

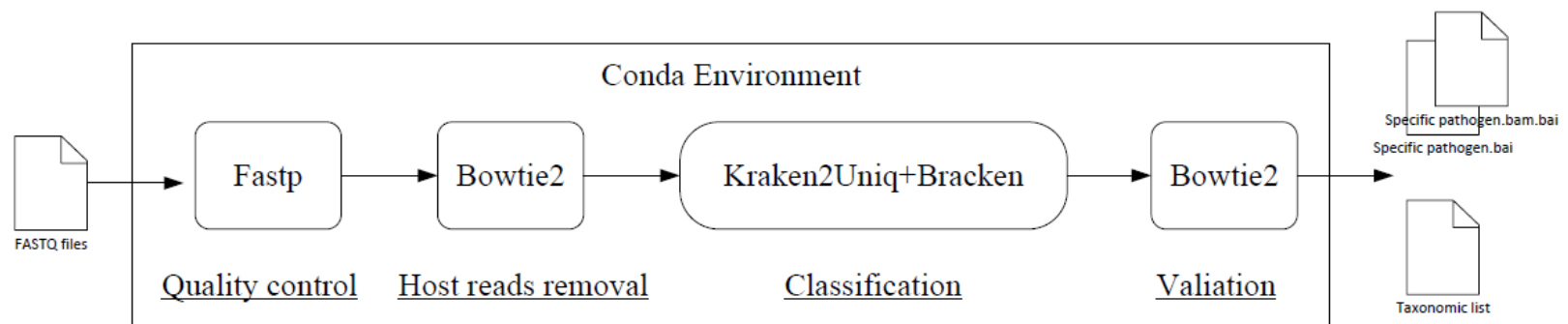


Fig S1 The analysis workflow for pathogen detection based on sequencing data.

pNGS v0.1.3

Summary Dataset

Analyze Dataset

Test report

Help

Experimental data

Flowcell ID

Sample ID

Nucleic Acid Concentration (ng/ul)

Adaptor

Library Concentration (ng/ul)

Sequencing File

Add

Edit

Delete

NGS Info

Chip_id

Sample_id

Extracted_NAC

Adaptor

Library_NAC

Rawfasta_id

Search

No data available in table

Showing 0 in 0 of 0 entries

Previous

Next

Confirm and Analyze

Processing Log

waiting for data analysis

Fig S2 Updating the information of each batch including flow cell id, sample id, nucleic acids concentration after extraction and library preparation steps, adaptor id, and file name of sequencing data. Then start the process step by click the process button.

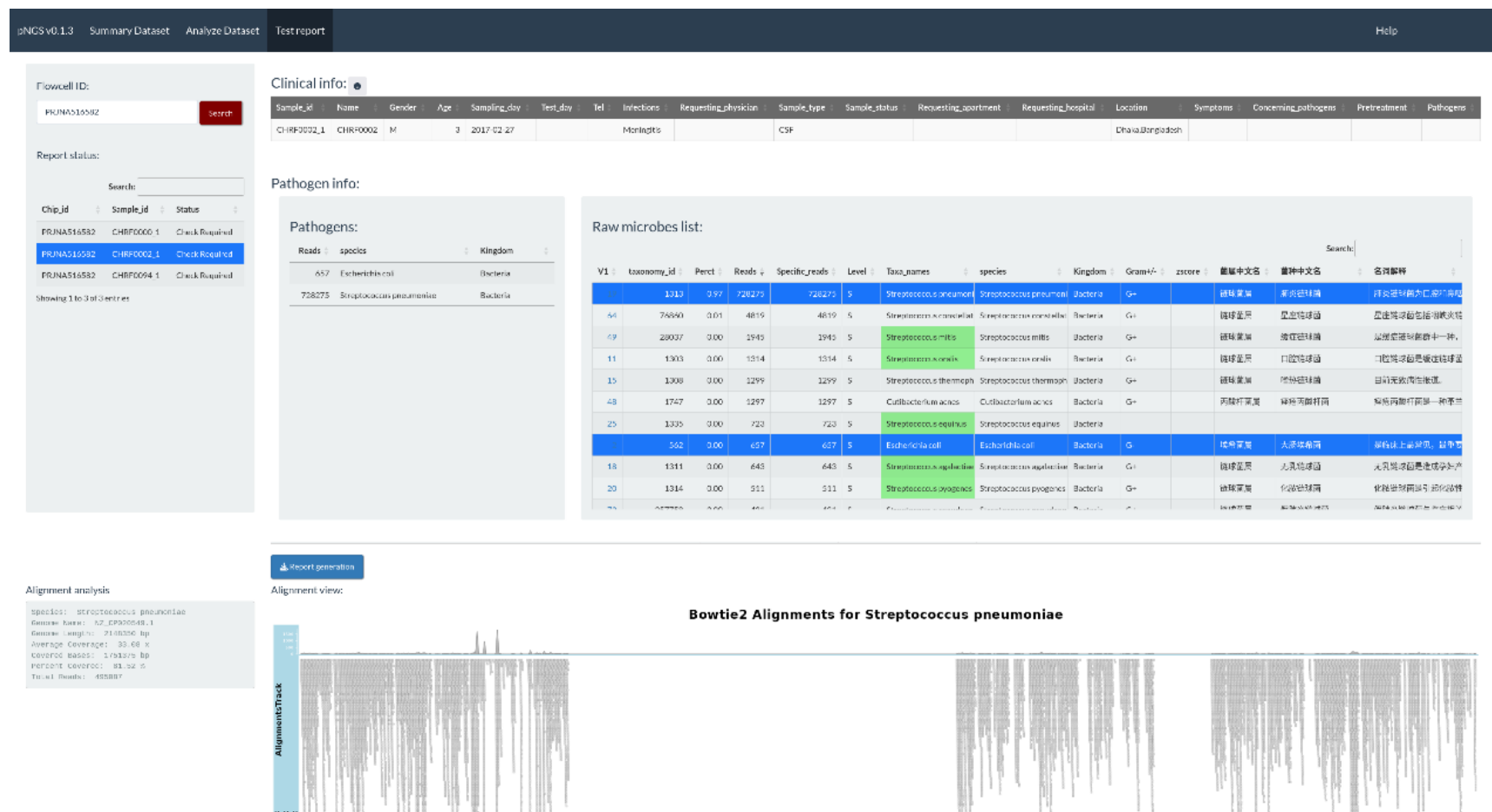


Fig S3 clinician or certified user, manually decide the pathogen for specific patient from microbes list based on multiple factors such as z-scores, mapped reads, average coverage for a specific pathogen, gender, age, sample types, clinical concerned pathogens types, anti-infection treatment. an mNGS detection report was generated with one click.

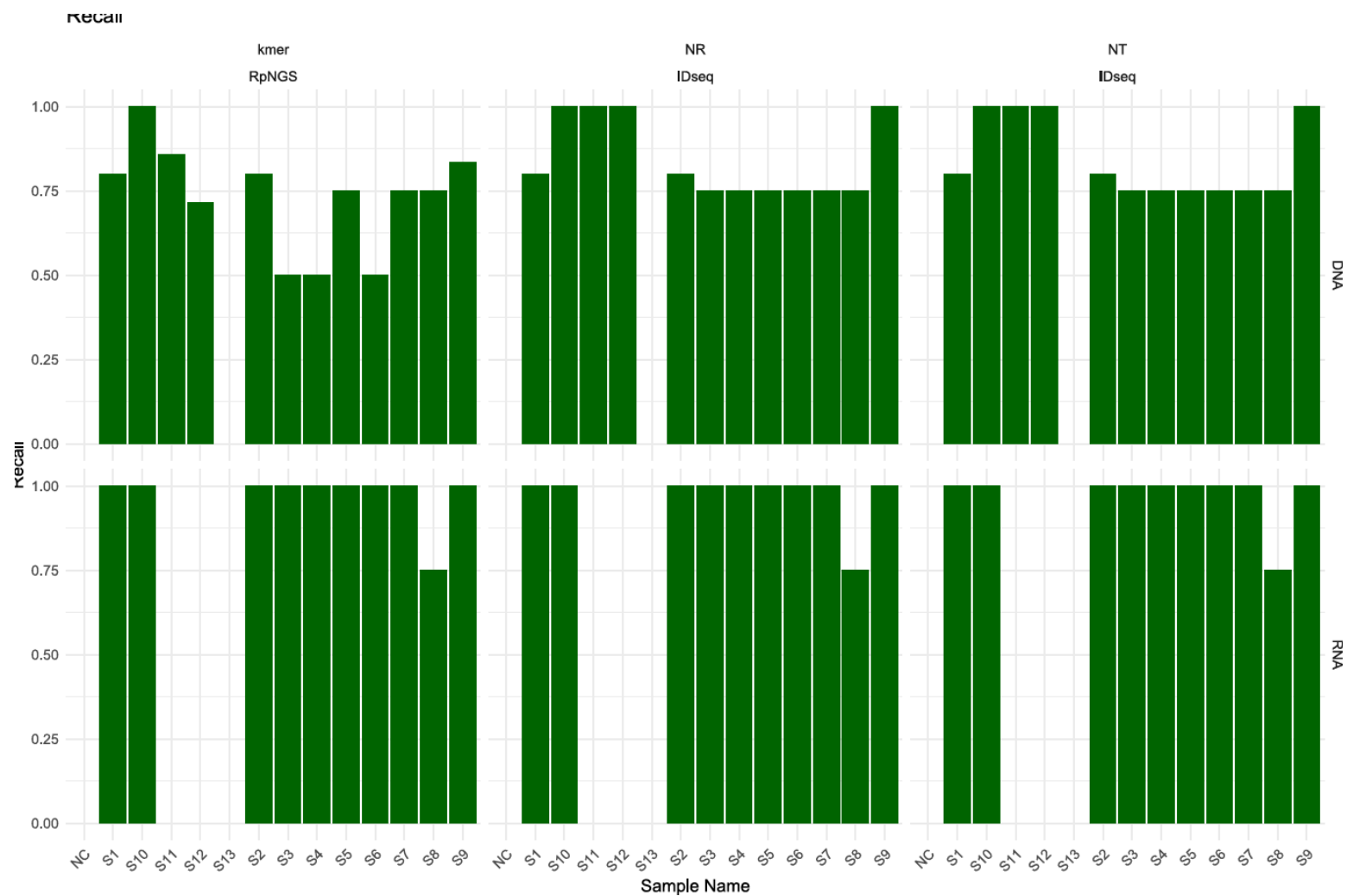


Fig S4 The Recall of IDseq_nt, IDseq_nr and RpNGS for DNA and RNA sequencing datasets among 14 samples.

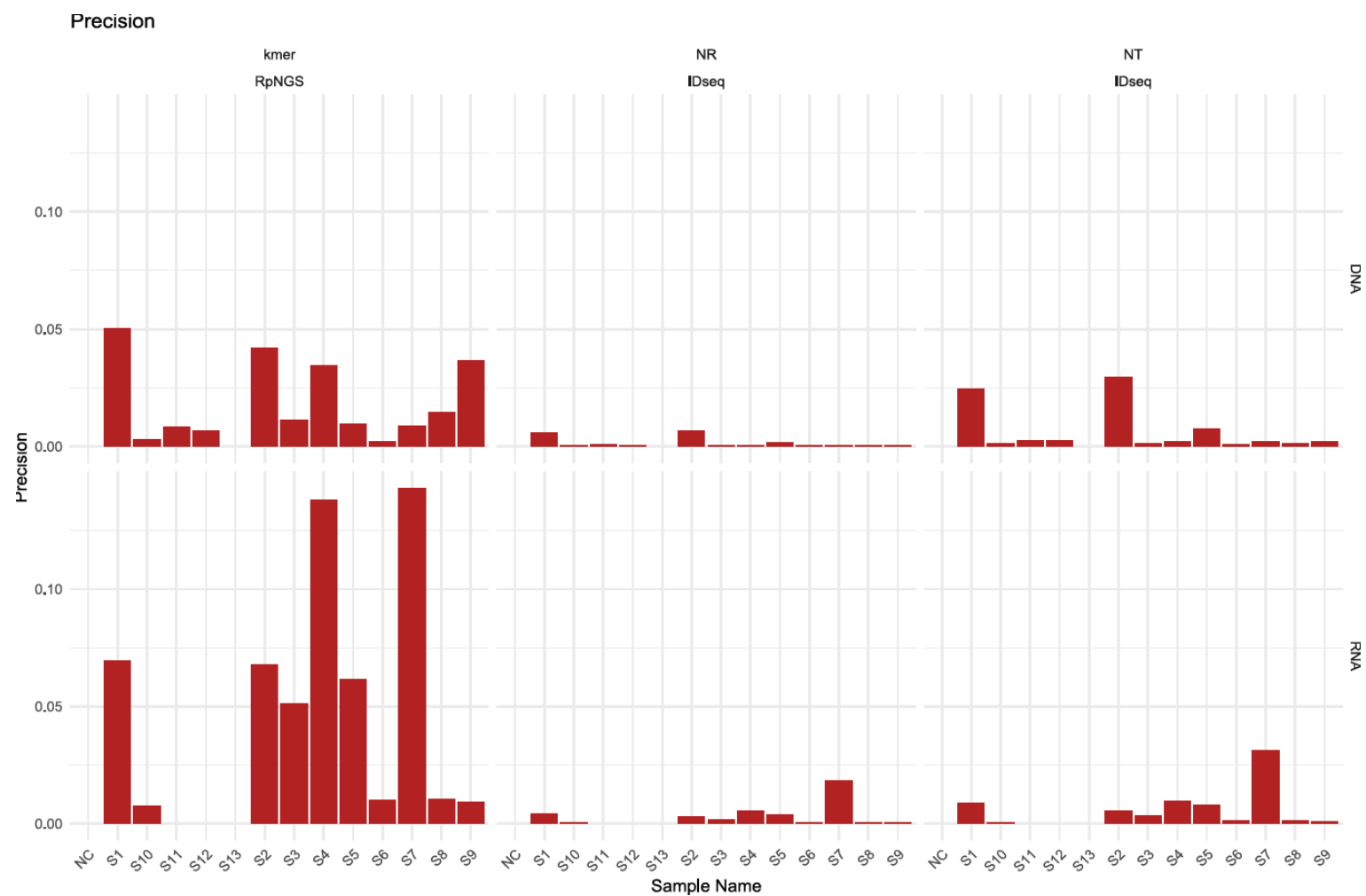


Fig S5 The precision of IDseq_nt, IDseq_nr and RpNGS for DNA and RNA sequencing datasets among 14 samples with z-score correction.

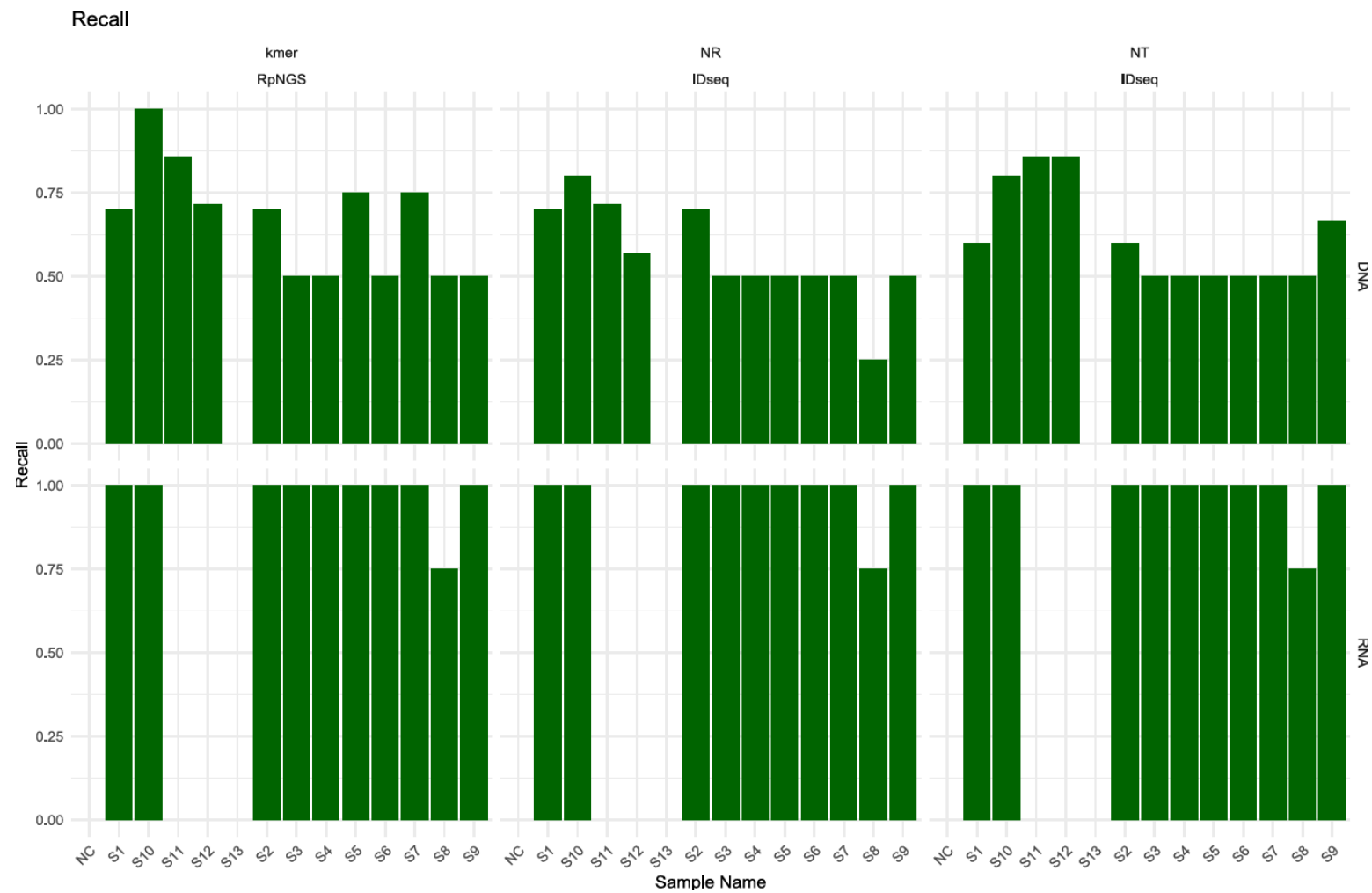


Fig S6 The Recall of IDseq_nt, IDseq_nr and RpNGS for DNA and RNA sequencing datasets among 14 samples with z-score correction.

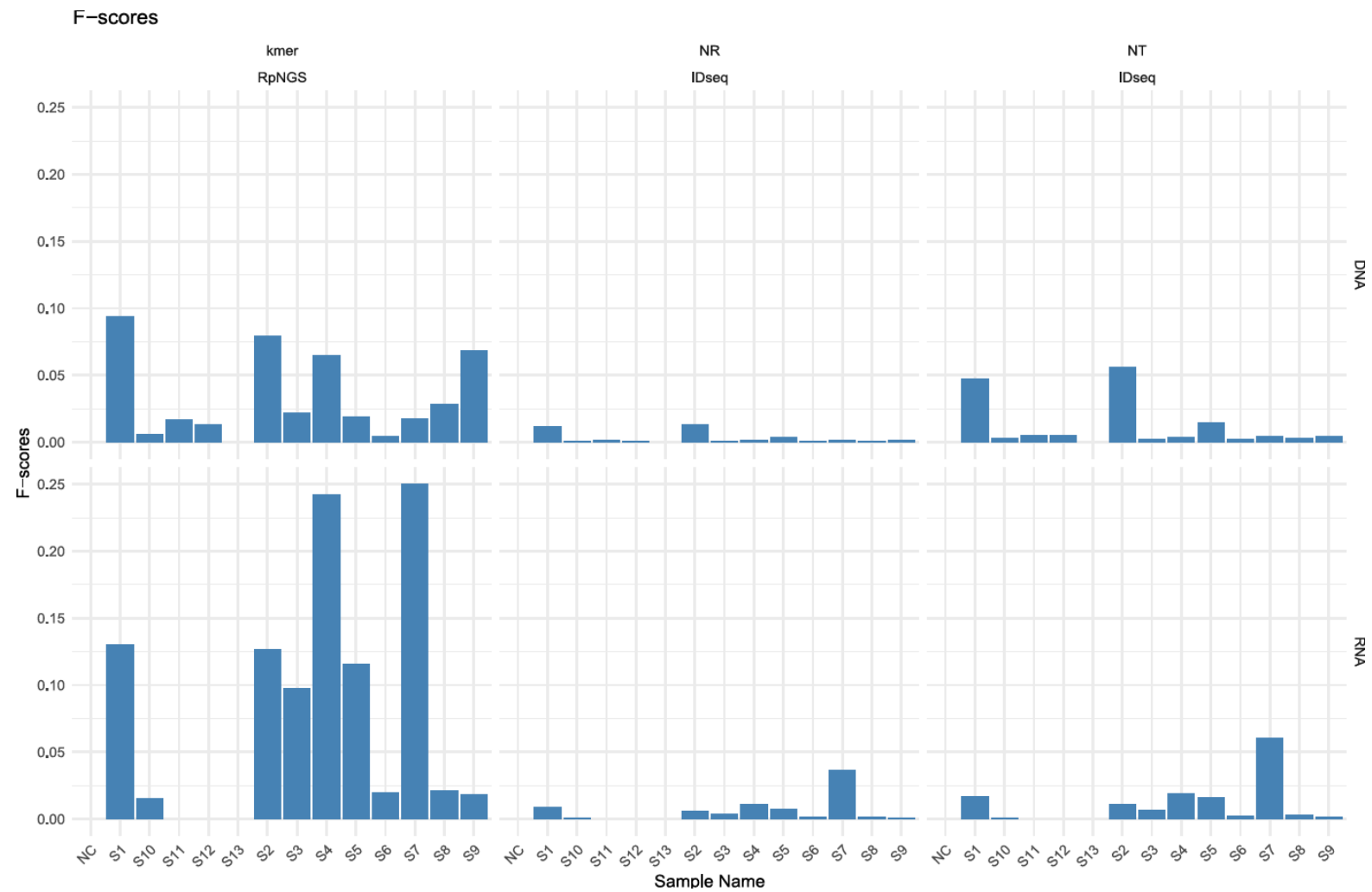


Fig S7 The F scores of IDseq_nt, IDseq_nr and RpNGS for DNA and RNA sequencing datasets among 14 samples with z-score correction.

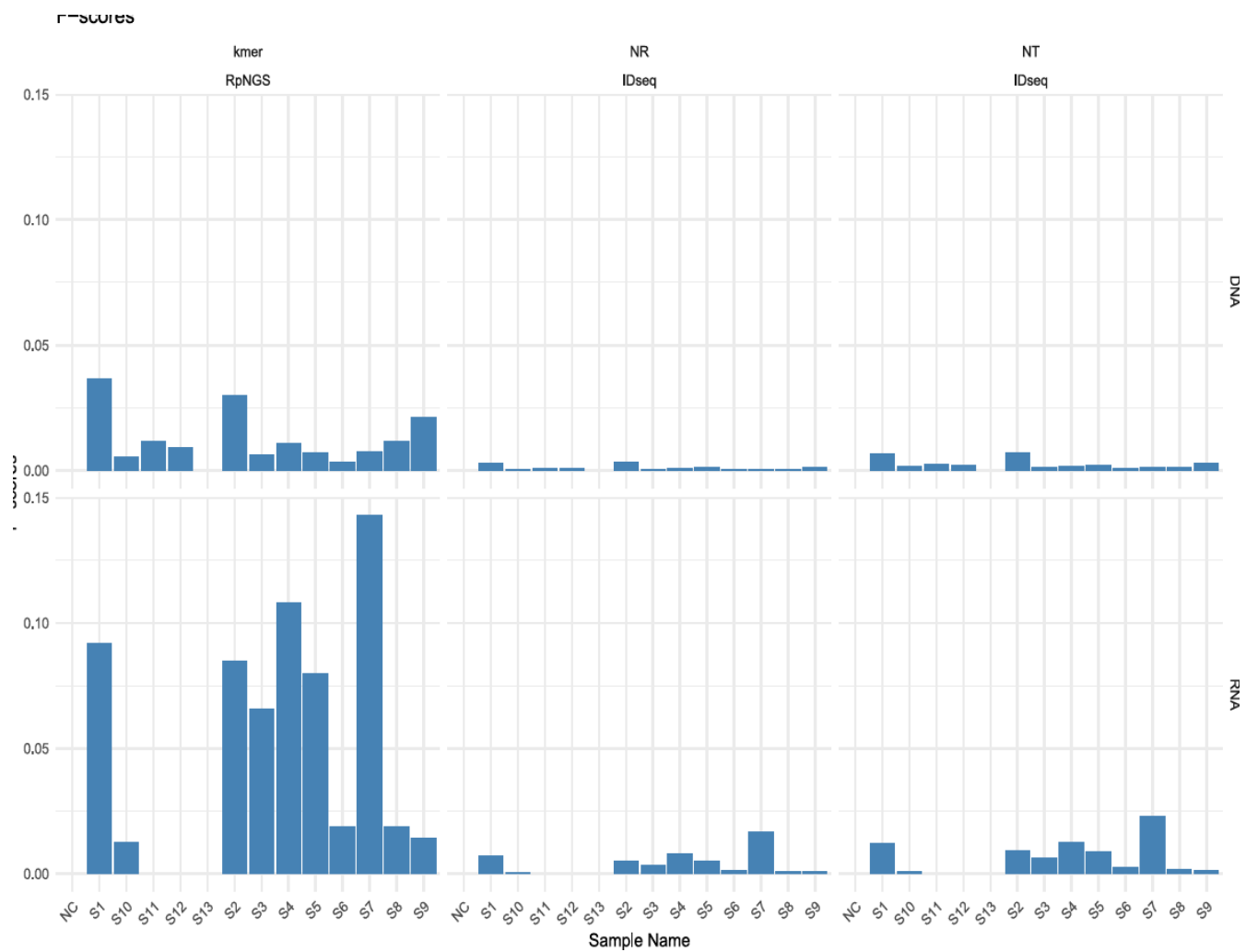


Fig S8 The F scores of IDseq_nt, IDseq_nr and RpNGS for DNA and RNA sequencing datasets among 14 samples without z-score correction.

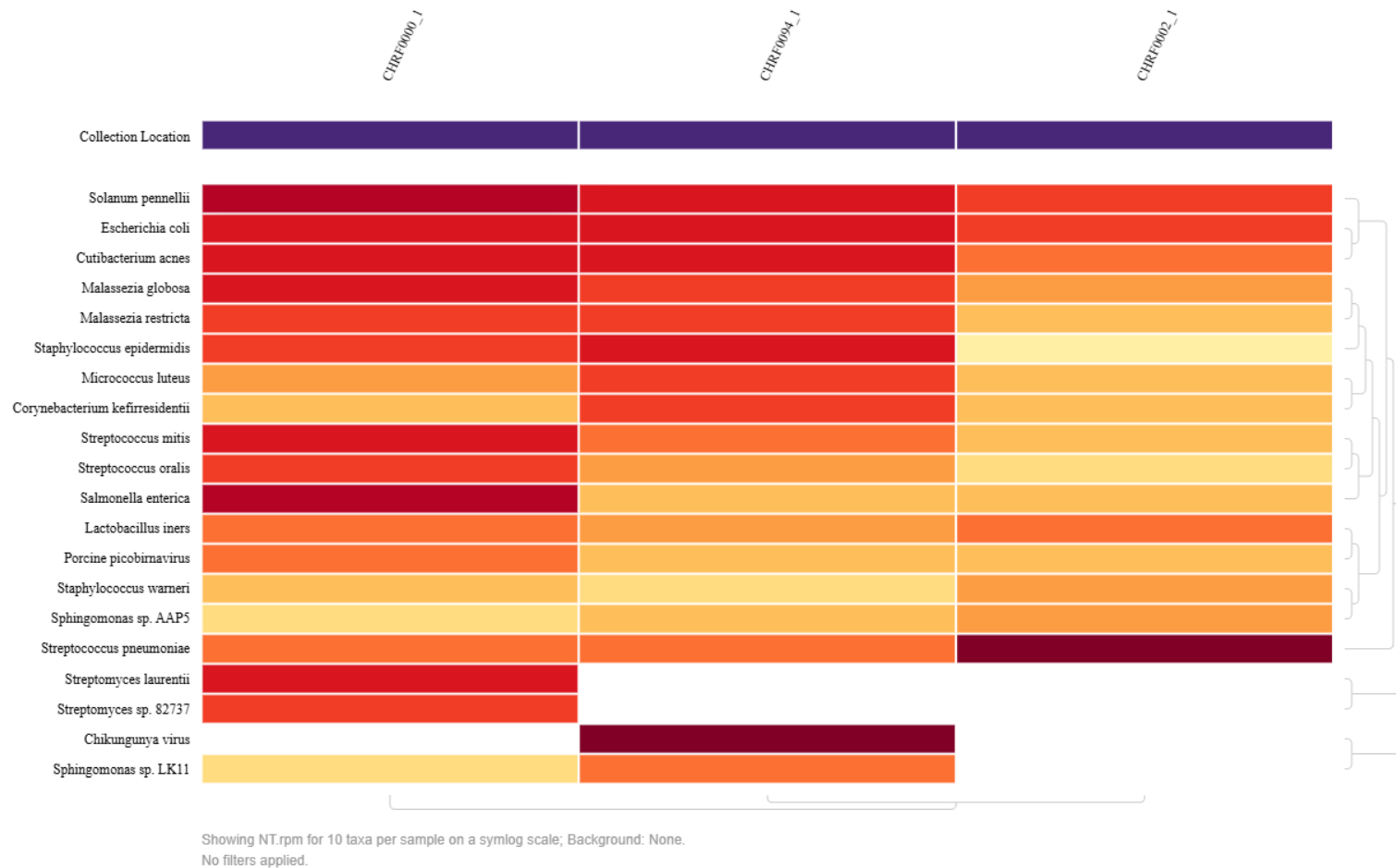


Fig. S9 Heatmap of species detected among these samples

Table S1 Basic characteristics of the microbes within the reference panel.

Microbes	Type	Genome Size (Mb)	GC %	Detailed information	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	NC
<i>Bacteroides fragilis</i>	Gram-bacteria	5.22248	43.4	ATCC 25285	1.0×10^2	1.0×10^2	\	\	\	\	\	\	\	1.0×10^2	1.0×10^2	1.0×10^2	\	\
<i>Betacoronavirus 1</i>	RNA virus	0.03088	37	Clinical strain	1.0×10^6	1.0×10^6	\	\	\	\	\	\	\	\	\	\	\	\
<i>Candida albicans</i>	Fungus	14.6999	33.6	ATCC 10231	2.0×10^4	2.0×10^4	1.0×10^2	1.0×10^3	1.0×10^4	1.0×10^3	1.0×10^3	1.0×10^3	\	\	\	\	\	\
<i>Cryptococcus neoformans</i>	Fungus	18.5736	48.2	ATCC 66031	\	\	\	\	\	\	\	\	\	\	\	2.0×10^4	\	\
<i>Enterococcus faecalis</i>	Gram+ bacteria	2.96381	37.4	Clinical strain	5.0×10^3	5.0×10^3	\	\	\	\	\	\	\	\	\	\	\	\
<i>Escherichia coli</i>	Gram-bacteria	5.11174	50.6	ATCC 25922	\	\	\	\	\	\	\	\	1.0×10^2	\	1.0×10^3	1.0×10^2	\	\
<i>Haemophilus influenzae</i>	Gram-bacteria	1.8477	38	ATCC 49247	1.0×10^5	1.0×10^5	\	\	\	\	\	\	\	\	1.0×10^2	1.0×10^4	\	\

[illegible]

<i>Negative BALF samples</i>	Host				\	\	\	\	\	\	\	2.0× 10 ⁶	\	\	\	\	\	\
--------------------------------------	------	--	--	--	---	---	---	---	---	---	---	-------------------------	---	---	---	---	---	---

Table S2 Fastq files of S1-S13 and 1 negative control generated by lab005.

Accession	Run title	Experiment accession	Experiment title	Sample Name	BioSample accession	Platform	Read length
HRR1209919	Lab005_NC_DNA	HRX955813	Lab005_NC_DNA.fq.gz	NC	HRS775964	Illumina Nextseq 550	75bp
HRR1209920	Lab005_NC_RNA	HRX955814	Lab005_NC_RNA.fq.gz	NC	HRS775964	Illumina Nextseq 550	75bp
HRR1209921	Lab005_S1_DNA	HRX955815	Lab005_S1_DNA.fq.gz	S1	HRS775951	Illumina Nextseq 550	75bp
HRR1209922	Lab005_S1_RNA	HRX955816	Lab005_S1_RNA.fq.gz	S1	HRS775951	Illumina Nextseq 550	75bp
HRR1209923	Lab005_S10_DNA	HRX955817	Lab005_S10_DNA.fq.gz	S10	HRS775960	Illumina Nextseq 550	75bp
HRR1209924	Lab005_S10_RNA	HRX955818	Lab005_S10_RNA.fq.gz	S10	HRS775960	Illumina Nextseq 550	75bp
HRR1209925	Lab005_S11_DNA	HRX955819	Lab005_S11_DNA.fq.gz	S11	HRS775961	Illumina Nextseq 550	75bp
HRR1209926	Lab005_S11_RNA	HRX955820	Lab005_S11_RNA.fq.gz	S11	HRS775961	Illumina Nextseq 550	75bp
HRR1209927	Lab005_S12_DNA	HRX955821	Lab005_S12_DNA.fq.gz	S12	HRS775962	Illumina Nextseq 550	75bp
HRR1209928	Lab005_S12_RNA	HRX955822	Lab005_S12_RNA.fq.gz	S12	HRS775962	Illumina Nextseq 550	75bp
HRR1209929	Lab005_S13_DNA	HRX955823	Lab005_S13_DNA.fq.gz	S13	HRS775963	Illumina Nextseq 550	75bp
HRR1209930	Lab005_S13_RNA	HRX955824	Lab005_S13_RNA.fq.gz	S13	HRS775963	Illumina Nextseq 550	75bp
HRR1209931	Lab005_S2_DNA	HRX955825	Lab005_S2_DNA.fq.gz	S2	HRS775952	Illumina Nextseq 550	75bp
HRR1209932	Lab005_S2_RNA	HRX955826	Lab005_S2_RNA.fq.gz	S2	HRS775952	Illumina Nextseq 550	75bp
HRR1209933	Lab005_S3_DNA	HRX955827	Lab005_S3_DNA.fq.gz	S3	HRS775953	Illumina Nextseq 550	75bp
HRR1209934	Lab005_S3_RNA	HRX955828	Lab005_S3_RNA.fq.gz	S3	HRS775953	Illumina Nextseq 550	75bp
HRR1209935	Lab005_S4_DNA	HRX955829	Lab005_S4_DNA.fq.gz	S4	HRS775954	Illumina Nextseq 550	75bp
HRR1209936	Lab005_S4_RNA	HRX955830	Lab005_S4_RNA.fq.gz	S4	HRS775954	Illumina Nextseq 550	75bp
HRR1209937	Lab005_S5_DNA	HRX955831	Lab005_S5_DNA.fq.gz	S5	HRS775955	Illumina Nextseq 550	75bp
HRR1209938	Lab005_S5_RNA	HRX955832	Lab005_S5_RNA.fq.gz	S5	HRS775955	Illumina Nextseq 550	75bp
HRR1209939	Lab005_S6_DNA	HRX955833	Lab005_S6_DNA.fq.gz	S6	HRS775956	Illumina Nextseq 550	75bp

HRR1209940	Lab005_S6_RNA	HRX955834	Lab005_S6_RNA.fq.gz	S6	HRS775956	Illumina Nextseq 550	75bp
HRR1209941	Lab005_S7_DNA	HRX955835	Lab005_S7_DNA.fq.gz	S7	HRS775957	Illumina Nextseq 550	75bp
HRR1209942	Lab005_S7_RNA	HRX955836	Lab005_S7_RNA.fq.gz	S7	HRS775957	Illumina Nextseq 550	75bp
HRR1209943	Lab005_S8_DNA	HRX955837	Lab005_S8_DNA.fq.gz	S8	HRS775958	Illumina Nextseq 550	75bp
HRR1209944	Lab005_S8_RNA	HRX955838	Lab005_S8_RNA.fq.gz	S8	HRS775958	Illumina Nextseq 550	75bp
HRR1209945	Lab005_S9_DNA	HRX955839	Lab005_S9_DNA.fq.gz	S9	HRS775959	Illumina Nextseq 550	75bp
HRR1209946	Lab005_S9_RNA	HRX955840	Lab005_S9_RNA.fq.gz	S9	HRS775959	Illumina Nextseq 550	75bp

Table S3 RpNGS did not detect certain pathogens

	Undetected pathogens	kraken2uniq(reads)	kraken2uniq+bracken(reads)	IDseq_nt(reads)	IDseq_nr(reads)
S11	<i>Haemophilus influenzae</i>	6	NA	6	14
S12	<i>Haemophilus influenzae</i>	8	NA	13	35
S12	<i>Streptococcus pneumoniae</i>	4	NA	222	376
S3	<i>Streptococcus pneumoniae</i>	1	NA	2	137
S3	<i>Human mastadenovirus C</i>	NA	NA	NA	NA
S4	<i>Streptococcus pneumoniae</i>	2	NA	2	70
S4	<i>Human mastadenovirus C</i>	NA	NA	NA	NA
S6	<i>Streptococcus pneumoniae</i>	9	NA	17	441
S6	<i>Human mastadenovirus C</i>	NA	NA	NA	NA
S9	<i>Haemophilus parainfluenzae</i>	2	NA	3	3

Table S4 species detected among these samples

	CHRF0000	CHRF0002	CHRF0094
Chikungunya virus	0	0	50506
Corynebacterium kefirresidentii	0	0	119
Cutibacterium acnes	3332	1297	13227
Escherichia coli	11161	657	3304
Lactobacillus iners	197	118	136
Malassezia restricta	1512	160	3749
Micrococcus luteus	248	0	717
Salmonella enterica	2124	183	1826
Sphingomonas sp. AAP5	0	117	0
Staphylococcus epidermidis	140	21	617
Streptococcus mitis	62	1945	155
Streptococcus oralis	18	1314	185
Streptococcus pneumoniae	882	728275	0