Microbial sources of exocellular DNA in the ocean

Morgan Linney (mlinney@hawaii.edu)
Daniel K. Inouye Center for Microbial Oceanography: Research and Education (C-MORE)

John Eppley
University of Hawaii at Manoa https://orcid.org/0000-0001-5334-1704

Anna Romano
University of Hawaii at Manoa

Elaine Luo
University of Hawaii at Manoa

Edward DeLong
Daniel K. Inouye Center for Microbial Oceanography: Research and Education, University of Hawai□i at Mānoa, Honolulu, HI 96822 https://orcid.org/0000-0002-3088-4965

David Karl
University of Hawaii at Manoa https://orcid.org/0000-0002-6660-6721

Article

Keywords: dissolved DNA (D-DNA), vesicles, viruses, exocellular free DNA (F-DNA)

Posted Date: May 5th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-431074/v1

Version of Record: A version of this preprint was published at Applied and Environmental Microbiology on March 21st, 2022. See the published version at https://doi.org/10.1128/aem.02093-21.

Abstract

Dissolved DNA (D-DNA) is operationally defined as DNA that passes through a 0.1 micrometer filter. It is composed of vesicles, viruses, and exocellular free DNA (F-DNA) and is ubiquitous in all aquatic systems, although the sources, sinks, and ecological consequences are largely unknown. Using a method that provides separation of vesicles, viruses, and F-DNA, we compared open ocean depth profiles of metagenomic DNA associated with each fraction. Pelagibacter-like DNA dominated the vesicle fractions for all samples examined over a depth range of 75–500 m. The viral DNA fraction was composed predominantly of myovirus-like and podovirus-like DNA, and contained the highest proportion of unannotated sequences. Euphotic zone F-DNA (75–125 m) contained mostly bacterial and viral sequences, with bacteria dominating in the mesopelagic (500–1000 m). A high proportion of mesopelagic (500 and 1000 m) F-DNA sequences appeared to originate from surface waters, including a large amount of DNA contributed by high-light *Prochlorococcus* species. Throughout the water-column, but especially in the mesopelagic, the composition of F-DNA sequences was not always reflective of co-occurring microbial communities that inhabit the same sampling depth. These results highlight the composition of F-DNA in different regions of the water-column (euphotic and mesopelagic zones), with implications for dissolved organic matter cycling and export (by way of sinking particles and/or migratory zooplankton) as a delivery mechanism.

Introduction

Nearly half of the DNA in the open ocean is present outside of living organisms (1-4). It is contained inside of protein-encapsidated viruses, extracellular vesicles, and as the naked molecule itself, free DNA (F-DNA). Together this DNA that passes a typical filter (0.2-0.1 μm) is dissolved (D-DNA), the cycling of labile DNA in the ocean has large-scale consequences for global nutrient cycling, export, and potential for genetic exchange. DNA in the ocean has been studied from multiple oceanographic perspectives (e.g. studying environmental DNA for fisheries management, sequencing cellular DNA to identify taxonomic relationships and to quantify population diversity, quantifying viral diversity and dynamics), making it challenging to disentangle the various overlapping terminology. While, environmental DNA (eDNA) can have similar motives of holistically studying an ecosystem, by collecting the remnants of large macro-organisms (e.g. scales, mucus, and reproductive cells) to better understand which organisms are present in a given study site (5), it employs the same collection method as cellular analyses (6). DNA such as this, that is captured on a filter has been termed, “particulate” DNA (P-DNA; 2), distinguishing it from DNA that passes a filter (D-DNA; 7). D-DNA can be thought of as predominantly, “microbial eDNA,” the manifestation of microbial dynamics (e.g. viral lysis, autolysis, and grazing) and activity in the ocean.

In the open ocean, subtropical gyres comprise 40% of the earth’s surface, yet surface waters there are deplete of essential elements like nitrogen and phosphorus (8), making D-DNA nutrient-rich organic nutrients in this system. It can be taken up as whole polydeoxyribonucleotides (>10 kb) (9-11), as well as, hydrolyzed and broken down to nucleotides and nucleic acid bases (12-14). Alternatively, the genomic information encoded by D-DNA has the potential to be integrated by natural transformation (15).

Streamlined genomes are common in oligotrophic systems (16), and are found in many of the numerically dominant microorganisms (e.g. photoautotrophic *Prochlorococcus*, heterotrophic *Pelagibacter*, UCYN-A, as well as others) inhabiting the North Pacific Subtropical Gyre (16-19). While these microorganisms have an advantage in a relatively stable and nutrient-deplete environment (20), they are more vulnerable to extinction in the face of environmental changes, than organisms with larger genomes (21). Therefore, genetic exchange is especially important to the survival of microorganisms like *Prochlorococcus* and *Pelagibacter*, evident by variable genomic islands and large pangenomes that are thought to have emerged by horizontal gene transfer (HGT; 22-23). However, the common mechanisms (e.g. conjugative systems, plasmids, or transposons) of HGT have yet to be confirmed in many of these organisms. In oligotrophic environments, D-DNA may be a source of nutrients or an agent of intracellular genetic exchange.

One of the largest ecosystems on earth, the North Pacific Subtropical Gyre (15°N to 35°N latitude and 135°E to 135°W) is home to one of the most well-studied water columns (8) and exhibits relatively low seasonality (24). Over thirty years of biogeochemical research provides a wealth of interdisciplinary baseline data: hydrographic data and its implications for thermohaline circulation, metabolic balances, nutrient cycling, cultured microbial isolates, complex microbial interactions, and export dynamics. More recently, gene catalogs of both cellular and viral communities throughout the water column at Station ALOHA have been reported (25), broadening understanding of the distributional patterns of diverse microbial and viral communities in the open ocean.

While vesicles (26) and viruses, two of the three D-DNA constituents have been previously examined in open ocean systems, the biological composition and depth origins of F-DNA is relatively uncharacterized. Primarily due to prior D-DNA isolation methods that utilized low-volume filtration (27) and chemical precipitation (7, 28) to isolate bulk D-DNA, these methods pose challenges in acquiring enough DNA to assemble metagenomic libraries for sequencing, leaving at least 25% of the open ocean DNA inventory uncharacterized. The average proportion of D-DNA that is either F-DNA, in viruses, or in vesicles is, 45, 35, and 20%, respectively (29).

To investigate all three D-DNA pools, and their potential contributions to biogeochemistry throughout the water column, we analyzed all three D-DNA pools throughout the water column of the North Pacific Subtropical Gyre. Employing a method (29) that separates the three pools by density, DNA from vesicle, virus, and F-DNA fractions was sequenced and compared. This study provides the first genetic characterization of the F-DNA pool, alongside the other D-DNA pools, and reveals viral lysis as a plausible source of F-DNA, and export as a delivery mechanism of DNA to mesopelagic depths.

Results

A. Microbial composition of exocellular free DNA and other D-DNA pools

Summed across six F-DNA samples collected in the North Pacific Subtropical Gyre, the majority of annotated genes in DNA-based metagenomic libraries were derived from bacteria and viruses (Figure 1),

with minimal contributions from Eukarya and Archaea (<1.0 and 1.1%, of all annotated sequences, respectively). The proportion of F-DNA virus sequences in surface waters ranged 10-35% above or at the deep chlorophyll maximum (75-125 m), about the same as the proportion of Bacteria (16-34%) over the same depth range. However, in the mesopelagic zone (250-1000 m), the proportion of F-DNA sequences from viruses was lower (2-16%) , whereas the proportion of bacterial-derived sequences was higher 38-45%. Across all F-DNA metagenomic libraries, the microbial assemblage was dominated (69-85% of family-level annotated sequences) by three main taxonomic groups: Pelagibacteraceae, Myoviridae, and Prochlorococcus (Figure 1); all of which have been documented as abundant microorganisms in the North Pacific Subtropical Gyre (6, 30, 31). Ubiquitous heterotrophs of the family Pelagibacteraceae contributed an average of 23% of F-DNA sequences over all depths sampled, with a range of 2-46%. Prochlorococcus represented 43, 42, and 53% of F-DNA sequences collected from 125, 500, and 1000 m, respectively, and represented an average 23% of family-annotated sequences. Of these Prochlorococcus sequences (125, 500, and 1000 m) more than 85% were from the high-light (HL) ecotypes (Table S2). The highest proportion of Myoviridae F-DNA sequences was observed in the upper euphotic zone, 45% (75 m) and 69% (100 m). Of the viral-derived sequences constituting the F-DNA samples, the majority (>90% of annotated viral sequences with known hosts) were dominated by viruses known to infect *Synechococcus* (38% ±4.6), *Prochlorococcus* (31% ±4.8), Pelagibacter (16% ±7.2), and other Cyanobacteria (7.9% ±1.3) (Figure S1). *Synechococcus* viruses peaked at the deep chlorophyll maximum on two occasions (43%). Other viruses known to infect SAR116 (5.4% ±2.2) and Vibrio (1.1% ±0.6), were present at much lower proportions in the F-DNA. F-DNA-derived Vibrio sequences were highest in mesopelagic samples (250; 2.3%, 500; 1.0%, 1000; 1.4%), but were only <1% in euphotic samples.

Among other taxonomic groups contributing to metagenomic F-DNA libraries, Archaeal contribution was minimal (ranged 0.18-2.03% for all samples). Of those, the most were most highly similar to those of ammonia-oxidizing Thaumarchaeota (0.02-1.6%), which were most prevalent at depths greater than 100 m. Of the three domains contributing to F-DNA metagenomes, Eukarya was the lowest (0.28-1.45%). The taxonomic families that were most abundant included heterokont Aureococcus, coccolithophore Noelaerhabdaceae, and Bathycoccaceae.

DNA from the vesicle D-DNA samples collected throughout the euphotic and mesopelagic zones (75-500 m; Figure 1) was also sequenced and compared. Summed across all samples, Bacteria contributed overwhelmingly to these metagenomic libraries (77% ±22 all annotated sequences). Viruses, Archaea, and Eukaryota contributed an average of 18.4, 3.7, 0.6%, respectively to all annotated sequences (Table S1). Of the bacterial sequences, at the family-level sequences were heavily dominated by *Pelagibacter ubique* (81% ±8 of family-level annotated sequences), with only 34% ±12 of sequences left unannotated, the lowest of all three D-DNA fractions. Other bacterial sequences that contributed to the vesicle samples, include those from Rhodospirillaceae (2.9%), Rhodobacteraceae (2.7%), Flavobacteriaceae (1.0%), and Prochlorococcus (0.5%). Viral sequences in the vesicle metagenomic DNA libraries were highest in the upper euphotic zone and were dominated by both Podoviruses (7%) and Myoviruses (6%).

Averaged across all virus fraction metagenomic libraries, these samples were dominated by annotated sequences derived from viruses (Figure 1), consistent with previous reports of this D-DNA fraction utilizing transmission electron micrographs and epifluorescence analyses (29). The viral metagenomic libraries had the lowest number of recovered sequences, many of which were novel and unannotated (61-73% unannotated across all samples). Recovered genes ranged from 13,080,771-15,551,348, and averaged at 13,822,361 (±1,071,408) (Table S1). For all samples, viral family-level annotated sequences were nearly split between Myoviruses (23-38%) and Podoviruses (19-32%), with minimal contributions from Siphoviruses (4-7%). Of the metagenomic libraries contributing to the viral libraries, Prochlorococcus phages, other Cyanophages and Pelagibacter phages were the most abundant (Figure S1). Synechococcus and Prochlorococcus phages dominated the euphotic samples (75-125 m). At the deep chlorophyll maximum, Synechococcus and Cyanophage sequences in the virus fraction metagemomes peaked. In mesopelagic samples, Pelagibacter phages and Vibro phages increased in proportion in the virus fraction. These depths are consistent with both cellular host and as well as virus abundances previously reported at Station ALOHA (6, 31, 32).

Overall, the DNA-based metagenomic libraries developed from the three dissolved DNA pools were distinct with respect to their microbial DNA compositions. The vesicle fraction was primarily dominated by a single taxonomic family (Pelagibacteraceae) across all depths, the viral fraction was dominated by bacteriophages, and lastly the exocellular F-DNA pool had both bacterial and viral derived DNA. While DNA from the former two pools have been previously described by metagenomic analyses, the composition DNA in the F-DNA fraction has not been previously reported.

B. Depth of origin of D-DNA throughout the water column

To infer the depths of origin of different D-DNA fractions, we mapped DNA sequences against a depth-resolved microbial gene catalogue from Station ALOHA (Figure 2; 6, 25, 31). The objective was to determine whether genes from the D-DNA fractions matched Station ALOHA genes recovered from the same sampling depths as the D-DNA, or whether the D-DNA was potentially transported from other depths.

The viral samples were most similar to Station ALOHA annotated genes, that matched the depth at which they were collected. This was particularly evident in the mesopelagic zone samples collected from 250-500 m (Figure 2). In these samples 25% of the genes were derived from their respective collection depths, with only 13% from the euphotic zone (5-200 m). Viral samples collected from the DCM had high contributions from typical DCM depths (100-175 m; 29-36%), as well as neighboring upper euphotic (5-75 m; 15%) and upper mesopelagic zones (200-250 m; 13%), with only 1% from the lower mesopelagic zone (500-1000 m). Of the three D-DNA fractions, the viral samples had the highest average of genes with unknown depths (43% ±4.5%).

In contrast to the viral samples, the vesicle and F-DNA samples appeared to contain both autochthonous and allochthonous DNA (Figure 2). In the euphotic zone samples (75-125 m), as expected, sequences were dominated (>50%) by surface-derived DNA (5-200 m), with minimal mesopelagic zone contributions

(<10%). However, in mesopelagic zone vesicle and F-DNA samples (250-1000 m), genes originated primarily from the upper euphotic zone (5-75 m; >30%), and to a lesser extent (<20%) the depth from which they were collected. Of these mesopelagic zone samples, the shallowest F-DNA sample (250 m) had the most depth-diverse genes, originating from depths throughout the euphotic and mesopelagic zones (5-500 m).

C. Size distributions of environmental F-DNA through the water column

The size spectra of recovered F-DNA was measured by capillary electrophoresis, following density gradient separation and buffer exchange. Seven F-DNA samples collected throughout the euphotic and mesopelagic zones (5-1000 m) were measured to assess the degradation of samples and molecular weight distributions prior to sequencing (Figure 3). Samples collected in the upper euphotic zone (5-100 m) had a distinct peak (<5000 bp peak width) of high molecular weight F-DNA (HMW; referred to here as >1,000 bp) and a lower proportion of low molecular weight (LMW; <1,000 bp) F-DNA, ranging between 24-38%, compared to mesopelagic zone (250-1000 m) samples which ranged between 33-65%. In these upper euphotic zone samples, there was a high proportion of F-DNA 1000-40,000 bp (62-73%), whereas in mesopelagic zone samples this HMW F-DNA tended to be lower (48% average, 35-66%). Lower euphotic (125 m) and mesopelagic zone samples (250-1000 m) tended to have a broader range of F-DNA sizes, suggesting that this DNA may have been degraded. In mesopelagic zone samples, the HMW F-DNA decreased, the maximum peaks of HMW-DNA in samples 250-1000 m were less distinct (>10,000 bp peak width) and there was more F-DNA between peaks. In samples 5-100 m, <25% of the DNA was <350 bp. At 1000 m the peaks were unpronounced, suggesting a notable level of degradation in this deep sample.

D. Comparing vesicle, viral, and F-DNA fractions by non-metric multidimensional scaling

To compare all dissolved DNA fractions (vesicles, viruses, and F-DNA) with each other and previously reported Station ALOHA viral and P-DNA metagenomic sequences, two-dimensional ordination methods were employed. Bray-Curtis dissimilarity based non-metric multidimensional scaling (NMDS) of dissolved DNA and the Station ALOHA gene catalogue (Figure S2a & b, respectively; 0.02 µm filtered “viral” and 0.2 µm filtered “cellular” communities) microbial communities were compared on family-level annotated metagenomes. This comparison confirmed that the viral D-DNA samples collected in this study were similar in composition to previously characterized Station ALOHA virioplankton communities recovered from the same respective depths (Figure S2a, stress = 0.11). As for F-DNA sequences, upper euphotic zone samples (75-100 m) clustered with their respective Station ALOHA viral samples, whereas lower euphotic (125 m) and mesopelagic zone (250-1000 m) samples clustered together and not with their respective sample depths. Similarly, vesicle D-DNA did not cluster with any Station ALOHA viral samples.

The same D-DNA sequences were compared to the cellular microbial community genes utilizing NMDS (stress = 0.07; Figure S2b). From this analysis two distinct cellular communities emerged, euphotic (75-125 m) and mesopelagic (250-1000 m) Station ALOHA samples clustered together, in general consistent with previous reports (6). Viral D-DNA did not cluster with any Station ALOHA cellular communities. Similarly, F-DNA fractions did not cluster with their respective Station ALOHA depths, deep F-DNA

samples (500 and 1000 m) clustered with cellular communities filtered from 75 m, revealing a potential cellular origin of this F-DNA. Upper mesopelagic (250 m) F-DNA clustered nearest cellular metagenomes collected from 125 m. Surface (75 m) vesicle D-DNA clustered with euphotic zone cellular communities, whereas lower euphotic (125 m) and mesopelagic (500 m) vesicles samples clustered with mesopelagic Station ALOHA cellular samples. D-DNA samples that had high proportions of viral sequences (Figure 1) clustered together (Figure S2b; viral D-DNA, 100 m F-DNA, and total dissolved DNA from 100 and 250 m).

Discussion

Describing the various forms of D-DNA in the ocean is important in order to better understand the diversity of marine life and microbial dynamics. Here we show that D-DNA is comprised of three distinct pools: vesicles, viruses, and exocellular F-DNA. Recent work has characterized viral communities are found throughout the water column (31-33) as well as vesicles as a potential mode of mobile gene element transfer (34). The vesicle D-DNA pool was anticipated to be well represented by DNA from the cyanobacterium *Prochlorococcus* (26, 35), yet Prochlorococcus appeared to contribute to only 0.5-1% of the annotated sequences. Instead, the vesicle-derived DNA sequences were dominated by DNA derived from the ubiquitous heterotrophic bacterial family Pelagibacteraceae. The predominance of *Pelagibacter* in the vesicle pool may have several explanations. This DNA fraction could be derived in part from intact ultra-small *Pelagibacter* cells (<0.1 µm), or directly from *Pelagibacter* vesicles. It has been recently documented that structures resembling vesicles were produced by *Pelagibacter* (36, 37). There were several differences in our study design compared to previous vesicle studies in marine plankton (26,35), that may account in part for these results. These differences include our use of a 0.1 µm prefilter (compared to 0.2 µm prefilters used in previous studies), and our use of CsCl density gradients, compared to Iodixanol gradients (26, 35). Consistent with previous virome reports from the North Pacific Subtropical Gyre (31, 33), the virus fraction was dominated by viruses related to those that infect Prochlorococcus and Pelagibacter (Figure S1).

Our results reveal the first metagenomic characterization of F-DNA, alongside two other D-DNA pools, vesicles and viruses. The importance of further characterizing the F-DNA fraction is compounded by recent discoveries of the other D-DNA pools (viruses and vesicles) as potential vectors of genetic exchange in the marine environment (34) and cyanobacterial pili capable of utilizing exogenous DNA (38, 39).

Exocellular microbial-derived DNA is not unique to the open ocean (3, 7, 29, 40). Terminology may vary, but F-DNA has also been reported in marine sediments (41) and terrestrial soils (42), even contributing 90% of the total DNA pool in marine sediments (43). *Vibrio cholerae* are known to take up exogenous DNA from dead cells, which is thought to shape antibiotic resistance, surface colonization, and intercellular communication (44). Cell-free DNA has also been documented in the human bloodstream, in some cases it originates from tumor cells, and can be used as a noninvasive cancer diagnosis (45). Free nucleic acids in the form of viroids (RNA) are even capable of causing infection in higher plants (46).

Across microbial systems, a variety of functions have been attributed to cell-free DNA, exemplifying the vast evolutionary and ecological potential F-DNA may have in the open ocean.

In the open ocean, F-DNA accounts for 25-50% of total D-DNA (29), and has been shown to be rapidly consumed by microorganisms (12, 14, 40, 47). Whether it is used primarily as a nutrient source or for genetic exchange remains largely unknown. Previous work comparing the turnover of D-DNA pools, indicates that F-DNA is taken up more quickly than DNA inside of viruses (40, 48), suggesting it may be more readily available than structurally enclosed D-DNA. Our investigation into F-DNA reveals that there are different microbial sources for this material at the surface (75-125 m) versus the mesopelagic (250-1000 m), implicating distinct ecological and evolutionary consequences.

At the surface, F-DNA sequences were dominated by viral and bacterial sequences – a possible manifestation of active viral lysis. Viral lysis is thought to be the dominant mode of viral replication at the surface, while there is evidence (viral gene markers – integrase, repressor protein cI, and excisionase) that temperate phage may be more prevalent below the deep chlorophyll maximum (33). At depth, cell abundances and primary production are lower than at the surface (8), providing less energy and fewer hosts for viral replication (31, 49, 50). This depth-distinction in viral replication at Station ALOHA coincides with the marked decrease in viral-derived F-DNA sequences in the mesopelagic, the same depths fewer free phage sequences (0.02 µm filtered) are observed (33). Conversely, when cells are lysed the contents of the cell and viral machinery are exposed and spill into the open-ocean. Thus all DNA that was enclosed by the cellular membrane, both viral and cellular DNA may become F-DNA after lysis. Indeed, viral lysis has been documented as a production mechanism of both total D-DNA (51) and F-DNA (40, 48). Furthermore, an experiment documenting the production of free ribosomes following viral lysis of *Synechococcus,* supports the possibility that viral lysis may be an important process introducing other free nucleic acids into the marine environment (52). Interrupted phage packaging has not yet been explicitly described in marine systems, but there is evidence of unpackaged DNA from nuclease treatments following cyanophage infection of dominant picocyanobacteria (53,54), as well as in other microbial systems (55, 56). Together, our results suggest that viral lysis of bacterial cells at the surface may be important to the production of F-DNA in the open-ocean. Protozoan grazing (12, 51) and cell exudates (47, 57) have also been shown to result in increased D-DNA and might also play important roles in the production of F-DNA in the open-ocean. More research, separating the D-DNA pools is needed to uncover the source and dynamics behind these production mechanisms.

The high proportion of surface-originated sequences in mesopelagic F-DNA sequences, suggests downward export of either cells, or free-DNA, as a delivery mechanism (Figure 2). The exact mechanisms of F-DNA delivery to the mesopelagic have yet to be documented. Potential mechanisms suggested by previous work include: sediment trap and water column metagenomic analyses (6, 58), flux calculations of sinking particle disaggregation and degradation (59, 60), and migratory zooplankton (61). All of these mechanisms begin by defining the source of F-DNA, which presumably originates from autochthonous particulate DNA (P-DNA; includes cellular DNA, detrital DNA, and eDNA), reflecting the microbial communities present at the collection depth, or from allochthonous P-DNA. Recently, surface-originated P-

DNA (as filtered cells) has been documented in the mesopelagic. One such report found *Prochlorococcus* sequences filtered from sediment traps at Station ALOHA (58), and was hypothesized to be due to entrainment on particles or in fecal pellets. Additionally, a recent investigation of core microbiome populations at Station ALOHA consistently found high-light *Prochlorococcus* ecotypes originating from the upper euphotic, at 500-1000 m in the P-DNA at Station ALOHA (6). Consistent with this observation, our work demonstrates that high-light *Prochlorococcus* also dominate the F-DNA metagenomes at similar depths (Table S2), implicating a surface water export.

With this evidence of surface P-DNA exported from the surface, rapid surface turnover of free DNA, and mobilization by migratory zooplankton, it is presumed that the majority of DNA leaves the surface in the particulate phase, rather than dissolved. Following the sinking of P-DNA, it is hypothesized that particles disaggregate biologically and/or mechanically rather than on-particle degradation (60), and remineralization occurs by free-living or suspended microorganisms in the water-column (62). This theory is supported by studies investigating carbon flux, nutrient content of particles, on-particle metabolic activity, and accompanying models. Measurements of biomass production (particulate ATP) found there was a net loss of living material on particles (62). Alongside organic nutrient measurements it was concluded that sinking particles were not likely to be sites of active microbial decomposition. Similarly, Collins et al. (2015) compared on-particle metabolic activity to the metabolic demand of the water column itself. They found that the measurement of on-particle degradation could not account for particle flux attenuation, suggesting that microbial decomposition occurs in the water column rather than on particles.

By estimating the flux of particulate DNA out of the euphotic zone, delivery of DNA to mesopelagic depths (by route of sinking particles) can be estimated (Supplementary equation 1). This calculation assumes: (1) P-DNA becomes F-DNA by disaggregation rather than on-particle degradation (60); (2) the flux conforms to a Martin Curve (63); (3) adheres to export efficiency and flux attenuation values determined from decades of production and export analyses at Station ALOHA (64); and (4) steady-state. This calculation predicts that the daily flux of DNA at 500 m is approximately 12-24% of the concentrations of F-DNA reported at 500 m and 3-8% at 1000 m (Figure 4; Supplementary equation 1; 29, 40). This calculation is consistent with our finding of 19% and 10% of the 500 m and 1000 m F-DNA annotated sequences were derived from surface-originated *Prochlorococcus*. Despite their small size, picophytoplankton like *Prochlorococcus* that dominate primary production in oligotrophic oceans are theorized to contribute proportional carbon export values (59). This suggests that carbon exported from the surface ocean is likely to be directly or indirectly derived from *Prochlorococcus* and other primary producers, therefore surface-derived sequences are likely to contribute significantly to mesopelagic F-DNA.

However, analyses revealed that F-DNA from 500 and 1000 m had more surface-derived sequences than the Martin curve and field data (65, 66) predicts, 67 and 46% euphotic-derived (0-200 m) sequences, respectively. If we assume that gene annotations accurately reflect D-DNA depth origins, there remains a significant fraction unaccounted for, suggesting other mechanisms may supplement F-DNA delivery to

the mesopelagic. Delivery and F-DNA standing stocks may also depend on depth-dependent turnover rates of F-DNA, which are still not well constrained. Another important consideration is water mass controls on the supply of surface F-DNA. It is theorized based upon decades of salinity and temperature measurements at Station ALOHA (67) that the water column is comprised of distinct water masses. From this, it is hypothesized that surface waters in the northwestern Pacific supply the North Pacific Intermediate Water which manifests at Station ALOHA around 500-770 m (68). This water mass is estimated to have been in contact with the atmosphere only 30-60 years ago according to a study (69) that used transient tracers to estimate apparent water mass age at Station ALOHA. This provides evidence for other additional delivery mechanisms of surface DNA to the mesopelagic.

Other possible mechanisms delivering surface DNA to the mesopelagic may be more episodic in nature, for example: the passage and disaggregation of organic matter like the summer export pulse (70, 71), migratory zooplankton (61), and aggregates formed by picophytoplankton transparent exopolymer particles (72) and/or clay-ballasted particles (73). Recent analyses of abyssal sinking particles (collected at 4000 m) at Station ALOHA found that the organic carbon-specific energy was relatively high and energy-replete suggesting that deep organic carbon may be surface-derived (64). Together, our findings and past reports highlight that surface and mesopelagic microbial communities may be more vertically connected than previously thought. Investigations of the mechanisms delivering F-DNA and other labile DOM constituents, is critical for expanding our understanding of water column dynamics and the biological carbon pump.

Our research utilized a new method (29) for separating ecologically important pools of D-DNA, and preliminary analyses of DNA sequences found in water column profiles of the different fractions of vesicles, viruses, and exocellular F-DNA. The results highlight the potential ecological contributions of D-DNA with respect to the cycling of limiting nutrients (nitrogen and phosphorus) in the open-ocean, as well as carrying mobile genetic elements. While these results have important implications, more research into the dynamics of F-DNA is required to understand how this molecule is produced, its connection to microbially-mediated food webs, and its capacity to carry surface microbial genes to mesopelagic depths.

Methods

Sample collection

Seawater samples were collected during three cruises near Station ALOHA (22°45’N, 158°00’W), *R/V* Kilo Moana cruise HOT297 (November 2017), *R/V* Falkor cruise FK180310 (April 2018), and *R/V* Ka’imikai-O-Kanaloa cruise HOT302 (May 2018) using standard Niskin-type bottles. Dissolved DNA samples (D-DNA; vesicles, viruses, and free DNA) were collected and separated following (Linney et al. 2021). This method includes three primary steps: (1) prefiltration directly from the Niskin-type bottles through a double-layered 0.1 µm polyethersulfone (PES) capsule filter, (2) concentration by tangential flow ultrafiltration (30 kDa) down to ~1 mL (3) D-DNA separation by density gradient ultracentrifugation in cesium chloride.

Following separation, D-DNA constituents are buffer exchanged with either TE (vesicles and F-DNA: 10 mM Tris-HCl, 1 mM disodium EDTA, pH 7.5) or SM (viruses: 100 mM sodium chloride, 8 mM magnesium sulfate, 50 mM Tris-HCl pH 7.5) buffers. This method ensures that all D-DNA constituents are collected from the same sample, enabling comparison.

Molecular weight determination of environmental F-DNA

The size distributions of isolated F-DNA was analyzed by capillary electrophoresis (*Fragment Analyzer*™ Automated CE System; Advanced Analytical Technologies, Incorporated) with a 33 cm capillary using the High Sensitivity Genomic DNA Analysis Kit. Samples were run following the manufacturer's instructions (protocol DNF-488-33). Digital sample peaks (electropherograms) were generated by ProSize 2.0 software.

DNA extraction and purification

DNA was extracted from viruses, vesicles, and exocellular "free" DNA (100 µL each sample) by first lysing and digesting membrane in sucrose lysis buffer (final concentrations: 40 mM EDTA, 50 mM Tris (pH 8.3) and 0.75 M sucrose, and lysozyme (0.5 mg ml^{-1} final concentration), at 37°C for 30 min. Proteinase K (0.8 mg ml^{-1} final concentration) and SDS (0.8% final concentration), and the lysate was further incubated at 55°C for 2 hours. The DNA was purified using a Chemagen MSM I instrument with the Saliva DNA CMG-1037 kit (Perkin Elmer, Waltham, MA). DNA and sequencing library sample quantity and quality were assessed by PicoGreen dsDNA quantitation (Invitrogen, Waltham MA), and by capillary electrophoresis using a Fragment Analyzer™ Automated CE System, protocol DNF-488-33.

Metagenomic library preparation

D-DNA (vesicles, viruses, and free DNA) metagenomic libraries were prepared using a semi-automated EPmotion instrument with TruSeq Nano DNA library preparation kit (15041110), DNA input per sample 2 ng/uL sheared to an average size of 350 bp utilizing a Covaris M220 focused ultrasonicator following manufacturer's recommendations with modifications to shear time to target 350 bp, and Microtube-50 AFA fiber tubes. (For some selected samples that had bimodal DNA size distributions two sequencing libraries were prepared, one with DNA shearing, and one without DNA shearing). Metagenome sequencing libraries were prepared using Illumina's TruSeq Nano LT library preparation kit, and sequenced on an Illumina Nextseq500 system, using V2 high output 300 cycle reagent kit, with addition of 1% of a PhiX control.

DNA sequence analysis and annotation

Sequenced D-DNA reads were filtered and trimmed for quality using the iu-filter-quality-minoche tool (74) from illumina-utils (75). Quality filtered reads were identified and counted by mapping against the ALOHA 2.0 metagenomic reference gene catalog (25, 31) using lastal (76). The output of read mapping against the ALOHA 2.0 gene catalog generated a gene count table for each D-DNA dataset (Dataset S1). Gene

counts in all samples were normalized by the total number of mapped genes. The family-level taxonomic abundance was measured by the number of gene counts assigned to each microbial family. The relative abundances of quality controlled D-DNA gene counts were used to assess variation among D-DNA sample types (vesicle, virus, and F-DNA) and across the water column (75-1000 m). The most abundant ("major") bacterial and viral families were determined by a contribution of >0.5% (cutoff) to all D-DNA samples (Dataset S2), and were compiled into proportional gene count tables. These assessments revealed the microbial sources of the D-DNA samples, as well as the viral hosts (Dataset S3) and *Prochlorococcus* HL ecotype proportion (Dataset S4).

To estimate the proportion of D-DNA sequences from each metagenomic sample that originated from a particular water column depth (0-4000 m), we compared the normalized coverage the ALOHA 2.0 catalog (25, 31) had at each depth in the ALOHA 2.0 survey to the normalized coverage of each gene in our samples. First we calculated the probability that a sequence matching a given gene originated at a given depth as the ratio of the coverage of that gene at that depth to the total coverage of the gene. Next, for all depths, we multiplied that probability by the normalized coverage for each gene in our metagenomic D-DNA samples to get the portion of that coverage that likely originated at each depth. The coverages, now portioned by probable depth origin, were then aggregated by domain (Bacteria, Archaea, Eukaryote, Viral, or unknown) to better understand potential source patterns.

Comparison to cellular and viral datasets collected from Station ALOHA

All D-DNA gene count tables (vesicles, viruses, and free DNA) were combined and mapped to previously reported cellular (25) and viral (31) ALOHA 2.0 metagenomic reference gene catalogs collected on multiple Hawaii Ocean Time-series (HOT) research cruises, from the same depths at Station ALOHA (75, 100, 125, 250, 500, 1000 m). Separate analyses were conducted to compare D-DNA with cellular and viral HOT samples. Read counts per taxonomic clade were normalized by calculating their proportions relative to the total number of mapped gene counts per sample. Normalized values were compiled in a proportional order-level count table and square root transformed. Two-dimensional ordination methods were used on normalized count tables to compare D-DNA datasets to cellular and viral samples from the same depths. To visualize distances, non-metric multidimensional scaling (NMDS) plots were generated using the metaMDS function and Bray-Curtis distance matrices constructed from normalized gene counts in the VEGAN R package.

References

1. Holm-Hansen, O, Sutcliffe Jr., W. H. & Sharp, J. Measurement of deoxyribonucleic acid in the ocean and its ecological significance. *Oceanogr.* **13**, 507-514 (1968).
2. Winn, C. D. & Karl, D. M. Diel nucleic acid synthesis and particulate DNA concentrations: conflicts with division rate estimates by DNA accumulation. *Oceanogr.* **31**, 637-645 (1986).
3. DeFlaun, M. F., Paul, J. H., & Jeffrey, W. H. Distribution and molecular weight of dissolved DNA in subtropical estuarine and oceanic environments. *Ecol. Prog. Ser.* **38**, 65-73 (1987).

4. Dell'Anno, A., Marrale, D., Pusceddu, A., Fabiano, M., & Danovaro, R. Particulate nucleic acid dynamics in a highly oligotrophic system: the Cretan Sea (Eastern Mediterranean). *Ecol. Prog. Ser.* **186**, 19-30 (1999).
5. Djurhuus, A. et al. Environmental DNA reveals seasonal shifts and spatial interactions in a marine community. *Commun.* **11**, 254 (2020).
6. Mende, D. R., Boeuf, D. & DeLong, E. F. Persistent core populations shape the microbiome throughout the water column in the North Pacific Subtropical Gyre. *Microbiol.* **10**, 2273 (2019).
7. Karl, D. M. & Bailiff, M. D. The measurement and distribution of dissolved nucleic acids in aquatic environments. *Oceanogr.* **34**, 543-558 (1989).
8. Karl, D. M. & Church, M. J. Microbial oceanography and the Hawaii Ocean Time-series programme. *Rev. Microbiol.* **12**, 699-713 (2014).
9. Paul, J. H, Jiang, S. C. & Rose, J. B. Concentration of viruses and dissolved DNA from aquatic environments by vortex flow filtration. *Environ. Microbiol.* **57**, 2197–2204 (1991).
10. Jeffrey, W. H, Paul, J. H. & Stewart, G. J. Natural transformation of a marine Vibrio species by plasmid DNA. *Ecol.* **19**, 259–268 (1990).
11. Frischer, M. E., Stewart, G. J. & Paul, J. H. Plasmid transfer to indigenous marine bacterial populations by natural transformation. *FEMS Microbiol. Ecol.* **15**, 127–136 (1994).
12. Turk, V., Rehnstam, A. S., Lundberg, E. & Hagström, Å. Release of bacterial DNA by marine nanoflagellates, an intermediate step in phosphorus regeneration. *Environ. Microbiol.* **58**, 3744–3750 (1992).
13. Jørgensen, N. O. G. & Jacobsen, C. S. Bacterial uptake and utilization of dissolved DNA. *Microb. Ecol.* **11**, 263-270 (1996).
14. Lennon, J. T. Diversity and metabolism of marine bacteria cultivated on dissolved DNA. *Environ. Microbiol.* **73**, 2799-2805 (2007).
15. Hermansson, M. & Linberg, C. Gene transfer in the marine environment. *FEMS Microbiol. Ecol.* **15**, 47-54 (1994).
16. Swan, B. K. et al. Prevalent genome streamlining and latitudinal divergence of planktonic bacteria in the surface ocean. *Natl Acad. Sci. USA* **110**, 11463-11468 (2013).
17. Giovannoni, S. J. et al. Genome streamlining in a cosmopolitan oceanic bacterium. *Science* **309**, 1242–1245 (2005).
18. Tripp, H. J. et al. Metabolic streamlining in an open- ocean nitrogen-fixing cyanobacterium. *Nature* **464**, 90–94 (2010).
19. Biller, S. J. et al. Genomes of diverse isolates of the marine cyanobacterium *Prochlorococcus*. *Data* **1**, 140034 (2014).
20. Braakman, R., Follows, M. J. & Chisholm, S. W. Metabolic evolution and the self-organization of ecosystems. *Natl. Acad. Sci USA* **114**, E3091-3100 (2017).

21. Bentkowski, P., Van Oosterhout, C. & Mock, T. A model of genome size evolution for prokaryotes in stable and fluctuating environments. *Genome Biol. Evol.* **7**, 2344-2351 (2015).
22. Coleman, M. L. et al. Genomic islands and the ecology and evolution of *Prochlorococcus*. *Science* **311**, 1768-1770 (2006).
23. Grote, J. et al. Streamlining and core genome conservation among highly divergent members of the SAR11 clade. *mBio* **3**, e00252-12 (2012).
24. Church, M. J., Lomas, M. W. & Muller-Karger, F. Sea change: charting the course for biogeochemical ocean time-series research in a new millennium. *Deep-Sea Res. Part II* **93**, 2-15 (2013).
25. Mende, D. R. et al. Environmental drivers of a microbial genomic transition zone in the ocean's interior. *Microbiol.* **2**, 1367-1373 (2017).
26. Biller, S. J. et al. Bacterial vesicles in marine ecosystems. *Science* **343**, 183-186 (2014).
27. Brum, J. R., Steward, G. F. & Karl, D. M. A novel method for the measurement of dissolved deoxyribonucleic acid in seawater. *Oceanogr.. Methods,* **2**, 248-255 (2004).
28. DeFlaun, M. F., Paul, J. H. & Davis, D. Simplified method for dissolved DNA determination in aquatic environment. *Environ. Microbiol.* **52**, 654-659 (1986).
29. Linney, M. D., Schvarcz, C. R., Steward, G. F., DeLong, E. F. & Karl, D. M. A method for characterizing dissolved DNA and its application to the North Pacific Subtropical Gyre. *Oceanogr.. Methods.*
30. Brum, J. R., Culley, A. I. & Steward, G. F. Assembly of a marine viral metagenome after physical fractionation. *PLoS One* **8**, e60604 (2013).
31. Luo, E., Eppley, J. M., Romano, A. E., Mende, D. R. & DeLong, E. F. Double-stranded DNA virioplankton dynamics and reproductive strategies in the oligotrophic open ocean water column. *ISME J***14**,1304–1315 (2020).
32. Aylward, F. O. et al. Diel cycling and long-term persistence of viruses in the ocean's euphotic zone. *Natl Acad. Sci. USA* **114**, 11446-11451 (2017).
33. Luo, E., Aylward, F. O., Mende, D. R. & DeLong, E. F. Bacteriophage distributions and temporal variability in the ocean's interior. *mBio* **8**, e01903-17 (2017).
34. Hackl, T. et al. Novel integrative elements and genomic plasticity in ocean ecosystems. *bioRxiv*, e424599 (2020).
35. Biller, S. J. et al. Membrane vesicles in sea water: heterogenous DNA content and implications for viral abundance estimates. *ISME J* **11**, 394-404 (2017).
36. Zhao, X. et al. Three-dimensional structure of the "*Candidatus* Pelagibacter ubique." *Environ. Microbiol.* **83**, e02807-16 (2017).
37. Morris, R. M., Cain, K. R., Hvorecny, K. L. & Kollman, J. M. Lysogenic host-virus interactions in SAR11 marine bacteria. *Microbiol.* **5**, 1011-1015 (2020).
38. Taton, A. et al. The circadian clock and darkness control natural competence in cyanobacteria. *Nat Commun***11,** 1688 (2020).

39. Aguilo-Ferretjans, M. d. M. et al. Pili allow dominant marine cyanobacteria to avoid sinking and evade predation. *Commun.* **12**, 1857 (2021).
40. Brum, J. R. Concentration, production and turnover of viruses and dissolved DNA pools at Stn ALOHA, North Pacific Subtropical Gyre. *Microb. Ecol.* **41**, 103-113 (2005).
41. Torti, A., Lever, M. A. & Jørgensen, B. B. Origin, dynamics, and implications of extracellular DNA pools in marine sediments. *Genomics* **24**, 185-196 (2015).
42. Pietramellara, G. et al.Extracellular DNA in soil and sediment: fate and ecological relevance. *Fertil. Soils***45,** 219–235 (2009).
43. Dell'Anno, A., Stefano, B. & Danovaro, R. Quantification, base composition, and fate of extracellular DNA in marine sediments. *Oceanogr.* **47**, 899-905 (2002).
44. Ellison, C. K., et al. Retraction of DNA-bound type IV competence pili initiates DNA uptake during natural transformation in *Vibrio cholerae*. *Microbiol.***3,** 773–780 (2018).
45. Stewart, C. M. & Tsui, D. W. Y. Circulating cell-free DNA for non-invasive cancer management. *Cancer Genet.* (2018).
46. Adkar-Purushothama, C. R. & Perreault, J-P. Current overview on viroid–host interactions. *WIREs RNA.* **11**, e1570 (2020).
47. Paul, J. H., Jeffrey, W. H. & DeFlaun, M. F. Dynamics of extracellular DNA in the marine environment. *Environ. Microbiol.* **53***,* 170–179 (1987).
48. Riemann, L., Holmfeldt, K. & Titelman, J. Importance of viral lysis and dissolved DNA for bacterioplankton activity in a P-limited estuary, Northern Baltic Sea. *Ecol.* **57**, 286-294 (2008).
49. Moebus, K. Marine bacteriophage reproduction under nutrient-limited growth of host bacteria. I. Investigations with six phage-host systems. *Ecol. Prog. Ser.* **144**, 1–12 (1996).
50. Weinbauer, M. G., Brettar, I. & Höfle, M. G. Lysogeny and virus-induced mortality of bacterioplankton in surface, deep, and anoxic marine waters. *Oceanogr.* **48**, 1457–65 (2003).
51. Alonso, M. O., Rodriguez, V., Rodriguez, J. & Borrego, J. J. Role of ciliates, flagellates and bacteriophages on the mortality of marine bacteria and on dissolved-DNA concentration in laboratory experimental systems. *Exp. Mar. Biol. Ecol.* **244**, 239–252 (2000).
52. Zhong, X., Wirth, J. & Suttle, C. Ribosomes in the sea: a window on taxon-specific lysis. (2016). *AGU Ocean Sciences.*
53. Baran, N., Goldin, S., Maidanik, I. & Lindell, D. Quantification of diverse virus populations in the environment using the polony method. *Microbiol.* **3**, 62-72 (2018).
54. Zborowsky, S. & Lindell, D. Resistance in marine cyanobacteria differs against specialist and generalist cyanophages. *Natl. Acad. Sci. USA* **116**, 16899-16908 (2019).
55. Liljemark, W. F. & Anderson, D. L. Morphology and physiology of the intracellular development of *Bacillus subtilis* bacteriophage ϕ25. *Virol.* **6**, 114-124 (1970).
56. Powell, I. B., Tulloch, D. L., Hillier, A. J. & Davidson, B. E. Phage DNA synthesis and host DNA degradation in the life cycle of *Lactococcus lactis* bacteriophage c6A. *Gen. Microbiol.* **138**, 945-950

(1992).

57. Paul, J. H., Jeffrey, W. H. & Cannon, J. P. Production of dissolved DNA, RNA, and protein by microbial populations in a Florida reservoir. *Environ. Microbiol.* **56**, 2957 – 2962 (1990).
58. Fontanez, K. M., Eppley, J. M., Samo, T. J., Karl, D. M. & DeLong, E. F. Microbial community structure and function on sinking particles in the North Pacific Subtropical Gyre. *Microbiol.* **6**, 469 (2015).
59. Richardson, T. L. & Jackson, G. A. Small phytoplankton and carbon export from the surface ocean. *Science* **315**, 838-840 (2007).
60. Collins, J. R. et al. The multiple fates of sinking particles in the North Atlantic Ocean. *Global Biogeochem. Cycles* **29**, 1471-1494 (2015).
61. Hannides, C. C. S., Popp, B. N., Choy, C. A. & Drazen, J. C. Midwater zooplankton and suspended particle dynamics in the North Pacific Subtropical Gyre: a stable isotope perspective. *Oceanogr.* **58**, 1931-1946 (2013).
62. Karl, D. M., Knauer, G. A. & Martin, J. H. Downward flux of particulate organic matter in the ocean: a particle decomposition paradox. *Nature* **332**, 438-441 (1988).
63. Martin, J. H., Knauer, G. A., Karl, D. M. & Broenkow, W. W. VERTEX: carbon cycling in the northeast Pacific. *Deep-Sea Research* **34**, 267-285 (1987).
64. Grabowski, E., Letelier, R. M., Laws, E. A. & Karl, D. M.Coupling carbon and energy fluxes in the North Pacific Subtropical Gyre. *Commun.***10**, 1895 (2019).
65. Karl, D. M. et al. Season and interannual variability in primary production and particle flux at Station ALOHA. *Deep-Sea Res. Part II* **43**, 539-568 (1996).
66. Hebel, D. V. & Karl, D. M. Seasonal, interannual and decadal variations in particulate matter concentrations and composition in the subtropical North Pacific Ocean. *Deep-Sea Res. Part II* **48**, 1669-1695 (2001).
67. Lukas, R. & Santiago-Mandujano, F. Interannual to interdecadal salinity variations observed near Hawaii: local and remote forcing by surface freshwater fluxes. *Oceanography***21**, 46–55 (2008).
68. Talley, L. D. Distribution and formation of North Pacific Intermediate Water. *Phys. Oceanogr.* **23**, 517-537 (1993).
69. Bullister, J. L., Wisegarver, D. P. & Sonnerup, R. E. Sulfur hexafluoride as a transient tracer in the North Pacific Ocean. *Res. Lett.* **33**, L18603 (2006).
70. Karl, D. M., Church, M. J., Dore, J. E., Letelier, R. M. & Mahaffey, C. Predictable and efficient carbon sequestration in the North Pacific Ocean supported by symbiotic nitrogen fixation. *Natl. Acad. Sci. USA* **109**, 1842-1849 (2012).
71. Poff, K. E., Leu, A. O., Eppley, J. M., Karl, D. M. & DeLong, E. F. Microbial dynamics of elevated carbon flux in the open ocean's abyss. *Natl. Acad. Sci USA,* **118**, e2018269118 (2021).
72. Cruz, B. N. & Neuer, S. Aggregation of the marine picocyanobacterial *Prochlorococcus* and *Synechococcus*. *Microbiol.* **10**, 1864 (2019).

73. Deng, W., Monks, L. & Neuer, S. Effects of clay minerals of the aggregation and subsequent settling of marine *Synechococcus*. *Oceanogr.* **60**, 805-816 (2015).
74. Minoche, A. E., Dohm, J. C. & Himmelbauer, H. Evaluation of genomic high-throughput sequencing data generated on Illumina HiSeq and Genome Analyzer systems. *Genome Biol.* **12**, R112 (2011).
75. Eren, A. M., Vineis, J. H., Morrison, H. G. & Sogin, M. L. A filtering method to generate high quality short reads using Illumina paired-end technology. *PLoS One* **8**, e66643 (2013).
76. Kielbasa, S. M., Wan, R., Sato, K., Horton, P. & Frith, M. C. Adaptive seeds tame genomic sequence comparison. *Genome Res.* **21**, 487-493 (2011).

Figures

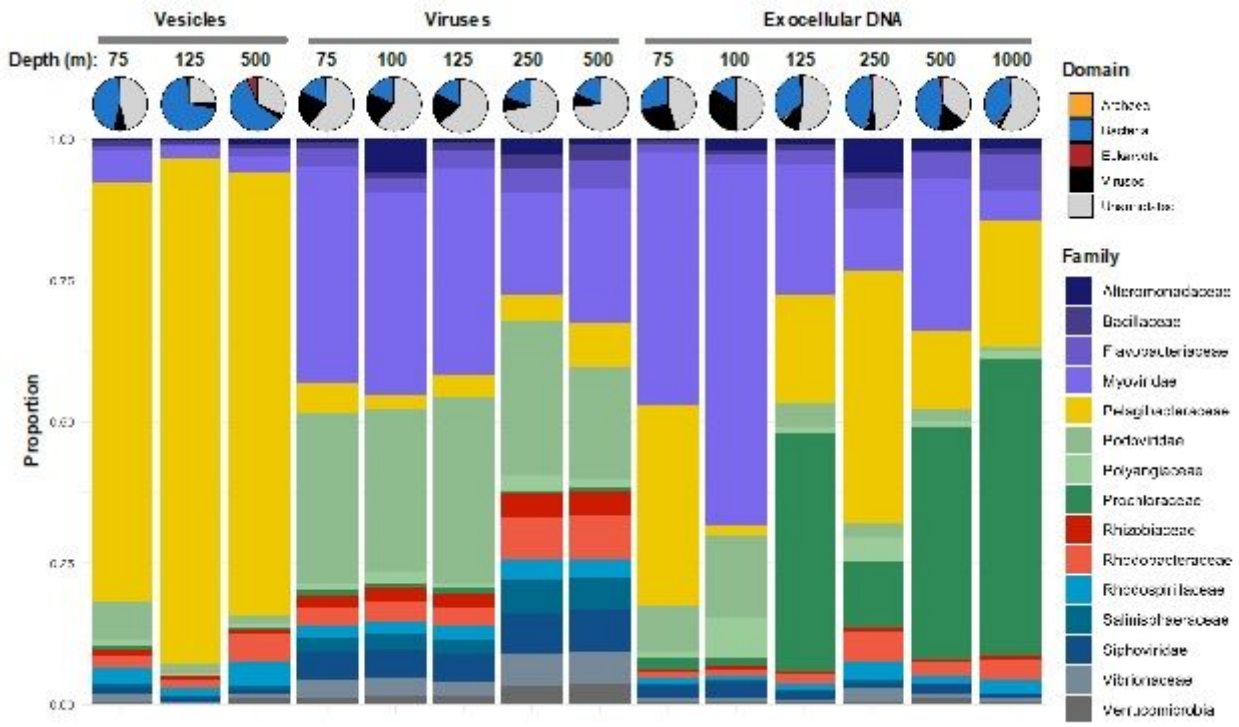

Figure 1

Taxonomic annotation of metagenomic sequences from the three dissolved DNA constituents collected from the North Pacific Subtropical Gyre. (A) Domain-level taxonomic composition and proportion of unannotated sequences for each sample. (B) Family-level taxonomic composition of three dissolved DNA constituents (vesicles, viruses, and exocellular free DNA). The low-density vesicle constituents were collected from three euphotic and mesopelagic depths (75, 125, and 500 m). Viral fractions were collected from five depths (75, 100, 125, 250, and 500 m) in the North Pacific Subtropical Gyre. Exocellular F-DNA samples were collected from six depths (75, 100, 125, 250, 500, and 1000 m).

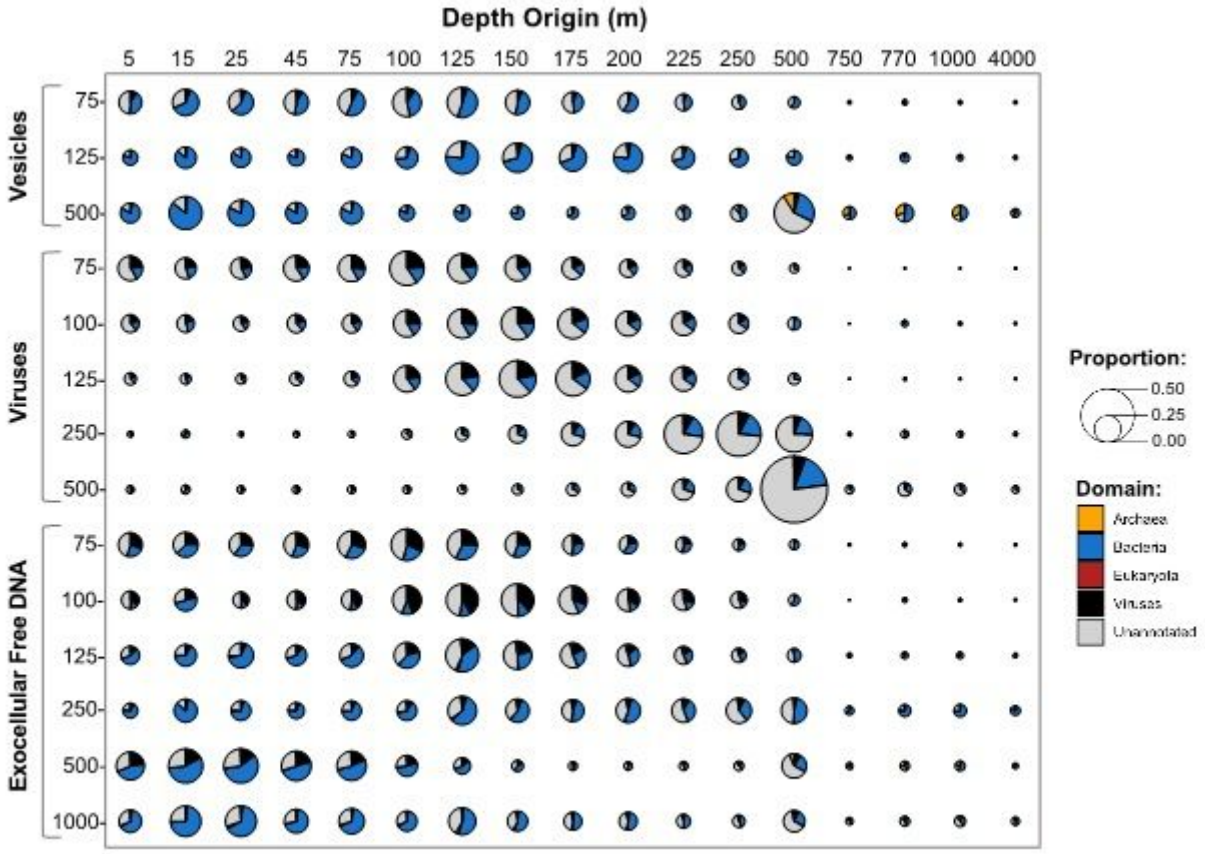

Figure 2

Annotation of D-DNA metagenomic sequences suggests their probable depths of origin. Best sequence matches to the ALOHA 2.0 gene catalogue (and their corresponding sampling depths) were used to assign the probable depth of origin to individual D-DNA metagenomic sequence reads. The size of the pie chart is proportional to the total number of D-DNA metagenomic reads, whose best match is to Station ALOHA gene originating from a given corresponding water depth.

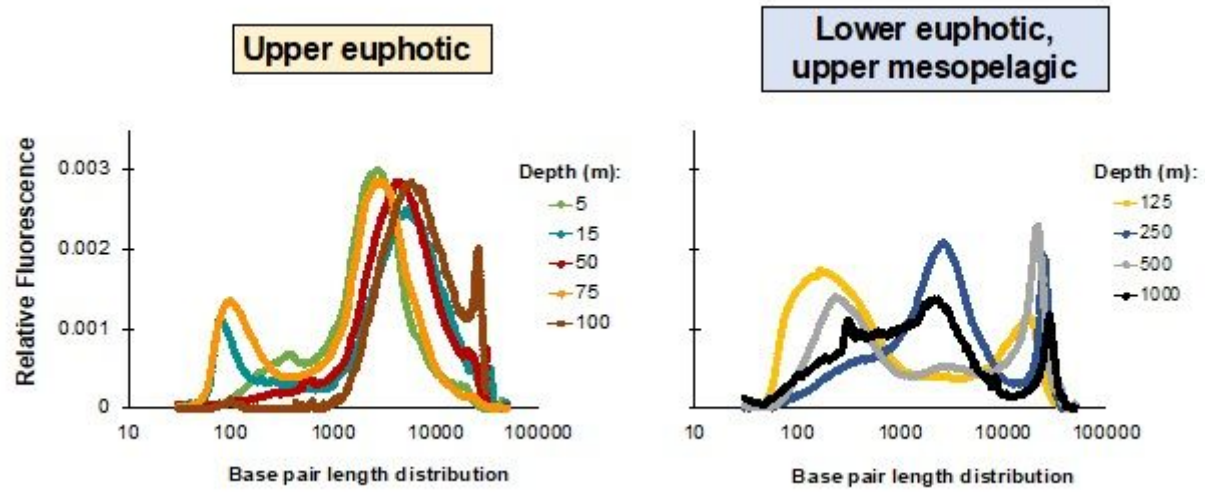

Figure 3

Fragment analysis of F-DNA samples from the North Pacific Subtropical Gyre. Relative fluorescence unit of each sample used to normalize and calculate proportion. Molecular weight values shown from 75-50,000 basepairs according to Agilent DNF-488-0500 protocol sizing range. Samples collected above the deep chlorophyll maximum (125 m) had more distinct peaks, samples below had a larger range of sizes indicative of degradation.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- DatasetS1.xlsx
- DatasetS2.xlsx
- DatasetS3.xlsx
- DatasetS4.xlsx
- SupplementaryMaterial.pdf