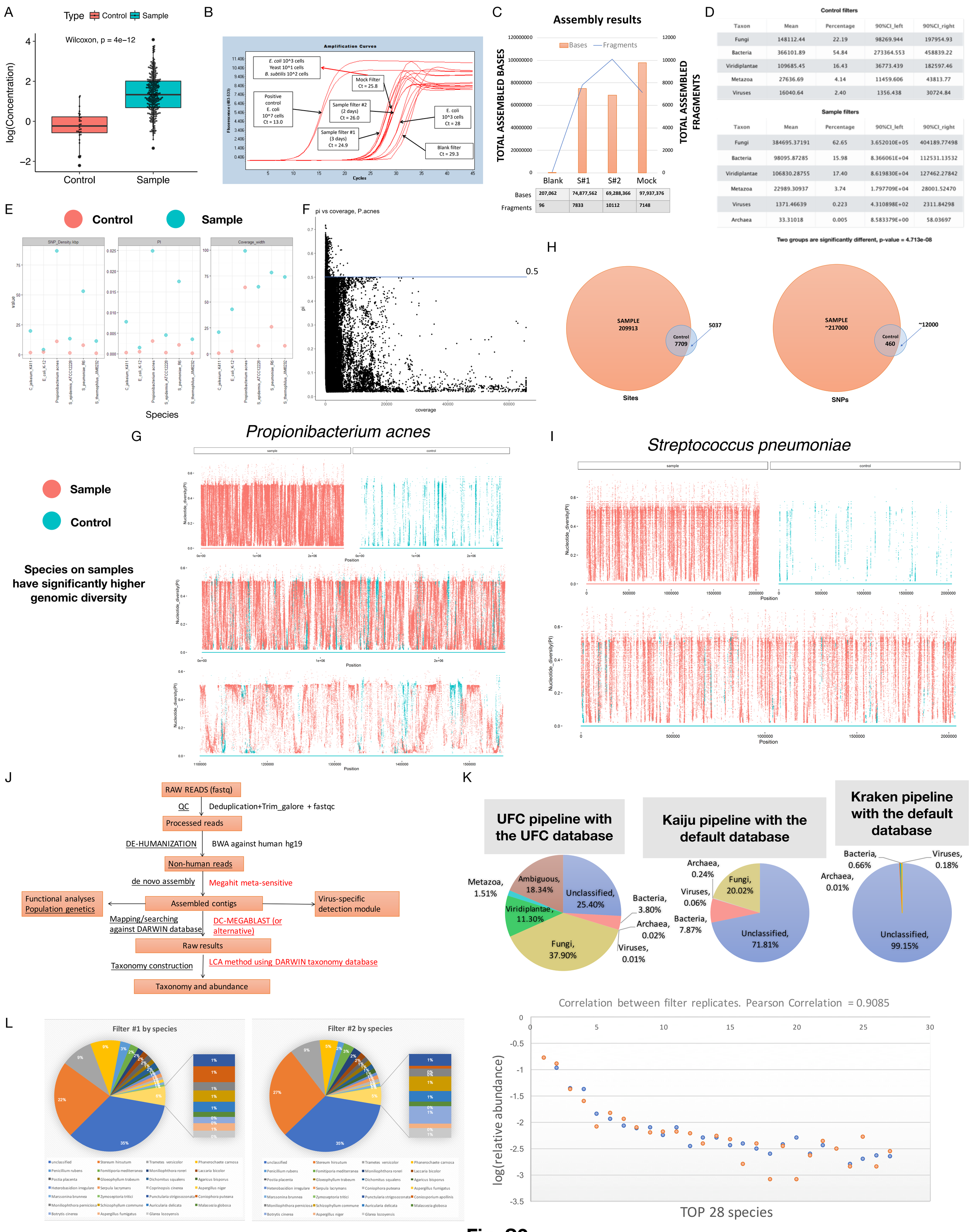
A
Type Control Sample
Wilcoxon, p = 4e−12
log(Concentration)
Control
Sample
B
Amplification Curves
E. coli 10^3 cells
Yeast 10^1 cells
B. subtilis 10^2 cells
Mock Filter Ct = 25.8
Positive control E. coli 10^7 cells Ct = 13.0
Sample filter #2 (2 days) Ct = 26.0
E. coli 10^3 cells Ct = 28
Sample filter #1 (3 days) Ct = 24.9
Blank filter Ct = 29.3
Fluorescence (483-533)
Cycles
C
Assembly results
Bases
Fragments
TOTAL ASSEMBLED BASES
TOTAL ASSEMBLED FRAGMENTS
Blank S#1 S#2 Mock
Bases 207,062 74,877,562 69,288,366 97,937,376
Fragments 96 7833 10112 7148
D
Control filters
Taxon Mean Percentage 90%CI_left 90%CI_right
Fungi 148112.44 22.19 98269.944 197954.93
Bacteria 366101.89 54.84 273364.553 458839.22
Viridiplantae 109685.45 16.43 36773.439 182597.46
Metazoa 27636.69 4.14 11459.606 43813.77
Viruses 16040.64 2.40 1356.438 30724.84
Sample filters
Taxon Mean Percentage 90%CI_left 90%CI_right
Fungi 384695.37191 62.65 3.652010E+05 404189.77498
Bacteria 98095.87285 15.98 8.366061E+04 112531.13532
Viridiplantae 106830.28755 17.40 8.619830E+04 127462.27842
Metazoa 22989.30937 3.74 1.797709E+04 28001.52470
Viruses 1371.46639 0.223 4.310898E+02 2311.84298
Archaea 33.31018 0.005 8.583379E+00 58.03697
Two groups are significantly different, p-value = 4.713e-08
E
Control
Sample
SNP_Density.kbp
PI
Coverage_width
value
Species
F
pi vs coverage, P.acnes
pi
coverage
0.5
H
SAMPLE 209913
Control 7709
5037
Sites
SAMPLE ~217000
Control 460
~12000
SNPs
G
Propionibacterium acnes
Sample
Control
Species on samples have significantly higher genomic diversity
sample
control
Nucleotide_diversity(PI)
Position
I
Streptococcus pneumoniae
sample
control
Nucleotide_diversity(PI)
Position
J
RAW READS (fastq)
QC
Deduplication+Trim_galore + fastqc
Processed reads
DE-HUMANIZATION
BWA against human hg19
Non-human reads
de novo assembly
Megahit meta-sensitive
Functional analyses
Population genetics
Assembled contigs
Virus-specific detection module
Mapping/searching against DARWIN database
DC-MEGABLAST (or alternative)
Raw results
Taxonomy construction
LCA method using DARWIN taxonomy database
Taxonomy and abundance
K
UFC pipeline with the UFC database
Metazoa, 1.51%
Ambiguous, 18.34%
Unclassified, 25.40%
Viridiplantae, 11.30%
Bacteria, 3.80%
Archaea, 0.02%
Viruses, 0.01%
Fungi, 37.90%
Kaiju pipeline with the default database
Archaea, 0.24%
Viruses, 0.06%
Fungi, 20.02%
Bacteria, 7.87%
Unclassified, 71.81%
Kraken pipeline with the default database
Bacteria, 0.66%
Viruses, 0.18%
Archaea, 0.01%
Unclassified, 99.15%
L
Filter #1 by species
Filter #2 by species
Correlation between filter replicates. Pearson Correlation = 0.9085
log(relative abundance)
TOP 28 species

Fig. S2