

Diversity and Population Structure of Uncultured Marine Viral Communities

Mya Breitbart¹, Ben Felts², Scott Kelley¹, Joseph M. Mahaffy², James Nulton², Peter Salamon², Forest Rohwer^{1,3}

¹Department of Biology, San Diego State University, San Diego, CA 92182-4614

²Department of Mathematical Sciences, San Diego State University, San Diego, CA 92182-7720

³Center for Microbial Sciences, San Diego State University, San Diego, CA 92182

Viruses, most of which are phage, are extremely abundant in the marine environment. However, very little is known about the identity or diversity of these viruses. Here we present the first metagenomic analyses of uncultured viral communities from two nearshore marine water samples and one marine sediment sample. In all three marine libraries, over 65% of the sequences were not significantly similar to previously reported sequences, suggesting that much of the diversity was novel. Among the sequences that could be identified, the majority belonged to dsDNA phage. Temperate phage were more common than lytic phage in the sediment community, suggesting that lysogeny may be an important lifestyle for viruses in this environment. Comparisons between the sediment and seawater viral communities showed that certain phage phylogenetic groups were abundant in all marine viral communities, while other phage groups were underrepresented. This "marineness" suggests that marine phage are derived from a set of common ancestors. Several independent mathematical models, based on the number of fragment overlaps from the shotgun library, were used to predict the population structure of the marine viral communities. Diversity of the viral communities was extremely high, with at least 10^3 viral genotypes in 100 liters of seawater and at least 10^4 viral genotypes per kg of sediment. Each of the viral communities had a Shannon Index of >7 , with the highest diversity observed in the sediment community. Overall, the marine viral communities were extremely novel, diverse, and likely represent one of the largest unexplored reservoirs of sequence space on the planet.