Supplementary Results

Metagenomic data reveal human skin-associated viruses, abundant enterophages, virulence genes, and *S. aureus*. There were no significant differences in the relative abundances of functional or taxonomic hits based on gender, usage, or time point (Bonferroni-corrected Kruskal-Wallis, $p > 0.2$). Eubacteria were the most abundant domain, followed by Eukaryota. Viruses and Archaea represented a minority of the community (**Fig. S4A**). Caudovirales (tailed bacteriophages), Herpesvirales (animal-associated), Picornavirales (eukaryote-associated), and Mononegavirales (eukaryote-associated) were the second, third, fourth, and fifth most abundant viral orders, respectively (**Fig. S4B**). However, the most abundant viral order was 'unclassified', and Microviridae was the most abundant family within this category (**Fig. S4C**).

Supplemental Figures and Tables

Pangenomes	Total size (Mbp)	No. of Contigs (>300bp)	%G+C	Avg. contig length (kbp)	Longest contig length (kbp)	No. of strains
SK1	2.98	777	33.0	3.8	115.8	3
SK2	1.90	844	32.5	2.3	16.8	3
SK3	2.67	1079	32.8	2.5	28.7	3
SK4	2.20	401	32.2	5.6	48.4	1
SK5	2.30	455	33.8	5.0	30.7	1
SK6	2.40	148	32.5	16.3	89.3	1

Table S1. Summary of *Staphylococcus* pan-genome assemblies.

		Staphylococcus Population Genomes (% of total reads mapping)						
Samples	**Total Reads (post-QC)**	**SK1**	**SK2**	**SK3**	**SK4**	**SK5**	**SK6**	**Other Taxa**
4pmD1	2768247	3.9	0.23	0.95	81.2	0.15	0.54	12.4
4pmD3	4946633	0.011	0.011	0.007	0.006	0.007	0.008	99.9
6pmD1	4137512	0.165	0.034	0.34	0.58	38.5	0.33	60.0
8pmD1	6560322	35.0	0.210	0.715	2.79	0.140	0.349	60.7
8pmD3	5035609	0.503	60.6	0.342	0.172	0.138	0.378	37.8
12pmD1	3058876	1.88	0.368	36.1	4.34	0.056	38.10	19.0
12pmD3	148282	0.273	0.325	0.577	0.336	0.194	0.632	97.6
An12pmD1	4045684	0.361	0.608	38.9	2.65	0.016	41.46	15.99

Table S2. Percent abundance of *Staphylococcus* pan-genomes across culture metagenome samples.

Metabolic systems	Sk1	Sk2	Sk3	Sk4	Sk5	Sk6
Cofactors, Vitamins, Prosthetic Groups, Pigments	241	146	119	152	123	144
Cell Wall and Capsule	95	90	63	56	87	93
Virulence, Disease and Defense	112	78	61	71	53	68
Potassium metabolism	11	10	4	4	9	4
Photosynthesis	0	0	0	0	0	0
Miscellaneous	43	30	26	33	33	24
Phages, Prophages, Transposable elements, Plasmids	33	12	31	34	15	27
Membrane Transport	79	58	57	49	32	52
Iron acquisition and metabolism	38	17	27	26	56	27
RNA Metabolism	196	140	116	114	110	117
Nucleosides and Nucleotides	158	83	91	88	86	89
Protein Metabolism	161	226	119	196	199	203
Cell Division and Cell Cycle	17	32	26	27	40	40
Motility and Chemotaxis	0	0	0	0	0	0
Regulation and Cell signaling	58	27	34	36	40	36
Secondary Metabolism	6	4	4	4	4	4
DNA Metabolism	97	112	71	65	74	106
Fatty Acids, Lipids, and Isoprenoids	124	92	83	92	67	86
Nitrogen Metabolism	28	23	26	23	11	21
Dormancy and Sporulation	22	10	4	10	10	10
Respiration	53	44	29	30	25	30
Stress Response	104	75	78	67	64	63
Metabolism of Aromatic Compounds	9	6	5	5	5	5
Amino Acids and Derivatives	345	321	317	309	248	305
Sulfur Metabolism	27	14	22	18	19	19
Phosphorus Metabolism	40	32	27	25	23	25
Carbohydrates	357	229	210	223	225	219

Table S3. Distribution of annotated protein-coding genes from culture metagenomes across SEED Subsystem categories, for each *Staphylococcus* pan-genome.

Study	# of Surfaces	# of Time points/Day	# of Days	# of Restrooms	# of Stalls/ Restroom	Total Surface Samples	Data Generated
Eight-Hour	1 (surface1)	8 (one collection per hour)	2	4 (male and female from high and low use restrooms)	1	64	Virus:Bacteria Microscopy, 16S sequencing
One-Month	3 (surface 1,2 & 3)	1	15	4 (male and female from high and low use restrooms)	1	180	Virus:Bacteria Microscopy, 16S sequencing
Eight-Week	3 (surface 1,2 & 3)	1	8	2 (one male and one female)	1	48	Virus:Bacteria Microscopy, 16S sequencing
Human-Free	1 (surface1)	5 (12, 2, 4, 6 & 8pm)	1	2 (one male and one female)	2	20	Virus:Bacteria Microscopy, 16S sequencing, 16S sequencing post aerobic & post anaerobic culture, and metagenomes

Figure S1. Sampling protocol for four restroom studies. The diagram shows the three surfaces sampled. Not all surfaces are sampled for each study as indicated in the table. The table includes the number of surfaces sampled, number of time points per day, number of sampling days, number of restrooms, and number of stalls per restroom for each study. Total number of surface samples collected is determined by multiplying all data in the corresponding row. Data originated from the corresponding surface samples after further processing as explained in the paper.

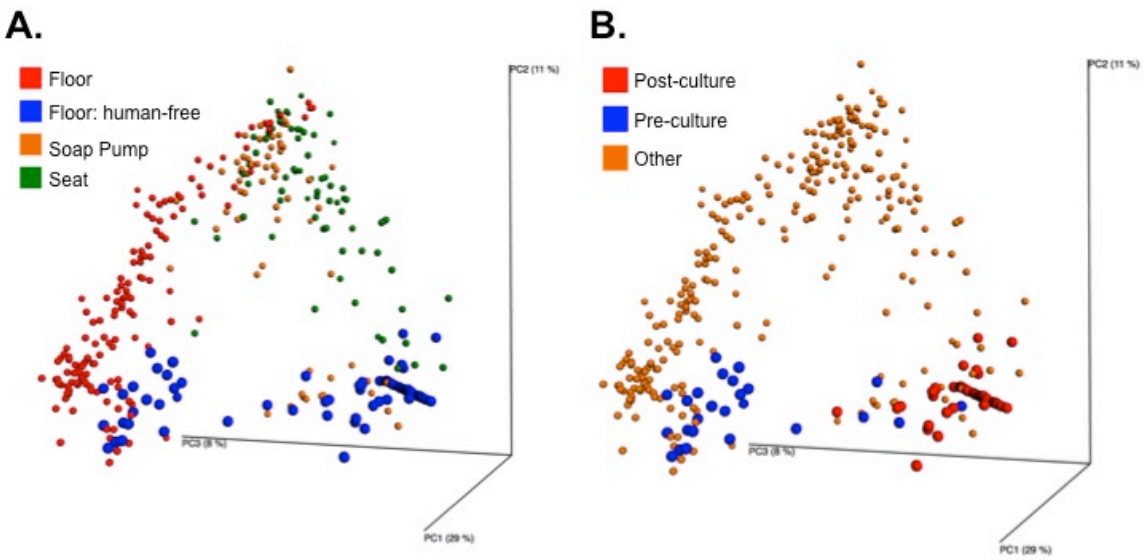
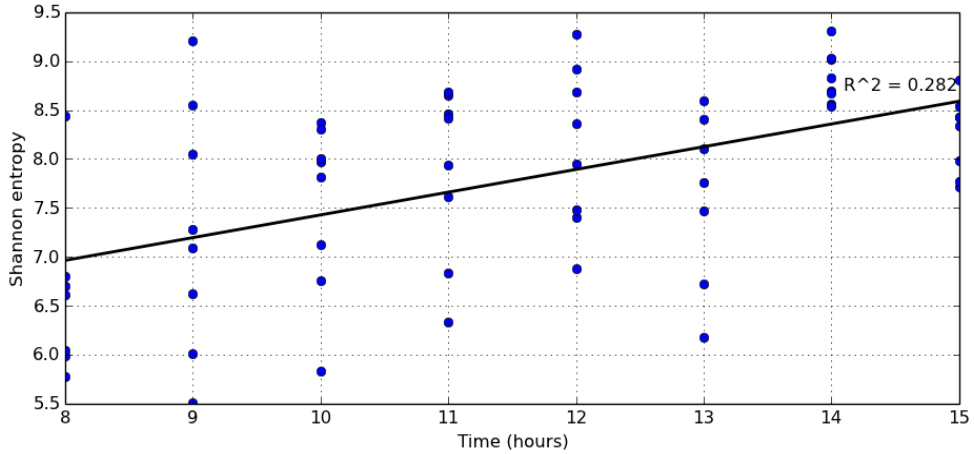


Figure S2. Principal coordinate plots (weighted UniFrac) showing samples from all four restroom studies along the first 3 axes. The enlarged spheres towards the bottom of the plots represent the samples from the human-free study. Panel A shows the samples colored by surface type (human-free samples indicated as 'na', but these samples were taken from the floor). Panel B shows samples colored by culture status ('na' denotes samples from all other studies except the human-free study; 'Post-culture' denotes swab samples that were sequenced after being cultured; 'Pre-culture' denotes swab samples that were sequenced prior to being cultured).

A.



B.

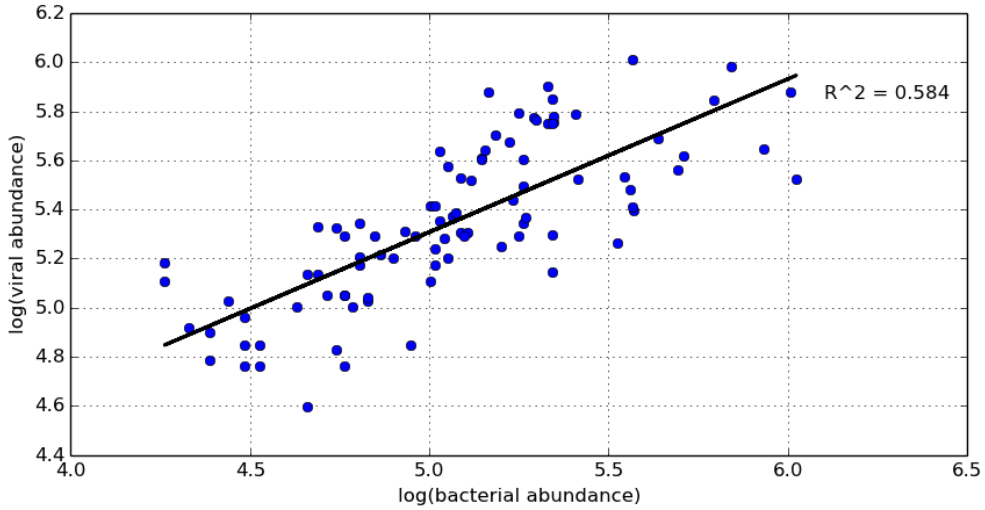


Figure S3. (A) $\log_{10}$ viral and bacterial abundances are positively correlated ($R^2 = 0.584$, $p < 0.0001$). (B) Shannon diversity is positively correlated with time since decontamination ($R^2 = 0.282$, $p < 0.0001$).

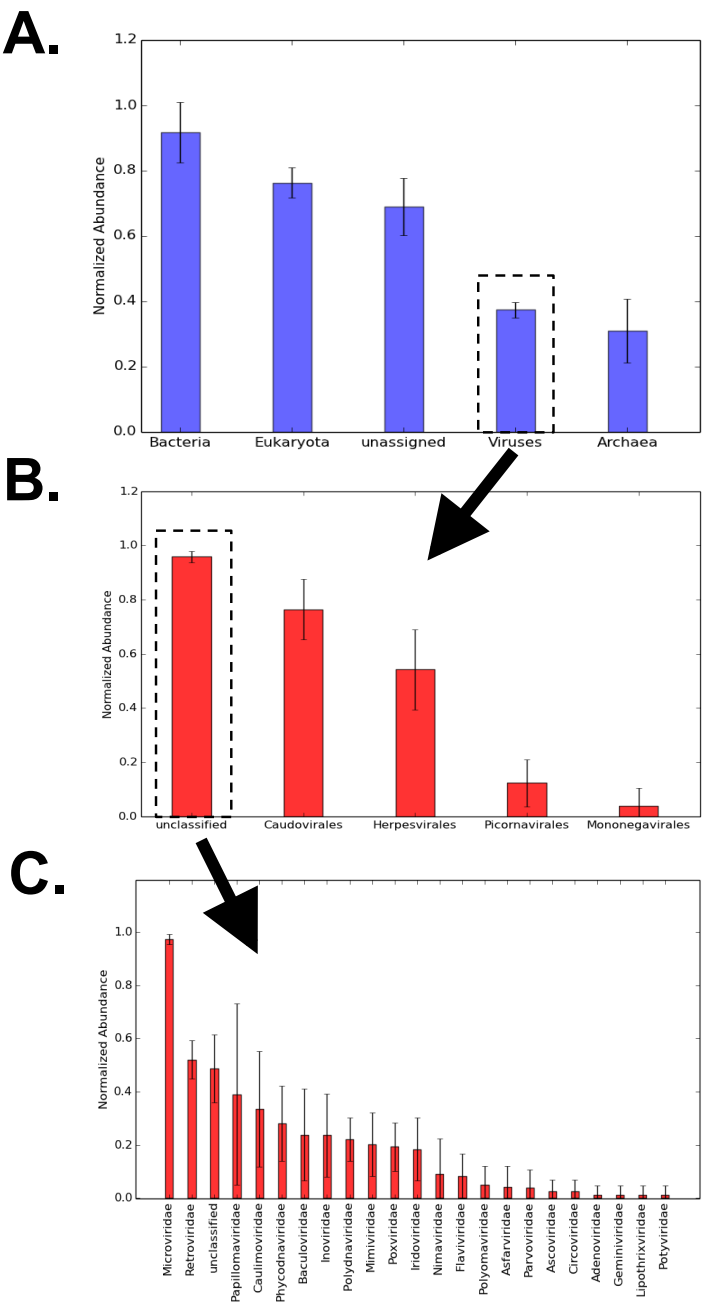


Figure S4. Normalized abundances of sequence reads annotated to hierarchical taxonomic categories (M5NR database) from the 7 floor metagenomes. Panel A shows domain-level abundances. B shows the prevalence of viral orders. C shows viral families within the 'unclassified' order. Dashed boxes show the categories that are expanded in the proceeding panel.

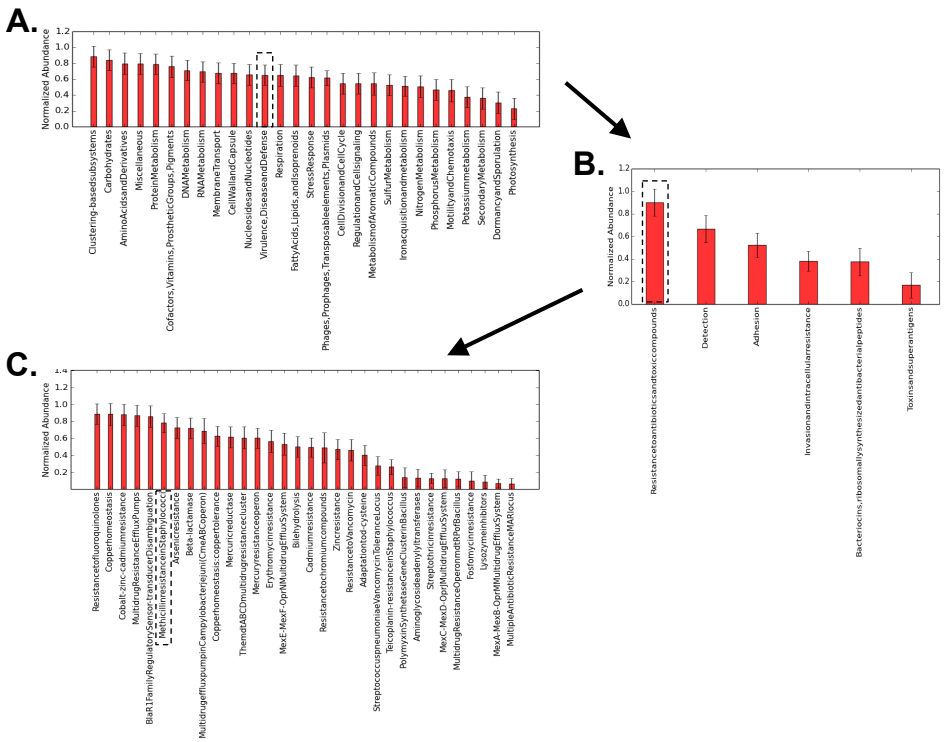


Figure S5. Normalized abundances of sequence reads annotated to hierarchical functional categories (SEED) from the 7 floor metagenomes. Panel A shows SEED L1 abundances. B shows the prevalence of virulence categories (SEED L2). C shows SEED L3 categories within the antibiotic resistance category from panel B. Dashed boxes show the categories that are expanded in the proceeding panel.