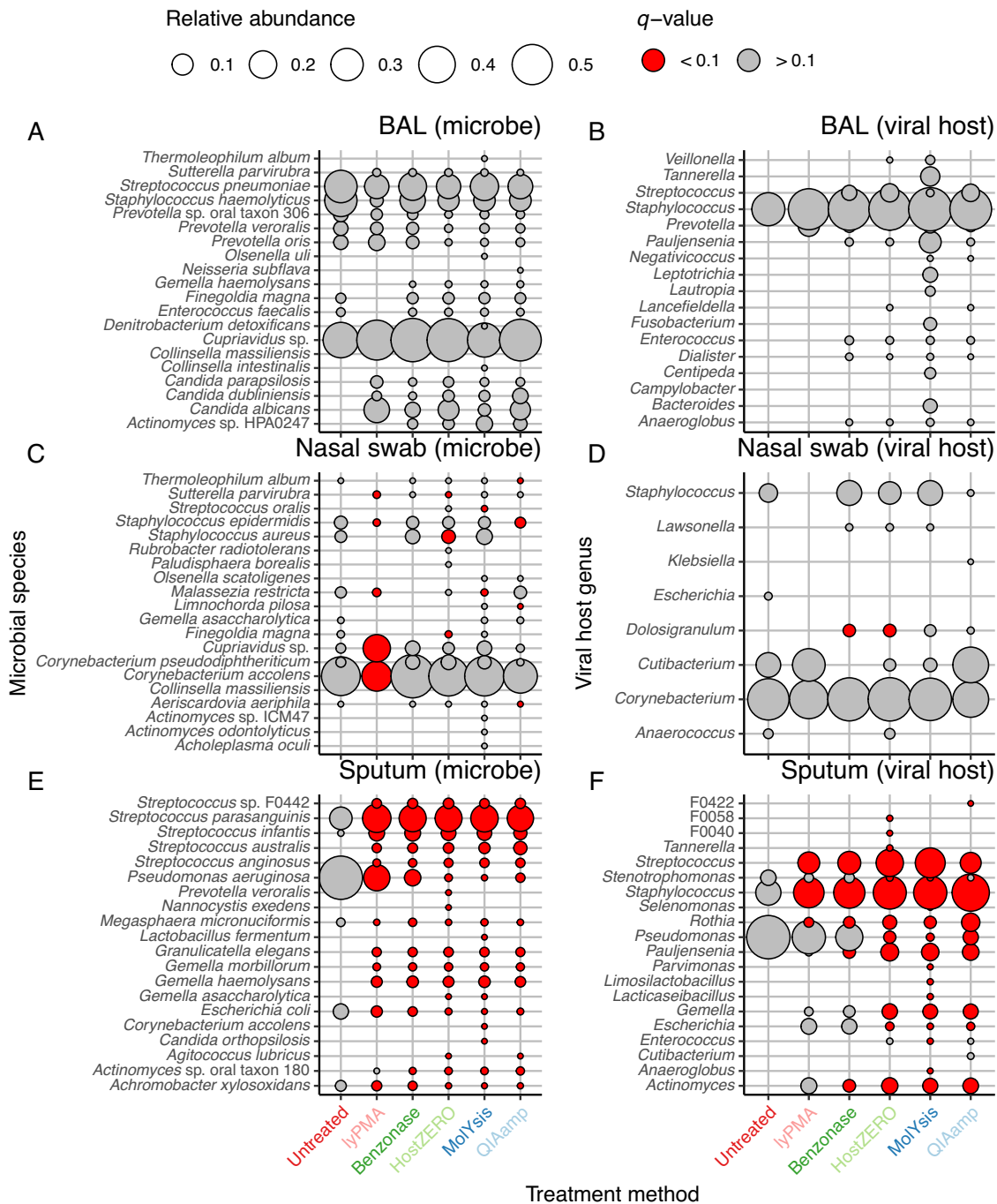
Fig. 5 | Mean relative abundance of top 20 significant taxa identified by differential abundance analysis using linear mixed-effect model (feature ~ lyPMA + Benzonase + HostZero + MolYsis + QIAamp + (1|subject id)) after centered log-ratio transformation. Analyses were stratified by sample type. A Non-viral microbial species and B viral host genus for bronchoalveolar lavage, C microbial species and D viral host genus for nasal swabs, and E microbial species and F viral host genus for sputum. Statistical significances threshold q-value < 0.1.

Extracellular microbial DNA production

To examine whether removal of extracellular DNA may contribute to the decrease in proportion of gram-negative species noted after host depletion, we quantified the proportion of extracellular DNA for isolates from sputum samples grown in overnight tryptic soy broth cultures (Table S15). The two tested gram-negative bacteria showed a higher proportion of extracellular DNA ($3.0 \pm 2.2\%$ for *Pseudomonas aeruginosa* and $4.2 \pm 3.8\%$ for *Achromobacter* spp.) compared to the gram-positive bacterium ($1.1 \pm 1.0\%$). In linear models, gram-negative *Achromobacter* spp. exhibited a significantly higher proportion of extracellular DNA relative to gram-positive *Staphylococcus aureus*, with an effect size of 3.1 [0.04, 6.1]%.

Discussion

Respiratory samples have host DNA content often exceeding 95%, making successful characterization of the respiratory microbiome using mNGS challenging even with deeper sequencing due to unobserved richness. We tested five host depletion approaches using published methods or commercial kits and showed that even in respiratory samples frozen without cryoprotectants, significant depletion of host DNA can be achieved. The increase in effective sequencing depth rather than contamination introduced by host depletion treatment explains the observed increase in species richness after host depletion. We saw similar findings when evaluating predicted microbial functional richness. Similarly, viral species richness was

Fig. 6 | Viability of pathogens isolated from sputum samples under different treatment conditions. For each species and experimental group, viabilities of 7 strains of *Staphylococcus aureus*, 6 strains of *Pseudomonas aeruginosa*, and 5 strains of *Achromobacter* spp. obtained from sputum cultures assessed using colony-forming unit (CFU) tests. A pseudocount of 1 was added prior to $\log_{10}$ transformation due to the presence of zero counts. MolYsis had a strong negative effect on viability for *Staphylococcus aureus* ($-4.3 \log_{10}$ cells/mL, p-value < 0.001) thus was not further tested in gram-negative isolates. DNA/RNA shield, when added to unfrozen culture, universally rendered all isolates non-viable.
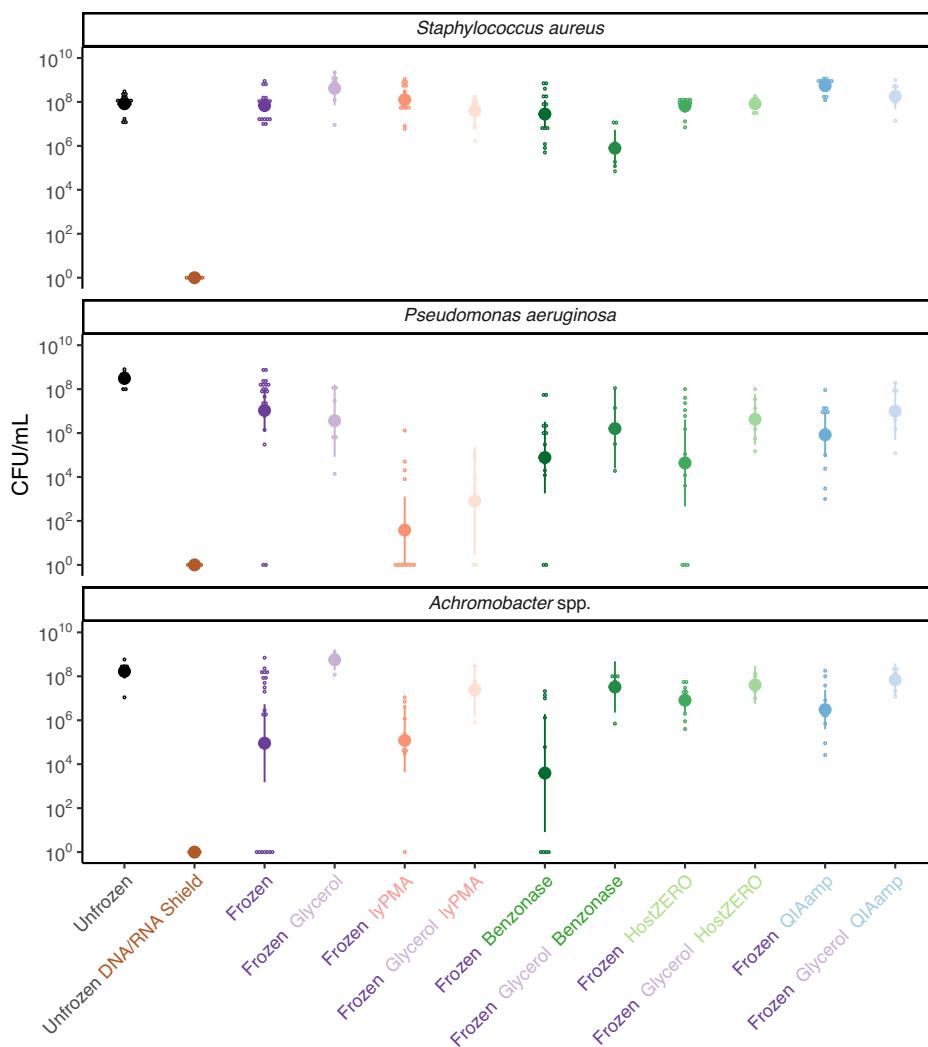
increased after adopting host depletion methods. In viability studies, the addition of DNA/RNA shield rendered all isolates non-viable and thus should not be used as a preservative if investigators wish to perform host depletion prior to mNGS. While freezing reduced the viability of gram-negative but not gram-positive bacteria, QIAamp treatment on frozen non-cryoprotected samples had the smallest effect on the viability of gram-negative bacteria. Adding a cryoprotectant minimized the differential effect of freezing and host depletion on the viability seen with gram-positive vs. gram-negative bacteria.

Previously published metagenomics sequencing studies of respiratory samples have high failure rates due to high host DNA content. For example, a study using nasopharyngeal swabs for COVID-19 testing found that 54.7% of samples had 100% human reads in mNGS without host DNA depletion[28,29]. Several studies have evaluated the efficacy of host depletion approaches on respiratory samples for mNGS, though they focused on a single type of respiratory sample, largely evaluated methods using fresh-unfrozen samples, and sequencing depth was significantly lower (32 million reads per sample or less) compared to our study. An earlier study tested MolYsis and a method using β-mercaptoethanol and DNase for the sputum of people living with cystic fibrosis for metagenomic microbial and viral profiling[30], however, the study used fresh-unfrozen samples and did not focus on the differences between different treatment methods. Marotz et al.[11] developed the lyPMA approach and found it superior to MolYsis and QIAamp when tested on unfrozen saliva samples. In this study, for frozen non-cryoprotected saliva samples, lyPMA was the least biased among all the methods tested though freezing decreased the efficacy and increased the variability of host depletion. Saliva contains high host DNA content but also

has higher microbial loads than respiratory samples from patients without infection[31]. Nelson et al.[17] developed the Benzonase method and found it superior to an alternate benzonase-based method, lyPMA, and the MolYsis Basic kit designed for small sample volumes of 0.2 mL or less (we tested the MolYsis Basic 5, which is designed for sample volumes up to 5 mL). Testing was performed on sputum frozen without cryoprotectants from children with cystic fibrosis; note sputum from patients with chronic infection often has paradoxically also higher host DNA content due to the influx of inflammatory immune cells. Benzonase led to a greater reduction in % host DNA than other methods, whereas we found that the MolYsis Basic 5 kit, followed by HostZERO and QIAamp were the most efficient. Similar to our results, they also noted a decrease in the relative abundance of certain gram-negative bacteria such as *Pseudomonas* and *Achromobacter* in Benzonase-treated samples compared to untreated samples. Based on viability studies using culture, Benzonase treatment did not decrease the viability of these gram-negative bacteria. Thus, the investigators concluded that the reduction in gram-negative bacteria was due to removal of extracellular DNA[17]. In contrast, we found that the viability of two gram-negative bacterial isolates was reduced with both freezing and most host depletion methods including Benzonase; this was mitigated with the addition of glycerol as a cryoprotectant. However, QIAamp, which also uses Benzonase, had the least impact on the viability of host-depletion-treated gram-negatives that were frozen without a cryoprotectant. Rajar et al.[21] evaluated frozen nasopharyngeal aspirates cryopreserved with 20% glycerol though their study design included combinations of different host depletion and extraction protocols (including spin column-based protocols, which led to high sample loss) thus limiting interpretability. QIAamp-based host depletion was extracted with a

spin column leading to insufficient nucleic acids for sequencing. They found that MolYsis performed the best in combination with an extraction protocol that did not use spin columns.

The optimal approach to host depletion depends on multiple factors, which may be weighted differently based on the investigator's priorities. These factors include the efficiency of host depletion (e.g., the greatest reduction in host DNA), potential bias (where some microbes such as gram-negative bacteria are more vulnerable to lysis than gram-positive bacteria), sequencing failure rates, as well as practical considerations such as workload (some host depletion methods, such as lyPMA, are long protocols that make this approach low throughput) and cost (QIAamp host depletion reagents at this time cannot be purchased separate from the larger microbiome nucleic acid extraction kit). In general, to avoid complete confounding of sample type with the potential bias introduced by specific host depletion treatments, it is best to apply the same approach to all sample types within a study. At this time, our results suggest QIAamp may be a good choice for diverse types of respiratory samples as it has effective host depletion efficiency for all sample types tested, has comparably less potential to introduce bias, particularly for easy-to-lyse gram-negative bacteria based on our viability studies, and is among the shortest protocols tested. The addition of a cryoprotectant such as glycerol at the time of freezing may mitigate the adverse effects of freezing on microbial cell viability and limit potential bias during host depletion. More broadly, we observed that the increase of species richness is saturated at 0.5–2 million microbial mapped reads. This finding, along with observed % host DNA for different respiratory sample types, offers a preliminary guide for targeting cost-efficient metagenomics sequencing respiratory samples. Our results show that for many respiratory samples, studies performing metagenomics without host depletion will not adequately describe the microbial communities and even if there is the potential for some bias introduced with host depletion, it is preferable to having the extremely low effective sequencing depth induced by high host DNA content.

Host depletion substantially reduced the proportion of gram-negative bacteria in sputum samples, which were obtained from persons with cystic fibrosis (PwCF); this was not noted in other sample types (nasal swabs and bronchoalveolar lavage). In some chronic lung diseases such as cystic fibrosis, higher levels of extracellular DNA are found in sputum compared to BAL, suggesting greater airway inflammation (including potential role of NET formation) in proximal airways compared to distal airways[23,25]. NETosis, the process of forming neutrophil extracellular traps, may increase the proportion of human DNA in sputum samples, and affect both the host depletion efficiency and microbial community composition. Our results show that certain gram-negative bacterial species in particular are capable of producing extracellular DNA; an important caveat is that certain stimuli, such as the presence of pyocyanin, may increase extracellular DNA production by *Pseudomonas aeruginosa*[32]. Further studies are needed to better understand the relative contribution of microbial cell lysis during host depletion versus the presence of extracellular microbial DNA to the changes in microbial community composition observed after host depletion.

Many ongoing large epidemiological studies have not added cryoprotectants prior to freezing respiratory specimens, nor is host depletion on freshly collected samples before freezing logistically possible as it requires additional trained personnel, equipment, and processing time compared to standard biobanking. Our study focused on non-cryoprotected frozen samples along the upper and lower respiratory tract and demonstrated that effective host depletion method can be performed. Our viability studies indicate that for non-cryoprotected frozen samples, QIAamp treatment had the lowest impact on the differential viability of gram-negative compared to gram-positive bacteria; the addition of a cryoprotectant may largely mitigate the effect of freezing on cell lysis, although it did not significantly change the effect of QIAamp-based host depletion on frozen non-cryo-preserved isolates.

Our study has several strengths. To our knowledge, there are no other methods papers evaluating the effect of host DNA depletion on viral as well as non-viral microbial communities in samples obtained along the respiratory tract. We showed that optimal methods based on one sample type cannot necessarily be extrapolated to another sample type. We focused on samples frozen without cryoprotectants, which is more generalizable to most respiratory sample collection methods for existing clinical studies. We performed deep metagenomics sequencing, 76.4 million reads per sample, which is approximately twice that of existing respiratory metagenomics studies. We show that even at this depth, without host depletion, there is inadequate characterization of respiratory microbial communities. We used mediation analysis to show that the deeper effective sequencing depth resulting from host depletion explains the majority of effects of host depletion on increased species richness. We performed careful sensitivity analyses to evaluate the potential contribution of contamination and show that even after the removal of potential contaminants, host depletion methods increased species richness. Viability studies also better characterize the potential bias introduced by the effect of various treatments including freezing, addition of DNA/RNA shield, and each host depletion treatment.

Nevertheless, there are some limitations to our study. While the respiratory samples tested were preserved "neat" (i.e., without cryoprotectants or nucleic acid stabilization solutions), our mock community was preserved in DNA/RNA shield, which contains a mild detergent that lyses microbial cells. Our viability studies indicate that DNA/RNA shield instantly killed all tested bacterial species. The manufacturer (Zymo) does not recommend host depletion on samples preserved in DNA/RNA shield for this reason, however, we chose this mock community to test because many ongoing respiratory microbiome studies, particularly since the COVID-19 pandemic, collected respiratory samples in additives such DNA/RNA shield for infection control reasons. Thus we felt it was important to demonstrate the degree to which samples collected in this fashion would bias sequencing results for all microbial species, as DNA/RNA shield lyses all microbial cells. To better assess for bias, untreated BAL, nasal, and sputum samples could have been sequenced much deeper than host-depleted samples to achieve the same effective sequencing depth after host read removal. Finally, the Marker-MAGu reference database is compiled from studies focused on the human gut microbiome, in part due to the limited number of mNGS studies performed in respiratory samples (most respiratory microbiome studies rely on amplicon sequencing). It is possible that novel phages resident to the respiratory tract do not yet exist in these databases. Future mNGS studies on the respiratory tract may address this challenge.

In summary, we show that host depletion treatment enables the characterization of the respiratory microbiome with mNGS, even in previously frozen samples. While some host depletion methods may shift resulting microbial composition, metagenomics sequencing without host depletion will severely underestimate microbial diversity of most respiratory samples due to shallow effective sequencing depth and is not recommended.

Methods
Sample collection
Anterior nasal swab samples were obtained from healthy adults according to a standardized protocol as previously described in an earlier study [33]. PBS was added (1 mL) to the nasal swab samples and vortexed briefly. Four aliquots were made with one nasal swab sample, and a swab and 200 μL of sample solution were utilized for each aliquot. Sputum was collected from adult patients with cystic fibrosis described in a previous study [34,35]. Briefly, adult persons over age 18 satisfying cystic fibrosis clinical diagnostic criteria and receiving routine care at the Massachusetts General Hospital Adult Cystic Fibrosis Center were recruited. The volume of sputum samples was supplemented with PBS to make 1 mL of 6 aliquots and gently homogenized by syringes to make evenly distributed aliquots. Bronchoalveolar lavage (BAL) fluid was collected from intubated patients for clinically indicated bronchoscopies with excess BAL. For BAL, each sample was evenly separated into 6 individual aliquots to have a volume of 200 μL. Ethical approval for this study was obtained by the Institutional Review Board of Mass General Brigham (Protocol #2018P002934, 2019P002868 and 2020P001761). All the samples described in this study were frozen without cryoprotectants within 1 h of sample collection and stored at −80 °C. None

of the samples were stored for more than 27 months before host depletion and DNA extraction.

Host DNA depletion treatments

Five different host DNA depletion methods (lyPMA, Benzonase, HostZERO, MolYSIS, and QIAamp) were tested for nasal swab, sputum, and BAL samples. The total number of treatment groups per sample was 6 (5 different treatments and 1 control group). Nasal swab samples were collected from 10 different subjects as it is not feasible to collect more than 4 nasal swabs per participant at any given time, thus the experimental design was modified to ensure an equal number of replicates for each host depletion treatment group resulting in a larger number of control samples for nasal swabs.

For lyPMA, the procedure followed a previously published protocol by Marotz et al.[11]. Briefly, samples were centrifuged to collect cells ($10,000 \times g$, 8 min). After carefully discarding the supernatant, the pellets were resuspended with 200 μL of nuclease-free water (129114, Qiagen, Germany) and mixed by Voltex-Gini2. Samples were left at room temperature for 5 min and 5 μL of PMA (40019, Biotium, USA) was added to the sample (10 μL of 1 mM PMA). After briefly vortexing, samples were incubated in the dark room at room temp (5 min). To bind PMA dyes to DNA, samples were placed horizontally on an orbital shaker and exposed to a light source with 2610 lumens (LED A21, GE, USA) at a 20 cm distance for 30 min, and rotated every 5 min.

Benzonase treatment method was conducted as described by Nelson et al.[17]. First, 7 mL of DNAse-free water was added to 200 μL samples, and then the samples were placed on an orbital shaker for 1 h at 60 RPM to lyse mammalian cells. 800 μL of 10x Benzonase buffer (200 mM Tris-HCl (15567027, Invitrogen, USA), 10 mM MgCl$_2$ (AM9530G, Invitrogen, USA)) and 250U of Benzonase (E1014-25KU, Sigma, USA) to each sample (1 μL) was added to the samples, and the mixtures were incubated for 2 h at 37 °C (120 rpm) in an incubator (New Brunswick Innova 42, Eppendorf, Germany). After centrifuging at $8000 \times g$ for 10 min, the pellets were resuspended with 1 mL PBS and moved to 1 mL tubes. The second centrifuge was conducted at $13,000 \times g$ for 3 min, the supernatants were removed, and the pellets were resuspended with 400 μL of TE (AM9849, Invitrogen, USA) + 5 mM EDTA (15575-020, Invitrogen, USA).

HostZERO was implemented according to the manufacturer's protocol (https://files.zymoresearch.com/protocols/d4310_hostzero_microbial_dna_kit.pdf). Briefly, 1 mL of host DNA depletion solution (D4310-1-20, Zymo, USA) was added per 200 μL of sample. The mixture was agitated by orbital shaking for 15 min at room temp at 180 rpm. After centrifuging the tube at $10,000 \times g$ for 5 min at room temperature, the supernatant was carefully removed. Then, 100 μL of microbial selection buffer (D4310-2-5, Zymo, USA) and 1 μL of microbial selection enzyme (D-4310-3-50, Zymo, USA) were added to the samples, and the samples were incubated at 600 rpm, 37 °C for 30 min in a thermomixer.

MolYsis Basic 5 (D301-050, Molzyme, Germany) was implemented following the manufacturer's protocol. Briefly, 250 μL buffer CM was added to the samples, and they were agitated by vortexing for 15 s and incubated at room temperature for 5 min. Reagents (250 μL buffer DB1, 10 μL MolDNase B) were added to the samples and briefly mixed by vortexing for 15 s. After an incubation process at room temp for 15 min, samples were centrifuged at $12,000 \times g$ for 10 min, and the supernatant was removed carefully. The pellet was resuspended with 1 mL buffer RS and centrifuged at $12,000 \times g$ for 5 min. Finally, 80 μL buffer RL was added to the pellets and mixed with pellets by Vortex Gini.

For QIAamp, the procedure followed the manufacturer's protocol (https://www.qiagen.com/us/resources/resourcedetail?id=c403392b-0706-45ac-aa2e-4a75acd21006&lang=en). Briefly, after adding 800 μL of PBS (MRGF-6230, Growcells, USA) to each sample to make the total reaction volume 1 mL, 500 μL Buffer AHL (1080302, Qiagen, Germany) was added to 1 mL of sample. Samples were incubated at room temperature for 30 min at 600 rpm. The pellet was collected by centrifuging the tube at $10,000 \times g$ for 10 min and removing the supernatant carefully. After adding 190 μL of Buffer RDD (1018702, Qiagen, Germany) and 2.5 μL of Benzonase (1038893, Qiagen, Germany), the samples were incubated at 37 °C for 30 min at 600 rpm.

20 μL of Proteinase K was added, and samples were incubated at 56 °C for 30 min at 600 rpm afterwards. All incubation processes were conducted with a thermomixer (Thermomixer C 5382, Eppendorf, Germany).

DNA extraction

The same nucleic acid extraction approach was applied to all sample types as previously described[10]. In brief, treated and untreated samples, reagent-only negative controls, and mock community positive controls (Zymo Research) were extracted using a protocol optimized for respiratory samples with a magnetic bead-based protocol using the Maxwell HT 96 gDNA Blood Isolation System (Promega) on a KingFisher Flex instrument. Briefly, cetyltrimethylammonium bromide (CTAB) is added to samples in individual Lysing Matrix E tubes (MP Biomedicals), incubated at 95 °C for 5 min followed by bead beating for three 30-s cycles at 7.0 m/s, incubated with proteinase K at 70 °C for 10 min, 300 sample μL lysate collected, additional bead beating for three 30 s cycles at 7.0 m/s with each cycle, and additional 300 sample μL lysate collected. Sample lysates are transferred to 96-well plates for binding, washing, and elution steps on the Kingfisher Flex sample purification system.

Quantitative polymerase reaction (qPCR)

Quantification of human DNA was determined by focusing on the LINE-1 region with the Femto human DNA quantification kit (Zymo E2005, USA) with standards. Bacterial DNAs were measured by targeting the 16S rRNA region with a set of universal primers (5′-CCTACGGGAGGCAGCAG-3′ and 5′-ATTACCGCGGCTGCTGG-3′) for bacterial 16S rRNA[20] and bacterial DNA standards (Zymo E2006-2, USA) for quantification. All reactions were performed in triplicate. Absolute quantification was determined using standard curves generated according to the manufacturer's protocol (https://files.zymoresearch.com/protocols/_e2006_femto_bacterial_dna_quantification_kit_ver.pdf and https://files.zymoresearch.com/protocols/_e2005_femto_human_dna_quantification_kit.pdf).

Metagenomics sequencing and data processing

PicoGreen dsDNA assay kits were used for DNA concentration measurement at library preparation (P7589, Invitrogen, USA). Due to low microbial DNA content, a DNA library prep kit (E6177L, New England Biolabs, USA) designed for low input (100–500 ng) was used. For all sample types, 1:25 diluted adapter was used during adapter ligation, and 12 cycles of PCR amplification were conducted. Success of library preparation was assessed with fragment analyzer (DNF-474-0500, Agilent, USA) and qubit (Invitrogen, USA). In total, 157 samples (30 BAL, 35 nasal, 30 sputum, 30 negative control from host depletion, 30 positive controls from host depletion, 1 extraction positive control, and 1 extraction negative control) were sequenced on the Illumina NovaSeq platform targeting 10 Gb/sample. Reads were processed with Casava (Illumina) and bbduk to retrieve sequences and remove Illumina adapters.

Profiling of metagenomes was processed with bioBakery 3[8] combined with bowtie2 with hg38 reference database for mapping and removing human reads[36]. Specifically, MetaPhlAn 3.0 and HUMAnN 3 were employed for taxonomical and functional profiling, respectively. Viral taxa and phage-bacteria dynamics were profiled with Marker-MAGu[37] with v1.1 database (https://github.com/cmmr/Marker-MAGu). Marker-MAGu was seen as ideal for profiling viruses in mNGS data as it uses a marker gene approach similar to MetaPhlAn. Community profiles, either microbial taxonomy or predicted function of genes, and their hierarchical structures were merged by 'phyloseq' package v1.41.1[38]. Outputs were normalized to relative abundances considering the length of core genomes used for the identification for taxa, and counts per million bases (CPM) for function, respectively. Proportions of host DNAs in a sample were calculated using both qPCR and mNGS results using the following equations.

$$\text{Host DNA}(\text{qPCR}) \, [\%] = H_q/(B_q + H_q) \qquad (1)$$

$$\text{Host DNA}(\text{sequencing}) \, [\%] = R_H/(R_C) \qquad (2)$$

Where H_q is the absolute amount of host DNA quantified by qPCR with LINE-1 region, B_q is the amount bacterial DNA quantified by qPCR with 16S rRNA region, R_H is the host reads identified by bowtie2 among R_C, and the R_C is the cleaned reads after removing low quality reads. Furthermore, low prevalent taxa were removed at a 5% threshold for statistical analyses, to avoid resulting in wrong association, i.e., detecting more taxa after host DNA depletion[39].

Microbial viability tests

From the CF sputum samples, bacterial colonies were isolated and taxonomy confirmed by MALDI-TOF. The viability of gram-negative (5 isolates of *Achromobacter* spp. and 6 isolates of *Pseudomonas aeruginosa*) and gram-positive (7 isolates of gram-positive *Staphylococcus aureus)* species were tested across a number of different treatment conditions to evaluate the effect of freezing without and with a cryoprotectant, addition of Zymo, DNA/RNA shield, and host depletion. Before treatment, the isolates were cultivated overnight (12–16 h) in tryptic soy broth (TSB, 211825, BD, USA) media at 37 °C with shaking at 200 rpm (New Brunswick Innova 42/42, Eppendorf, Germany).

The overnight culture broth was subsequently aliquoted, and the viability of each isolate was tested under 12 different experimental conditions by colony-forming unit (CFU) testing. The experimental groups were as follows: unfrozen, unfrozen with DNA/RNA shield, frozen, frozen with 20% glycerol, frozen lyPMA, frozen glycerol lyPMA, frozen Benzonase, frozen glycerol Benzonase, frozen HostZERO, frozen glycerol HostZERO, frozen QIAamp, frozen glycerol QIAamp, frozen Molysis, and frozen glycerol Molysis. Given the substantial negative impact of Molysis treatment on *Staphylococcus aureus* viability, this method was not carried forward in experiments using the gram-negative isolates. During host DNA depletion treatments, unfrozen control samples were kept in a 4 °C refrigerator to retard growth until plating. For the unfrozen with DNA/RNA shield condition, liquid culture aliquots were centrifuged at $13,000 \times g$ at 4 °C for 3 min and then the pellet resuspended with DNA/RNA shield (R1100-50, Zymo, USA). For all glycerol groups, sample aliquots were mixed with glycerol (356350, Millipore, USA) to reach a final glycerol concentration of 20% before freezing. For all frozen and glycerol frozen groups, 220 μL aliquots were frozen in a −80 °C freezer for minimum 1 h and then thawed at 4 °C before implementing host DNA depletion methods. All the host DNA depletion methods were carried out as described earlier.

After host DNA depletion treatment, bacterial pellets were resuspended in TSB, and CFU testing was conducted as described by Jean-Pierre et al.[40]. Briefly, all samples were serial-diluted up to a 10^7-fold dilution in 96-well plates. Using a 96-well replicator (140500, Boekel Scientific, USA), 1 μL of each sample was transferred onto tryptic soy agar (DF0369176, BD, USA) and incubated overnight at 37 °C. The number of colonies of each dilution series at the highest dilution factor with at least 3 colonies was counted and then the CFU counts were calculated using the following equation.

$$\text{Colony forming unit (CFU/mL)} = \frac{\text{Number of colonies}}{\text{Spot volume (μL)}} \times 1000 \left(\frac{\text{μL}}{\text{mL}}\right) \quad (3)$$
$$\times \text{Dilution factor}$$

Extracellular DNA quantification

Overnight liquid cultures were prepared by inoculating single colonies into tryptic soy broth. To determine the amount of microbial DNA in intact microbial cells versus extracellular DNA, 1 mL of each liquid culture was centrifuged at $10,000 \times g$ for 3 min. The supernatant was syringe-filtered through a sterile 0.22 μm filter, and a 200 μL aliquot representing extracellular DNA taken for nucleic acid extraction. The pellet was resuspended in 1 mL TSB, and a 200 μL aliquot representing intracellular DNA was taken for nucleic acid extraction. 16S rRNA qPCR was performed in triplicate, along with a standard curve to determine 16S rRNA gene copy number. The genome copy number of each species was calculated using Eqs. 4 and 5, and the relative contribution of intracellular versus extracellular DNA was

calculated using the subsequent equation.

$$16S \text{ rRNA copies} \left(\frac{\text{copies}}{\text{μL}}\right) =$$
$$\frac{\text{DNA estimate}\left(\frac{\text{ng}}{\text{μL}}\right) \times \text{ng per bp constant}\left(\frac{\text{bp}}{\text{ng}}\right)}{E. \, coli \text{ genome size} \left(\frac{\text{bp}}{\text{genome}}\right)} \times E. coli \text{ 16S rRNA(copies/genome)} \quad (4)$$

$$\text{Genome copy number} \left(\frac{\text{genome}}{\text{μL}}\right) =$$
$$\frac{16S \text{ rRNA copies} \left(\frac{16S \text{ rRNA copies}}{\text{μL}}\right)}{\text{Average 16S rRNA copies per genome} \left(\frac{16S \text{ rRNA copies}}{\text{genome}}\right)} \quad (5)$$

$$\text{Extracellular DNA (\%)} =$$
$$\frac{\text{Genome copies in supernatant} \left(\frac{\text{genome}}{\text{μL}}\right)}{\text{Genome copies in pellet} \left(\frac{\text{genome}}{\text{μL}}\right) + \text{Genome copies in supernatant} \left(\frac{\text{genome}}{\text{μL}}\right)} \quad (6)$$

Where the ng per bp constant was 9.26×10^{11} (bp/ng), the 16S rRNA copies of *E. coli* JM109 were 7 (copies/genome), the genome size was 4.6×10^6 (bp/genome), and the average 16S rRNA copies per genome for *Staphylococcus aureus*, *Pseudomonas aeruginosa*, and *Achromobacter* spp. were 5.7, 4.0, and 3.5, respectively[41].

Statistical analyses

Statistical analyses were pre-registered on the Open Science Foundation (https://osf.io/2jtc5)[42]. All statistical analyses were conducted in R version 4.3.1 (https://www.r-project.org). Library preparation success rates were assessed with a logistic mixed-effect model using the glmer function from 'lme4' R package v1.1.34[43]. Final reads, % host DNA, and other continuous outcomes were assessed by linear mixed-effect models using the lme4::lmer function. Predictors of the models were sample type, treatment method, and an interaction term for sample type x treatment method. Repeated measurements from one participant were accounted for with a random effect term. Alpha diversity indices were calculated with 'vegan' package v2.6.4[44]. Beta diversity was calculated with the vegan::vegdist, where Morisita-Horn dissimilarity index was calculated by subtracting the Morisita-Horn similarity index from 1, ordinated with the 'phyloseq', and paired distances extracted using 'harrietr' package v0.2.3 (https://github.com/andersgs/harrietr). Stratified analyses by sample type were conducted when the sample type * treatment method interaction term was significant. Predictors of overall microbial community structure were evaluated on beta diversity using permutational analysis of variance (PERMANOVA). All the predictors were tested as fixed effects to compare the effect size between them (feature ~ subject + sample type + treatment method + sample type * treatment method). After confirming sample-type-treatment specific effects with the sample type * treatment method interaction term, analyses stratified by sample type were used (feature ~ lyPMA + Benzonase + HostZERO + MolYsis + QIAamp, strata = subject id). To quantify the potential bias of treatment compared to controls, paired beta diversity indices between each treated and untreated sample were extracted and used for a subsequent linear mixed effects model.

The effect of treatment on alpha diversity and species-specific differential abundance was identified by linear mixed-effect model implemented in lme4::lmer (feature ~ sample type + lyPMA + Benzonase + HostZERO + MolYsis + QIAamp + (1|subject id)). Analyses performed for differential abundance first implemented a centered log-ratio transformation to account for compositionality prior to regression modeling using the 'microbiome' R package v1.22.0[45]. The false discovery rate (FDR) was calculated using the 'qvalue' R package v2.32.0[46].

Mediation analysis was conducted using the 'mediation' package v4.5.0[47] with species richness as outcomes, treatment method as exposures, final reads as mediators, and sample type as mediator-outcome confounders. Mixed effects linear regression was used for both outcome and mediator models, and the analysis was stratified by each treatment method.

All the data were visualized using R packages ggplot2 v3.4.4[48] and ggpubr v0.6.0 (https://github.com/kassambara/ggpubr/).

Sensitivity analysis

For further identification of contaminants, two different statistical decontamination methods were employed. First, the 'decontam' package[26] was used with the 'combined' method based on bacterial DNA from the qPCR result as total DNA concentration. The analysis employed negative controls, BAL, nasal, and sputum without negative controls after prevalence and abundance filtering for microbial datasets (at least 5% prevalent, except features with high abundance more than 0.75 quantile). No filtering was conducted for viral datasets. Second, we estimated decontaminated genus-level relative abundances using 'tinyvamp'[27] based on the MetaPhlAn read count table. The community compositions of the negative control and mock community samples were treated as known, and the relative abundance profiles were estimated for each of the remaining samples. A single contaminant relative abundance profile was estimated for each protocol. Detection efficiencies were estimated for taxa in the mock community relative to *E. faecalis*. The expected number of reads attributable to contamination was assumed to be inversely proportional to the bacterial DNA concentration in each sample (parameter 'Z_tilde'). Optimization of the unweighted Poisson criterion was performed until convergence was reached. Optimization was performed separately for each of the six protocols. Despite its presence in the mock community, *S. enterica* was not detected in any of the mock community samples in the HostZERO protocol. Therefore, to enable estimation, a pseudocount of 1 was imputed for a single positive control sample for this protocol only. For downstream analysis, we only considered the estimated relative abundances of the samples of unknown composition (output matrix P).

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All raw sequencing data are available under BioProject accession number PRJNA1019400 at the NCBI Sequencing Read Archive (SRA).

Code availability

Full documentation including data wrangling, exploratory data analyses, data processing, statistical modeling, and code for figure and table generation is available at a GitHub repository and archived via Zenodo (https://doi.org/10.5281/zenodo.14228803)[49].

References

1. Di Simone, S. K., Rudloff, I., Nold-Petry, C. A., Forster, S. C. & Nold, M. F. Understanding respiratory microbiome-immune system interactions in health and disease. *Sci. Transl. Med.* **15**, eabq5126 (2023).
2. Janda, J. M. & Abbott, S. L. 16S rRNA gene sequencing for bacterial identification in the diagnostic laboratory: pluses, perils, and pitfalls. *J. Clin. Microbiol.* **45**, 2761–2764 (2007).
3. Johnson, J. S. et al. Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. *Nat. Commun.* **10**, 5029 (2019).
4. Anyansi, C., Straub, T. J., Manson, A. L., Earl, A. M. & Abeel, T. Computational methods for strain-level microbial detection in colony and metagenome sequencing data. *Front. Microbiol.* **11**, 1925 (2020).
5. Beck, L. C. et al. Strain-specific impacts of probiotics are a significant driver of gut microbiome development in very preterm infants. *Nat. Microbiol.* **7**, 1525–1535 (2022).
6. Kieft, K. & Anantharaman, K. Deciphering active prophages from metagenomes. *mSystems* **7**, e00084-22 (2022).
7. Qin, J. et al. A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* **464**, 59–65 (2022).
8. Beghini, F. et al. Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with bioBakery 3. *eLife* **10**, e65088 (2021).
9. The Human Microbiome Project Consortium. Structure, function and diversity of the healthy human microbiome. *Nature* **486**, 207–214 (2012).
10. Sui, H. Y. et al. Impact of DNA extraction method on variation in human and built environment microbial community and functional profiles assessed by shotgun metagenomics sequencing. *Front. Microbiol.* **11**, 953 (2020).
11. Marotz, C. A. et al. Improving saliva shotgun metagenomics by chemical host DNA depletion. *Microbiome* **6**, 42 (2018).
12. Lande, R. Statistics and partitioning of species diversity, and similarity among multiple communities. *Oikos* **76**, 5–13 (1996).
13. Whelan, F. J. et al. Culture-enriched metagenomic sequencing enables in-depth profiling of the cystic fibrosis lung microbiota. *Nat. Microbiol.* **5**, 379–390 (2020).
14. Comstock, W. J. et al. The WinCF model—an inexpensive and tractable microcosm of a mucus plugged bronchiole to study the microbiology of lung infections. *J. Vis. Exp.* **123**, e55532 (2017).
15. Thoendel, M. et al. Comparison of microbial DNA enrichment tools for metagenomic whole genome sequencing. *J. Microbiol. Methods* **127**, 141–145 (2016).
16. Horz, H.-P., Scheer, S., Huenger, F., Vianna, M. E. & Conrads, G. Selective isolation of bacterial DNA from human clinical specimens. *J. Microbiol. Methods* **72**, 98–102 (2008).
17. Nelson, M. T. et al. Human and extracellular DNA depletion for metagenomic analysis of complex clinical infection samples yields optimized viable microbiome profiles. *Cell Rep.* **26**, 2227–2240.e5 (2019).
18. Amar, Y. et al. Pre-digest of unprotected DNA by Benzonase improves the representation of living skin bacteria and efficiently depletes host DNA. *Microbiome* **9**, 123 (2021).
19. Cheng, W. Y. et al. High sensitivity of shotgun metagenomic sequencing in colon tissue biopsy by host DNA depletion. *Genomics Proteomics Bioinformatics* **21**, 1195–1205 (2023).
20. Heravi, F. S., Zakrzewski, M., Vickery, K. & Hu, H. Host DNA depletion efficiency of microbiome DNA enrichment methods in infected tissue samples. *J. Microbiol. Methods* **170**, 105856 (2020).
21. Rajar, P. et al. Microbial DNA extraction of high-host content and low biomass samples: optimized protocol for nasopharynx metagenomic studies. *Front. Microbiol.* **13**, 1038120 (2022).
22. Shu, Z. et al. Cryopreservation of *Mycobacterium tuberculosis* complex cells. *J. Clin. Microbiol.* **50**, 3575–3580 (2012).
23. Keir, H. R. & Chalmers, J. D. Neutrophil extracellular traps in chronic lung disease: implications for pathogenesis and therapy. *Eur. Respir. Rev.* **31**, 210241 (2022).
24. Whitchurch, C. B., Tolker-Nielsen, T., Ragas, P. C. & Mattick, J. S. Extracellular DNA required for bacterial biofilm formation. *Science* **295**, 1487 (2002).
25. Papayannopoulos, V. Neutrophil extracellular traps in immunity and disease. *Nat. Rev. Immunol.* **18**, 134–147 (2018).
26. Davis, N. M., Proctor, D. M., Holmes, S. P., Relman, D. A. & Callahan, B. J. Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. *Microbiome* **6**, 226 (2018).
27. Clausen, D. S. & Willis, A. D. Modeling complex measurement error in microbiome experiments. Preprint at *arXiv* https://doi.org/10.48550/arXiv.2204.12733 (2022).
28. Nagy-Szakal, D. et al. Targeted hybridization capture of SARS-CoV-2 and metagenomics enables genetic variant discovery and nasal microbiome insights. *Microbiol. Spectr.* **9**, e0019721 (2021).
29. Pereira-Marques, J. et al. Impact of host DNA and sequencing depth on the taxonomic resolution of whole metagenome sequencing for microbiome analysis. *Front. Microbiol.* **10**, 1277 (2019).

30. Lim, Y. W. et al. Metagenomics and metatranscriptomics: windows on CF-associated viral and microbial communities. *J. Cyst. Fibros.* **12**, 154–164 (2013).

31. Charlson, E. S. et al. Topographical continuity of bacterial populations in the healthy human respiratory tract. *Am. J. Respir. Crit. Care Med.* **184**, 957–963 (2011).

32. Das, T. & Manefield, M. Pyocyanin promotes extracellular DNA release in *Pseudomonas aeruginosa*. *PLoS ONE* **7**, e46718 (2012).

33. Lai, P. S. et al. Alternate methods of nasal epithelial cell sampling for airway genomic studies. *J. Allergy Clin. Immunol.* **136**, 1120–1123.e4 (2015).

34. Vieira, J. et al. Design and development of a model to study the effect of supplemental oxygen on the cystic fibrosis airway microbiome. *J. Vis. Exp.* **174**, e62888 (2021).

35. Vieira, J. et al. Supplemental oxygen alters the airway microbiome in cystic fibrosis. *mSystems* **7**, e0036422 (2022).

36. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).

37. Tisza, M. et al. Phage-bacteria dynamics during the first years of life revealed by trans-kingdom marker gene analysis. Preprint at *bioRxiv* https://doi.org/10.1101/2023.09.28.559994 (2023).

38. McMurdie, P. J. & Holmes, S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS ONE* **8**, e61217 (2013).

39. Nearing, J. T. et al. Microbiome differential abundance methods produce different results across 38 datasets. *Nat. Commun.* **13**, 342 (2022).

40. Jean-Pierre, F. et al. Community composition shapes microbial-specific phenotypes in a cystic fibrosis polymicrobial model system. *eLife* **12**, e81604 (2023).

41. Stoddard, S. F., Smith, B. J., Hein, R., Roller, B. R. K. & Schmidt, T. M. rrnDB: improved tools for interpreting rRNA gene abundance in bacteria and archaea and a new foundation for future development. *Nucleic Acids Res.* **43**, D593–D598 (2015).

42. Lai, P. Benchmarking host-DNA depletion methods on frozen respiratory samples for metagenomic whole genome sequencing. OSF https://doi.org/10.17605/OSF.IO/2JTC5 (2023).

43. Bates, D., Mächler, M., Bolker, B. & Walker, S. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67**, 1–48 (2015).

44. Dixon, P. VEGAN, a package of R functions for community ecology. *J. Veg. Sci.* **14**, 927–930 (2003).

45. Leo Lahti, S. S. microbiome. Bioconductor https://doi.org/10.18129/b9.bioc.microbiome (2012).

46. Storey, J. D., Bass, A. J. & Robinson, D. qvalue: Q-value estimation for false discovery rate control. Bioconductor https://doi.org/10.18129/b9.bioc.qvalue (2023).

47. Tingley, D., Yamamoto, T., Hirose, K., Keele, L. & Imai, K. mediation: *R* package for causal mediation analysis. *J. Stat. Softw.* **59**, 1–38 (2014).

48. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis* 2nd edn, 241–253 (Springer, 2016).

49. Kim, M. Non-viral and viral microbial communities in host DNA depleted respiratory samples. Zenodo https://doi.org/10.5281/zenodo.14228803 (2024).

Acknowledgements

Author contributions

P.S.L. conceived of the study. P.S.L. and M.K. designed the study. M.K. wrote the main manuscript text and prepared all figures and tables. R.C.P. and V.S.S. helped with respiratory sample collection and manuscript revision. T.T. performed the viability studies. M.R. and J.C. conducted library preparation and sequencing. A.B. helped visualization of data. M.J.T., C.-Y.H., L.B., I.N., K.W., and J.K.H. helped run analyses and revise the manuscript. A.D.W. ran decontamination analysis and revised the manuscript. P.S.L. acquired funding, administrated the whole project, and revised the figures, tables, and manuscript. Artificial intelligence-assisted technology was not used in the production of the submitted work.

Competing interests

The authors declare no competing interests.

Ethical approval

All ethical regulations relevant to human research participants were followed. Ethical approval for this study was obtained by the Institutional Review Board of Mass General Brigham (Protocol #2018P002934, 2019P002868 and 2020P001761).

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s42003-024-07290-3.

Correspondence and requests for materials should be addressed to Peggy S. Lai.

Peer review information *Communications Biology* thanks the anonymous reviewers for their contribution to the peer review of this work. Primary handling editor: Tobias Goris.

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.