

# **Sequencing and analysis of crAssphage regions from around the globe: New York NY, Davis CA, and San Juan PR samples**

Samuel L. Díaz Muñoz\*, Kristen Gulino, Alan Twaddle, Elodie Ghedin  
Center for Genomics and Systems Biology, Department of Biology, New York University

\*Contributed to NY, CA, and PR data collection. All other authors contributed to NY data collection.

## **Summary:**

- New York, NY: 1 sewer sample positive at all loci; 1 Gowanus Canal sample negative
- Davis, CA: 2 sewer samples positive at all loci
- San Juan, PR: 4 samples negative

## **I. Sample Information**

### A. New York, NY

1. Gowanus Canal: pooled sample from 3 different depths/locations on sampled on 12/14/14:

Approximate address	lat (N)	long(W)	depth (ft)
213 Douglass St Brooklyn, NY 11217	40.6814191	-73.98708344	-2.5'
510 Degraw St Brooklyn, NY 11217	40.68045903	-73.98761988	-7.5'
482 Union St Brooklyn, NY 11231	40.67935248	-73.98839235	-8.5'

2. NYC Sewer: Single sample pooled from 14 locations. Using city center coordinates

NYC Sewage sampling, virome Date: 11/20/14					
Location	Address	Alt (m)	Lat.	Lon.	
26th Ward WWTP	12266 Flatlands Ave, New York, 11207	3.532	40.652416	-73.879078	
Bowery Bay WWTP	4301 Berrian Blvd, Astoria, NY 11105	4.818	40.778919	-73.895319	
Coney Island WWTP	2591 Knapp St., Brooklyn, NY 11235	2.841	40.590609	-73.932227	
Hunt's Point WWTP	1270 Ryawa Ave, Bronx, NY 10474	5.185	40.804645	-73.885766	
Jamaica WWTP	150-20 134th Street, Queens, NY 11430	4.265	40.663297	-73.806852	
Newtown Creek WWTP	327 Greenpoint Ave, Brooklyn, NY 11222	2.822	40.732809	-73.947231	
North River WWTP	725 W 135 STREET, New York, NY 10031	19.042	40.821329	-73.957729	
Oakwood Beach WPCP	751 Mill Rd, Staten Island, NY 10306	2.364	40.553765	-74.117267	
Bay Park STP	106-21 Beach Channel Dr, Rockaway, NY 11694	0.065	40.581799	-73.848493	
Owls Head WWTP	6700 Shore Road Brooklyn, NY 11220	1.659	40.641886	-74.034062	
Wards Island WWTP	7 Wards Island, New York, NY 10035	11.506	40.788997	-73.926867	
Tallman Island WWTP	128th St, College Point, NY 11356	4.377	40.791890	-73.839846	
Port Richmond WWTP	1801 Richmond Terrace, Staten Island, NY 10310	3.929	40.637227	-74.125905	
Red Hook WWTP	63 Flushing Ave, Brooklyn, NY 11205	0.247	40.703456	-73.971780	
NYC center	NYC center	10.4440.7128° N	74.0059° W		

## B. Davis, CA

1. City of Davis WWTP: Closest address 45452 Co Rd 28H, California 95776, Latitude 38.590894, Longitude: -121.666779, altitude 8.659 m
2. I (SLDM) was provided a large sample by WWTP staff collected on Thursday April 21, 2016 1:12pm
3. The two individual Davis samples are each a ~5mL subsample of this larger initial ~1L sample.

## C. San Juan, PR

1. I (SLDM) could not secure access to Puerto Nuevo WWTP, so I collected in areas suspected or known to have high coliform levels. All samples were negative for all crAssphage primers.

Collection Location	lat	long	date	time	ID
Laguna de Condado	18.4538138	-66.0741903	12/26/2016	7:04 am	PR1
Caño Martín Peña West	18.4316494	-66.059226	12/26/2016	7:24 am	PR2
Caño Martín Peña East	18.4262696	-66.0299169	12/26/2016	7:50 am	PR3
Water channel adj to Puerto Nuevo WWTP	18.4343084	-66.0864768	12/26/2016	8:20 am	PR4

## **II. Methods**

PCR amplicons were generated as described in the standard project protocol, using the same master mix (Apex RED Taq DNA Polymerase Master Mix, 2.0X, Genesee Scientific San Diego CA) as used at SDSU (Alejandro Vega, pers. comm.).

Sequencing was conducted in an ABI 3730 automated capillary sequencer. Sequencing was primed using both the forward and reverse primers for each locus. Each sequence was the manually curated consensus of at least 2 sequences.

## **III. Results**

### A. New York NY

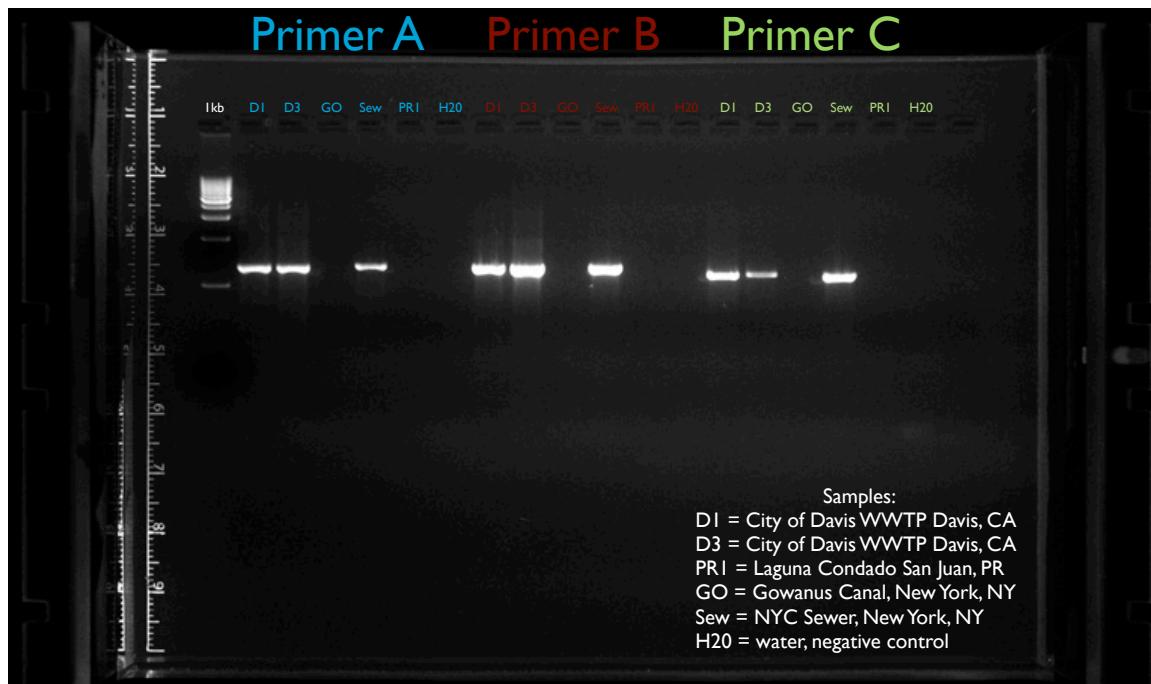
1. One sequence is reported; it was generated by sequencing two separate PCR amplicons. One was the NYC Sewage sample and the second was an identical sample spiked with hRV14 (human rhinovirus) as a control for a separate metagenomic sequencing project. Thus, the reported sequence is the consensus of 4 Sanger sequences.
2. Because a metagenome was available for the NYC Sewage sample, I mapped Illumina 101bp PE reads to the Sanger sequence from the PCR amplicons to get an idea of the variability within the sample. The SNP's are reported in three VCF files, one per locus. However, coverage was not deep (at best 40, mostly 20's and teens) or uniform across the amplicon (some gaps with no coverage) for all three loci.

### B. Davis CA

1. Two sequences are reported, but please note these were generated from samples gathered from a larger sample (see Sample Information).

### C. San Juan PR

1. No sequences are reported, as there was no successful amplification.



Gel of all three crAssphage loci on samples from Davis CA, New York NY, and San Juan PR. Only PR1 sample shown, but PR1-4 were all PCR negative.

### **IV. Observation**

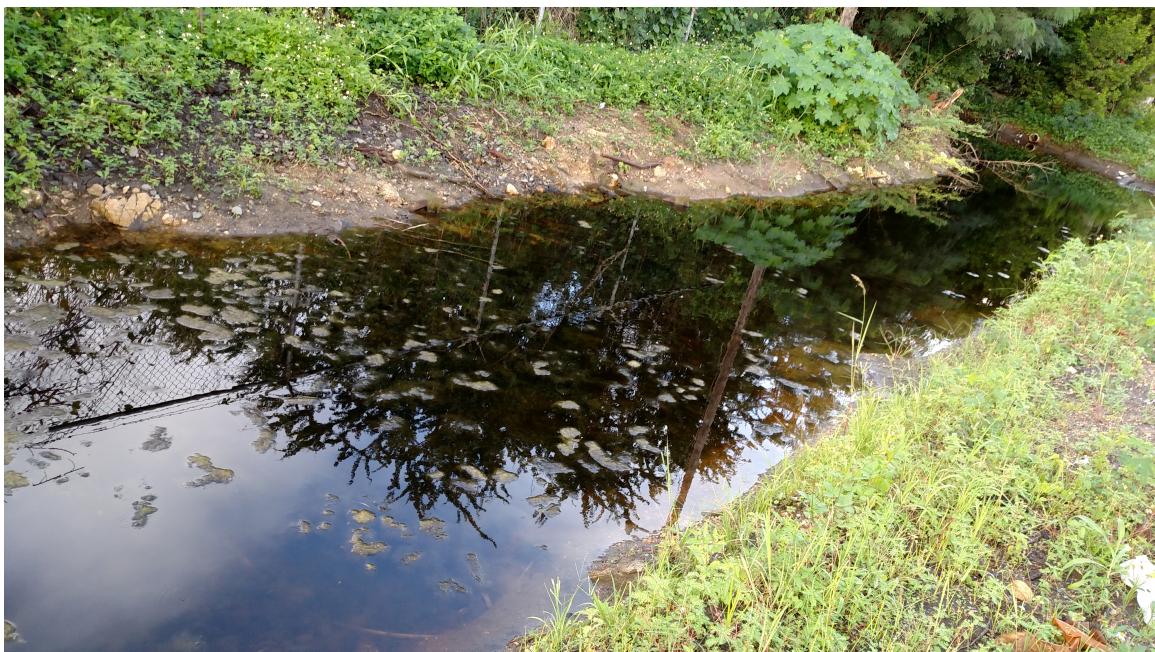
I (SLDM) note that no amplification was found on samples collected from the environment, even when the water bodies were strongly suspected of having fecal contamination. The Gowanus Canal in New York, NY is known to receive millions of gallons of sewage. The Caño Martín Peña in San Juan, PR routinely has elevated levels of fecal coliforms due to human settlements along the channel that are not connected to the sewage system. Samples from both of these locations were negative. See Puerto Rico collection site pictures below.

However, all sewage samples tested were positive.

If this observation holds across samples in different parts of the world, it may be useful for hypothesizing about the natural host, biology, and ecology of crAssphage.



Martín Peña West sampling site in San Juan, PR



Channel next to Puerto Nuevo WWTP sampling site in San Juan, PR

### **III. Acknowledgments**

We thank: John Alexander, Robert Matthews, and Mitchell Wright for providing the Davis CA samples; Gary Toranzos and Emmanuel Carrasquillo for providing lab facilities and assistance at the University of Puerto Rico, Rio Piedras; Lydia Smith for sequencing assistance at UC Berkeley. NYC sewage sample collection was made possible by the NYC Department of Environmental Protection.