The Composition of Diesel Fuel Influences the Structure of Microbiological Assemblages in Contaminated Storage Tanks

James G. Floyd¹, Blake W. Stamps^{2,3}, Caitlin L. Bojanowski⁴, Wendy J. Goodson⁴, Bradley S. Stevenson¹

1 – University of Oklahoma, Department of Microbiology and Plant Biology, Norman OK

2 – 711th Human Performance Wing, Airman Systems Directorate, Air Force Research Laboratory, Wright-Patterson AFB, OH, United States

3 – UES Inc., Integrative Health and Performance Sciences Division, Dayton, OH, United States

4 – Soft Matter Material Branch, Materials and Manufacturing Directorate, Air Force Research Laboratory, Wright-Patterson AFB, OH

Key Words: Biofouling, Diesel, Biodiesel, Redundancy Analysis

Abstract

Microbiological contamination in diesel storage tanks is widespread and under-reported. The contaminating microorganisms can degrade components of the fuel, which contributes to fouling and corrosion. To better understand the connection between the microorganisms that are responsible for this contamination, the composition of affected fuels, and the resulting impact on fuel quality, we conducted a survey of 106 fuel tanks at 17 military bases across the continental United States. Fuel was sampled from each tank and the resident microbial communities were characterized using high throughput sequencing of small subunit ribosomal RNA gene libraries. Fatty acid methyl esters (FAME) and n-alkanes were characterized and quantified using GC-MS to determine their correlation with the presence of microbial taxa. Redundancy Analyses identified which microbial taxa were more prominent in contaminated fuels. Members of the fungal family *Trichomaceae* were found to be prominent in fuels containing more FAME. Members of the yeast family *Debaryomycetaceae* were found to be prominent in fuels containing more pentadecanoic and oleic acid methyl esters. These relationships between fungal taxa and fuel components were directly tested in growth experiments with representative isolates of the *Trichomaceae* (*Paecilomyces* AF001) and *Debaryomycetaceae* (*Wickerhamomyces* SE3) families. *Paecilomyces* was capable of growth on linoleic acid methyl ester but unable to grow on pentadecanoic acid methyl ester, while *Wickerhamomyces* was able to grow on both substrates. Fuel composition may provide some insight into which microorganisms can proliferate but other factors like competition and symbiosis may also drive microbial proliferation, fouling, degradation, and corrosion in diesel fuels.

Importance

Biodiesel, widely used as an additive or extender of ultra-low sulfur diesel, can increase the potential for microorganisms to proliferate in storage tanks. It is important to know how the composition of diesel fuels can influence the growth of organisms linked to fuel degradation and microbiologically influenced corrosion. This research describes how certain populations of fungi and bacteria can prevail in fuels of different composition, which can be helpful in predicting biodegradation and biocorrosion, and formulating fuels less susceptible to the growth of problematic organisms.

Introduction

The U.S. Energy Information Administration predicts that petroleum and other fuel consumption can be expected to rise from the current consumption of 16 million barrels per day to just under 25 million barrels per day by the year 2050 (1). One common transportation fuel is diesel blended with biodiesel (*i.e.*, biodiesel blends). In recent years, the U.S. Department of Energy has incentivized the use of biodiesel through December 2022 in efforts to offset carbon emissions (2). One common biodiesel blend in the U.S. is B20, containing 20% biodiesel and 80% petroleum diesel. B20 represents a balance of costs, emissions control, cold-weather performance, and engine material compatibility (3). Additionally, biodiesel is readily added to ultralow sulfur diesel (ULSD) at up to 5% v/v to compensate for the loss of lubricity from the desulfurization process (4). Biodiesel (and its' blends) is also more easily degraded by endogenous environmental microorganisms and has a much higher flash point when compared to ULSD, making it much safer for the environment and those handling the fuel (4–6); However, the high biodegradability of B20 is problematic for operators that store biodiesel containing fuels long term (7).

Biodiesel is composed of fatty acid methyl esters (FAMES) that are produced through a transesterification reaction with feedstocks of animal, plant, or microbial lipids (8). The FAMES in biodiesel make the fuel more susceptible to microbial contamination since FAME compounds exist naturally in the environment and are readily degraded by many different microorganisms (9, 10). Additionally, hydrocarbons (*e.g.* n-alkanes) in diesel fuel can also be used as an oxidizable substrate for many microorganisms (11). The microbial metabolism of both FAMES and n-alkanes in fuels manifests itself as microbial contamination within fuel storage infrastructure. Microbial contamination and proliferation in storage tanks containing biodiesel and biodiesel blends leads to increased operating costs through mechanical cleaning of affected storage tanks. Additionally, the contaminating microbial populations can increase corrosion of storage tank materials because their metabolism of the fuel produces acidic metabolic byproducts and their production of biofilms on surfaces can generate oxygen corrosion cells under aerobic conditions (12). Fungi and bacteria grown on B20 biodiesel can not only degrade FAME and hydrocarbons, but also increase corrosion rates and pitting corrosion of carbon steel (12–14).

All biodiesels or biodiesel blends are not identical. When different feedstocks (*e.g.* animal fats, soy, etc.) are used to produce biodiesel, they result in unique FAME profiles with variations in fatty acid chain lengths and degrees of saturation (15). The varying profiles of each biodiesel, therefore, may selection for differing communities of microorganisms that are adapted to preferentially consume specific FAMES. Microorganisms oxidize FAME using lipases to first hydrolyze the ester bond, converting the FAME to methanol and the corresponding fatty acid, which is then further oxidized to yield ATP (9,16). Different microorganisms have differing capabilities of fatty acid metabolism that depend on fatty acid chain length. Some metabolic pathways even require the use of syntrophic communities to utilize long chain fatty acids (17).

Microorganisms exhibit substrate specificity towards certain FAMEs, which is why each fuel may influence the composition of the microbial community composition within each fuel storage tank.

The major components of most biodiesel blends, except for B100, come from the petroleum itself, which varies in composition based on its provenance and distillation processes used to formulate the final product (18). Diesel fuel primarily consists of hydrocarbons between 7 and 24 carbon atoms in length and is typically comprised of 64% alkanes (including n-, iso-, and cyclo-alkanes), 1-2% alkenes and 34% aromatic hydrocarbons (19, 20). In general, microorganisms preferentially oxidize shorter chained hydrocarbons (21), although there are some exceptions of microorganisms preferentially metabolizing longer chain alkanes (C₂₀+) (22). The combination of compositional differences in diesel fuels and differential metabolic capabilities among microorganisms would represent a highly selective framework that could explain the structure of microbial communities in contaminated fuel.

Microbial community assemblies can shift based on the carbon substrates that are available in the environment (23–26). To determine the relationship between fuel composition and microbial communities in contaminated fuels, we conducted a survey of diesel fuels (ULSD and biodiesel blends) from military installations across the continental United States. Microbial community composition was analyzed via SSU rRNA gene sequencing and fuel composition was analyzed using gas chromatography coupled to mass spectrometry (GC-MS). Previous work identified both fungi and bacteria linked to elevated rates of corrosion when grown on biodiesel and diesel fuels (12, 13, 27). Correlations between fuel substrates and taxa from RDA analyses were tested using the fungal isolates *Paecilomyces* AF001 (Family - *Trichocomaceae*) and *Wickerhamomyces* SE3 (Family - *Debaryomycetaceae*), which have both been previously linked

to corrosion risks when grown using biodiesel as a sole carbon and energy source (13). Herein we identify correlative links between the bacterial and fungal microbial communities present within fuel storage tanks across the U.S. and the concentrations of specific FAME and hydrocarbon components of stored fuel. With a better understanding of the relationships between fuel composition and their contaminating microbial communities, we can alert fuel storage operators to the potential risks associated with fouling and microbiologically influenced corrosion for diesel fuel based on its composition.

Results

Characterization of n-alkanes and FAME components in fuels.

Biodiesel blend and diesel fuels were found to have different profiles of n-alkanes and FAME components. When visualized by PCA, fuels separated by geographic region instead of fuel type (FIG 1A). The largest drivers of alkane variability in the collected fuels were from C9 (nonane)-C20 (eicosane) alkanes and were responsible for much of the separation of samples in the PCA (FIG 1A).

FIG 1A. PCA ordination of fuel samples based on n-alkane composition. Biodiesel (circles) and diesel (triangles) fuel samples from different geographical areas in the continental U.S. are represented with different colors with Central (Orange), Northwest (Green), Southeast (Blue), and Southwest (Purple). Blue vectors represent the different n-alkanes and their contribution to the separation of the fuel samples in this ordination space. **B.** Blue vectors represent the different FAME components and their contribution to the separation of the fuel samples in this ordination space. Many FAME components had low eigenvectors in this PCA, and their chemical names are excluded for simplicity; however, their vectors remain in this PCA to give a better overview of how fuel samples are separated.

In contrast to n-alkanes in which fuel region appeared to primarily distinguish fuel composition, fuel type (biodiesel blend or diesel) primarily separated samples, with diesel fuels clustering tightly together when visualized by PCA. The biggest drivers of FAME variability were linoleic, oleic, and palmitic acid methyl esters concentrations and to a lesser extent steric and linoleic acid methyl ester (FIG 1B).

Microbial Community Analysis of Contaminated Fuels.

A total of 150 samples produced 3,834,984 sequence reads for 16S and 523,852 sequence reads for 18S, resulting in an average of 25,567 reads per sample for 16S and 3492 for 18S. Bacterial communities were more diverse than the fungi in contaminated fuels based on the Faith diversity index, where Bacteria had a Faith PD median of 14 in biodiesel samples and 10 in diesel samples, and fungal samples had a Faith PD median of 1.4 in biodiesel and 2.0 in diesel (Supplemental Fig. 1). Some exceptions were detected that appeared to be a bacterial “bloom”, where bacterial communities were composed of more than 80% of the bacterial families *Acetobacteriaceae* and *Moraxellaceae* (FIG 2). Fuels contaminated with *Acetobacteriaceae* were not constrained by geography as these contamination events occurred in both the Northwest and Southeast U.S. Blooms where *Moraxcellaceae* were the most abundant population were only observed in fuel samples from the Southeast U.S.

FIG 2. Heatmap representing the relative abundance of bacterial populations in microbial communities from contaminated fuels. Relative abundance of 100% is represented by dark blue, while gray represents 0% of the reads. Samples are sorted by the dendrogram on the left y-axis, which was based on bacterial population similarity. Bacterial families are denoted along the x-axis and sample names are denoted on the right y-axis.

In contrast to the diversity observed among the bacteria, the fungal populations present in microbial communities from contaminated fuels were primarily comprised of a single family, the *Trichocomaceae*. While the *Trichocomaceae* were usually the most abundant family in both diesel and biodiesel blends sequenced, samples within the Southeast U.S. primarily contained members of the family *Debaryomycetaceae* (FIG 3).

FIG 3. Heatmap representing the relative abundance of fungal populations in microbial communities from contaminated fuels. Relative abundance of 100% is represented by dark blue, while gray represents 0% of the reads. Samples are sorted by the dendrogram on the left y-axis, which was based on fungal population similarity. Fungal families are denoted along the x-axis and sample names are denoted on the right y-axis.

Redundancy Analysis Correlating Fuel Components to Taxonomy.

Redundancy analyses (RDA) were conducted to extract and summarize the variation in the microbial taxonomic data (both bacterial and fungal) and derive correlations with the fuel composition. The bacterial family *Acetobacteriaceae* were positively correlated to fuels containing more long chain n-alkanes including eicosane, tricosane, and tetracosane (FIG 4). The *Acetobacteriaceae* were also positively correlated with increases in linoleic acid methyl ester. In

contrast, the bacterial family *Moraxellaceae* was positively correlated with fuels containing shorter n-alkanes and higher concentrations of myrisoleic acid methyl ester. The RDA model from the bacterial taxa and fuel composition had a R^2 value of 0.33 with $p=0.003$. Fuel compounds found to significantly impact microbial communities were used in the bacterial RDA using an adonis test on the Hellinger transformed taxonomic data can be seen in Table 1.

FIG 4. Redundancy analysis (RDA) of the Hellinger-transformed bacterial populations constrained by the fuel composition. Bacterial families are denoted in red text. In this RDA the red crosses are bacterial families that still contribute to this ordination space, but not specifically named due to their central clustering indicating a low correlation coefficient. Fuel component variables are represented by blue vectors. Blue vectors lengths indicate the relative weight of a fuel component in the ordination.

Table 1. Forward selection of significant ($\text{Pr}(> F) \leq 0.05$) fuel variables based on an adonis permutational multivariate analysis of variance on the bacterial Hellinger transformed taxonomic data.

Fuel Component	Mean Squares	F. Model	R ²	Pr(>F)
Octane	0.656	4.076	0.042	0.001
Nonane	0.505	3.137	0.032	0.003
Decane	0.297	1.849	0.019	0.045
Dodecane	0.307	1.907	0.020	0.044
Tridecane	0.342	2.121	0.021	0.041
Tetradecane	0.617	3.835	0.040	0.001
Pentadecane	0.313	1.946	0.020	0.050
Hexadecane	0.389	2.415	0.025	0.019
Docosane	0.466	2.896	0.030	0.005
Tetracosane	0.329	2.040	0.021	0.039
Pentacosane	0.366	2.274	0.024	0.020
Hexacosane	0.489	3.033	0.031	0.004
Myrisoleic Acid ME	0.328	2.035	0.021	0.025
Tetradecanoic Acid ME	0.365	2.264	0.023	0.018
Linoleic Acid ME	0.519	3.222	0.033	0.002

An RDA was also generated to identify how fungal communities may be correlated with fuel composition (FIG 5). The abundant fungal family *Trichocomaceae* was positively correlated

204 with fuels containing higher concentrations of FAME. *Debaryomycetaceae* was found to be
 205 abundant in fuels containing higher concentrations of pentadecanoic and oleic acid methyl acids.
 206 The *Debaryomycetaceae* were also found in fuels containing more mid-chain n-alkanes (e.g.
 207 pentadecane and hexadecane). The fungal RDA had a R^2 value of 0.45 with $p = 0.005$. Fuel
 208 compounds found to significantly impact fungal communities used in the fungal RDA
 209 determined via a permutational multivariate analysis of variance using distance matrices on the
 210 Hellinger transformed taxonomic data can be seen in Table 2. The fungal family
 211 *Trichocomaceae* had no correlation with the FAME pentadecanoic acid methyl ester which was
 212 tested later as a sole carbon and energy source for the representative isolate *Paecilomyces* AF001
 213 for growth.

FIG 5. Redundancy analysis (RDA) of the Hellinger-transformed fungal communities constrained by the fuel composition. Fungal families are denoted as red text. In this RDA the red crosses are fungal families that still contribute to this ordination space, but not specifically named due to their central clustering indicating a low correlation coefficient. Fuel component variables are represented by blue vectors. Blue vectors lengths indicate the relative weight of a fuel component in the ordination.

227 **Table 2.** Forward selection of significant ($\text{Pr}(> F) \leq 0.05$) fuel variables based on an adonis
228 permutational multivariate analysis of variance on the fungal Hellinger transformed taxonomic
229 data.

Fuel Component	Mean Squares	F. Model	R²	Pr(>F)
Octane	0.389	4.269	0.040	0.003
Nonane	0.257	2.819	0.026	0.019
Decane	0.258	2.823	0.027	0.025
Pentadecane	0.223	2.438	0.023	0.049
Hexadecane	0.276	3.031	0.029	0.018
Octadecane	0.458	5.019	0.047	0.002
Eicosane	0.544	5.963	0.056	0.001
Docosane	0.309	3.384	0.032	0.012
Tricosane	0.337	3.698	0.035	0.005
Myrisoleic Acid ME	0.543	5.962	0.056	0.001
Pentadecenoic Acid ME	0.059	4.132	0.038	0.006
Palmitoleic Acid ME	0.295	3.236	0.030	0.011
Palmitic Acid ME	0.240	2.982	0.027	0.046
Oleic Acid ME	0.345	3.782	0.036	0.010
Stearic Acid ME	0.262	2.877	0.027	0.020
Eicosadenoic Acid ME	0.306	3.354	0.032	0.006
Arachidic Acid ME	0.224	2.460	0.023	0.041
Heneicosanoic Acid ME	0.231	2.530	0.024	0.044

Physiological Characterization of Isolates to Test Correlations Predicted from Fungal RDA.

The RDA identified many correlations between both bacterial and fungal communities and fuel composition. Two isolates representative of abundant fungal genera from contaminated fuels were used to test correlations identified from via RDA (13). The filamentous fungus *Paecilomyces* was able to grow successfully on palmitoleic acid methyl ester as a sole carbon and energy source (strong correlation); however, it was unable to grow using pentadecanoic acid methyl ester (no correlation). *Wickerhamomyces*, which had a strong correlation with both palmitoleic and pentadecanoic acid methyl esters, was able to use them both as sole carbon and energy sources (FIG 6).

FIG 6 Growth of the fungal isolates *Paecilomyces* AF001 (Filled circles; *Trichocomaceae*) and *Wickerhamomyces* SE3 (Open circles; *Debaryomycetaceae*) palmitoleic acid methyl ester as a sole carbon and energy source. Error bars represent standard deviation for mean *Paecilomyces* AF001 or *Wickerhamomyces* SE3 (n=3). **B:** Growth of *Paecilomyces* AF001 (Filled circles; *Trichocomaceae*) and *Wickerhamomyces* SE3 (Open circles; *Debaryomycetaceae*) on pentadecanoic acid methyl ester as a sole carbon and energy source. Error bars represent standard deviation for mean *Paecilomyces* AF001 or *Wickerhamomyces* SE3 (biological replicates n=3).

Discussion

The composition of microbial communities can be influenced by numerous factors including the availability of oxidizable substrates that can be differentially consumed (28). We considered this type of selection true in fuel storage tanks where the composition of fuel components may play a significant role in dictating which microorganisms may be prevalent. Correlations were identified between the chemical composition of both B20 biodiesel and diesel fuels and their resident microbial communities. Two positive correlations were between members of the fungal families *Trichocomaceae* and *Debaryomycetaceae* and palmitoleic and pentadecanoic acid methyl esters, respectively. Two representative fungi, *Paecilomyces* AF001 and *Wickerhamomyces* SE3, previously isolated from contaminated fuel storage tanks, were then used to test the correlations. Growth studies with palmitoleic and pentadecanoic acid methyl esters as a sole carbon and energy source verified the link predicted between these abundant members the fuels they inhabited.

Regional differences in fuel composition were identified via GC-MS, in addition to differences more clearly attributable to the type of fuel analyzed (B20 biodiesel vs. ULSD). Most of the differences in n-alkanes of biodiesel and diesel fuels were attributed to the presence of higher chain alkanes (C15-24) vs. lower chain alkanes (C9-C14, Fig 3). Most biodiesel fuel samples were similar to one another, with smaller differences in their n-alkane composition compared to the diesel fuels. As predicted, diesel fuels had more similar FAME profiles, likely because FAME was in much lower abundance compared to the B20 biodiesel blends. Among the B20 biodiesel fuels, the variation in FAME composition was largely based on the concentration of linoleic, palmitic, oleic, stearic, and linolelaidic acid methyl esters. When different feedstocks are used, the resulting biodiesel will have different total percent weights of these FAME (15).

The fuels analyzed in this study were chemically diverse, exhibited differences based on region, but were also correlated with the composition of their microbial contaminants.

The bacterial populations found in contaminated fuels were much more diverse in terms of richness than the fungal populations (Supplementary FIG 1). However, there were some exceptions as some contaminated fuels contained higher relative abundances of bacteria from the families *Acetobacteriaceae* and *Moraxellaceae*. Both bacterial families have been implicated in diesel and biodiesel contamination. Additionally, their metabolism of fuel components produces organic acids, causing increased rates of corrosion (12, 29).

Fungal populations were not as richly diverse as the bacteria in the contaminated fuels. Fungal contamination in sampled fuels was largely comprised of members of the family *Trichocomaceae*. Although the *Trichocomaceae* are ubiquitous in nature and normally associated with the spoilage of fruit juices and foods, they have also been linked to biodiesel degradation and MIC (13, 30). Members of the fungal family *Debaryomycetaceae* were also a primary contaminant in fuels although to a lower extent than *Trichocomaceae*. Members of this yeast family are used in fermentation processes in the food industry but have also been linked to fuel degradation and MIC (12, 13, 31).

Different microbial populations in fuels can have different fouling properties, different metabolic processes, and pose different risks to infrastructure as they metabolize fuels (32, 33). If fuel composition selected for distinct microbial community assemblages, we expected correlations between fuel composition and the bacterial or fungal populations found in contaminated fuel. Members of the bacterial family *Acetobacteraceae* were positively correlated to fuels that contained higher alkane chain lengths than others as well as with linoleic acid methyl ester. Some bacteria, such as *Pseudomonas* and *Acinetobacter*, primarily degrade

alkanes with higher chain lengths (C19+) (22, 34). Although it is still not clearly understood which hydrocarbon substrates *Acetobacteriaceae* may preferentially use, they are commonly found in contaminated diesel and biodiesel fuels (12, 35, 36). Their correlation with fuels containing longer chain alkanes (C19-24) would suggest that they are able to use these alkanes, representing a potential niche other bacteria or fungi cannot occupy. Additionally, the bacterial family *Moraxellaceae* was positively correlated with smaller alkane chains (C8-14) and myristoleic acid methyl ester (Fig 5). Members of this bacterial family have also been observed in contaminated fuels; however, like the *Acetobacteraceae*, little is known about which fuel compounds they specifically metabolize (37, 38).

Correlations were also found between the abundance of fungal populations and certain fuel components. *Trichocomaceae* was the prominent fungal population in contaminated fuels and correlated with fuels containing more fatty acid methyl esters including palmitoleic acid methyl ester. Members of the yeast family *Debaryomycetaceae* were more abundant in fuels containing higher concentrations of Cis-10-Pentadecenoic Acid Methyl Esters and Cis-9-Oleic Acid Methyl Esters (Fig 3.6). The correlations of these two fungal families suggested that they may have differential capabilities in degrading these fuel components. Both members of the *Trichocomaceae* and *Debaryomycetaceae* have been implicated in contamination of diesel and biodiesel fuels previously and appear to be a common fuel contaminant across the continental U.S. when fungal contamination is present (12, 39). Representative organisms from both fungal families have also recently been implicated in corrosion when grown on B20 biodiesel (13).

The predicted difference in fuel component preference between the *Trichocomaceae* and *Debaryomycetaceae* was directly tested *in vitro* using two representative taxa, *Paecilomyces* AF001 and *Wickerhamomyces* SE3, previously isolated from contaminated B20 biodiesel (13).

Wickerhamomyces SE3 grew well on both palmitoleic and pentadecanoic acid methyl ester, but *Paecilomyces* could only grow on palmitoleic acid methyl ester (Fig 7). Both FAME compounds contain one degree of unsaturation, but palmitoleic acid methyl ester has an even number carbon chain (C16) while pentadecanoic acid methyl ester has an odd number carbon chain (C15). Even though both have one degree of saturation, the double bond in palmitoleic acid methyl ester is at C9, while pentadecanoic acid methyl ester has a double bond at C10, an odd and an even numbered carbon in the fatty acid chains. Not all organisms are able to metabolize odd numbered fatty acids and a more diverse set of enzymes are needed to metabolize the double bonds in the fatty acid chain (40). Odd numbered bonds are typically activated by isomerases while the double bonds on even numbered chains must be activated by isomerases and reductases, which *Paecilomyces* may not have. This would explain why it was not able to grow with pentadecanoic acid methyl ester as a sole carbon and energy source.

The availability of different carbon substrates can influence microbial community composition and differences have been observed in diesel and biodiesel fuels as well (41–45). Correlations between both bacterial and fungal populations in contaminated fuels and the composition of FAME and alkanes in those fuels suggest that fuel composition influences the composition of contaminating microorganisms. A few of these distinct correlations were supported by growth studies with two representative fungi. While carbon substrates in an environment can drive community structure, it is important to note that this is not the only driver. Biotic and abiotic factors such as interactions between microorganisms, the availability of nutrients like phosphorous and nitrogen, or the amount of water present in a contaminated fuel tank can also influence which microorganisms will proliferate. Our findings described here link fuel composition to the microorganisms found in contaminated storage tanks, providing insight

as to which feedstocks are best to use when producing biodiesel. For instance, limiting the presence of certain FAMES, such as those readily consumed by *Paecilomyces* may strongly control biocontamination in tanks. Further, identifying fuel components that encourage the growth of microorganisms that do not cause pervasive biofouling or corrosion would be useful in developing potentially probiotic fuel communities.

Materials and Methods

Fuel Tank Sampling Protocol.

Fuel was collected from 17 sites across the continental United States. Site selection was contingent upon obtaining access to contaminated fuel storage tanks. A map containing an overview of general site region and the number of fuel samples taken is described in Supplementary FIG 2 and Supplementary Tables 1-3. Fuel from storage tanks was collected using a stainless-steel Bacon Bomb (Koehler™ Instrument Petroleum Bacon Bomb Sampler, Fisher Scientific) that was field disinfected using 100% isopropanol prior to insertion into the tanks. Once field sterilized the Bacon Bomb was inserted into the fuel storage tanks and 500 mL of fuel was collected from the bottom of the tanks and transferred into sterile 1L bottles. Immediately after collection, the fuel was filtered through a polyether sulfone (PES) Steritop Filter Unit™ (Millipore) with a nominal pore size of 0.45 µm affixed to the top of a second sterile 1L bottle. The filter was retained for DNA extraction and the fuel was retained for chemical analysis, both described below.

DNA Extraction, SSU rRNA Gene Library Preparation, and Sequencing.

Following fuel filtration, the PES filter was removed and sectioned into four equal quarters using sterile disposable scalpels. Three sections from each sample were individually placed into Zymo Quick-DNA Fecal/Soil Microbe ZR BashingBead™ Lysis Tubes filled with

DNA/RNA Shield™ (Zymo Research Co, Irvine, CA, United States). Homogenization and physical lysis of biomass samples was achieved in the field using a custom tube holder attached to a reciprocating saw, agitated at full speed for 1 min. After homogenization, the samples were shipped back to lab, and stored at -20°C until (≤ 3 days) the DNA was extracted via the Zymo Quick-DNA Fecal/Soil Microbe kit (Zymo Research) according to the manufacturer's protocol.

Small subunit (SSU) rRNA gene libraries were generated by amplifying the extracted DNA using PCR primers that spanned the V4 and V5 hypervariable regions of the 16S ribosomal RNA gene between position 515 and 926 (*Escherichia coli* numbering), which produced a ~400 bp fragment for Bacteria and Archaea and ~600 bp fragments for the Eukarya 18S rRNA gene. These primers were selected to amplify a broad distribution of SSU rRNA genes from all three domains of life (44). The forward primer 515F-Y (5'-**GTA AAA CGA CGG CCA GCCG** TGY CAG CMG CCG CGG TAA-3') contains the M13 forward primer (in bold) fused to the gene-specific forward primer (non-bold). The reverse primer 926R (5'-CCG YCA ATT YMT TTR AGT TT-3') was unmodified from Parada et al. (46). 5 PRIME HotMasterMix (Quanta Biosciences, Beverly, MA, United States) was used for all library prep PCR reactions with a final reaction volume of 50 μ L. Thermocycler conditions included a 3 min initial denaturation at 95 °C followed by 30 cycles of 95 °C for 45 s, 50 °C for 45 s, 68 °C for 90 s, and a final extension of 68 °C for 5 min. Reactions were then purified with Sera-Mag™ paramagnetic beads (MilliporeSigma, St. Louis, MO) at a final concentration of 0.8x. Following purification, 4 μ L of PCR product was used in a barcoding reaction affixing a unique 12 bp barcode with a complementary M13 sequence to those found on the forward primer of the initial reaction (Integrated DNA Technology) to each library in 50 μ L reactions. A 6-cycle reaction was carried out with identical cycling conditions to the initial amplification. The resulting barcoded PCR

products were purified using Sera-Mag™ beads at a final concentration of 0.8 x and eluted in 50 µL of nuclease free water. The cleaned, barcoded PCR products were quantified using the QuBit™ dsDNA HS assay kit (Thermo Fisher Scientific Inc., Waltham, MA, United States) and pooled in equimolar amounts before concentration using an Amicon® Ultra 0.5 mL centrifugal filter with Ultracel-30K membrane (Millipore Sigma, Billerica, MA, United States) to a final volume of 80 µL. To detect and mitigate the effects of reagent contamination, triplicate extraction blanks (DNA extraction with no sample addition) and negative PCR controls (PCR with no template DNA added) were sequenced as well. The pooled prepared libraries were submitted for sequencing on an Illumina MiSeq (Illumina Inc., San Diego, CA, United States) using V2 PE250 chemistry at the University of Oklahoma Consolidated Core Lab.

GC-MS Quantification and Characterization of FAME and Alkanes in Fuels.

Sampled biodiesel and diesel fuels were analyzed to determine the amount and composition of FAME and n-alkanes present. All fuel samples were filtered prior to shipment to the University of Oklahoma to remove biomass. Fuels were analyzed using a Shimadzu GCMS-QP2010 coupled to a AOC-20i autosampler (Shimadzu, Kyoto, Japan). Fuel was diluted prior to injection using GC grade hexane in a 1:10 ratio. All fuel samples were run in triplicate (*i.e.* technical replicate). The autosampler injection syringe was rinsed after each 1 µL injection with 100% ethanol three times. After the ethanol rinse the syringe was rinsed once with the diluted fuel sample prior to the injection. A 1:10 split injection was used, representing a final dilution of all fuel samples of 1:100.

A 30.0 m Rtx-5MS column (Restek) with a nominal thickness and diameter of 0.25 µm was used. High purity helium at a linear velocity of 36.8 cm per second was used as the carrier gas. GC injection temperature was set at 300°C while the initial column temperature was 40.0

°C. Column temperature was held at 40.0 °C for 1.5 minutes upon injection. After an initial hold the temperature of the column increased by 10 °C per minute to a final temperature of 300 °C, which was held for 2.75 minutes. Mass spectra were analyzed in scan mode with event times of 0.25 seconds with a scan speed of 2000 and a solvent cut time of 2.75 min, an ion source temperature of 200 °C and an interface temperature of 300 °C. Spectra between m/z 35 and 500 were used for all downstream analyses.

External standards were obtained for FAME (Supelco 37 FAME Mix) and n-alkanes (C7-C40 Saturated Alkanes Standard) to quantify the concentration of fuel components in all samples (in parts per million, or ppm). A three-point standard curve for each reference compound was generated. All standards were diluted in hexane in ratios of 1:50, 1:25, and 1:10. The limit of detection of reference standards was 0.2 ppm. Ion intensities in each fuel sample were then compared to the generated standard curves to produce concentration estimates of FAME and n-alkane components.

Analysis of SSU rRNA Gene Sequencing Libraries.

Small subunit rRNA gene analyses were carried out in QIIME2 version 2019.10.0 (47). Briefly, barcodes from the sequences were extracted using fastx truncate in the usearch64 software package, and samples were demultiplexed prior to operational taxonomic unit (OTU) clustering. Chimeric sequences and singletons were filtered out prior to clustering using the USEARCH64 reference database. Representative sequences for each OTU were assigned a taxonomic identity with the QIIME2 sklearn classifier against the SILVA 132 database clustered to a 97% identity (48, 49). Following this, sequences were separated to generate feature tables that contained separated 16S and 18S rRNA gene data. After OTU generation and sequence alignment using MAFFT, 16S rRNA and 18S rRNA phylogenetic trees were generated using

FastTree 2 (50). Datasets were rarefied to account for variable library size between samples to 1000 sequences per sample for bacteria based on rarefaction curves and 250 sequences per sample for Eukarya due to the lower sequence and diversity in these samples (54). These data were then converted to BIOM files and exported for further analyses in R (51).

Redundancy Analyses (RDA).

Fuel and taxonomic relative abundance data were used to generate redundancy analyses to determine how fuel components correlated to microbial community composition. Taxonomic data generated in QIIME 2 were imported into R and transformed using the Hellinger method to account for the sparsity of amplicon data (55). Transformed taxonomy data were then used in a forward selection of fuel variables, carried out using a permutational multivariate analysis of variance using distance matrices test in R using the function ‘adonis’ within the vegan package to determine which fuel components were significantly correlated ($p < 0.05$) with community composition. An RDA was then generated by using the transformed community data and those fuel components that were significantly correlated to microbial community composition.

Validation of Fungal RDA Model

FAME components that were strongly correlated with *Trichocomaceae* and *Debaryomycetaceae* from the RDA model were used to determine how well representative organisms isolated from contaminated fuels could grow on individual FAMES as a sole carbon and energy substrate. Prior work from our lab yielded the fungal isolates *Paecilomyces* AF001 and *Wickerhamomyces* SE3 which are representative of the fungal families *Trichocomaceae* and *Debaryomycetaceae* respectively (13). A spore suspension of *Paecilomyces* AF001 and a cell suspension of *Wickerhamomyces* SE3 were used to inoculate a 1:20 FAME substrate in artificial

461 sump water (ASW, per L: 0.015g NaCl, 0.035g NaF, 0.02g CaCl₂, 0.018g KNO₃, 0.01g Na₂SO₄,
462 0.015g (NH₄)₂SO₄, and 0.017g K₂HPO₄) (51).

463 To prepare a spore suspension, a glycerol stock of *Paecilomyces* AF001 was struck onto
464 Hestrin Schramm (HS) agar medium (per L: 20g glucose, 5g yeast extract, 5g peptone, 2.7g
465 Na₂HPO₄, 1.15g citric acid, 7.5g Agar; pH adjusted to 6.0 with diluted HCl or NaOH) and
466 incubated at 25 °C for 7 days (51) After 7 days 5mL of phosphate buffered saline was added
467 overtop the HS agar containing hyphal growth of *Paecilomyces* AF001 and an inoculating loop
468 was used to scrape off the fungal growth to produce a suspension of cells and spores. The fungal
469 suspension was then collected from the agar plate and filtered through a 10 µm nominal pore size
470 polyether sulfone filter to separate spores from the hyphal biomass. The filtrate containing spores
471 was then centrifuged at 10,000 x RCF for 1 minute. The supernatant was decanted, and sterile
472 PBS was added back to the spore pellet and vortexed to resuspend the spores. This wash step was
473 repeated for a total of three washes. *Paecilomyces* spore concentrations were determined using a
474 Petroff-Hausser counting chamber and then diluted in ASW to adjust the final inoculum
475 concentration to 1x10⁴ spores/mL.

476 To produce a suspension of yeast cells *Wickerhamomyces* SE3 was grown in HS broth for
477 48 hours and centrifuged at 10,000 RCF to pellet cell mass. Following centrifugation, the
478 supernatant was decanted, sterile PBS was added to the cell pellet, and finally vortexed to
479 resuspend the cells. This was step was repeated for a total of three washes. Cell concentrations
480 were again determined using a Petroff-Hausser counting chamber and diluted to adjust the initial
481 inoculum concentration to 1x10⁴ cells/mL.

482 *Paecilomyces* AF001 and *Wickerhamomyces* SE3 were grown on palmitoleic and
483 pentadecanoic acid methyl esters (Fisher Scientific, Waltham, MA) to determine their capacity

on using these substrates for growth. A total volume of 5 mL 1:20 FAME and artificial sump water mixture was made and inoculated with 10^4 spores or cells per mL for T=0. Growth was measured by destructively sampling triplicate test tubes at time points 0, 6, 6.25, 6.5, 6.75, 7, 7.25, 7.5, 7.75, 8, and 14 days at room temperature and generating MPNs.

Statistical Analyses and Data Visualization.

Statistical analyses and figure generation was carried out in R version 3.3.3 and GraphPad Prism 8.3.0. Significant differences from fuel components that contributed to microbial community structure were determined via a permutational multivariate analysis of variance using distance matrices on the Hellinger transformed taxonomic data.

Acknowledgements

We acknowledge the men and women of the US Air Force and Civilian Personnel at US Air Force bases; their cooperation and assistance were critical to this research. Additionally, we would like to thank Dr. Emily Junkins for insightful comments during the development of this manuscript. This work was supported by the Air Force Research Laboratory Biological Materials and Processing Research Team, Materials and Manufacturing Directorate and the U.S. Department of Defense Office of Corrosion Policy & Oversight Technical Corrosion Collaboration grant to B.S.S. (Grant # FA7000-15-2-0001).

Data Availability Statement

The raw sequence data generated for this study can be found in the Sequence Read Archive under the accession numbers SRR5826605–SRR5826609.

506 **Works Cited**

- 507 1. U.S. Energy Information Administration. 2021. Annual Energy Outlook 2021 Petroleum PDF.
- 508 2. U.S. Department of Energy. 2021. Biodiesel Income Tax Credit.
- 509 3. U.S. Department of Energy. 2021. Alternative Fuels Data Center.
- 510 4. U.S. Department of Energy. 2021. Biodiesel Benefits and Considerations.
- 511 5. Demirbaş A. 2008. Biodegradability of biodiesel and petrodiesel fuels, Energy Sources, Part
512 A: Recovery, Utilization, and Environmental Effects, 31:169-174.
- 513 6. Mariano A, Tomasella R, Oliveira L, Angelis D. 2008. Biodegradability of diesel and
514 biodiesel blends. Afr J Biotechnol 7:1323-1328.
- 515 7. Dodos GS, Zannikos F. 2013. Microbiological Growth Study of Biodiesel Fuel. Sae Int J Fuels
516 Lubr 6:419–429.
- 517 8. Moser B. 2009. Biodiesel production, properties, and feedstocks. In Vitro Cell Dev-Pl 45:229-
518 266.
- 519 9. Fuller S, Spikmans V, Vaughan G, Guo C. 2013. Effects of Weathering on Sterol, Fatty Acid
520 Methyl Ester (FAME), and Hydrocarbon Profiles of Biodiesel and Biodiesel/Diesel Blends.
521 Environ Forensics 14:42–49.
- 522 10. Owsianiak M, Chrzanowski Lukasz, Szulc A, Staniewski J, Olszanowski A, Olejnik-Schmidt
523 AK, Heipieper HJ. Biodegradation of diesel/biodiesel blends by a consortium of hydrocarbon

524 degraders: Effect of the type of blend and the addition of biosurfactants. *Bioresource Technol*
525 100:1497-1500.

526 11. Rojo F. 2010. Enzymes for aerobic degradation of alkanes, p 781-797. In Timmis KN (eds),
527 Handbook of hydrocarbon and lipid microbiology, Springer, Berlin, Heidelberg.

528 12. Stamps BW, Bojanowski CL, Drake CA, Nunn HS, Lloyd PF, Floyd JG, Emmerich KA,
529 Neal AR, Crookes-Goodson WJ, Stevenson BS. 2020. In situ Linkage of Fungal and Bacterial
530 Proliferation to Microbiologically Influenced Corrosion in B20 Biodiesel Storage Tanks. *Front*
531 *Microbiol* 11:197.

532 13. Floyd JG, Stamps BW, Goodson WJ, Stevenson BS. 2021. Locating and Quantifying Carbon
533 Steel Corrosion Rates Linked to Fungal B20 Biodiesel Degradation. *Appl Environ Microb*
534 *AEM0117721*.

535 14. Lee J, Ray R, Little BJ. 2010. An assessment of alternative diesel fuels: microbiological
536 contamination and corrosion under storage conditions. *Biofouling* 6:623-635.

537 15. Hoekman KS, Broch A, Robbins C, Cenicerros E, Natarajan M. 2012. Review of biodiesel
538 composition, properties, and specifications. *Renew Sustain Energy Rev* 16:143–169.

539 16. Peraza-Reyes L, Berteaux-Lecellier V. 2013. Peroxisomes and sexual development in fungi.
540 *Front Physiol* 4:244.

541 17. Sousa DZ, Balk M, Alves M, Schink B, McInerney MJ, Smidt H, Plugge CM, Stams AJM.
542 2010. Handbook of Hydrocarbon and Lipid Microbiology 963–980.

- 543 18. National Academics of Sciences, Engineering, and Medicine. Spills of Diluted Bitumen from
544 Pipelines: A Comparative Study of Environmental Fate, Effects, and Response, p. 21–33. *In*
545 Chemical and Physical Properties of Crude Oils. The National Academies Press, Washington,
546 DC.
- 547 19. Demirbas A. 2008. Distillation Properties of Various Diesel Oils. Energy Sources, Part A:
548 Recovery, Utilization, and Environmental Effects 30:1484–1490.
- 549 20. Pitz WJ, Mueller CJ. 2011. Recent progress in the development of diesel surrogate fuels.
550 Prog Energ Combust 37:330–350.
- 551 21. Hasinger M, Scherr K, Lundaa T, Brauer L, Zach C, Loibner A. 2012. Changes in iso-and n-
552 alkane distribution during biodegradation of crude oil under nitrate and sulphate reducing
553 conditions. J Biotechnol 157:490-498.
- 554 22. Ciric L, Philp J, Whiteley AS. 2010. Hydrocarbon utilization within a diesel-degrading
555 bacterial consortium. FEMS Microbiol Lett 2:116-122.
- 556 23. Konopka A, Oliver L, Turco, Jr. RF. 1998. The Use of Carbon Substrate Utilization Patterns
557 in Environmental and Ecological Microbiology. Microbial Ecol 35:103–115.
- 558 24. Rinnan R, Bååth E. 2009. Differential Utilization of Carbon Substrates by Bacteria and Fungi
559 in Tundra Soil. Appl Environ Microb 75:3611–3620.
- 560 25. Xing W, Lu X, Xu F, Ying J, Chen D, Bai Y. 2019. Linking microbial community structure
561 to carbon substrate chemistry in soils following aboveground and belowground litter additions.
562 Appl Soil Ecol 141:18–25.

563 26. Liu Q, Tang J, Gao K, Gurav R, Giesy JP. 2017. Aerobic degradation of crude oil by
564 microorganisms in soils from four geographic regions of China. *Scientific Reports* 7:14856.

565 27. Jia R, Unsal T, Xu D, Lekbach Y, Gu T. 2019. Microbiologically influenced corrosion and
566 current mitigation strategies: A state of the art review. *Int Biodeter Biodegr* 137:42–58.

567 28. Goldfarb KC, Karaoz U, Hanson CA, Santee CA, Bradford MA, Treseder KK, Wallenstein
568 MD, Brodie EL. 2011. Differential Growth Responses of Soil Bacterial Taxa to Carbon
569 Substrates of Varying Chemical Recalcitrance. *Front Microbiol* 2:94.

570 29. Williamson CHD, Jain LA, Mishra B, Olson DL, Spear JR. 2015. Microbially influenced
571 corrosion communities associated with fuel-grade ethanol environments. *Appl Microbiol Biot*
572 99:6945–6957.

573 30. Kotzekidou P. 2014. *Paecilomyces*. In Batt C and Tortorello ML (eds) *Encyclopedia of Food*
574 *Microbiology* (Second Edition). Article Titles: B 344–350.

575 31. Wang W, Fan G, Li X, Fu Z, Liang X, Sun B. 2020. Application of *Wickerhamomyces*
576 *anomalous* in Simulated Solid-State Fermentation for Baijiu Production: Changes of Microbial
577 Community Structure and Flavor Metabolism. *Front Microbiol* 11:598758.

578 32. Passman FJ. 2013. Microbial contamination and its control in fuels and fuel systems since
579 1980 – a review. *Int Biodeter Biodegr* 81:88–104.

580 33. Thomas AO, Leahy MC, Smith JWN, Spence MJ. 2017. Natural attenuation of fatty acid
581 methyl esters (FAME) in soil and groundwater. *Q R Eng Geol Hydroge* 50:301-317.

582 34. Xu X, Liu W, Tian S, Wang W, Qi Q, Jiang P, Gao X, Li F, Li H, Yu H. 2018. Petroleum
583 Hydrocarbon-Degrading Bacteria for the Remediation of Oil Pollution Under Aerobic
584 Conditions: A Perspective Analysis. *Front Microbiol* 9:2885.

585 35. U.S. Environmental Protection Agency. 2016. Investigation of Corrosion-Influencing Factors
586 In Underground Storage Tanks With Diesel Service.

587 36. Xue J, Wu Y, Liu Z, Li M, Sun X, Wang H, Liu B. 2017. Characteristic Assessment of
588 Diesel-degrading Bacteria Immobilized on Natural Organic Carriers in Marine Environment: the
589 Degradation Activity and Nutrient. *Sci Rep-uk* 7:8635.

590 37. Révész F, Figueroa-Gonzalez PA, Probst AJ, Kriszt B, Banerjee S, Szoboszlai S, Maróti G,
591 Tánácsics A. 2020. Microaerobic conditions caused the overwhelming dominance of
592 *Acinetobacter* spp. and the marginalization of *Rhodococcus* spp. in diesel fuel/crude oil mixture-
593 amended enrichment cultures. *Arch Microbiol* 202:329–342.

594 38. Fuentes S, Barra B, Caporaso JG, Seeger M. 2016. From Rare to Dominant: A Fine-Tuned
595 Soil Bacterial Bloom during Petroleum Hydrocarbon Bioremediation. *Appl Environ Microb*
596 82:888–896.

597 39. Ye C, Ching TH, Yoza B, Masutani S, Li Q. 2017. Cometabolic degradation of blended
598 biodiesel by *Moniliella wahieum* Y12T and *Paecilomyces nivea* M1. *Int Biodeter Biodegr*
599 125:166-169.

600 40. Berg JM, Tymoczko JL, Stryer L. 2002. Certain Fatty Acids Require Additional Steps for
601 Degradation. *In Biochemistry* 5th edition. 22.3.

- 602 41. Poret-Peterson AT, Albu S, McClean AE, Kluepfel DA. 2019. Shifts in Soil Bacterial
603 Communities as a Function of Carbon Source Used During Anaerobic Soil Disinfestation.
604 *Frontiers Environ Sci* 6:160.
- 605 42. Ghimire R, Norton JB, Stahl PD, Norton U. 2014. Soil Microbial Substrate Properties and
606 Microbial Community Responses under Irrigated Organic and Reduced-Tillage Crop and Forage
607 Production Systems. *Plos One* 9:103901.
- 608 43. Liu M, Sui X, Hu Y, Feng F. 2019. Microbial community structure and the relationship with
609 soil carbon and nitrogen in an original Korean pine forest of Changbai Mountain, China. *Bmc*
610 *Microbiol* 19:218.
- 611 44. Mitter E.K., J.J. & de Freitas, J.R. 2021. Impact of diesel and biodiesel contamination on soil
612 microbial community activity and structure. *Sci Rep* 11:10856.
- 613 45. Andreolli, M., Albertarelli, N., Lampis, S., Brignoli P., Khoei N.S., Vallini G. 2016.
614 Bioremediation of diesel contamination at an underground storage tank site: a spatial analysis of
615 the microbial community. *World J Microbiol Biotechnol* 32:6.
- 616 46. Parada A, Needham DM, Fuhrman JA. 2016. Every base matters: assessing small subunit
617 rRNA primers for marine microbiomes with mock communities, time series and global field
618 samples. *Environ Microbiol* 5:1403-1414.
- 619 47. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, Alexander H,
620 Alm EJ, Arumugam M, Asnicar F, Bai Y, Bisanz JE, Bittinger K, Brejnrod A, Brislawn CJ,
621 Brown CT, Callahan BJ, Caraballo-Rodríguez AM, Chase J, Cope EK, Silva RD, Diener C,

622 Dorrestein PC, Douglas GM, Durall DM, Duvallet C, Edwardson CF, Ernst M, Estaki M,
623 Fouquier J, Gauglitz JM, Gibbons SM, Gibson DL, Gonzalez A, Gorlick K, Guo J, Hillmann B,
624 Holmes S, Holste H, Huttenhower C, Huttley GA, Janssen S, Jarmusch AK, Jiang L, Kaehler
625 BD, Kang KB, Keefe CR, Keim P, Kelley ST, Knights D, Koester I, Kosciulek T, Kreps J,
626 Langille MGI, Lee J, Ley R, Liu Y-X, Loftfield E, Lozupone C, Maher M, Marotz C, Martin
627 BD, McDonald D, McIver LJ, Melnik AV, Metcalf JL, Morgan SC, Morton JT, Naimey AT,
628 Navas-Molina JA, Nothias LF, Orchanian SB, Pearson T, Peoples SL, Petras D, Preuss ML,
629 Priesse E, Rasmussen LB, Rivers A, Robeson MS, Rosenthal P, Segata N, Shaffer M, Shiffer A,
630 Sinha R, Song SJ, Spear JR, Swafford AD, Thompson LR, Torres PJ, Trinh P, Tripathi A,
631 Turnbaugh PJ, Ul-Hasan S, Hooft JJJ van der, Vargas F, Vázquez-Baeza Y, Vogtmann E, Hippel
632 M von, Walters W, Wan Y, Wang M, Warren J, Weber KC, Williamson CHD, Willis AD, Xu
633 ZZ, Zaneveld JR, Zhang Y, Zhu Q, Knight R, Caporaso JG. 2019. Reproducible, interactive,
634 scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol* 37:852–857.

635 48. Edgar RC. 2013. uParse: highly accurate otu sequences from microbial amplicon reads.
636 *Nature Methods* 10:996-998.

637 49. Yilmaz P, Parfrey L, Yarza P, Gerken J, Priesse E, Quast C, Schweer T, Peplies J, Ludwig
638 W, Glockner F. 2013. The SILVA ribosomal RNA gene database project: improved data
639 processing and web-based tools. *Nucl. Acids Res.* 41:590-596.

640 50. Price MN, Dehal PS, Arkin AP. 2010. FastTree 2 – Approximately Maximum-Likelihood
641 Trees for Large Alignments. *Plos One* 5:9490.

- 642 51. R Core Team. 2018. A language and environment for statistical computing. R Foundation for
643 Statistical Computing. Vienna, Austria.
- 644 52. McNamara CJ, Thomas PD, Leard R, Bearce K, Dante J, Mitchell R. 2006. Corrosion of
645 aluminum alloy 2024 by microorganisms isolated from aircraft fuel tanks. *Biofouling* 21:257–
646 265.
- 647 53. Hestrin S, Schramm M. 1954. Synthesis of cellulose by *Acetobacter xylinum*. 2. Preparation
648 of freeze-dried cells capable of polymerizing glucose to cellulose. *Biochem J* 58:345-352.
- 649 54. Cameron E, Schmidt P, Trmblay B, Emelko M, Muller K. 2020. To rarefy or not to rarefy:
650 Enhancing diversity analysis of microbial communities through next-generation sequencing and
651 rarefying repeatedly. *bioRxiv* 2020.09.09.290049.
652
- 653 55. Legendre P, Gallagher ED. 2001. Ecologically meaningful transformations for ordination of
654 species data. *Oecologia*. 129(2): 271-280

Color Key

A **B**