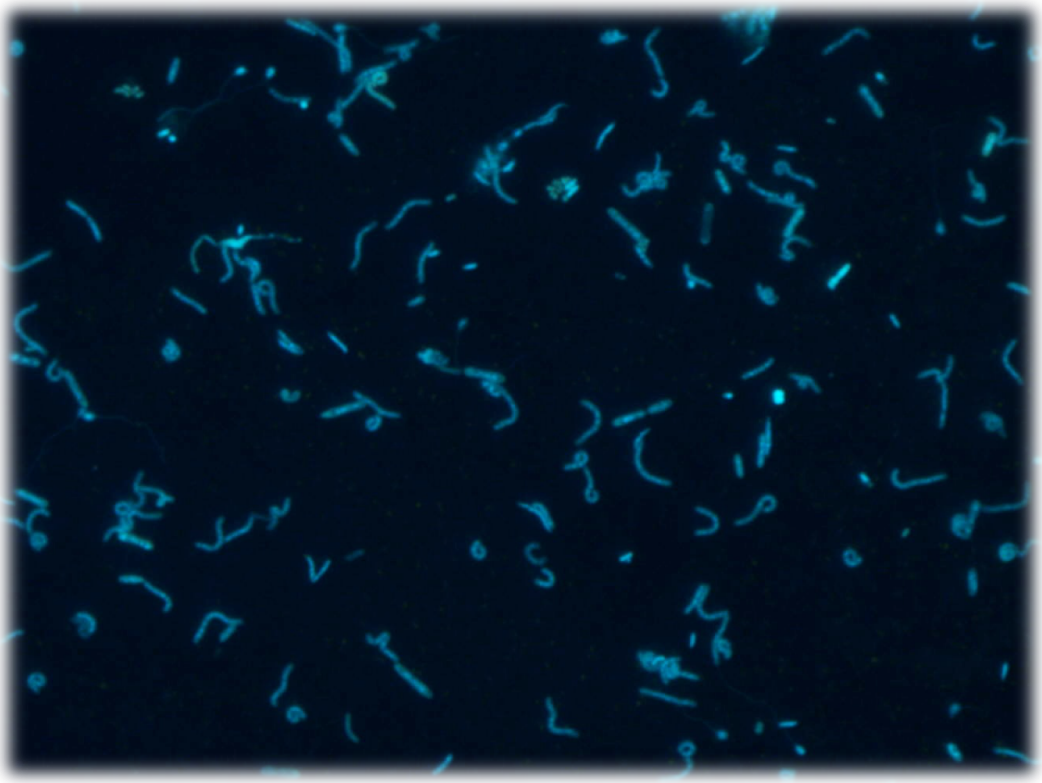
Description and isolation of marine Protist Viruses

Eliana Ruiz Martínez

Directors: Dolors Vaqué and Ramon Massana

Oceanography Master Tesis July 2013



CSIC Institut de Ciències del Mar

FACULTAD DE CIENCIAS DEL MAR
UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA

UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA

Index

Introduction

Viruses are the most abundant biological entities in the ocean and are the reservoir of most of the genetic diversity in the sea (Suttle 2007). Typically, the concentration of marine viruses range from 10^5 to 10^8 ml^{-1}, and that abundance decreases with depth and distance from the shore (Bratbak et al., 1990; Suttle, 2007).

They seem to be responsible for significant mortality in marine microbial communities (Proctor and Fuhrman, 1990) and in phytoplankton populations (Suttle & Chan 1994), on average causing the lysis of about 20% of bacteria and around 3% of phytoplankton on a daily basis (Suttle, 1994).

When host organisms are lysed, nutrients are released into the surrounding environment and thus influence biogeochemical and ecological processes (Fuhrman, 1999). Viral lysis affects the efficiency of the biological pump (the combination of processes that leads to the sequestration of carbon in the deep ocean as a result of the sinking of particulate organic matter from surface waters) by increasing or decreasing the relative amount of carbon in exported production (Suttle, 2007). The viral shunt, then, moves material from heterotrophs and photoautotrophs as particulate organic matter (POM) to dissolved organic matter (DOM) as seen in the image below.

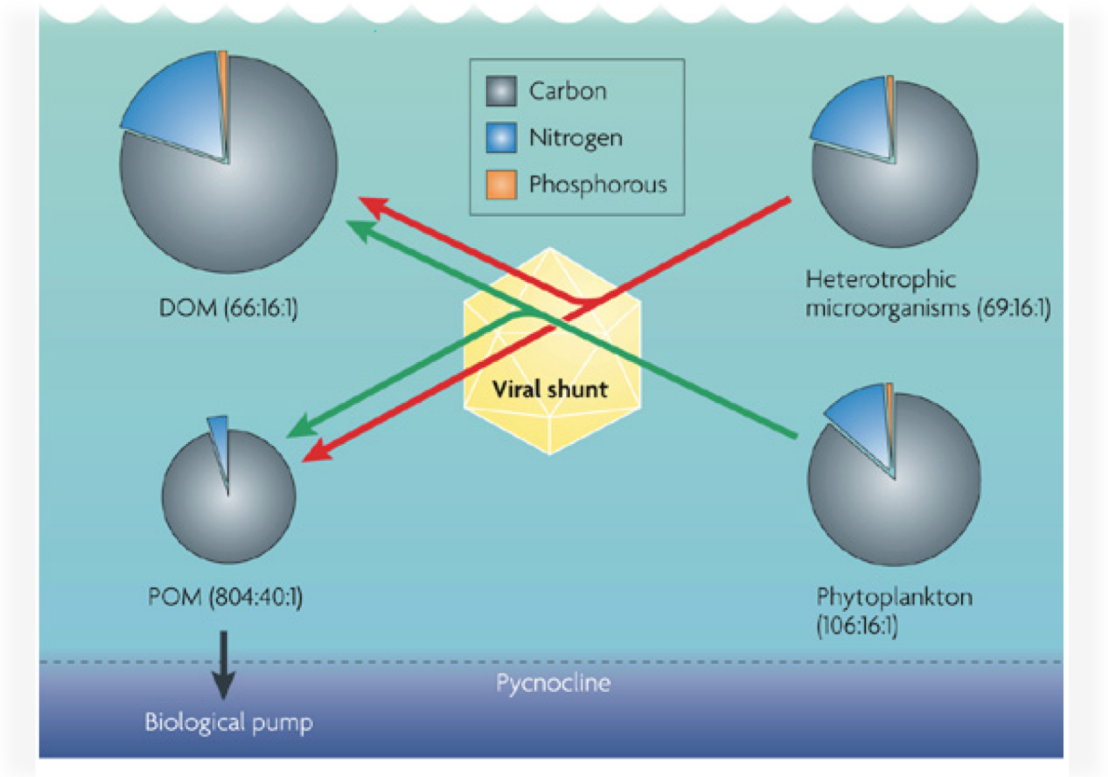


Figure 1: Shunt and pump. Suttle 2007 "Marine viruses – major player s in the global ecosystem" – Figure 2, page 805

The most part of this organic matter will be converted to carbon dioxide by respiration and photodegradation in the photic zone, thereby decreasing the trophic transfer efficiency of nutrients and energy through the marine food web. (Suttle, 2005 and 2007)

For the reasons explained above, we can say that viruses can exert significant control on marine bacteria and phytoplankton communities, with respect to both biological production and species composition, influencing the pathways of matter and energy transfer in the system (Fuhrman, 1999).

Moreover, viruses play a significant role in the transfer of information encoded in DNA, and that is because viral genes and viral activity generate genetic variability of prokaryotes and are a driving force for ecological functioning and evolutionary change (Weinbauer and Rassoulzadegan, 2004).

Viruses found in aquatic systems have different morphologies, including regular structures in the capsid such as icosahedrons (with or without a "tail"), filaments, or other morphologies, and their capsid diameters range from 20 to 300nm (Fuhrman and Suttle, 1993). The smallest ones are considered to infect mostly bacteria (Weinbauer 2004), but there is another group, typically larger and less abundant, that infects a number of important phytoplankton taxa (Suttle et al., 1990).

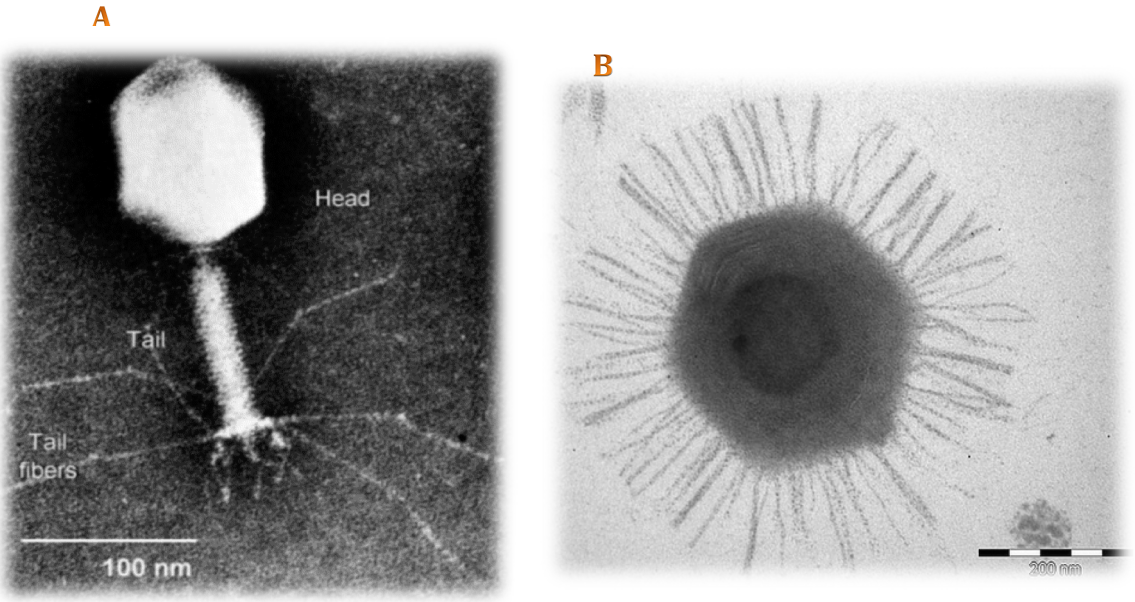


Figure 2: Diferences between a bacteriophage (A) and an eukaryotic virus (B)

Marine eukaryotic phytoplanktonic viruses are being studied from the past years, and we now have more information about their host-interactions, ecology, distribution and metagenomics (Garza and Suttle, 1995; Massana et al., 2004; Logares et al, 2012). But, currently, we don't have enough knowledge about viruses that infect heterotrophic protists, particularly, pico/nanoflagellates that have a size

between 2 and 5µm and are quite abundant in marine systems. They have a role in the microbial food webs as bacterial grazers, trophic linkers and nutrient remineralizers (Pernthaler, 2005). Hence the main goal in this study is to identify viruses of pico/nanoflagellates from different oceans based in a metagenomic approach, and to try isolating them by using cultured pico/nanoflagellates strains.

For our purpose, we divided this study in two parts: one is mostly experimental, in which we tried to infect a culture of the heterotrophic picoflagellate *Developayella sp.* with a virus concentrate, and the other is a metagenomic approach, in which we searched at different databases for genes of the known virus infecting pico/nanoflagellates.

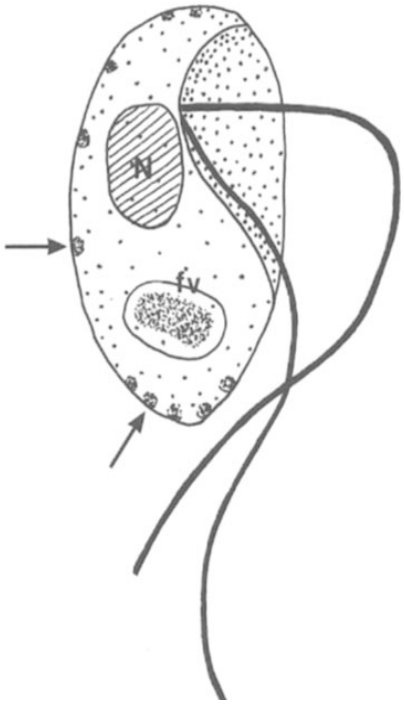


Figure 3: Drawing of Developayella sp. Tong (1995)

The specific objectives of this study are:

1) To gather together all the present information about viruses of marine protists, trying to quantify the gene abundances of the most representative protist viruses in the oceans and their geographic distributions.

2) To try to isolate a virus from the heterotrophic flagellate *Developayella sp.* strain.

Materials and Methods

- Metagenomic's research

Genomic sequences

Nucleotide sequences of all genomes were downloaded from NCBI/GenBank database, except for the *Bathycoccus*, *Micromonas* and *Ostreococcus lucimarinus* and OtV5 genomes, which were provided by Nigel Grimsley from the Observatoire Océanologique de Banyuls/Mer, and the *Chlorella* virus genomes, that were retrieved from Greengene.

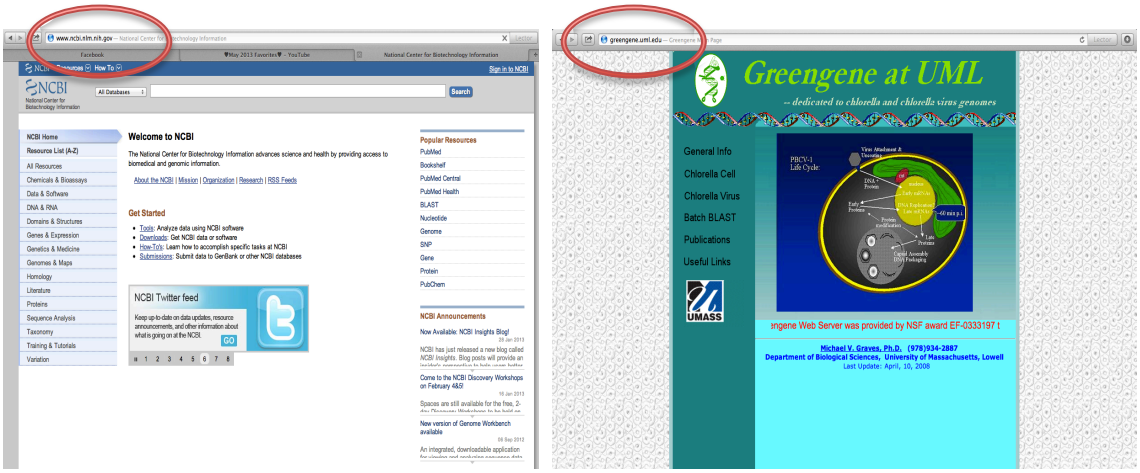


Figure 4: NCBI/GenBank and Greengene web sites.

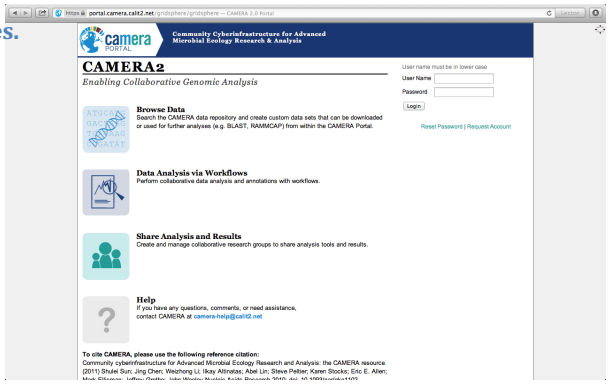


Figure 5: CAMERA web site

Bioinformatics analysis

The available genomes of protists viruses were used as templates for recruiting metagenomic reads from already published metagenomes. Database searches were performed against CAMERA GOS Reads database (E value: 10^{-10}). Several metagenome collections within CAMERA were inspected (see Results).

After the previous search, where we retrieved metagenomic reads related

to protist viral genomes, we neded to double check the closest genome for each read. For that, we created a database with the genomes of all viruses, and we performed a Local Blast of the metagenomic reads against this genome viral database.

Mapping

We used MatLab for the elaboration of the viral reads distribution. We exported a whorksheet with our results and the coordinates of each sample location and we got a file with the entire variable matrix ("databasevirus.mat"). We used "programa_mapa_virus.m" as MatLab's script.

- Experiments

Cultures of *Developayella* sp. and obtention of virus

The heterotrophic picoflagellate *Developayella sp.* (JC09) isolated from the Marine Microbial Blanes Bay Observatory (MMBO, Catalonia Spain) and maintained in culture in the ICM was used for the experiments. The picoflagellate cultures were grown in aged seawater medium with the addition of a concentrate of marine bacteria (MED134) at a 10^7 cells ml^{-1}. They were maintained in 30ml flasks and transferred every 3 to 4 weeks to fresh media at 1/10 dilution and incubated at 20°C in the dark.

Viruses were obtained from the MMBO seawater. Samples were taken in two different seasonal times (May and October 2012), with the aim of gather the highest viral diversity.

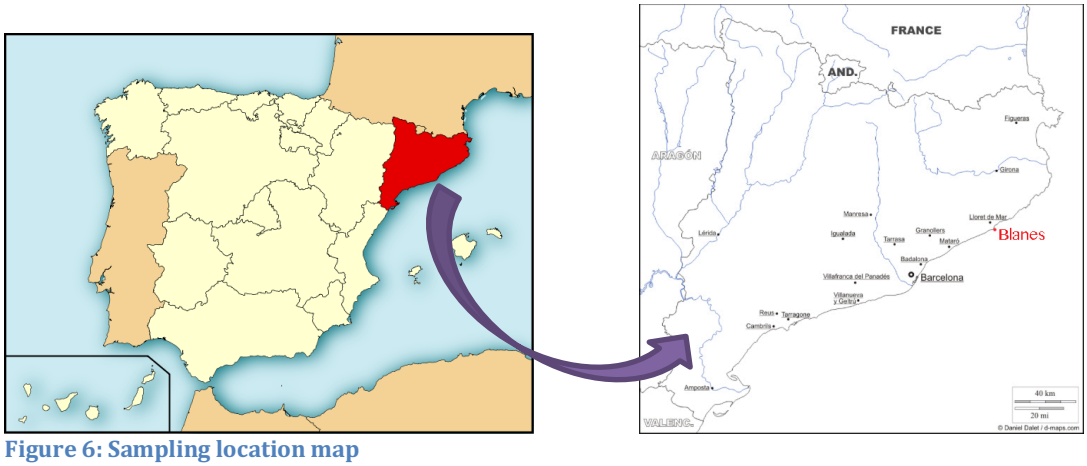


Figure 6: Sampling location map

Seawater samples were collected in a 50 L carboy, previously filtered through 200 µm, and then carried to the laboratory. Once there, 4 L of this water was filtered again through the following filters: 20, 3 and 0.45 µm respectively; thus, we made sure of keeping in the filtrate the 3-0.45 µm content of the sample. We didn't use the 0.2µm filter because we were looking for large viruses. Finally this filtered seawater was used to concentrate all viruses by tangential flow filtration (TFF, cartridge VIVAFLOW, 30 Kd) in a final volume of 10 ml.

Enrichments

Before our experiments had taken place, we decided to concentrate even more our virus concentrate taken from the TFF, by two previous enrichment experiments (see Diagrams 1 and 2).

We re-inoculated the 30ml *Developayella* sp. culture in a 150 ml flask, maintaining the same bacterial abundance as described above and followed daily its dynamics by epifluorencence microscopy after DAPI (4',6-diamidino-2-phenylindole) staining (Porter and Feig, 1980) waiting for the exponential phase of *Developayella* sp growth. We considered this phase of the picoflagellate development as the right moment to proceed with the viral infection, on condition that our culture reached a concentration of 10^4 flagellates ml^{-1} at least.

Before to inoculate the virus concentrate to the picoflagellate culture, we took 5ml from each virus concentrate sample, March and October 2012, and we filtered them through 0.45µm (to eliminate any bacteria). Then, they were added to the 150 ml flask,

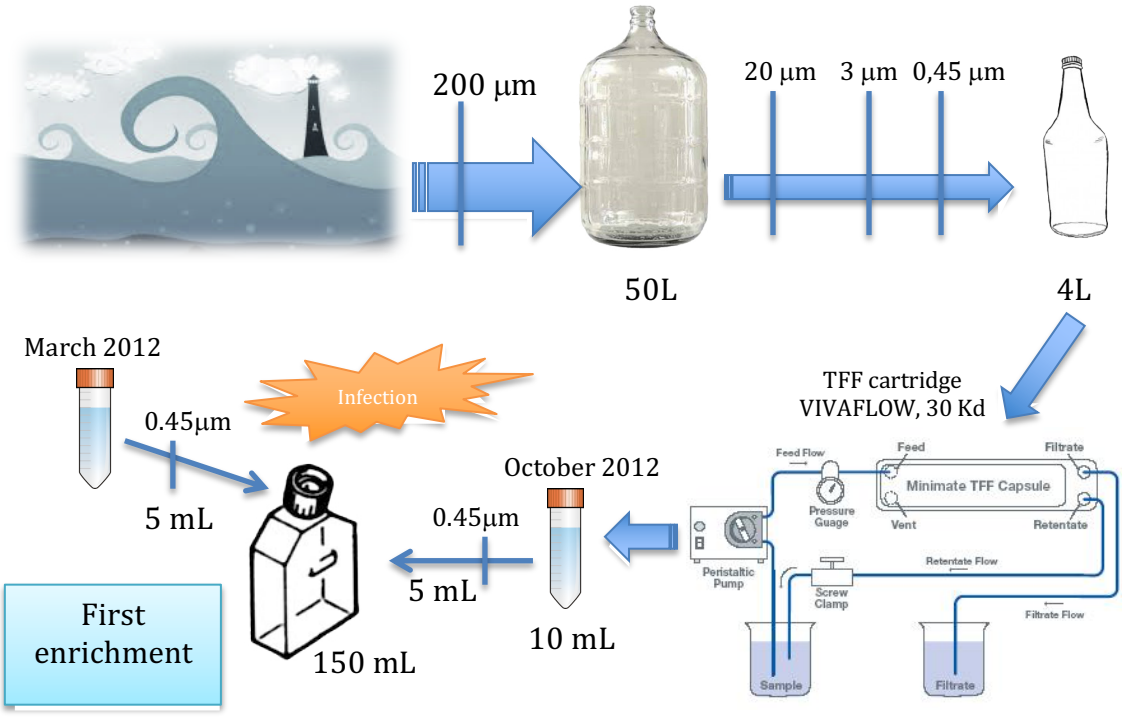


Diagram 1: Sampling procedure, from sea sampling to First Enrichment experiment.

The abundances of bacteria and flagellates were followed once a day by epifluorencence microscopy (EFM) after DAPI (4', 6-diamidino-2-phenylindole) staining (2.7ml/sample), according the method described in Porter and Feig (1980). We followed this procedure until the flagellate's concentration decreased to 10^3, which it corresponds to 12 (more or less) days according to the development of our flagellate.

At this point of the experiment, we recovered the virus concentrate of the enrichment flask by centrifugation of 12ml of the culture with the following parameters:

Temperature	15°C
Volume	12ml
Time	10 minutes
RPM	1800

Table 1: centrifugation parameters

Then, we filtered through 0.45µm a10ml supernatant, and stored it at 4°C until its use.

Two enrichment experiments were carried out with the same characteristics and procedure explained above. The first one took place from 15/10/2012 to 25/10/2012, and the second one from 29/10/2012 to 5/11/2012.

It has to be said, that the virus concentrate recovered from the first enrichment was used to infect the second one (we have always filtered it through 0.45µm before its addition), and so on…

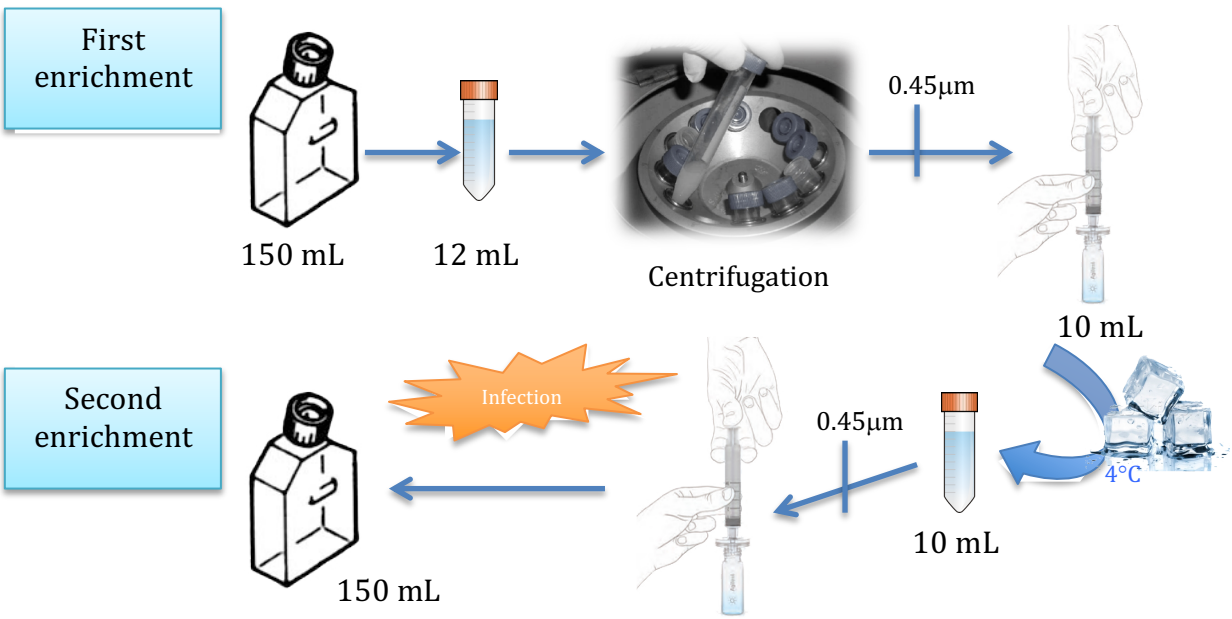


Diagram 2: From first to second Enrichment procedure

First experiment

The main procedure for the rest of experiments was the same as the enrichment ones, although we modified culture volumes in some cases and we worked with two cultures at the same time (control and experiment). We followed the bacteria and flagellate abundances twice a day (morning and evening), and we got one more sample per day (1ml/sample) to count virus by flow cytometry (FCM, Brussaard 2004) from each culture bottle, and incorporated a new virus concentration method that will be explained below.

The first experiment was carried out as the enrichment ones but as said before, we worked with two different cultures (see Diagram 3). We added the virus concentrate (recovered from the second enrichment) to one 150ml flask, and we called it the "experiment" flask. We added the same volume of aged water to the other 150 ml flask, called "control". This culture was called "virus-free," because we did not add virus concentrate, so it let us compare the behaviour of both cultures.

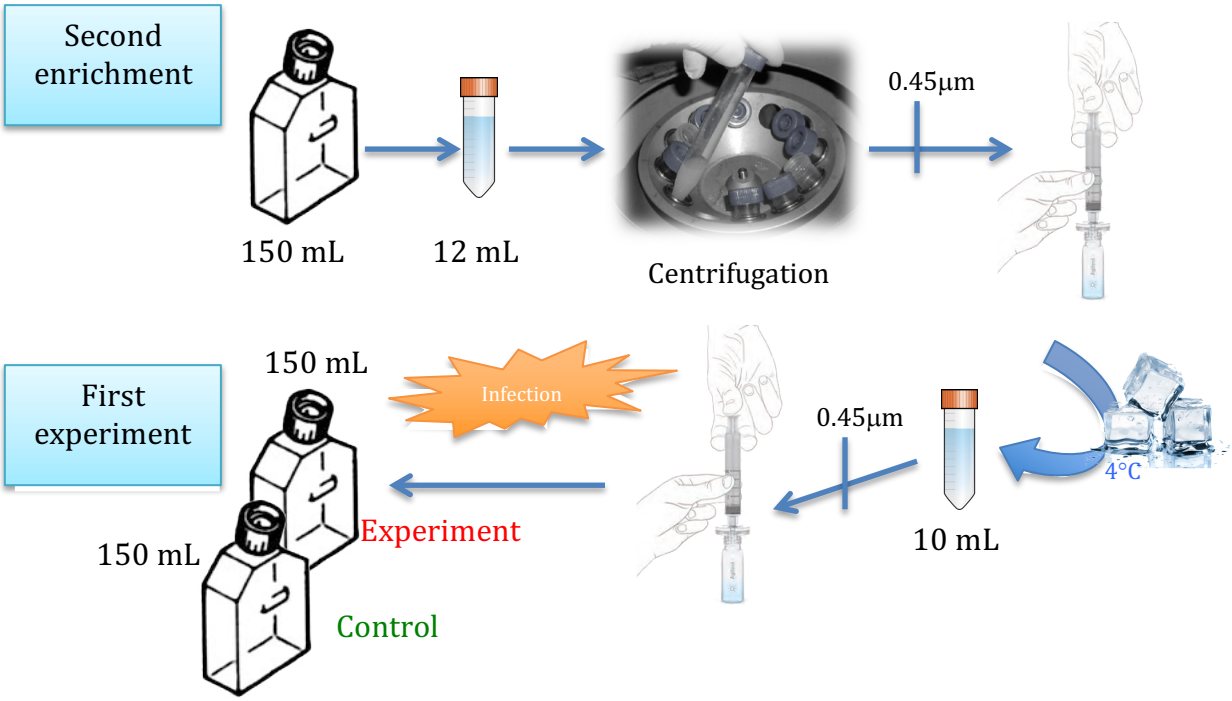


Diagram 3: From second enrichment to first experiment procedure.

Second and third experiments

For the second experiment, we recovered all viruses from the first "experiment" culture as we did with the enrichment experiments (see Diagram 4). Then, we filtered the obtained 10ml supernatant through 0.8 µm, and we re-centrifuged it using the Amicon Ultra-15 Centrifugal Filter Units. It took several centrifuge rounds to finally get a 3 ml virus concentrate; which was used to infect the second "experiment" culture (as always, we used the same volume of aged seawater for the

control one). The volume of these cultures was of 100 ml each and the centrifuge parameters were these:

Temperature	15°C
Volume	4ml
Time	5-6 minutes/round
RPM	4500

Table 2: centrifugation parameters II

The third experiment was carried out like the second one, but we filtered the virus concentrate through 0.6 µm instead of 0.8µm, before re-centrifuged it with the Amicons.

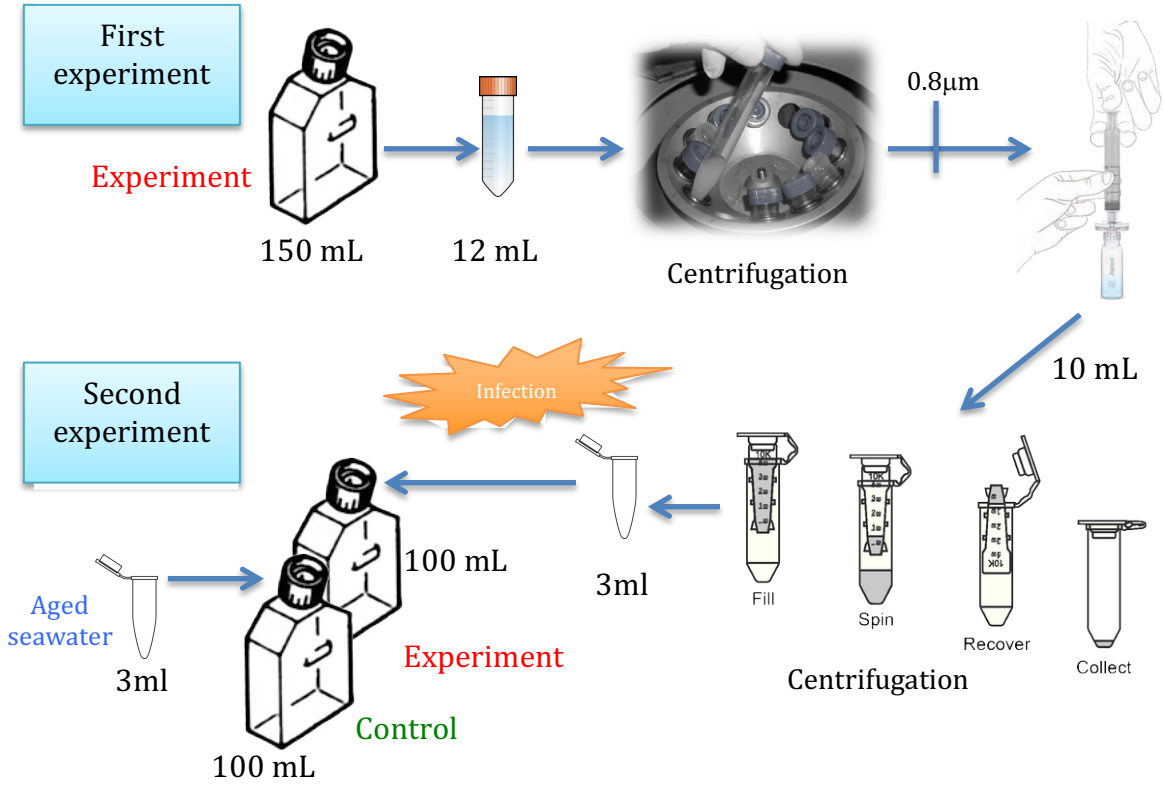


Diagram 4: From first experiment to second experiment procedure.

EFM and FCM procedures

Daily sub-samples were taken from each culture during all experiments, from the start until the end of them. Aliquots for epifluorescence microscopy (1.8ml) were fixed with 200 µl of glutaraldehyde (10% final concentration), stained with DAPI and filtered through 0.2 µm pore-size black polycarbonate filters for counting bacteria and heterotrophic flagellates.

For bacteria, we counted 40 fields with a total of 1000 bacteria and for flagellates we counted 2 transects of 10 mm each achieving from 200 to 400 cells.

Viral abundance was determined following Brussaard (2004). Aliquots for viruses were fixed with 20 µl of glutaraldehyde (25% final concentration), kept at 4°C in the dark for 15 min and deep-frozen in liquid nitrogen. Then, they were kept at -80°C. Fixed samples were stained with SYBR Green I, and run at an optimal event rate (between 100 and 800 events per second) (Marie et al. 1999), which in our cytometer corresponds to the high flow speed (Brussaard 2004). Samples were analysed on a FACSCalibur flow cytometer *(Beckton Dickinson)*, with a blue laser emitting at 488 nm, at the Institut de Ciències del Mar (CSIC) of Barcelona.

Transmission electron microscopy (TEM)

We took 5ml from each exponentially growing culture (control and experiment) from the third experiment, filtered them through 0,6µm and we stored it at 4°C. These samples were negatively stained with uranyl acetate and were observed using TEM (JEOL 1010), at the laboratory of Centres Científics i Tecnològics de la UB (CCiTUB), in Barcelona.

Firstly, we charged the grids with UVA light during 30 seconds in a Glow Discharge CTA 005
BAL-TEC, in order to get a better attached and dispersion of the samples on it.

Figure 7: JEOL 1010 TEM.

Then, we put some distilled water over a table, and we laid out a piece of parafilm on it. We put 5µl of each culture sample on the parafilm piece, and 5 drops per sample of 50µl of uranyl acetate as stain.

We placed the grid over the sample drop during 30 seconds (we repeated this procedure, but holding the grid during 60 seconds) and right after, we put the grid above the first drop of uranyl acetate, and we moved on it during 10 seconds. We repeated this procedure four times more, moving from one drop to another, and finally we removed the excess dye with a filter paper. After the grid was dried in a desiccator for 2h, negatively stained VLP were observed using TEM at an acceleration voltage of 80kV. Particle diameters were estimated using the negatively stained images.

We carried out two more experiments, with the same procedure as the third one, in order to get some ultrathin sections of the flagellates, but we never reached the minimum volume required of cell suspension.

Results

- Metagenomic's reseach

Our first investigations were focused on three viruses: *Cafeteria roenbergensis, Micromonas pusilla* and *Ostreococcus tauri* 1 and 2. We downloaded the genome of these viruses, and extracted the following gene sequences, considered to be very conserved (Colson et al., 2011):

- DNA polymerase family B (DNA pol)
- Proliferating cell nuclear antigen (PCNA)
- Ribonucleotide reductase (RNR)
- Transcription factor II (TFIIB)
- Topoisomerase II A (TopoIIB)

Then, we searched against different databases in CAMERA, and we obtained the following results:

	BLASTn						BLASTx	
C. roenbergensis	GOS contigs	GOS reads	HOT	Marine viromes	Monterey Bay	Viral metagenome in MB	Viral proteïns	NCBI ESP
Genome	100	100	15	0	9	0	100	100
DNApol	1	5	0	0	1	0	100	100
PCNA	0	0	0	0	0	0	33	100
RNR	2	3	0	0	0	0	100	100
TFIIB	1	2	0	0	0	0	13	100
TopoIIA	2	3	0	0	0	0	100	100
M. pusilla								
Genome	100	100	2	11	56	0		
DNApol	100	100	0	1	3	0		
PCNA	54	0	0	0	0	0		
RNR	100	100	0	0	1	0		
TFIIB	45	100	0	0	4	0		
TopoIIA	100	100	0	0	3	0		
O. tauri 1								
Genome	100	100	2	12	89	0		
DNApol	95	100	0	0	1	0		
PCNA	54	100	1	0	0	0		
RNR	100	100	0	0	1	0		
TFIIB	22	56	0	0	0	0		
TopoIIA	100	100	0	0	3	0		
O. tauri 2								
Genome	100	100	3	12	100	0		
DNApol	95	100	0	0	5	0		
PCNA	57	100	1	0	0	0		
RNR	100	100	0	0	1	0		
TFIIB	27	67	0	0	1	0		
TopoIIA	100	100	100	1	5	0		

Table 3: CAMERA hits for the different viral genes and genomes (E value= 10^{-10}; db alignments per query=100)

As we can see in Table 3, the number of hits was low for the searches with BLASTx (the program that compares the six-frame conceptual translation products of a nucleotide query sequence, both strands, against a protein sequence database), and very low for some BLASTn (the program that, given a DNA query, returns the most

similar DNA sequences from the DNA database that the user specifies) base dates. In this initial attempt, the number of reads was limited to 100, but even with this limitation it was obvious that in most cases the detected reads did not arrive to this value.

For this reason we decided to work only with GOS contigs and reads databases using the BLASTn routine. We also decided to analyse only the viral genomes, because they always produced the largest number of hits but, as we fixed the number of alignments in 100, we increased it to the maximum allowed by CAMERA (50000).

We repeated the search using these parameters, and we added the *Emiliania huxleyi* virus' genome. As we can see in the results (see Table 4), we obtained more number of hits than in the first search, being *M. pusilla* virus the genome with the greatest number of hits.

	BLASTn	
C. roenbergensis	GOS contigs	GOS reads
Genome	202	368
E.Huxleyi		
Genome	27	30
M. pusilla		
Genome	3214	6408
O. tauri 1		
Genome	2223	4223

Table 4: CAMERA GOS Contigs and Reads results for the different virus genomes. (E value= 10^{-10}; db alignments per query=50000)

After that, we finally decided to work with the genomes of all the marine protist viruses known, with double-stranded DNA. In Table 10 we can see all of them and we can appreciate that they infect different kind of marine protists, like amoebae from the genus *Acanthamoeba (*APMV, MGVC), chlorophyta from the genus Bathycoccus (BpV1, BpV2), *Micromonas* (MpV1), *Ostreococcus* (OlV1, OtV1, OtV2, OtV5) and *Chlorella* (PBCV-1, PBCV-AR158, PBCV-FR483, PBCV-MT325, PBCV-NY2A, TN603, CVM-1), flagellates like *Cafeteria roenbergensis* (CroV), diatoms from the genus *Chaetoceros* (ClorDNAV01, CsalDNAV, CtenDNAV06), brown algae like *Ectocarpus siliculosus* (EsV_1), Coccolithophyceae like *Emiliania huxleyi* (EhV-84, EhV-86) and the rare eukaryotic algae *Picobiliphyta* (MS584-5). We have also provided genome and virus sizes (capsid diameter) in the same table, to get used to the fact of size rank we are managing with.

With these twenty-six genomes, we searched against CAMERA GOS Reads, and the results are also represented in Table 10, in the column "Reads num. (initial)". As every search is independent, lower hits may appear to match more than one virus; and that is the reason why we created another column called "Reads num. (final)", with the number of hits obtained from a Local Blast, performed against a base date that we created with the genome of all these protist viruses.

The latest three columns of Table 10 are very important. We represented the percentage of identical matches, or "pident", (the percent identity over the alignment length), and the alignment length (the overall length of the alignment including any gaps) for each search; both calculated as an average. These parameters give us an idea of how similar is the read to each virus genome sequence. In the last column we represent the product of the number of Reads (the final one) and the alignment length. This allows us to get some idea of the role that every single virus plays in the ocean.

All results were filtered with excel, with the aim of eliminating read duplicates, and we finally found that *Micromonas pusilla* virus was most represented in the ocean, followed by *Bathycoccus* 1 and 2, *Ostreococcus sp*. and *Cafeteria roenbergensis*.

We also represented the reads distribution of these main protist viruses (see Table 11 and Annexes 12 to 18), and the number of total reads per sample location (see Table 12). In table 11 (and the other tables for all the main viruses) we represented the number of reads per sample, with the size fraction of which they were extracted from; mostly from the 0.1-0.8 μm size range. We also represented the "sequencing effort", extracted from Rusch *et al*. (2007). This parameter allows us to know the degree of assembly of each metagenomic sample and it gave us the results from the last column of the tables. These results were obtained by dividing the number of reads from each sample by its sequencing effort; we called them the "Relative virus importance", and we represented them in a world map, one per virus (see Figure 8, a to h). These maps give us an easy vision of the viral distribution around the world.

As we can see in Table 11 (and Annexes 12 to 18), the sample locations with more reads for each virus are the following ones:

Virus	Sample	Location	Num. Reads
MpV1	GS014	South of Charleston, SC	436
BpV1	GS002	Gulf of Maine	385
BpV2	GS006	Bay of Fundy, Nova Scotia	195
OtV2	GS013	Off Nags Head, NC	74
OlV1	GS007	Northern Gulf of Maine	49
OtV5	GS012	Chesapeake Bay, MD	47
OtV1	GS012	Chesapeake Bay, MD	26
CroV	GS021	Gulf of Panama	5

Table 5: Sample Locations with more Reads for each virus

The sample locations with a major sequencing effort for each virus are:

Virus	Sample	Location	Seq. effort
BpV1	GS000c	Sargasso Stations 3	1382197
BpV2	GS108	Coccos Keeling, Inside Lagoon	1382197
CroV	GS112	Indian Ocean	1156475

MpV1	GS000a	Sargasso Station 11	644551
OlV1	GS000a	Sargasso Station 11	644551
OtV2	GS000a	Sargasso Station 11	644551
OtV5	GS000c	Sargasso Stations 3	368835
OtV1	GS117a	St. Anne Island, Seychelles	346952

Table 6: Sample locations with a major sequencing effort for each virus.

And the sample locations with a major relative virus importance for each virus are:

Virus	Sample	Location	Relative virus importance
BpV2	GS006	Bay of Fundy, Nova Scotia	$3.27*10^{-3}$
MpV1	GS007	Northern Gulf of Maine	$4.12*10^{-3}$
BpV1	GS007	Northern Gulf of Maine	$6.45*10^{-3}$
OlV1	GS007	Northern Gulf of Maine	$9.61*10^{-4}$
OtV1	GS007	Northern Gulf of Maine	$3.73*10^{-4}$
OtV5	GS007	Northern Gulf of Maine	$6.08*10^{-4}$
OtV2	GS007	Northern Gulf of Maine	$9.02*10^{-4}$
CroV	GS021	Gulf of Panama	$3.79*10^{-5}$

Table 7: Sample locations with a major relative virus importance

As we can see in Table 14, the three locations with more reads are the following ones:

All samples	All locations	Num. Reads
GS002	Gulf of Maine	847
GS007	Northern Gulf of Maine	730
GS003	Browns Bank, Gulf of Maine	550

Table 8: Three main Reads locations.

We counted a number of 5048 reads total, distributed in the following mode:

Virus	Num. Reads Total
MpV1	1694
BpV1	1630

BpV2	450
OtV2	450
OlV1	326
OtV5	271
OtV1	188
CroV	39

Table 9: Number of reads total for each virus.

As we said before, in Figure 8 we can see the Relative virus importance per sample, represented in a world map for each virus. All results are distributed along the Indic and Atlantic Ocean (East coast of North America).

Base Date Position	Full name	GenBank Accession	Abbreviation	Host	Genome size (nt)	Virus size (capsid diameter in nm)	READS num (initial)	READS num (final)	Percentage of identical matches (avg)	Alignment length (avg)	READS num (final)*Alignment length
1	*Acanthamoeba polyphaga* mimivirus	NC_014649	APMV	*Acanthamoeba polyphaga*	1181549	350	818	37	8,22E+01	1,91E+02	7,08E+03
2	*Acanthocystis turfacea* chlorella virus 1	NC_008724	ATCV-1	*Chlorella*	288047	160 ± 60 nm (Phycodnaviruses mean diameter)	44	10	8,07E+01	2,30E+02	2,30E+03
3	*Bathycoccus* sp. RCC1105 virus BpV1	NC_014765	BpV1	Bathycoccus	198519		3439	1630	9,02E+01	7,69E+02	1,25E+06
4	*Bathycoccus* sp. RCC1105 virus BpV2	HM004430	BpV2	Bathycoccus	187069		3111	1052	8,99E+01	8,66E+02	9,11E+05
5	*Cafeteria roenbergensis* virus	NC_014637	CroV	Cafeteria roenbergensis	617453	300	336	39	8,08E+01	5,74E+02	2,24E+04
6	*Chaetoceros lorenzianus* DNA virus	NC_015211	ClorDNAV01	*Chaetoceros lorenzianus*	5813		0	-	-	-	-
7	*Chaetoceros salsugineum* DNA virus	NC_007193	CsalDNAV	*Chaetoceros salsugineum*	6000	38	0	-	-	-	-
8	*Chaetoceros tenuissimus* DNA virus	NC_014748	CtenDNAV06	*Chaetoceros tenuissimus*	5639	31	0	-	-	-	-
9	*Ectocarpus siliculosus* virus 1	NC_002687	EsV-1	*Ectocarpus siliculosus*	335593	100-220 (phycodnaviridae mean diameter)	0	-	-	-	-
10	*Emiliania huxleyi* virus 84	JF974290	EhV-84	*Emiliania huxleyi*	395820	160-180	0	-	-	-	-
11	*Emiliania huxleyi* virus 86	NC_007346	EhV-86	*Emiliania huxleyi*	407339	160-180	8	1	85.45	8,59E+02	8,59E+02
12	*Feldmannia* species virus	NC_011183	FsV-158	*Feldmannia*	154641	150	6	0	0,00E+00	0,00E+00	0,00E+00
13	*Megavirus chiliensis*	NC_016072	MGVC	*Megavirus chiliensis*	1259197	400	95	6	8,53E+01	1,24E+02	7,45E+02
14	*Micromonas* sp. RCC1109 virus MpV1	NC_014767	MpV1	Micromonas	184095	-	4850	1694	8,07E+01	7,42E+02	1,26E+06
15	*Ostreococcus lucimarinus* virus OlV1	NC_014766	OlV1	Ostreococcus lucimarinus	194022	120 (Ostreococcus mean diameter)	3421	326	7,80E+01	7,57E+02	2,47E+05
16	*Ostreococcus tauri* virus 1	NC_013288	OtV1	Ostreococcus tauri	191761	121 (Ostreococcus mean diameter)	3194	188	7,90E+01	7,37E+02	1,39E+05
17	*Ostreococcus tauri* virus 2	NC_014789	OtV2	Ostreococcus tauri	184409	122 (Ostreococcus mean diameter)	3399	450	8,09E+01	8,03E+02	3,61E+05
18	*Ostreococcus* virus OsV5	NC_010191	OtV5	Ostreococcus	185373	122	3191	271	7,95E+01	7,25E+02	1,96E+05
19	*Paramecium bursaria* chlorella virus 1c	NC_000852	PBCV-1	*Chlorella*	330611	174	104	63	9,32E+01	7,25E+01	4,57E+03
20	*Paramecium bursaria* chlorella virus AR158	NC_009899	PBCV-AR158	*Chlorella*	344691	160 ± 60 nm (Phycodnaviruses mean diameter)	16	1	8,65E+01	5,20E+01	5,20E+01
21	*Paramecium bursaria* chlorella virus NY2A	NC_009898	PBCV-FR483	*Chlorella*	321240	161 ± 60 nm (Phycodnaviruses mean diameter)	41	0	-	-	-
22	*Paramecium bursaria* chlorella virus FR483	NC_008603	PBCV-MT325	*Chlorella*	314335	162 ± 60 nm (Phycodnaviruses mean diameter)	44	2	9,51E+01	6,16E+02	1,23E+03
23	*Paramecium bursaria* chlorella virus MT325	DQ491001	PBCV-NY2A	*Chlorella*	368683	163 ± 60 nm (Phycodnaviruses mean diameter)	17	0	-	-	-
24	*Phycobiliphytes* virus	HQ322117	MS584-5	*Picobiliphytes*	1832	-	0	-	-	-	-
25	*Acanthocystis turfacea* chlorella virus	-	TN603	*Chlorella*	328767	160 ± 60 nm (Phycodnaviruses mean diameter)	87	2	8,69E+01	2,06E+02	4,11E+02
26	*Paramecium bursaria* Chlorella virus	-	CVM-1	*Chlorella*	301079	160 ± 60 nm (Phycodnaviruses mean diameter)	38	4	9,37E+01	3,20E+02	1,28E+03

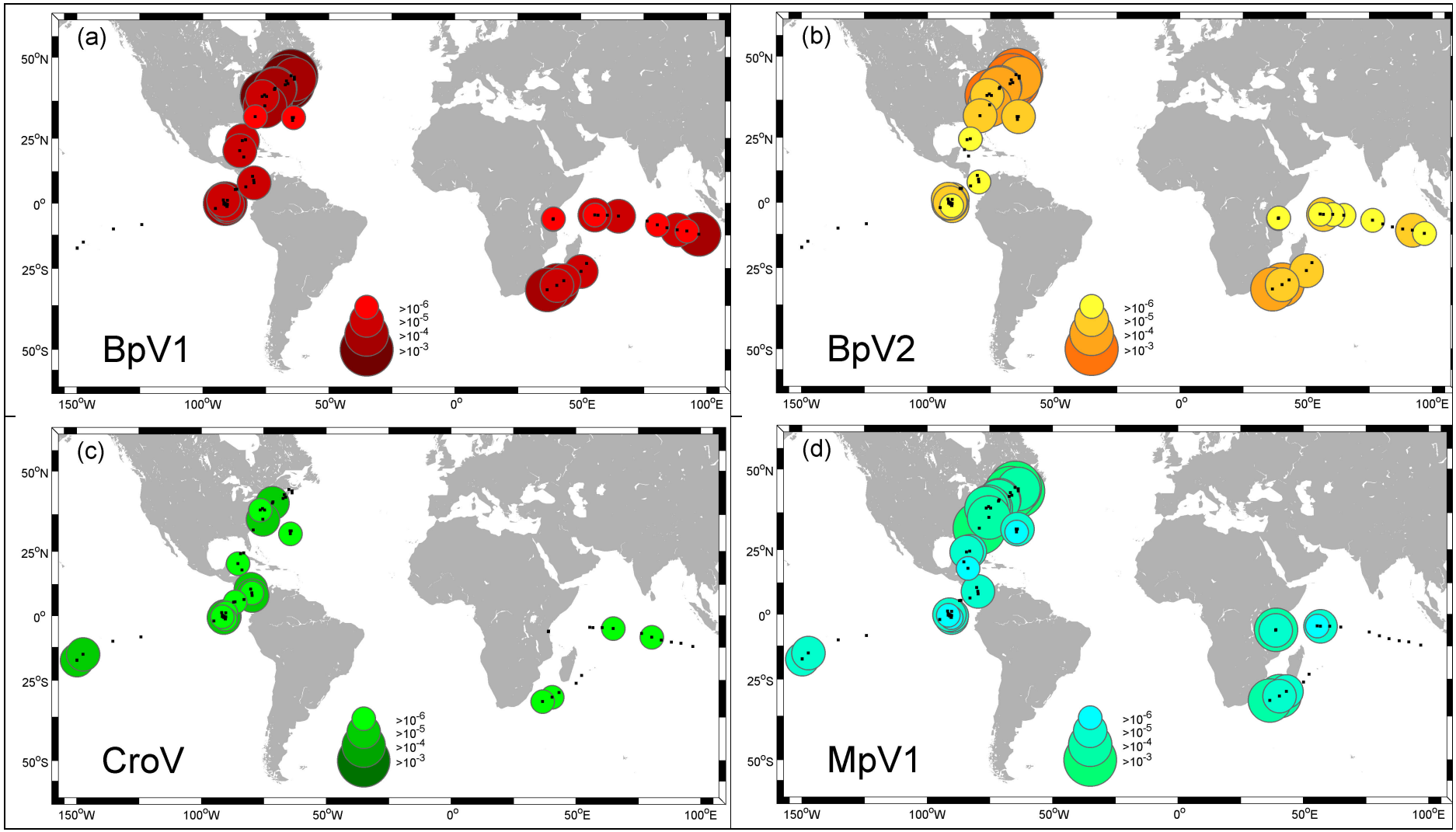
Table 10: Results of the metagenomic's research (Viruses marked in yellow are the most represented in the oceans and in red, the second ones).

Sample	Localitation	num Reads	Filter size	Seq. effort	Relative virus importance
GS002	Gulf of Maine	385	0.1-0.8	121590	3,17E-03
GS007	Northern Gulf of Maine	329	0.1-0.8	50980	6,45E-03
GS003	Browns Bank, Gulf of Maine	274	0.1-0.8	61605	4,45E-03
GS006	Bay of Fundy, Nova Scotia	218	0.1-0.8	59679	3,65E-03
GS010	Cape May, NJ	121	0.1-0.8	78304	1,55E-03
GS013	Off Nags Head, NC	73	0.1-0.8	138033	5,29E-04
GS031	Upwelling, Fernandina Island	30	0.1-0.8	436401	6,87E-05
GS123	International water between Madagascar and South Africa	30	0.1-0.8	107966	2,78E-04
GS122a	International waters between Madagascar and South Africa	29	0.1-0.8	52959	5,48E-04
GS004	Outside Halifax, Nova Scotia	27	0.1-0.8	112278	2,40E-04
GS008	Newport Harbor, RI	15	0.1-0.8	79303	1,89E-04
GS009	Block Island, NY	15	0.1-0.8	129655	1,16E-04
GS035	Wolf Island	11	0.1-0.8	140814	7,81E-05
GS027	Devil's Crown, Floreana Island	8	0.1-0.8	222080	3,60E-05
GS029	North James Bay, Santigo Island	8	0.1-0.8	131529	6,08E-05
GS036	Cabo Marshall, Isabella Island	8	0.1-0.8	77538	1,03E-04
GS108	Coccos Keeling , Inside Lagoon	8	0.8 - 3.0	50095	1,60E-04
GS122b	International waters between Madagascar and South Africa	7	0.22-0.8	368835	1,90E-05
GS000c	Sargasso Stations 3	6	0.1-0.8	1382197	4,34E-06
GS012	Chesapeake Bay, MD	3	0.1-0.8	189052	1,59E-05
GS028	Coastal Floreana	3	0.1-0.8	126162	2,38E-05
GS121	International water between Madagascar and South Africa	3	0.1-0.8	257581	1,16E-05
GS016	Outside Seychelles, Indian ocean	2	0.8 - 3.0	49597	4,03E-05
GS017	Yucatan Channel	2	0.8 - 3.0	52118	3,84E-05
GS021	Gulf of Panama	2	0.1-0.8	110720	1,81E-05
GS034	North Seamore Island	2	0.1-0.8	128885	1,55E-05
GS117a	St. Anne Island, Seychelles	2	0.1-0.8	127122	1,57E-05
GS014	South of Charleston, SC	1	0.1-0.8	131798	7,59E-06
GS109	Indian Ocean	1	0.1-0.8	134347	7,44E-06
GS110b	Indian Ocean	1	0.1-0.8	59812	1,67E-05
GS112b	Indian Ocean	1	0.1-0.8	346952	2,88E-06
GS114	500 Miles west of the Seychelles in the Indian Ocean	1	0.1-0.8	46052	2,17E-05
GS117b	St. Anne Island, Seychelles	1	0.1-0.8	348823	2,87E-06
GS120	Madagascar Waters	1	0.8 - 3.0	50609	1,98E-05
GS148	East coast Zanzibar (Tanzania), offshore Paje lagoon	1	0.1-0.8	107741	9,28E-06
GS149	West coast Zanzibar (Tanzania), harbour region	1	0.1-0.8	110984	9,01E-06

Table 11: BpV1 Reads distribution.

All samples	All Locations	BpV1	BpV2	CroV	MpV1	OlV1	OtV1	OtV2	OtV5	Total Reads/sample
GS002	Gulf of Maine	385	56		264	45	16	56	25	847
GS007	Northern Gulf of Maine	329	46		210	49	19	46	31	730
GS003	Browns Bank, Gulf of Maine	274	36		133	42	9	36	20	550
GS014	South of Charleston, SC	1	16		436	18	15	16		502
GS013	Off Nags Head, NC	73	74	2	108	23	26	74	15	395
GS006	Bay of Fundy, Nova Scotia	218	9		81	9	9	9	7	342
GS010	Cape May, NJ	121	19		63	10	10	19	28	270
GS012	Chesapeake Bay, MD	3	70	1	26	37	26	70	24	257
GS008	Newport Harbor, RI	15	40		77	8	20	40	26	226
GS011	Delaware Bay, NJ		41		21	43	20	41	47	213
GS004	Outside Halifax, Nova Scotia	27	11		56	3	2	11	6	116
GS123	International water between Madagascar and South Africa	30	4	1	51	7	1	4	5	103
GS122a	International waters between Madagascar and South Africa	29	1	1	41	8	4	1	4	89
GS009	Block Island, NY	15	15	2	28	5		15	8	88
GS000c	Sargasso Stations 3	6			26	4			1	37
GS005	Bedford Basin, Nova Scotia		6		7	4	6	6	6	35
GS031	Upwelling, Fernandina Island	30		4	1					35
GS149	West coast Zanzibar (Tanzania), harbour region	1	1		17	1		1	2	23
GS036	Cabo Marshall, Isabella Island	8			5	1			4	18
GS148	East coast Zanzibar (Tanzania), offshore Paje lagoon	1	1		8	2		1	2	15
GS020	Lake Gatun			1	6	1	2		1	11
GS027	Devil's Crown, Floreana Island	8		1	2					11
GS035	Wolf Island	11								11
GS122b	International waters between Madagascar and South Africa	7			1	1			2	11
GS029	North James Bay, Santigo Island	8		2						10
GS117a	St. Anne Island, Seychelles	2			2		3		3	10
GS015	Off Key West, FL		1		5	2		1		9
GS021	Gulf of Panama	2	1	5						8
GS108	Coccos Keeling , Inside Lagoon	8								8
GS034	North Seamore Island	2			3	2				7
GS016	Outside Seychelles, Indian ocean	2	1		2			1		6
GS028	Coastal Floreana	3			3					6
GS121	International water between Madagascar and South Africa	3			2			1		6
GS051	Rangirora Atoll			2	3					5
GS000a	Sargasso Station 11		1		1	1		1		4
GS019	Northeast of Colon			4						4
GS017	Yucatan Channel	2		1						3
GS115	Indian Ocean	1							2	3
GS116	Outside Seychelles, Indian Ocean				1				2	3
GS000b	Sargasso Station 11			1	1					2
GS032	Mangrove on Isabella Island			2						2
GS049	Moorea, Outside Cooks Bay			2						2
GS112	Indian Ocean			2						2
GS114	500 Miles west of the Seychelles in the Indian Ocean			2						2
GS000d	Sargasso Station 13				1					1
GS018	Rosario Bank				1					1
GS023	30 miles from Cocos Island			1						1
GS026	134 miles NE of Galapagos			1						1
GS048a	Moorea, Cooks Bay				1					1
GS109	Indian Ocean	1								1
GS110b	Indian Ocean	1								1
GS112a	International waters between Madagascar and South Africa			1						1
GS112b	Indian Ocean	1								1
GS117b	St. Anne Island, Seychelles	1								1
GS120	Madagascar Waters	1								1
GS113	Indian Ocean									0
	TOTAL	1630	450	39	1694	326	188	450	271	5048

Table 12: Number of Reads per sample location.


(a)
BpV1
(b)
BpV2
(c)
CroV
(d)
MpV1
50°N
25°N
0°
25°S
50°S
150°W
100°W
50°W
0°
50°E
100°E
>10-6
>10-5
>10-4
>10-3

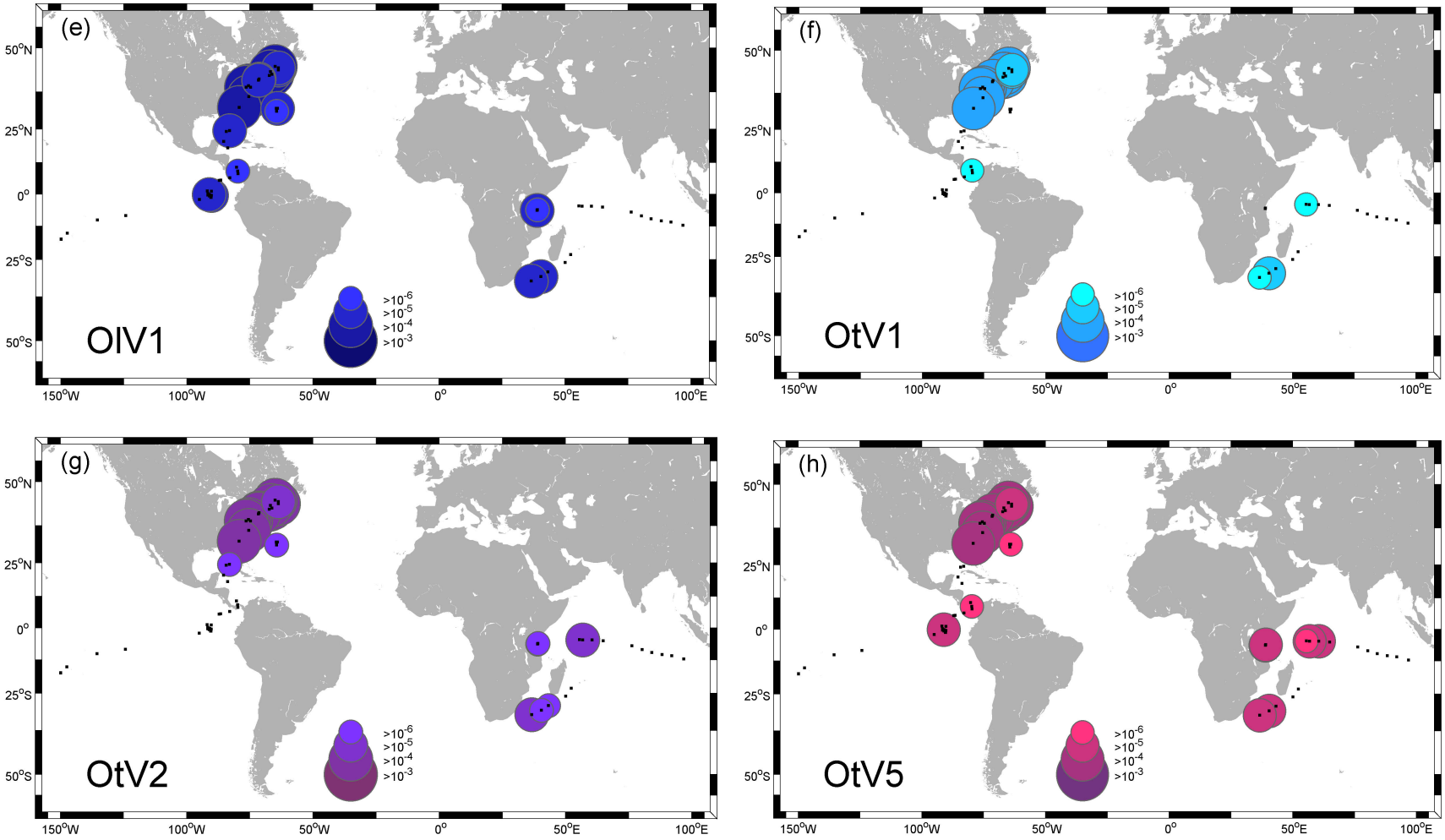


Figure 8: Viruses World distribution calculated on the basis of their "Relative virus importance". Black dots are sampling sites.

- Dynamics of bacteria, flagellates and viruses.

Specifications of all experiments, including the previous enrichments before infection and batches to get material for TEM observations are shown in Table 13.

Initial date	Duration (d)	Experiment	Variables	Inoculum's day
15/10/12	10	Enrichment I	FLV, BAC, VIR	0
29/10/12	7	Enrichment II	FLV, BAC, VIR	0
31/10/12	15	I	FLV, BAC, VIR	5.81
14/1/13	7	II	FLV, BAC, VIR	0
6/3/13	12	II	FLV, BAC, VIR	0
23/4/13	6	TEM I	FLV, BAC, VIR	0
8/5/13	9	TEM II	FLV, BAC, VIR	0

Table 3: Specifications of all experiments. FLV: flagellate's, BAC: bacteria, VIR: viruses

During experiment I, in the control culture *Developayella* sp grew exponentially up to a maximum abundance of $5.9*10^4$ flag ml^{-1} at the fifth day (Fig. 9A). Then, decreased gradually reaching an abundance of $1.1*10^4$ flag ml^{-1} at the end of the experiment (day 15). Bacteria, which had an initial concentration of $1.2*10^7$ cells ml^{-1}, decreased in number as the flagellates were growing up until they reached a minimum of $3.4*10^6$ cells ml^{-1} at the seventh day and reaching an abundance of $5.3*10^6$ bact ml^{-1} at the end of the experiment.

Regarding to the experiment culture (Fig. 9B), we observed a very similar trend, with an exponential growth of the flagellates until a maximum of $5.56*10^4$ flag. ml^{-1} at the sixth day (Fig. 9B) and a minimum of $8.06*10^3$ flag ml^{-1} at day fifteen. The same bacteria as the control culture were present in the experiment one ($1.2*10^7$ bact ml^{-1}). This number decreased until day six, with $3.7*10^7$ bact ml^{-1}, and they recovered a little more than the control culture, with a final concentration of $7.23*10^7$ bact ml^{-1} (See also Annex 3). However, 24 h after to add the viral concentrate there was a sudden decrease of *Developayella sp*. of -0.51 (d^{-1}) higher than in the control (-0.23 d^{-1}), in the same period (Table 4).

In Experiment I, the first value for viral abundance was collected at time 5d, and viral concentrate was inoculated during the exponential face of *Developayella sp.*, at day 6. However, as is shown in Fig. 9 (A and B), the abundances of viruses from control and experiment I, presented similar values in the day that we add the viruses concentrate (Fig. 9 B). Although, there is a difference, in viral abundance between the control and the experiment, coinciding with the higher decrease of *Developayella sp.* in the experiment than in the control (Fig. 9 A,B). Also the final concentration of virus in the experiment ($1.24*10^6$ virus ml^{-1}) was higher than in the control. For more details of viral abundance dynamics see annex 4 and 5.

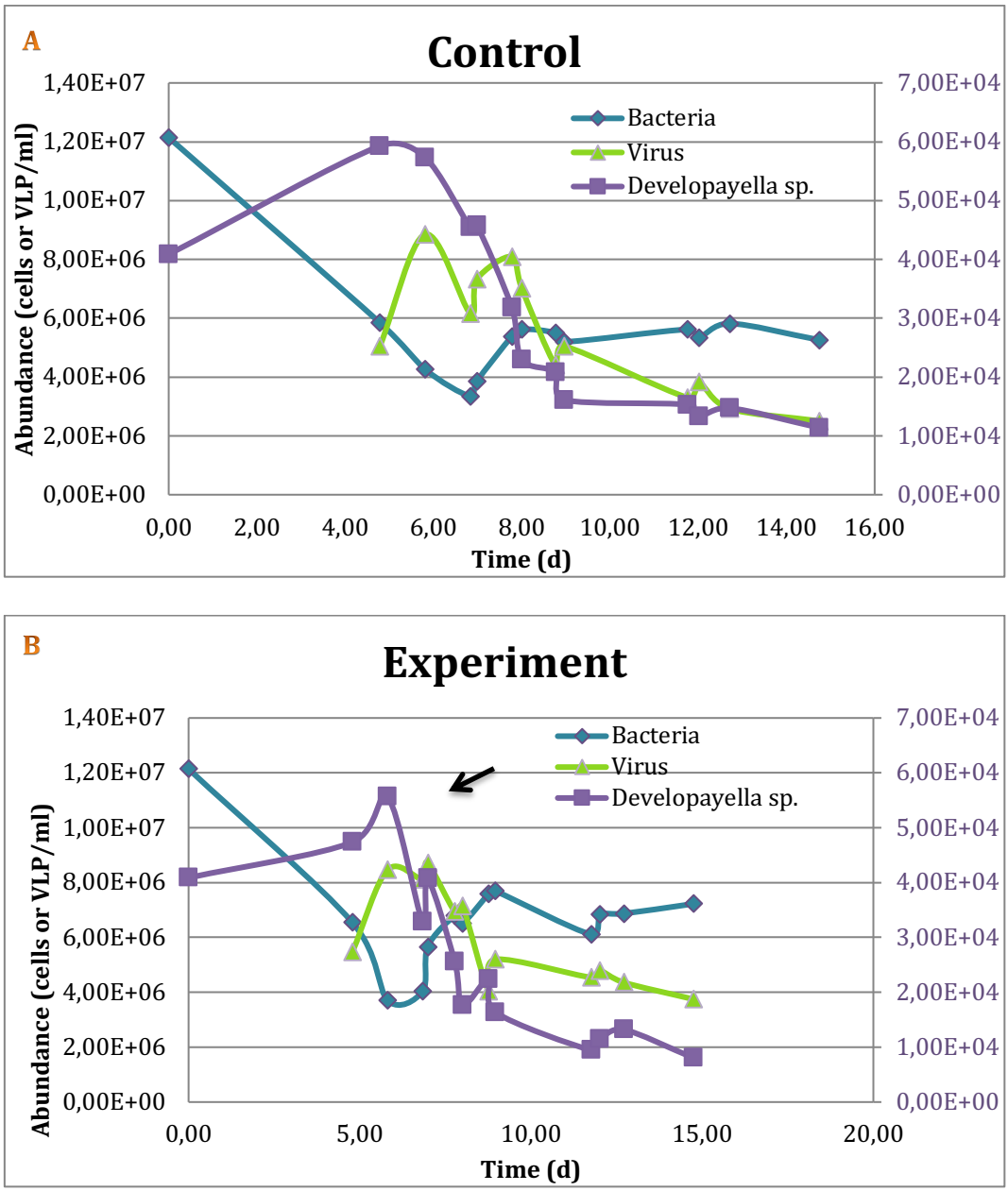


Figure 9: graphics of the cells and VLP abundances of the first experiment (A: control; B: experiment). Arrow indicates the moment of the infection.

The results from the second experiment were quite different from the previous ones (see also Annex 6). Both cultures, control (Fig. 10 A) and experiment (Fig. 10 B), started with a flagellate concentration of $4.55*10^4$ flag ml^{-1}. They decreased to $1.70*10^4$ flag ml^{-1} and $7.1*10^3$ flag ml^{-1}, respectively, at the end of the process (seventh day), but the experiment culture did it quickly and the control one presented a peak at day 3, with $5.35*10^4$ cells ml^{-1}.

Nevertheless, bacteria from the control culture didn't change very much; its abundance always ranged around 10^7 bact ml^{-1} while the ones in the experiment culture grew up more ($1.91*10^7$ bact ml^{-1} at day 7).

In fact, they reached a peak of $2.07*10^7$ cells ml^{-1} (day 3) that fits with a decrease in the flagellate's abundance (day 2), and a high peak of viruses of $2.29*10^7$ viruses ml^{-1} at day 2 (see Annexes 7 and 8).

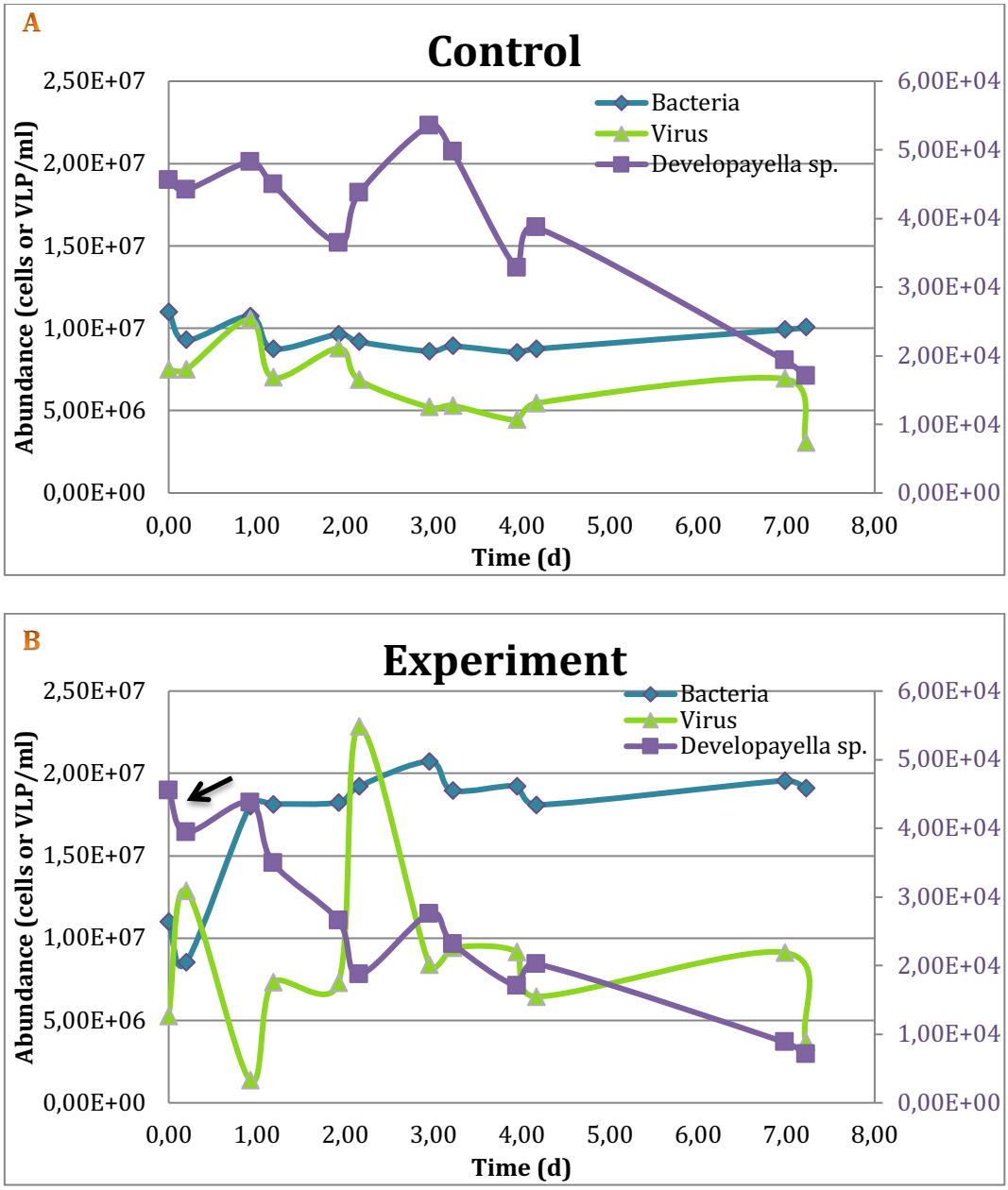


Figure 10: graphics of the cells and VLP abundances of the second experiment (A: control; B: experiment). Arrow indicates the moment of the infection.

Viruses from the control tended to decrease with time since day 1, and from the experiment culture decreased right after they were inoculated, at day 1. Two days after (day 3) viral abundance reached a peak just when, a minimum flagellate abundance was detected, and finally they decreased again (Fig. 10B). Comparatively the decreasing rates of *Developayella sp.*, in both control and experiment, since virus inoculum until viruses reached a peak (day 3) (Fig. 10B) were -0.13 d^{-1}, and -0.60 d^{-1}, respectively (Table 14).

In the third experiment, abundance of flagellates in the control (Fig. 11 A) and in the experiment (Fig. 11 B) cultures, showed an exponential growth curve achieving a peak at day 2. Both cultures started with $9.71*10^3$ flag ml^{-1}, the control one reached a maximum of $2.04*10^4$ flag ml^{-1} at day 2 and this value was maintained until days 5, to decrease up to $9.33*10^3$ flag ml^{-1}; whereas *Developayella sp*. in the experiment (Fig. 11B) decreased after day 2 until it reached an abundance of $7.51*10^3$ flag ml^{-1} at the end of the experiment (see also Annex 9).

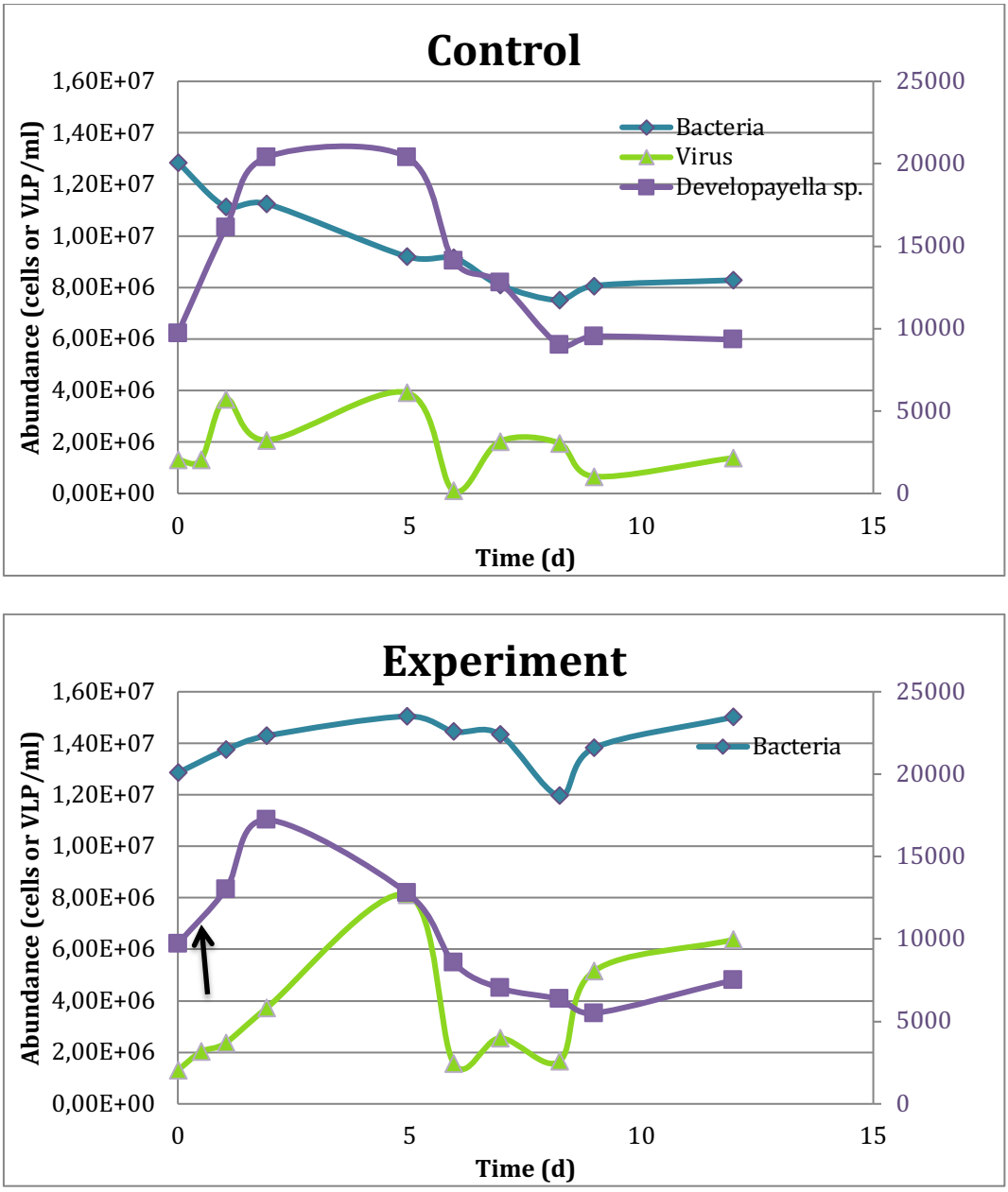


Figure 11: graphics of the cells and VLP abundances of the third experiment (A: control; B: experiment). Arrow indicates the moment of the infection.

On the other hand, bacteria from both cultures started with $1.29*10^7$ bact ml^{-1}, but they presented different behaviours. In the control culture bacteria decreased until they reached a final concentration of $8.28*10^6$ cells ml^{-1} at day 12 (with a slight

increase as the flagellates tended to decrease at the end of the experiment), but from the experiment culture bacteria grew up until day 12, with a final concentration of $1.50*10^7$ cells ml^{-1}. Despite they suffered a minimum at the ninth day ($1.20*10^7$ cells ml^{-1}), they were always around 10^7 cells ml^{-1}.

The abundance of viruses from the control culture suffered little oscillations, but always tended to be around 10^6 viruses ml^{-1}. It reached a maximum at the fifth day with $3.91*10^6$ virus ml^{-1} with a tendency to decrease until the end. Viral abundance from the experiment culture, after inoculation increased until reaching a peak at day 5 with $8.12*10^6$ viruses ml^{-1}, that coincides with the quickly decrease of the flagellate's abundance. Then, the growth rates of *Developayella sp.* in both cultures within the same period (3 days) after viral inoculation was higher in the control (0.04) than in the experiment (-0.17) (Table 14). For more details on virus dynamics see annex 10 and 11.

Experiment		Flagellate's growth decrease (d^{-1})	Period Virus inoc. (d^{-1})
I	Control	-0.23	1
	Experiment	-0.51	1
II	Control	-0.14	2.5
	Experiment	-0.60	2.5
III	Control	0.05	4
	Experiment	-0.17	4

Table 4: Flagellate growth's decrease and days after virus inoculation for the three main experiments.

In summary these results indicate that flagellate's from the experiment cultures tended to decrease when virus increase.. In addition bacterial abundance was also maintained at higher values in the experiment than in the control cultures when flagellates abundance was depleted, presumably due to viruses. This let us believe that viruses infected the flagellate's experiment culture.

Discussion

With this study we wanted to increase the knowledge of the virus protists diversity by looking at the global distribution of the main marine protist viruses and trying to isolate in the lab specific viruses for *Developayella sp.*

The first objective was achieved by the existent metagenomic database, specifically with the CAMERA GOS Reads database. Despite there is so much to discover, and lots of genomes to sequence, we could determine that MpV1 and BpV1 and BpV2 (chlorophyta viruses) are the main viruses represented in the oceans with the information that we have until now.

Even though it was the database that brought us more hits, we can't forget that all the metagenomes were obtained during the Global Ocean Sampling (GOS) expedition (see Figure 12), in which the surface water samples were collected across several-thousand km transect, from the North Atlantic through the Panama Channel and ending in the South Pacific (Rusch et al, 2007). It was an extensive work but we can't be completely sure that our distribution is faithful with the real one, because we will be always subjected to where the samples were taken from (and this is evident in our viral distributions from Figure 8).

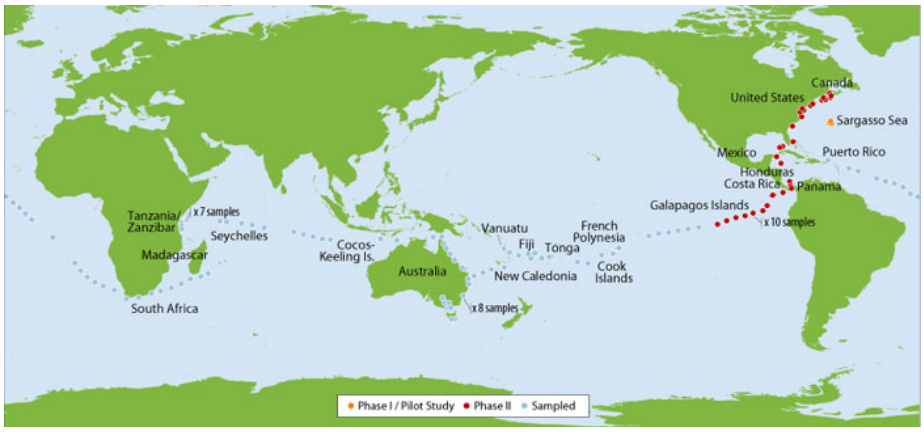


Figure 12: Sorcerer II Expedition circumnavigation route and analysis progress as of January 2007.

Nevertheless, we know that our Relative virus importance (RVI) depends on the number of reads, and their sequencing effort. One could think that these distributions may be distorted because of the human manipulation, RVI will be reduced by increasing the sequencing effort, and the opposite thing will happen by reducing it; but the sampling site with higher RVI coincide with the second one with more number of reads. In this case, we can conclude that RVI depends on the number of reads founded in each sampling site, and this presumably corresponds to its natural abundance in seawater. For example, MpV1 is the most represented virus in the ocean (in our study), and is the same one with higher RVI and more number of reads.

Moreover, we can be reassured that we worked with the correct material, because the size fraction sampled was the appropriate to get viral metagenomes (0.1-0.8 μm; mostly).

Regarding to the second objective, we found that in our experiment cultures *Developayella* sp. presented a major decrease than the control ones, presumably, caused by the lysis of the heterotrophic flagellate (see Table 14). Our results agree with the ones found by Bratbak et al. (1998), which worked with the haptophyta *Phaeocystis pouchetii*. They studied the carbon flow and population dynamics in a phytoplankton-DOC-bacteria food chain during viral lysis of the phytoplankton population, and found that viral infection perturbed the exponential growth and decimated the *Phaeocystis* SP. population within 3 days while in the non-infected culture growth continued undisturbed. If we compare their population dynamics graphics with ours, tendencies of protists, bacteria and viruses, follow a similar pattern that is: phytoplankton abundance dropped sharply after the viral addition, and right after, they obtained a peak of DOC, followed by an increase of viruses and bacteria. On the other hand, in the non-infected culture there was no substantial increase in DOC or bacterial biomass compared to the algal biomass, just like our experiments took place.

Despite we tried to maintain our cultures without other organisms that could interfere with our experiment, we suspect that there could have been some bacteriophages. One can thought that they were the reason why flagellates decreased their abundance but, as we can see in Figures 9 to 11, bacteria was always higher in the experiment cultures than in the control ones, so we think that was a heterotrophic flagellate virus the source of the flagellate's decrease.

With the aim to confirm this hypothesis, we tried to observe by TEM infected flagellates and presumably free eukaryotic viruses. For that we collected samples from the third experiment and control cultures along the exponential growth phase of flagellates. Unfortunately, we did not have enough material to run ultra-thin sections to observe if they were infected or not. . We also tried a couple of more time but again, our flagellates abundance was always lower than that needed to get a pellet of cells. Without this proof, we can't confirm that we were in front of a heterotrophic flagellate virus. However when examining the morphology of free viruses present in the sample (Fig. 13 A-D) with sizes between 43 nm (Fig. 13 C) and 100 nm (Fig. 13 B) it seems to correspond to the ones described for other protist viruses as is observed in Wilson et al. (2006). These authors isolated viruses from *Phaeocystis* sp., which are also untailed and with sizes near 100 nm.

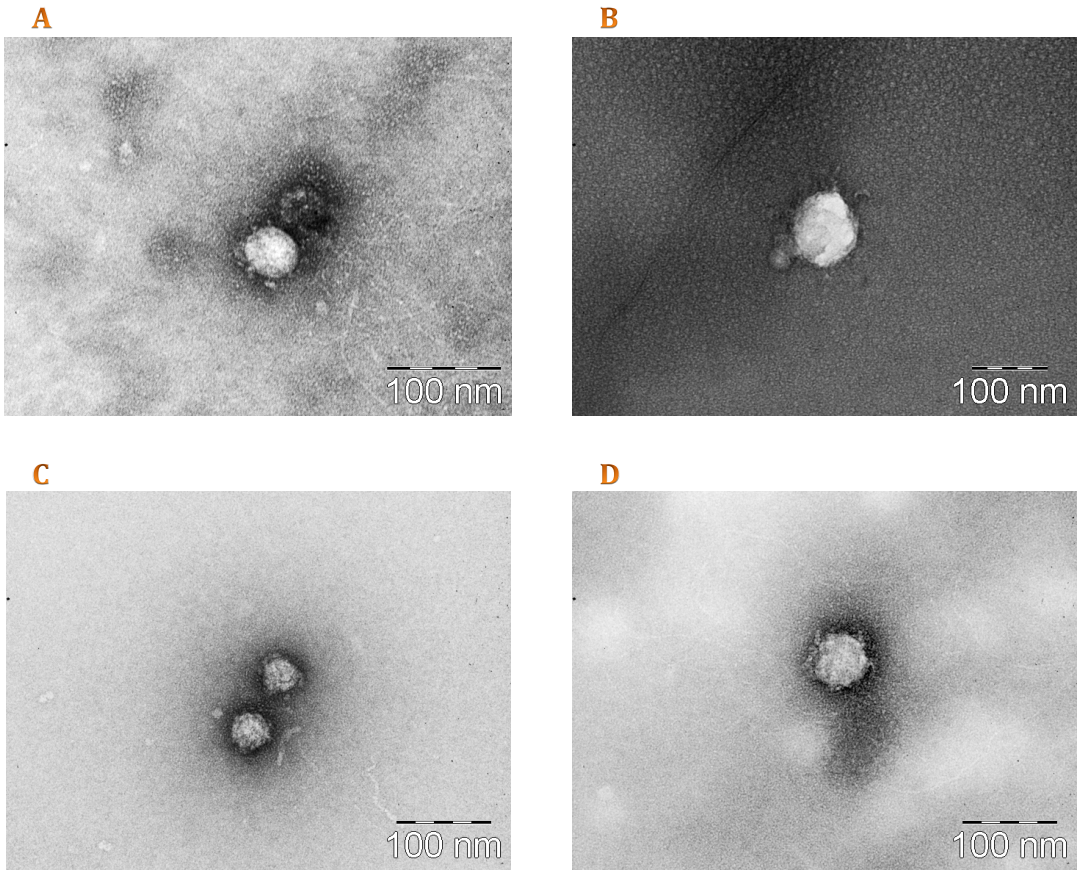


Figure 13: Transmission electron micrographs of the third experiment culture. VLPs (A,B,C and D).

Thanks to this work, we have now a wide vision of the actual protist viruses in our oceans. We know that we are subject to our limitations, like advanced technology, new laboratory procedures and, why not, some luck.

Maybe we couldn't complete all the purposes, with which we started the investigation, but we are proud to say that we done it well, but it didn't happen.

As we said before, there is so much to discover, some methods, techniques and instruments to improve and, little by little, everything comes to light.

Bibliography

Bratbak, G., Heldal, M., Norland, S., Thingstad, T.F. (1990) Viruses as a partners in sping bloom microbial trophodynamics. Applied and Environmental Microbiology. Volume 56, No.5: 1400-1405

Breibart, M., Thomson, L.K., Suttle, C.A., and Sullivan, M.B. (2007) Exploring the vast diversity of marine viruses. Oceanography: 135-139

Brussaard, C., Payet, J., Winter, C. and Weinbauer, M. (2010). Quantification of aquatic viruses by flow cytometry. MAVE Chapter 11: 102-109.

Colson, P., Gimenez, G., Boyer, M., Fournous, G. and Raoult, D. (2011). The giant Cafeteria roenbergensis virus that infects a widespread marine phagocytic protest is a new member of the fourth domain of life. PLos ONE. Volume 6. Issue 4: 1-11.

Del Campo, J., Not, F., Forn, I., Sieracki, M.E., Massana, R. Taming the smallest predators of the oceans. The ISME Journal: 1-8

Fischer, M., Allen, M., Wilson, W. and Suttle, C. (2010). Giant virus with a remarkable complement of genes infects marine zooplankton. PNAS. Volume 107, Num. 45: 19508-19513.

Fuhrman, J.A. (1999) Marine viruses and their biogeochemical and ecological effects. Nature. Volume 399: 541-548

Fuhrman, J.A. and Suttle, C.A. (1993) Viruses in marine planktonic systems. Oceanography. Volume 6, No. 2: 51-63

Garza, D.R., and Suttle, C.A. (1995) Large double-stranded DNA viruses wich cause the lysis of a marine heterotrophic nanoflagellate (*Bodo* sp.) occur in natural marine viral communities. Aquatic Microbial Ecology. Volume 9: 203-210

Logares, R., S. Audic, S. Santini, M.C. Pernice, C. de Vargas and R. Massana (2012) Diversity patterns and activity of uncultured marine heterotrophic flagellates unveiled with pyrosequencing. ISME Journal. Volume 6: 1823-1833.

Massana, R., Balagué, V., Guillou, L., Pedrós-Alió, C. (2004) Picoeukaryotic diversity in an oligotrophic coastal site studied by molecular and culturing approaches. FEMS Microbiology Ecology. Volume 50: 231-243.

Massana, R., del Campo, J., Dinter, C., Sommaruga, R. (2007) Crash on population of the marine heterotrophic flagellate Cafeteria roenbergensis by viral infcetion. Environmental Microbiology. Volume 9. No. 9: 2660-2669

Massana, R., del Campo, J., Dinter, C. and Sommaruga, R. (2007). Environmental Microbiology. Volume 9(11): 2660-2669.

Nagasaki, K., Ando, M., Imai, I., Itakura, S. and Ishida, Y.(1993). Virus-like particles in an apochlorotic flagellate in Hiroshima Bay, Japan. Marine Ecology Progress Series. Volume 96:307-310.

Nagasaki, K. and Bratbak, G. (2010). Isolation of viruses infecting photosynthetic and nonphotosynthetic protists. MAVE Chapter 10: 92-101.

Ogata, H., Monier, A. and Claverie, J-M. (2011). Distribution of Giant Viruses in Marine Environments. Global Change: Mankind-Marine Environmental Interactions, Proceedings of the 13th French-Japanese Oceanography Symposium: 157-162.

Proctor , L.M., and Fuhrman, J.A. (1990) Viral mortality of marine bacteria and cyanobacteria. Nature. Volume 343: 60-62

Proctor , L.M., and Fuhrman, J.A. (1992) Mortality of marine bacteria in response to enrichment of the virus size fraction from seawater. Marine ecology progressive series. Volume 87: 283-293.

Rusch, D., *et al.* (2007) The *Sorcerer II* Global Sampling Expedition: Northwest Atlantic through Eastern Tropical Pacific. PLoS Biology. Volume 5, Issue 3: 398-431.

Suttle, C.A. (1994) The significance of viruses to mortality in aqatic microbial communities. Microbial Ecology. Volume 28: 237-243

Suttle, C.A. (2005) Viruses in the sea. Nature. Volume 437: 356-361

Suttle, C.A. (2007) Marine viruses-major players in the global ecosystem. Nature Reviews. Microbiology. Volume 5: 801-809

Suttle, C.A., and Chan, A.M. (1994) Dynamics and distribution of cyanophagues and their effect on marine Synechococcus spp.

Tomaru, Y., Shirai, Y., Suzuki, H., Nagumo, T. and Nagasaki, K. (2008). Isolation and characterization of a new single-stranded DNA virus infecting the cosmopolitan marine diatom *Chaetoceros debilis*. Aqatic Microbial Ecology. Volume 50: 103-112.

Tomaru, Y., Toyoda, K., Kimura, K., Hata, N., Yoshida, M. and Nagasaki, M. (2012). The ISME Journal. Volume 6: 1445-1448.

Yoon, H.S. et al. (2011). Single-Cell Genomics Reveals Organismal Interactions in uncultivated marine protists. Science. Volume 332: 714-717.

Weinbauer, M.G. and Rassoulzadegan, F. (2004) Environmental microbiology. Volume 6, No.1: 1-11

ANNEXES

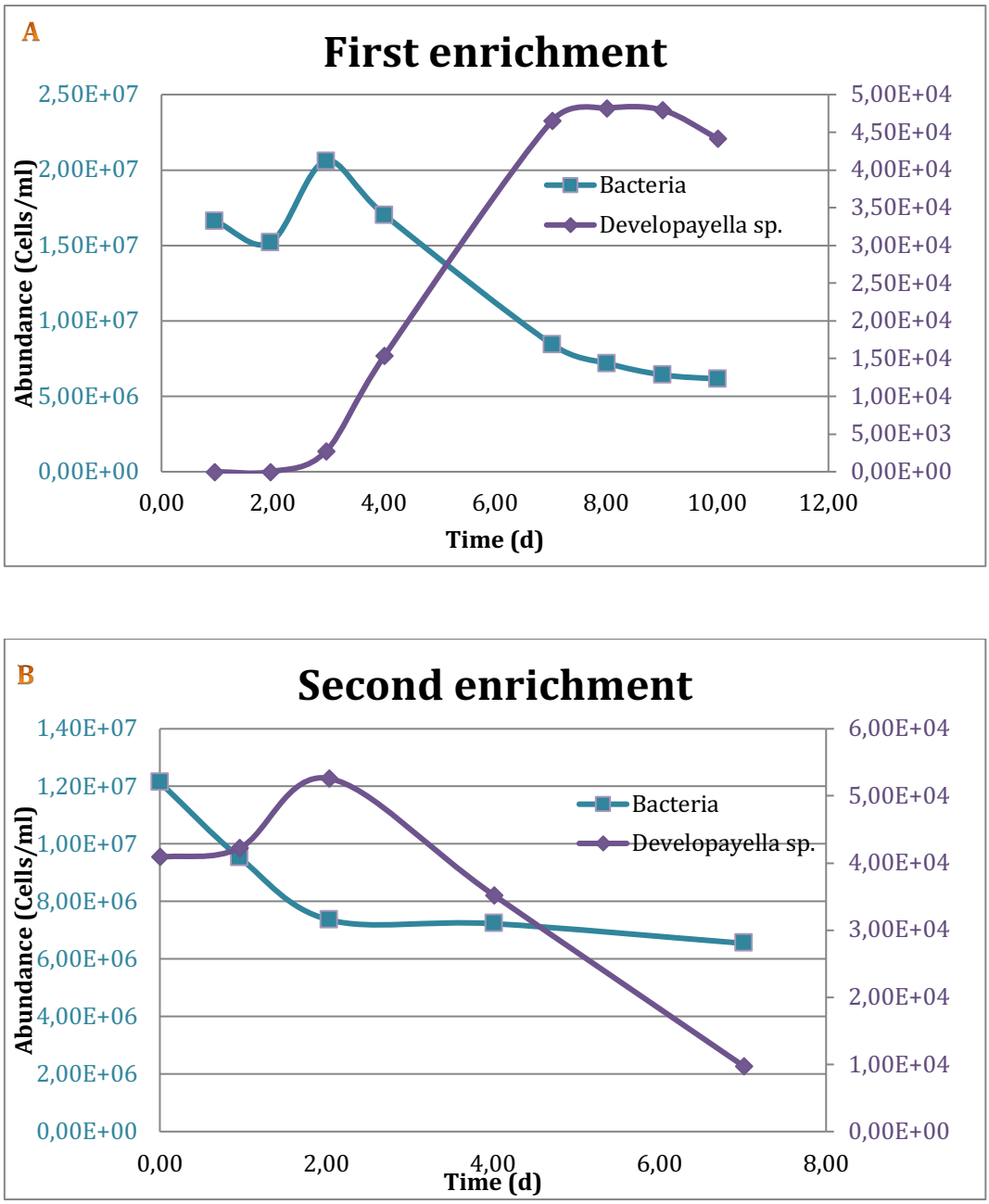
A

Culture Vol.	Sampling date	Sampling	Time (d)	Filter	vol filt. (ml)	ransect (mm	flag/transect	flag/ml	Fields	Grid	bact	bact/ml
150ml	15/10/12 11:30	T0	0,00	0,2	2	27	58	**4,13E+03**	20	10	916	8812937,78
150ml	16/10/12 10:30	T1	0,96	0,2	0,5	23	nd		20	100	4320	**1,66E+07**
150ml	17/10/12 10:30	T2	1,96	0,2	0,5	21	nd		10	100	1974	**1,52E+07**
150ml	18/10/12 10:30	T3	2,96	0,2	2	20	28	**2,69E+03**	20	10	2142	**2,06E+07**
150ml	19/10/12 11:50	T4	4,01	0,2	2	21	168	**1,54E+04**	20	10	1769	**1,70E+07**
150ml	22/10/12 12:15	T5	7,03	0,2	2	21	508	**4,66E+04**	20	10	876	**8,43E+06**
150ml	23/10/12 11:45	T6	8,01	0,2	2	22	551	**4,82E+04**	20	10	748	**7,20E+06**
150ml	24/10/12 11:55	T7	9,02	0,2	2	22	548	**4,79E+04**	20	10	667	**6,42E+06**
150ml	25/10/12 11:45	T8	10,01	0,2	2	21	482	**4,42E+04**	20	10	639	**6,15E+06**

B

Culture Vol.	Sampling date	Sampling	Time (d)	Filter	vol filt. (ml)	ransect (mm	flag/transect	flag/ml	Fields	Grid	bact	bact/ml
150ml	29/10/12 11:30	T0	0,00	0,2	2	21	446	**4,09E+04**	20	10	1262	**1,21E+07**
150ml	30/10/12 10:30	T1	0,96	0,2	2	21	461	**4,22E+04**	20	10	990	**9,52E+06**
150ml	31/10/12 12:15	T2	2,03	0,2	2	21	574	**5,26E+04**	20	10	765	**7,36E+06**
150ml	02/11/12 11:50	T3	4,01	0,2	2	21	384	**3,52E+04**	20	10	751	**7,23E+06**
150ml	05/11/12 11:50	T4	7,01	0,2	2	20	101	**9,72E+03**	20	10	680	**6,54E+06**

Annex 1: EFM tables from the first (A) and second (B) enrichment experiments



Annex 2: graphics of the flagellagets and bacteria abundances of the first (A) and second (B) enrichments.

A

Culture Vol.	Sampling date	Sampling	Time (d)	Filter	vol filt. (ml)	Transect (mm)	flag/transect	flag/ml	Fields	Grid	bact	bact/ml
150ml	31/10/12 17:00	T0	0,00	0,2	2	21	446	**4,09E+04**	20	10	1262	12141842,2
150ml	05/11/12 11:50	T1	4,78	0,2	2	20	616	**5,93E+04**	20	10	608	**5,85E+06**
150ml	06/11/12 12:26	T2	5,81	0,2	2	20	596	**5,73E+04**	20	10	443	**4,26E+06**
150ml	07/11/12 12:58	T3	6,83	0,2	2	20	473	**4,55E+04**	20	10	348	**3,35E+06**
150ml*	07/11/12 16:36	T4	6,98	0,2	2	22	524	**4,58E+04**	20	10	401	**3,86E+06**
150ml	08/11/12 11:35	T5	7,77	0,2	2	25	414	**3,19E+04**	20	10	559	**5,38E+06**
150ml	08/11/12 17:00	T6	8,00	0,2	2	21	251	**2,30E+04**	20	10	585	**5,63E+06**
150ml	09/11/12 11:20	T7	8,76	0,2	2	21	227	**2,08E+04**	20	10	572	**5,50E+06**
150ml	09/11/12 16:00	T8	8,96	0,2	2	21	176	**1,61E+04**	20	10	540	**5,20E+06**
150ml	12/11/12 11:15	T9	11,76	0,2	2	20	159	**1,53E+04**	20	10	584	**5,62E+06**
150ml	12/11/12 17:17	T10	12,01	0,2	2	21	145	**1,33E+04**	20	10	554	**5,33E+06**
150ml	13/11/12 10:00	T11	12,71	0,2	2	23	176	**1,47E+04**	20	10	605	**5,82E+06**
150ml	15/11/12 10:47	T12	14,74	0,2	2	26	154	**1,14E+04**	20	10	547	**5,26E+06**

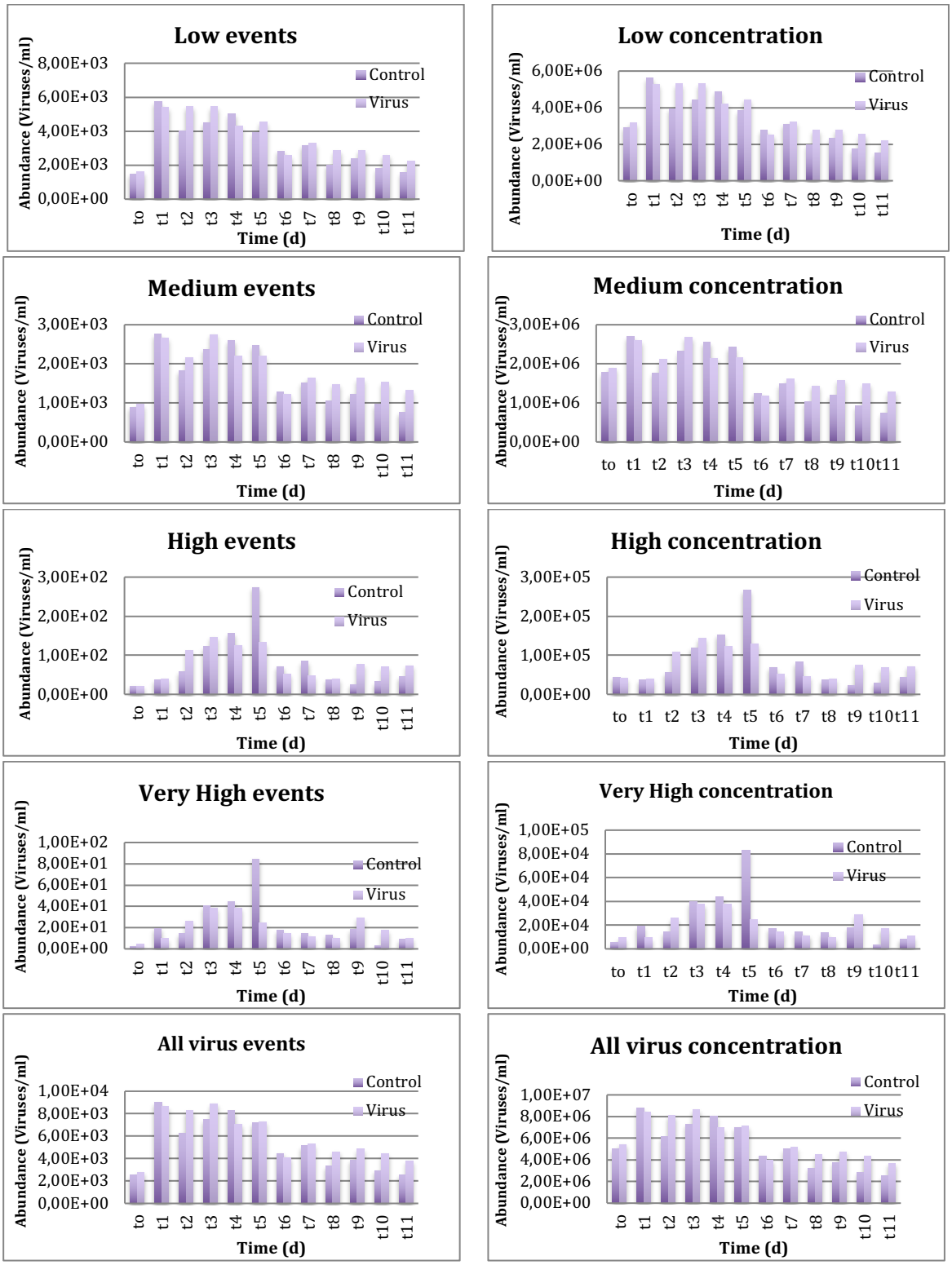
B

Culture Vol.	Sampling date	Sampling	Time (d)	Filter	vol filt. (ml)	Transect (mm)	flag/transect	flag/ml	Fields	Grid	bact	bact/ml
150ml	31/10/12 17:00	T0	0,00	0,2	2	21	446	**4,09E+04**	20	10	1262	12141842,2
150ml	05/11/12 11:50	T1	4,78	0,2	2	21	517	**4,74E+04**	20	10	680	**6,54E+06**
150ml	06/11/12 12:26	T2	5,81	0,2	2	21	607	**5,56E+04**	20	10	386	**3,71E+06**
150ml	07/11/12 12:58	T3	6,83	0,2	2	21	359	**3,29E+04**	20	10	419	**4,03E+06**
150ml	07/11/12 16:36	T4	6,98	0,2	2	20	424	**4,08E+04**	20	10	588	**5,66E+06**
150ml	08/11/12 11:35	T5	7,77	0,2	2	21	279	**2,56E+04**	20	10	705	**6,78E+06**
150ml	08/11/12 17:00	T6	8,00	0,2	2	20	183	**1,76E+04**	20	10	677	**6,51E+06**
150ml	09/11/12 11:20	T7	8,76	0,2	2	20	233	**2,24E+04**	20	10	788	**7,58E+06**
150ml	09/11/12 16:00	T8	8,96	0,2	2	20	170	**1,64E+04**	20	10	799	**7,69E+06**
150ml	12/11/12 11:15	T9	11,76	0,2	2	22	108	**9,45E+03**	20	10	635	**6,11E+06**
150ml	12/11/12 17:17	T10	12,01	0,2	2	22	131	**1,15E+04**	20	10	710	**6,83E+06**
150ml	13/11/12 10:00	T11	12,71	0,2	2	21	144	**1,32E+04**	20	10	714	**6,87E+06**
150ml	15/11/12 10:47	T12	14,74	0,2	2	21	88	**8,06E+03**	20	10	751	**7,23E+06**

Annex 3: EFM tables from the control (A) and experiment (B) cultures from the first experiment.

File Name	Tipo	Fecha muestra	Fecha	Speed	Speed µL · min-1	TIME O	TIME F	TIME (s)	Vol. (µL)	Factor Fijación	Vol(µl) mostra	Vol (µl) SG1	Dilution (1/100)	low Events	medium Events	high Events	very high Events	all virus Events	low Conc.	medium Conc.	high Conc.	very high Conc.	all virus Conc.		all virus - control
CONTROL TE.001	CONTROL TE	21/11/12	21-de nov-12	HI	55,9			60	55,90	1	500	5	1,00	325	194	8	1	545	5,87E+03	3,51E+03	1,45E+02	1,81E+01	9,85E+03		0,00E+00
1r enriq_100x.002	1r enriq_100x	30/10/12	21-de nov-12	HI	55,9			60	55,90	0,98039216	500	5	0,01	8874	4795	53	10	14504	1,64E+07	8,84E+06	9,77E+04	1,84E+04	2,67E+07		2,67E+07
2n enriq pre-inocul_100x.003	2n enriq pre-inocul_100x	30/10/12	21-de nov-12	HI	55,9			60	55,90	0,98039216	500	5	0,01	7886	3525	58	11	12182	1,45E+07	6,50E+06	1,07E+05	2,03E+04	2,25E+07		2,24E+07
2n enriq post-inocul_100x.004	2n enriq post-inocul_100x	30/10/12	21-de nov-12	HI	55,9			60	55,90	0,98039216	500	5	0,01	7377	4101	48	6	12048	1,36E+07	7,56E+06	8,85E+04	1,11E+04	2,22E+07		2,22E+07
1r enriq pre-inocul_100x.005	1r enriq pre-inocul_100x	19/10/12	21-de nov-12	HI	55,9			60	55,90	0,98039216	500	5	0,01	10825	6622	107	16	18396	1,99E+07	1,22E+07	1,97E+05	2,95E+04	3,39E+07		3,39E+07
1r enriq 24nov_100x.006	1r enriq 24nov_100x	24/10/12	21-de nov-12	HI	55,9			60	55,90	0,98039216	500	5	0,01	4157	3291	131	15	7865	7,66E+06	6,07E+06	2,41E+05	2,76E+04	1,45E+07		1,45E+07
1r enriq 25nov_100x.007	1r enriq 25nov_100x	25/10/12	21-de nov-12	HI	55,9			60	55,90	0,98039216	500	5	0,01	3623	2818	146	17	6796	6,68E+06	5,19E+06	2,69E+05	3,13E+04	1,25E+07		1,25E+07
1r enriq 25nov_100x.008 (AIGUA)	1r enriq 25nov_100x	21/11/12	21-de nov-12	HI	55,9			60	55,90	0,98039216	500	5	0,01	4	0	0	0	5	7,37E+03	0,00E+00	0,00E+00	0,00E+00	9,21E+03		-6,32E+02
eliana.009	blanc te	21/11/12	22-de nov-12	HI	52,75			60	52,75	1	500	5	1,00	185	67	3	0	276	3,54E+03	1,28E+03	5,74E+01	0,00E+00	5,28E+03		0,00E+00
to control_100x.010	to control_100x	06/11/12	22-de nov-12	HI	52,75			60	52,75	0,98039216	500	5	0,010	1512	910	23	3	2585	2,95E+06	1,78E+06	4,49E+04	5,86E+03	5,05E+06		5,04E+06
to virus_100x.011	to virus_100x	06/11/12	22-de nov-12	HI	52,75			60	52,75	0,98039216	500	5	0,010	1644	975	22	5	2805	3,21E+06	1,90E+06	4,30E+04	9,76E+03	5,48E+06		5,47E+06
t3 control_50x.017	t3 control_50x	07/11/12	22-de nov-12	HI	52,75			60	52,75	0,98039216	500	5	0,02	4545	2388	124	41	7512	4,44E+06	2,33E+06	1,21E+05	4,00E+04	7,34E+06		7,33E+06
t3 virus_50x.018	t3 virus_50x	07/11/12	22-de nov-12	HI	52,75			60	52,75	0,98039216	500	5	0,02	5470	2762	148	39	8922	5,34E+06	2,70E+06	1,45E+05	3,81E+04	8,71E+06		8,71E+06
t1control_50x.019	t1control_50x	06/11/12	22-de nov-12	HI	52,75			60	52,75	0,98039216	500	5	0,02	5768	2768	39	20	9065	5,63E+06	2,70E+06	3,81E+04	1,95E+04	8,85E+06		8,85E+06
t1virus_50x.020	t1virus_50x	06/11/12	22-de nov-12	HI	52,75			60	52,75	0,98039216	500	5	0,02	5435	2673	42	10	8671	5,31E+06	2,61E+06	4,10E+04	9,76E+03	8,47E+06		8,46E+06
t2 control_50x.021	t2 control_50x	07/11/12	22-de nov-12	HI	52,75			60	52,75	0,98039216	500	5	0,02	4058	1828	59	15	6321	3,96E+06	1,79E+06	5,76E+04	1,46E+04	6,17E+06		6,17E+06
t2 virus_50x.022	t2 vius_50x	07/11/12	22-de nov-12	HI	52,75			60	52,75	0,98039216	500	5	0,02	5460	2175	113	27	8322	5,33E+06	2,12E+06	1,10E+05	2,64E+04	8,13E+06		8,12E+06
t4 control_50x.023	t4 control_50x	08/11/12	22-de nov-12	HI	52,75			60	52,75	0,98039216	500	5	0,02	5033	2617	158	45	8298	4,91E+06	2,56E+06	1,54E+05	4,39E+04	8,10E+06		8,10E+06
t4 virus_50x.024	t4 virus_50x	08/11/12	22-de nov-12	HI	52,75			60	52,75	0,98039216	500	5	0,02	4323	2210	127	39	7117	4,22E+06	2,16E+06	1,24E+05	3,81E+04	6,95E+06		6,94E+06
t5 control_50x.025	t5 control_50x	08/11/12	22-de nov-12	HI	52,75			60	52,75	0,98039216	500	5	0,02	3952	2478	275	85	7199	3,86E+06	2,42E+06	2,69E+05	8,30E+04	7,03E+06		7,02E+06
t5 virus_50x.026	t5 virus_50x	08/11/12	22-de nov-12	HI	52,75			60	52,75	0,98039216	500	5	0,02	4575	2212	134	25	7332	4,47E+06	2,16E+06	1,31E+05	2,44E+04	7,16E+06		7,15E+06
t6 control_50x.027	t6 control_50x	09/11/12	22-de nov-12	HI	52,75			60	52,75	0,98039216	500	5	0,02	2871	1291	73	18	4518	2,80E+06	1,26E+06	7,13E+04	1,76E+04	4,41E+06		4,41E+06
t6 virus_50x.028	t6 virus_50x	09/11/12	22-de nov-12	HI	52,75			60	52,75	0,98039216	500	5	0,02	2599	1232	54	15	4138	2,54E+06	1,20E+06	5,27E+04	1,46E+04	4,04E+06		4,04E+06
t7 control_50x.029	t7 control_50x	09/11/12	22-de nov-12	HI	52,75			60	52,75	0,98039216	500	5	0,02	3192	1534	86	15	5180	3,12E+06	1,50E+06	8,40E+04	1,46E+04	5,06E+06		5,05E+06
t7 virus_50x.030	t7 virus_50x	09/11/12	22-de nov-12	HI	52,75			60	52,75	0,98039216	500	5	0,02	3319	1660	49	12	5330	3,24E+06	1,62E+06	4,78E+04	1,17E+04	5,20E+06		5,20E+06
t8 control_50x.031	t8 control_50x	12/11/12	22-de nov-12	HI	52,75			60	52,75	0,98039216	500	5	0,02	2061	1066	40	14	3388	2,01E+06	1,04E+06	3,91E+04	1,37E+04	3,31E+06		3,30E+06
t8 virus_50x.032	t8 virus_50x	12/11/12	22-de nov-12	HI	52,75			60	52,75	0,98039216	500	5	0,02	2879	1486	42	10	4656	2,81E+06	1,45E+06	4,10E+04	9,76E+03	4,55E+06		4,54E+06
t9 control_50x.033	t9 control_50x	12/11/12	22-de nov-12	HI	52,75			60	52,75	0,98039216	500	5	0,02	2401	1245	26	19	3927	2,34E+06	1,22E+06	2,54E+04	1,86E+04	3,83E+06		3,83E+06
t9 virus_50x.034	t9 virus_50x	12/11/12	22-de nov-12	HI	52,75			60	52,75	0,98039216	500	5	0,02	2891	1628	78	30	4916	2,82E+06	1,59E+06	7,62E+04	2,93E+04	4,80E+06		4,80E+06
t10 control_50x.035	t11 control_50x	15/11/12	22-de nov-12	HI	52,75			60	52,75	0,98039216	500	5	0,02	1596	772	47	9	2571	1,56E+06	7,54E+05	4,59E+04	8,79E+03	2,51E+06		2,51E+06
t10 virus_50x.036	t11 virus_50x	15/11/12	22-de nov-12	HI	52,75			60	52,75	0,98039216	500	5	0,02	2263	1332	75	11	3845	2,21E+06	1,30E+06	7,32E+04	1,07E+04	3,75E+06		3,75E+06
t11 control_50x.037	t10 control_50x	13/11/12	22-de nov-12	HI	52,75			60	52,75	0,98039216	500	5	0,02	1827	966	32	4	3002	1,78E+06	9,43E+05	3,12E+04	3,91E+03	2,93E+06		2,93E+06
t11 virus_50x.038	t10 virus_50x	13/11/12	22-de nov-12	HI	52,75			60	52,75	0,98039216	500	5	0,02	2630	1531	71	18	4471	2,57E+06	1,50E+06	6,93E+04	1,76E+04	4,37E+06		4,36E+06

Annex 4: First experiment FCM table.



Annex 5: First experiment FCM graphics.

A

Culture Vol.	Sampling date	Sampling	Time (d)	Filter	vol filt. (ml)	Transect (mm)	flag/transect	flag/ml	Fields	Grid	bact	bact/ml
200ml	14/01/13 12:20	T0	0,00	0,2	2	20	473	**4,55E+04**	20	10	1143	**1,10E+07**
100ml	14/01/13 17:07	T1	0,20	0,2	2	20	459	**4,42E+04**	20	10	964	**9,27E+06**
100ml	15/01/13 10:25	T2	0,92	0,2	2	20	501	**4,82E+04**	20	10	1115	**1,07E+07**
100ml	15/01/13 16:40	T3	1,18	0,2	2	22	514	**4,50E+04**	20	10	910	**8,76E+06**
100ml	16/01/13 10:30	T4	1,92	0,2	2	25	473	**3,64E+04**	20	10	1001	**9,63E+06**
100ml	16/01/13 16:10	T5	2,16	0,2	2	21	477	**4,37E+04**	20	10	954	**9,18E+06**
100ml	17/01/13 11:15	T6	2,95	0,2	2	21	584	**5,35E+04**	20	10	894	**8,60E+06**
100ml	17/01/13 17:35	T7	3,22	0,2	2	21	542	**4,97E+04**	20	10	926	**8,91E+06**
100ml	18/01/13 11:00	T8	3,94	0,2	2	20	341	**3,28E+04**	20	10	887	**8,53E+06**
100ml	18/01/13 16:15	T9	4,16	0,2	2	21	422	**3,87E+04**	20	10	909	**8,75E+06**
100ml	21/01/13 11:50	T10	6,98	0,2	2	21	211	**1,93E+04**	20	10	1030	**9,91E+06**
100ml	21/01/13 17:35	T11	7,22	0,2	2	21	186	**1,70E+04**	20	10	1044	**1,00E+07**

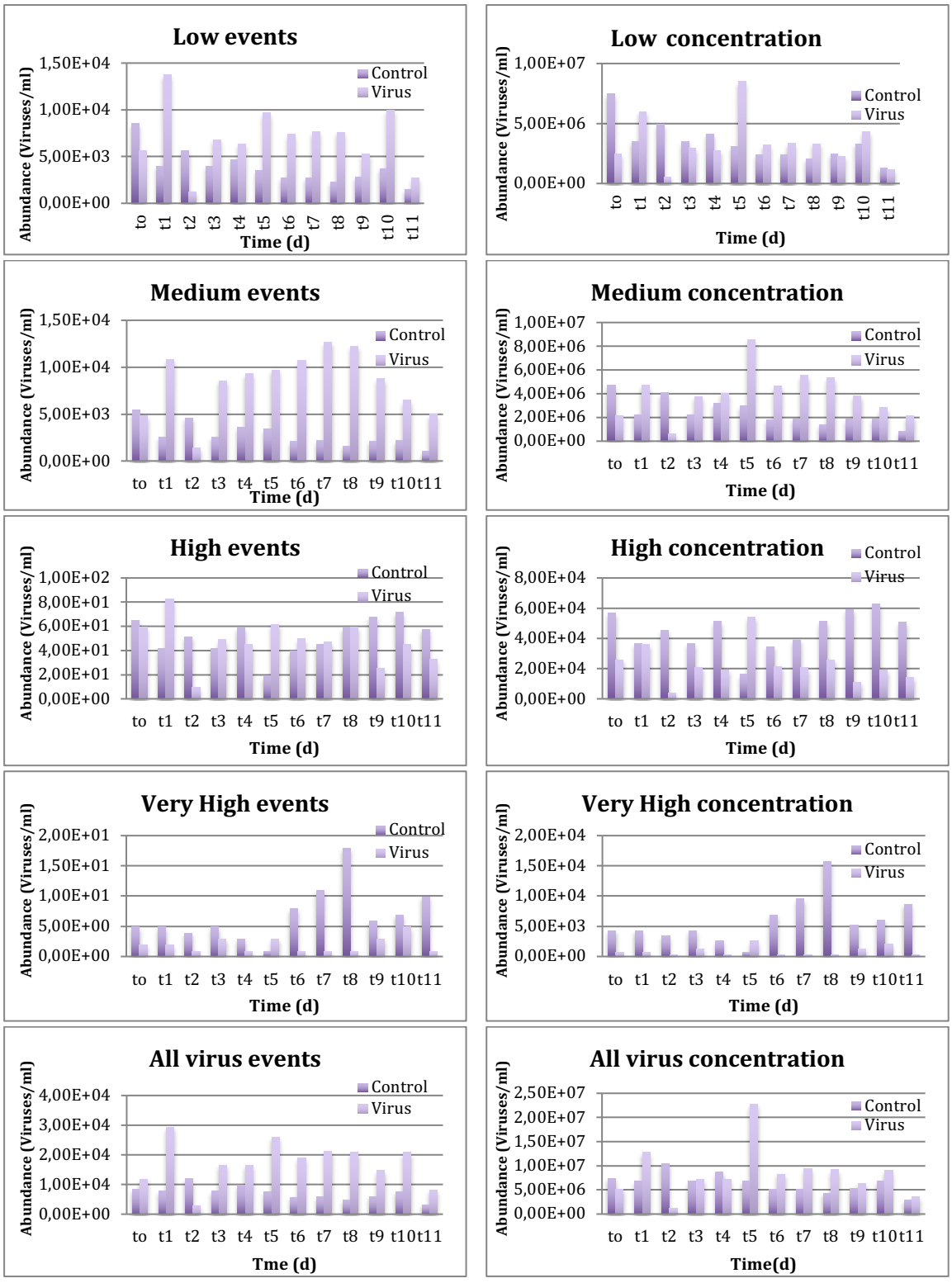
B

Culture Vol.	Sampling date	Sampling	Time (d)	Filter	vol filt. (ml)	Transect (mm)	flag/transect	flag/ml	Fields	Grid	bact	bact/ml
200ml	14/01/13 12:20	T0	0,00	0,2	2	20	473	**4,55E+04**	20	10	1143	**1,10E+07**
100ml	14/01/13 17:07	T1	0,20	0,2	2	20	410	**3,95E+04**	20	10	887	**8,53E+06**
100ml	15/01/13 10:25	T2	0,92	0,2	2	20	455	**4,38E+04**	20	10	1874	**1,80E+07**
100ml	15/01/13 16:40	T3	1,18	0,2	2	22	400	**3,50E+04**	20	10	1885	**1,81E+07**
100ml	16/01/13 10:30	T4	1,92	0,2	2	25	345	**2,66E+04**	20	10	1896	**1,82E+07**
100ml	16/01/13 16:10	T5	2,16	0,2	2	21	204	**1,87E+04**	20	10	1998	**1,92E+07**
100ml	17/01/13 11:15	T6	2,95	0,2	2	21	301	**2,76E+04**	20	10	2153	**2,07E+07**
100ml	17/01/13 17:35	T7	3,22	0,2	2	21	253	**2,32E+04**	20	10	1972	**1,90E+07**
100ml	18/01/13 11:00	T8	3,94	0,2	2	20	177	**1,70E+04**	20	10	1998	**1,92E+07**
100ml	18/01/13 16:15	T9	4,16	0,2	2	21	221	**2,03E+04**	20	10	1880	**1,81E+07**
100ml	21/01/13 11:50	T10	6,98	0,2	2	22	101	**8,83E+03**	20	10	2033	**1,96E+07**
100ml	21/01/13 17:35	T11	7,22	0,2	2	23	85	**7,11E+03**	20	10	1988	**1,91E+07**

Annex 6: EFM tables from the control (A) and experiment (B) cultures from the second experiment (red marks mean a value that was obtained from the average of the previous and subsequent values because the sample was lost).

File Name	Tipo	Fecha muestra	Fecha	Speed	Speed µL·min-1	TIME 0	TIME F	TIME (s)	Vol. (µL)	Factor Fijación	Vol(µl) mostra	Vol (µl) SG1	Dilution (1/100)	low Events	medium Events	high Events	very high Events	all virus Events	low Conc.	medium Conc.	high Conc.	very high Conc.	all virus Conc.	all virus - control
2n_exp.004	CONTROL TE	30/01/13	30/01/13	HI	58,80			60	58,80	1	500	5	1,00	334	149	2	0	660	5,74E+03	2,56E+03	3,44E+01	0,00E+00	1,13E+04	0,00E+00
2n_exp.005	concentrat virus ex_100x	09/01/13	30/01/13	HI	58,80			60	58,80	0,98039216	500	5	0,01	2682	2276	16	0	5876	4,70E+06	3,99E+06	2,80E+04	0,00E+00	1,03E+07	1,03E+07
exp2_100x.006	concentrat virus_100x	11/01/13	30/01/13	HI	58,80			60	58,80	0,98039216	500	5	0,01	21281	14968	69	14	44325	3,73E+07	2,62E+07	1,21E+05	2,45E+04	7,77E+07	7,76E+07
exp3_100x.007	t5 exp	16/01/13	30/01/13	HI	58,80			60	58,80	0,98039216	500	5	0,01	9728	9728	62	3	26097	1,70E+07	1,70E+07	1,09E+05	5,26E+03	4,57E+07	4,57E+07
exp5_100x.008	t1 control	14/01/13	30/01/13	HI	58,80			60	58,80	0,98039216	500	5	0,01	8566	5409	65	5	8566	1,50E+07	9,48E+06	1,14E+05	8,76E+03	1,50E+07	1,50E+07
exp6_100x.009	t11 control	21/01/13	30/01/13	HI	58,80			60	58,80	0,98039216	500	5	0,01	1569	1007	58	10	3481	2,75E+06	1,76E+06	1,02E+05	1,75E+04	6,10E+06	6,09E+06
exp8_100x.010	t3 control	15/01/13	30/01/13	HI	58,80			60	58,80	0,98039216	500	5	0,01	4041	2617	42	5	7997	7,08E+06	4,59E+06	7,36E+04	8,76E+03	1,40E+07	1,40E+07
control te_1x.011	CONTROL TE	30/01/13	30/01/13	HI	58,80			60	58,80	1	500	5	1,00	174	99	4	0	373	2,99E+03	1,70E+03	6,87E+01	0,00E+00	6,41E+03	0,00E+00
exp9_100x.012	t5 control	16/11/13	30/01/13	HI	58,80			60	58,80	0,98039216	500	5	0,01	3568	3435	19	1	7828	6,25E+06	6,02E+06	3,33E+04	1,75E+03	1,37E+07	1,37E+07
exp10_100x.013	t7 control	17/11/13	30/01/13	HI	58,80			60	58,80	0,98039216	500	5	0,01	2819	2214	45	11	6052	4,94E+06	3,88E+06	7,88E+04	1,93E+04	1,06E+07	1,06E+07
exp11_100x.014	t9 control	18/11/13	30/01/13	HI	58,80			60	58,80	0,98039216	500	5	0,01	2892	2168	68	6	6228	5,07E+06	3,80E+06	1,19E+05	1,05E+04	1,09E+07	1,09E+07
exp12_100x.015	t6 control	17/11/13	30/01/13	HI	58,80			60	58,80	0,98039216	500	5	0,01	2781	2156	40	8	5937	4,87E+06	3,78E+06	7,01E+04	1,40E+04	1,04E+07	1,04E+07
exp2_200x.016	concentrat virus	11/01/13	30/01/13	HI	58,80			60	58,80	0,98039216	500	5	0,02	8061	5964	23	0	17362	7,06E+06	5,22E+06	2,01E+04	0,00E+00	1,52E+07	1,52E+07
exp4_200x.017	t3 exp	15/01/13	30/01/13	HI	58,80			60	58,80	0,98039216	500	5	0,02	6849	8609	49	3	16719	6,00E+06	7,54E+06	4,29E+04	2,63E+03	1,46E+07	1,46E+07
exp7_200x.018	t7 exp	17/01/13	30/01/13	HI	58,80			60	58,80	0,98039216	500	5	0,02	7728	12691	47	1	21505	6,77E+06	1,11E+07	4,12E+04	8,76E+02	1,88E+07	1,88E+07
exp13_200x.019	t10 exp	21/01/13	30/01/13	HI	58,80			60	58,80	0,98039216	500	5	0,02	9985	6472	45	5	20833	8,75E+06	5,67E+06	3,94E+04	4,38E+03	1,83E+07	1,82E+07
exp14_200x.020	t6 exp	17/01/13	30/01/13	HI	58,80			60	58,80	0,98039216	500	5	0,02	7426	10718	50	1	19136	6,51E+06	9,39E+06	4,38E+04	8,76E+02	1,68E+07	1,68E+07
exp15_200x.021	t8 exp	18/01/13	30/01/13	HI	58,80			60	58,80	0,98039216	500	5	0,02	7651	12240	60	1	20923	6,70E+06	1,07E+07	5,26E+04	8,76E+02	1,83E+07	1,83E+07
exp16_200x.022	t10 control	21/01/13	30/01/13	HI	58,80			60	58,80	0,98039216	500	5	0,01	3802	2201	72	7	7915	6,66E+06	3,86E+06	1,26E+05	1,23E+04	1,39E+07	1,39E+07
exp17_100x.023	t8 control	18/01/13	30/01/13	HI	58,80			60	58,80	0,98039216	500	5	0,01	2392	1664	59	18	5065	4,19E+06	2,92E+06	1,03E+05	3,15E+04	8,87E+06	8,87E+06
exp18_100x.024	t4 control	16/01/13	30/01/13	HI	58,80			60	58,80	0,98039216	500	5	0,01	4768	3679	59	3	10032	8,35E+06	6,45E+06	1,03E+05	5,26E+03	1,76E+07	1,76E+07
exp19_100x.025	t2 control	15/01/13	30/01/13	HI	58,80			60	58,80	0,98039216	500	5	0,01	5683	4666	52	4	12023	9,96E+06	8,18E+06	9,11E+04	7,01E+03	2,11E+07	2,11E+07
exp20_100x.026	t9 exp	18/01/13	30/01/13	HI	58,80			60	58,80	0,98039216	500	5	0,02	5331	8818	26	3	14728	4,67E+06	7,72E+06	2,28E+04	2,63E+03	1,29E+07	1,29E+07
exp21_200x.027	t11 exp	21/01/13	30/01/13	HI	58,80			60	58,80	0,98039216	500	5	0,02	2757	4971	33	1	8464	2,42E+06	4,35E+06	2,89E+04	8,76E+02	7,41E+06	7,41E+06
exp22_200x.028	t1 exp	14/01/13	30/01/13	HI	58,80			60	58,80	0,98039216	500	5	0,02	5715	4908	60	2	12067	5,01E+06	4,30E+06	5,26E+04	1,75E+03	1,06E+07	1,06E+07
exp23_200x.029	t4 exp	16/01/13	30/01/13	HI	58,80			60	58,80	0,98039216	500	5	0,02	6380	9355	45	1	16620	5,59E+06	8,20E+06	3,94E+04	8,76E+02	1,46E+07	1,46E+07
exp24_200x.030	t2 exp	15/01/13	30/01/13	HI	58,80			60	58,80	0,98039216	500	5	0,02	1342	1420	10	1	3132	1,18E+06	1,24E+06	8,76E+03	8,76E+02	2,74E+06	2,74E+06

Annex 7: Second experiment FCM table.



Annex 8: Second experiment FCM graphics.

A

Culture Vol.	Sampling date	Sampling	Time (d)	Filter	vol filt. (ml)	Transect (mm)	flag/transect	flag/ml	Fields	Grid	bact	bact/ml
200ml	06/03/13 11:30	T0	0,00	0,2	2	21	106	**9,71E+03**	20	10	1336	**1,29E+07**
100ml	07/03/13 12:20	T1	1,03	0,2	2	21	176	**1,61E+04**	20	10	1157	**1,11E+07**
100ml	08/03/13 9:30	T2	1,92	0,2	2	20	212	**2,04E+04**	20	10	1169	**1,12E+07**
100ml	11/03/13 10:10	T3	4,94	0,2	2	20	212	**2,04E+04**	20	10	956	**9,20E+06**
100ml	12/03/13 10:20	T4	5,95	0,2	2	21	154	**1,41E+04**	20	10	951	**9,15E+06**
100ml	13/03/13 10:10	T5	6,94	0,2	2	23	153	**1,28E+04**	20	10	841	**8,09E+06**
100ml	14/03/13 17:00	T6	8,23	0,2	2	28	131	**9,00E+03**	20	10	780	**7,50E+06**
100ml	15/03/13 10:56	T7	8,98	0,2	2	20	99	**9,53E+03**	20	10	838	**8,06E+06**
100ml	18/03/13 11:00	T8	11,98	0,2	2	20	97	**9,33E+03**	20	10	861	**8,28E+06**

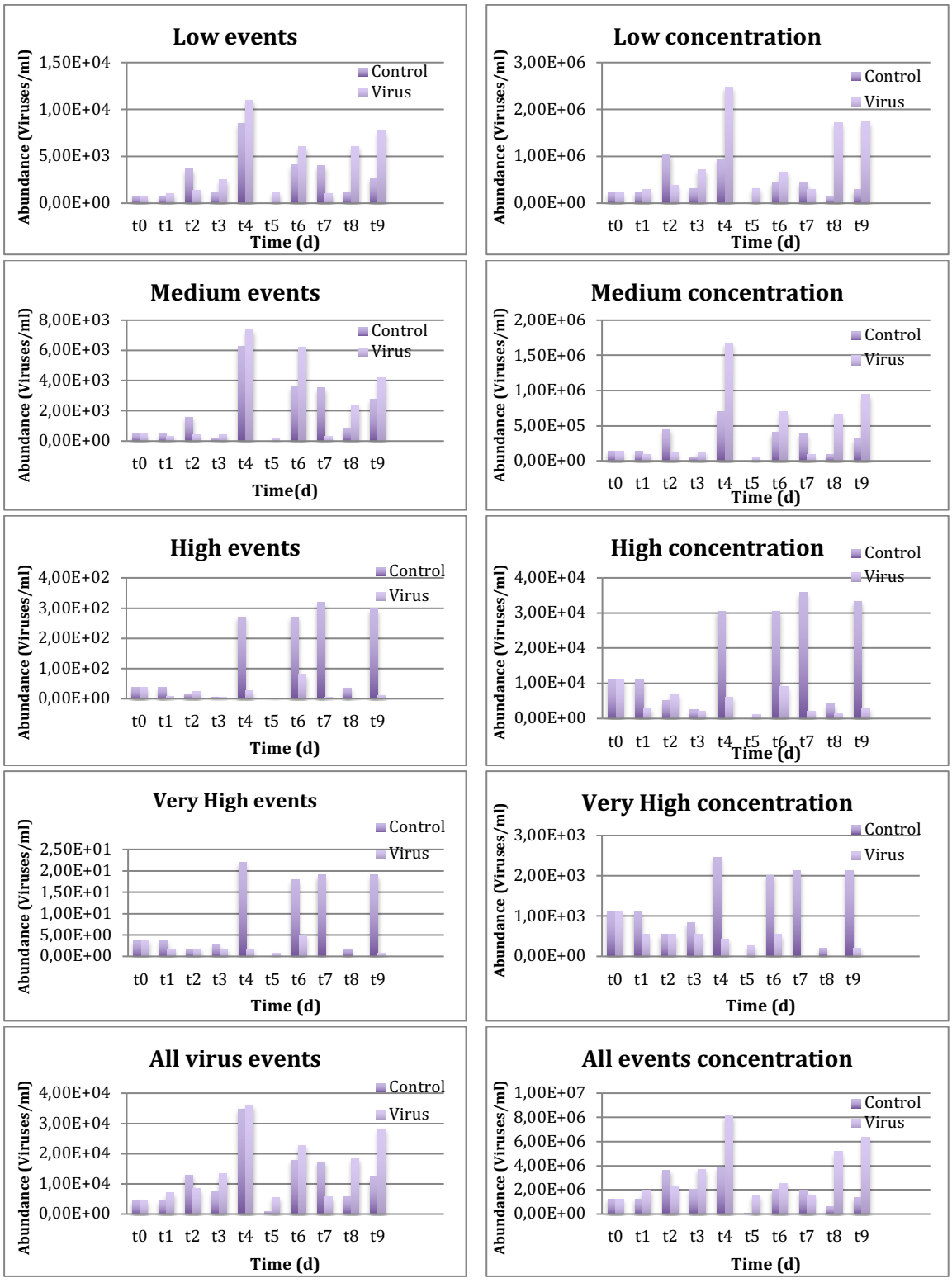
B

Culture Vol.	Sampling date	Sampling	Time (d)	Filter	vol filt. (ml)	Transect (mm)	flag/transect	flag/ml	Fields	Grid	bact	bact/ml
200ml	06/03/13 11:30	T0	0,00	0,2	2	21	106	**9,71E+03**	20	10	1336	**1,29E+07**
100ml	07/03/13 12:20	T1	1,03	0,2	2	21	142	**1,30E+04**	20	10	1429	**1,37E+07**
100ml	08/03/13 9:30	T2	1,92	0,2	2	20	179	**1,72E+04**	20	10	1485	**1,43E+07**
100ml	11/03/13 10:10	T3	4,94	0,2	2	20	133	**1,28E+04**	20	10	1564	**1,50E+07**
100ml	12/03/13 10:10	T4	5,94	0,2	2	20	89	**8,56E+03**	20	10	1502	**1,45E+07**
100ml	13/03/13 10:10	T5	6,94	0,2	2	20	73	**7,02E+03**	20	10	1491	**1,43E+07**
100ml	14/03/13 17:00	T6	8,23	0,2	2	20	66	**6,35E+03**	20	10	1243	**1,20E+07**
100ml	15/03/13 10:56	T7	8,98	0,2	2	21	60	**5,50E+03**	20	10	1436	**1,38E+07**
100ml	18/03/13 11:00	T8	11,98	0,2	2	20	78	**7,51E+03**	20	10	1561	**1,50E+07**

Annex 9: EFM tables from the control (A) and experiment (B) cultures from the third experiment.

File Name	Tipo	Fecha muestra	Fecha	Speed	Speed µL·min-1	TIME 0	TIME F	TIME (s)	Vol. (µL)	Factor Fijación	Vol(µl) mostra	Vol (µl) SG1	Dilution (1/100)	low Events	medium Events	high Events	very high Events	all virus Events	low Conc.	medium Conc.	high Conc.	very high Conc.	all virus Conc.		all virus - control
Control TE.001	CONTROL TE	04/04/13	04/04/13	HI	183,30			60	183,30	1	500	5	1,000	331	86	2	2	1818	1,82E+03	4,74E+02	1,10E+01	1,10E+01	1,00E+04		0,00E+00
e1_50x.002	t2 control 50x	07/03/13	04/04/13	HI	183,30			60	183,30	0,98039216	500	5	0,02	3735	1627	19	2	13079	1,05E+06	4,57E+05	5,34E+03	5,62E+02	3,68E+06		3,67E+06
e2_50x.004	t8 exp 50x	15/03/13	04/04/13	HI	183,30			60	183,30	0,98039216	500	5	0,02	6140	2380	6	0	18380	1,73E+06	6,69E+05	1,69E+03	0,00E+00	5,17E+06		5,16E+06
e3_20x.013	t8 control 20x	15/03/13	04/04/13	HI	183,30			60	183,30	0,98039216	500	5	0,05	1327	909	38	2	5984	1,49E+05	1,02E+05	4,27E+03	2,25E+02	6,73E+05		6,63E+05
e4_50x.006	t3 control 50x	08/03/13	04/04/13	HI	183,30			60	183,30	0,98039216	500	5	0,02	1191	235	9	3	7411	3,35E+05	6,60E+04	2,53E+03	8,43E+02	2,08E+06		2,07E+06
e5_50x.007	concentrat virus 50 en 1 50x	06/03/13	04/04/13	HI	183,30			60	183,30	0,98039216	500	5	0,02	7619	1192	1	0	40137	2,14E+06	3,35E+05	2,81E+02	0,00E+00	1,13E+07		1,13E+07
e6_50x.008	cultiu 200 virus-free 50x	04/03/13	04/04/13	HI	183,30			60	183,30	0,98039216	500	5	0,02	2176	652	8	5	9432	6,11E+05	1,83E+05	2,25E+03	1,41E+03	2,65E+06		2,64E+06
e7_50x.009	t3 exp 50x	08/03/13	04/04/13	HI	183,30			60	183,30	0,98039216	500	5	0,02	2568	495	8	2	13288	7,22E+05	1,39E+05	2,25E+03	5,62E+02	3,73E+06		3,72E+06
e8_50x.010	t2 exp 50x	07/03/13	04/04/13	HI	183,30			60	183,30	0,98039216	500	5	0,02	1438	465	26	2	8458	4,04E+05	1,31E+05	7,31E+03	5,62E+02	2,38E+06		2,37E+06
e9_50x.011	t1 post-infecció 50x	06/03/13	04/04/13	HI	183,30			60	183,30	0,98039216	500	5	0,02	1091	370	11	2	7229	3,07E+05	1,04E+05	3,09E+03	5,62E+02	2,03E+06		2,02E+06
e10_50x.012	t0 pre-infecció 50x	06/03/13	04/04/13	HI	183,30			60	183,30	0,98039216	500	5	0,02	873	511	40	4	4680	2,45E+05	1,44E+05	1,12E+04	1,12E+03	1,32E+06		1,31E+06
e15_50x.018	t7 exp 50x	14/03/13	04/04/13	HI	183,30			60	183,30	0,98039216	500	5	0,02	1124	343	8	0	5901	3,16E+05	9,64E+04	2,25E+03	0,00E+00	1,66E+06		1,65E+06
e16_50x.019	t5 exp 50x	12/03/13	04/04/13	HI	183,30			60	183,30	0,98039216	500	5	0,02	1175	219	5	1	5620	3,30E+05	6,15E+04	1,41E+03	2,81E+02	1,58E+06		1,57E+06
control te2.020	CONTROL TE	04/04/13	04/04/13	HI	183,30			60	183,30	1	500	5	1,000	371	162	3	1	2242	2,04E+03	8,93E+02	1,65E+01	5,51E+00	1,24E+04		0,00E+00
e12_20x.022	t6 control 20x	13/03/13	04/04/13	HI	183,30			60	183,30	0,980392157	500	5	0,05	4137	3632	273	18	18084	4,65E+05	4,08E+05	3,07E+04	2,02E+03	2,03E+06		2,02E+06
e13_20x.023	t9 control 20x	18/03/13	04/04/13	HI	183,30			60	183,30	0,98039216	500	5	0,05	2734	2836	298	19	12425	3,07E+05	3,19E+05	3,35E+04	2,14E+03	1,40E+06		1,38E+06
e18_20x.024	t5 control 20x	12/03/13	04/04/13	HI	183,30			60	183,30	0,98039216	500	5	0,05	40	18	0	0	1016	4,50E+03	2,02E+03	0,00E+00	0,00E+00	1,14E+05		1,02E+05
e20_20x.025	t4 control 20x	11/03/13	04/04/13	HI	183,30			60	183,30	0,98039216	500	5	0,05	8521	6305	273	22	34908	9,58E+05	7,09E+05	3,07E+04	2,47E+03	3,92E+06		3,91E+06
e19_20x.026	t6 exp 20x	13/03/13	04/04/13	HI	183,30			60	183,30	0,98039216	500	5	0,05	6100	6265	83	5	22828	6,86E+05	7,04E+05	9,33E+03	5,62E+02	2,57E+06		2,55E+06
e16_20x.027	t5 exp 20x	12/03/13	04/04/13	HI	183,30			60	183,30	0,98039216	500	5	0,05	1392	621	16	0	7932	1,56E+05	6,98E+04	1,80E+03	0,00E+00	8,92E+05		8,79E+05
e17_20x.028	t7 control 20x	14/03/13	04/04/13	HI	183,30			60	183,30	0,98039216	500	5	0,05	4085	3566	319	19	17484	4,59E+05	4,01E+05	3,59E+04	2,14E+03	1,97E+06		1,95E+06
e11_40x.031	t4 exp 40x	11/03/13	04/04/13	HI	183,30			60	183,30	0,98039216	500	5	0,025	11024	7444	28	2	36176	2,48E+06	1,67E+06	6,29E+03	4,50E+02	8,13E+06		8,12E+06
e14_40x.032	t9 exp 40x	18/03/13	04/04/13	HI	183,30			60	183,30	0,98039216	500	5	0,025	7743	4205	14	1	28405	1,74E+06	9,45E+05	3,15E+03	2,25E+02	6,39E+06		6,37E+06

Annex 10: Third experiment FCM table.



Annex 11: Third experiment FCM graphics.

Sample	Localitation	num Reads	Filter size	Seq. effort	Relative virus importance
GS006	Bay of Fundy, Nova Scotia	195	0.1-0.8	59679	3,27E-03
GS002	Gulf of Maine	177	0.1-0.8	121590	1,46E-03
GS003	Browns Bank, Gulf of Maine	152	0.1-0.8	61605	2,47E-03
GS007	Northern Gulf of Maine	142	0.1-0.8	50980	2,79E-03
GS013	Off Nags Head, NC	115	0.1-0.8	138033	8,33E-04
GS010	Cape May, NJ	112	0.1-0.8	78304	1,43E-03
GS004	Outside Halifax, Nova Scotia	25	0.1-0.8	52959	4,72E-04
GS008	Newport Harbor, RI	21	0.1-0.8	129655	1,62E-04
GS009	Block Island, NY	19	0.1-0.8	79303	2,40E-04
GS123	International water between Madagascar and South Africa	15	0.1-0.8	107966	1,39E-04
GS122a	International waters between Madagascar and South Africa	14	0.1-0.8	112278	1,25E-04
GS014	South of Charleston, SC	9	0.1-0.8	128885	6,98E-05
GS012	Chesapeake Bay, MD	8	0.1-0.8	126162	6,34E-05
GS031	Upwelling, Fernandina Island	8	0.1-0.8	436401	1,83E-05
GS035	Wolf Island	5	0.1-0.8	140814	3,55E-05
GS122b	International waters between Madagascar and South Africa	5	0.8 - 3.0	50095	9,98E-05
GS000c	Sargasso Stations 3	4	0.22-0.8	368835	1,08E-05
GS116	Outside Seychelles, Indian Ocean	3	0.1 - 0.8	60932	4,92E-05
GS026	134 miles NE of Galapagos	2	0.1-0.8	102708	1,95E-05
GS027	Devil's Crown, Floreana Island	2	0.1-0.8	222080	9,01E-06
GS028	Coastal Floreana	2	0.1-0.8	189052	1,06E-05
GS029	North James Bay, Santigo Island	2	0.1-0.8	131529	1,52E-05
GS036	Cabo Marshall, Isabella Island	2	0.1-0.8	77538	2,58E-05
GS108	Coccos Keeling , Inside Lagoon	2	0.1-0.8	1382197	1,45E-06
GS117a	St. Anne Island, Seychelles	2	0.1-0.8	346952	5,76E-06
GS015	Off Key West, FL	1	0.1-0.8	127362	7,85E-06
GS021	Gulf of Panama	1	0.1-0.8	131798	7,59E-06
GS109	Indian Ocean	1	0.1-0.8	59812	1,67E-05
GS113	Indian Ocean	1	0.1 - 0.8	109700	9,12E-06
GS114	500 Miles west of the Seychelles in the Indian Ocean	1	0.1-0.8	348823	2,87E-06
GS115	Indian Ocean	1	0.1 - 0.8	127362	7,85E-06
GS120	Madagascar Waters	1	0.1-0.8	46052	2,17E-05
GS148	East coast Zanzibar (Tanzania), offshore Paje lagoon	1	0.1-0.8	107741	9,28E-06
GS149	West coast Zanzibar (Tanzania), harbour region	1	0.1-0.8	110984	9,01E-06

Annex 12: BpV2 Reads distribution.

Sample	Localitation	num Reads	Filter size	Seq. effort	Relative virus importance
GS021	Gulf of Panama	5	0.1 - 0.8	131798	3,79E-05
GS019	Northeast of Colon	4	0.1 - 0.8	135325	2,96E-05
GS031	Upwelling, Fernandina Island	4	0.1 - 0.8	436401	9,17E-06
GS009	Block Island, NY	2	0.1 - 0.8	79303	2,52E-05
GS013	Off Nags Head, NC	2	0.1 - 0.8	138033	1,45E-05
GS029	North James Bay, Santigo Island	2	0.1 - 0.8	131529	1,52E-05
GS032	Mangrove on Isabella Island	2	0.1 - 0.8	148018	1,35E-05
GS049	Moorea, Outside Cooks Bay	2	0.1 - 0.8	92501	2,16E-05
GS051	Rangirora Atoll	2	0.1-0.8	128982	1,55E-05
GS112	Indian Ocean	2	0.1 - 0.8	1156475	1,73E-06
GS114	500 Miles west of the Seychelles in the Indian Ocean	2	0.1 - 0.8	348823	5,73E-06
GS000b	Sargasso Station 13	1	0.22 - 0.8	317180	3,15E-06
GS012	Chesapeake Bay, MD	1	0.1 - 0.8	126162	7,93E-06
GS017	Yucatan Channel	1	0.1 - 0.8	257581	3,88E-06
GS020	Lake Gatun	1	0.1 - 0.8	296355	3,37E-06
GS023	30 miles from Cocos Island	1	0.1 - 0.8	133051	7,52E-06
GS026	134 miles NE of Galapagos	1	0.1 - 0.8	102708	9,74E-06
GS027	Devil's Crown, Floreana Island	1	0.1 - 0.8	222080	4,50E-06
GS112a	International waters between Madagascar and South Africa	1	0.1 - 0.8	99781	1,00E-05
GS122a	International waters between Madagascar and South Africa	1	0.1 - 0.8	112278	8,91E-06
GS123	International water between Madagascar and South Africa	1	0.1 - 0.8	107966	9,26E-06

Annex 13: CroV Reads distribution.

Sample	Localitation	num Reads	Filter size	Seq. effort	Relative virus importance
GS014	South of Charleston, SC	436	0.1-0.8	128885	3,38E-03
GS002	Gulf of Maine	264	0.1-0.8	121590	2,17E-03
GS007	Northern Gulf of Maine	210	0.1-0.8	50980	4,12E-03
GS003	Browns Bank, Gulf of Maine	133	0.1-0.8	61605	2,16E-03
GS013	Off Nags Head, NC	108	0.1-0.8	138033	7,82E-04
GS006	Bay of Fundy, Nova Scotia	81	0.1-0.8	59679	1,36E-03
GS008	Newport Harbor, RI	77	0.1-0.8	129655	5,94E-04
GS010	Cape May, NJ	63	0.1-0.8	78304	8,05E-04
GS004	Outside Halifax, Nova Scotia	56	0.1-0.8	52959	1,06E-03
GS123	International water between Madagascar and South Africa	51	0.1-0.8	107966	4,72E-04
GS122a	International waters between Madagascar and South Africa	41	0.1-0.8	112278	3,65E-04
GS009	Block Island, NY	28	0.1-0.8	79303	3,53E-04
GS000c	Sargasso Stations 3	26	0.22-0.8	368835	7,05E-05
GS012	Chesapeake Bay, MD	26	0.1-0.8	126162	2,06E-04
GS011	Delaware Bay, NJ	21	0.1-0.8	124435	1,69E-04
GS149	West coast Zanzibar (Tanzania), harbour region	17	0.1-0.8	110984	1,53E-04
GS148	East coast Zanzibar (Tanzania), offshore Paje lagoon	8	0.1-0.8	107741	7,43E-05
GS005	Bedford Basin, Nova Scotia	7	0.1-0.8	61131	1,15E-04
GS020	Lake Gatun	6	0.1-0.8	296355	2,02E-05
GS015	Off Key West, FL	5	0.1-0.8	127362	3,93E-05
GS036	Cabo Marshall, Isabella Island	5	0.1-0.8	77538	6,45E-05
GS028	Coastal Floreana	3	0.1-0.8	189052	1,59E-05
GS034	North Seamore Island	3	0.1-0.8	134347	2,23E-05
GS051	Rangirora Atoll	3	0.1-0.8	128982	2,33E-05
GS016	Outside Seychelles, Indian ocean	2	0.1-0.8	127122	1,57E-05
GS027	Devil's Crown, Floreana Island	2	0.1-0.8	222080	9,01E-06
GS117a	St. Anne Island, Seychelles	2	0.1-0.8	346952	5,76E-06
GS121	International water between Madagascar and South Africa	2	0.1-0.8	110720	1,81E-05
GS000a	Sargasso Station 11	1	0.1-0.8	644551	1,55E-06
GS000b	Sargasso Station 11	1	0.22-0.8	317180	3,15E-06
GS000d	Sargasso Station 13	1	0.22-0.8	332240	3,01E-06
GS018	Rosario Bank	1	0.1-0.8	142743	7,01E-06
GS031	Upwelling, Fernandina Island	1	0.1-0.8	436401	2,29E-06
GS048a	Moorea, Cooks Bay	1	0.1-0.8	90515	1,10E-05
GS116	Outside Seychelles, Indian Ocean	1	0.1 - 0.8	60932	1,64E-05
GS122b	International waters between Madagascar and South Africa	1	0.8 - 3.0	50095	2,00E-05

Annex 14: MpV1 Reads distribution.

Sample	Localitation	num Reads	Filter size	Seq. effort	Relative virus importance
GS007	Northern Gulf of Maine	49	0.1-0.8	50980	9,61E-04
GS002	Gulf of Maine	45	0.1-0.8	121590	3,70E-04
GS011	Delaware Bay, NJ	43	0.1-0.8	124435	3,46E-04
GS003	Browns Bank, Gulf of Maine	42	0.1-0.8	61605	6,82E-04
GS012	Chesapeake Bay, MD	37	0.1-0.8	126162	2,93E-04
GS013	Off Nags Head, NC	23	0.1-0.8	138033	1,67E-04
GS014	South of Charleston, SC	18	0.1-0.8	128885	1,40E-04
GS010	Cape May, NJ	10	0.1-0.8	78304	1,28E-04
GS006	Bay of Fundy, Nova Scotia	9	0.1-0.8	59679	1,51E-04
GS008	Newport Harbor, RI	8	0.1-0.8	129655	6,17E-05
GS122a	International waters between Madagascar and South Africa	8	0.1-0.8	112278	7,13E-05
GS123	International water between Madagascar and South Africa	7	0.1-0.8	107966	6,48E-05
GS009	Block Island, NY	5	0.1-0.8	79303	6,30E-05
GS000c	Sargasso Stations 3	4	0.22-0.8	368835	1,08E-05
GS005	Bedford Basin, Nova Scotia	4	0.1-0.8	61131	6,54E-05
GS004	Outside Halifax, Nova Scotia	3	0.1-0.8	52959	5,66E-05
GS015	Off Key West, FL	2	0.1-0.8	127362	1,57E-05
GS034	North Seamore Island	2	0.1-0.8	134347	1,49E-05
GS148	East coast Zanzibar (Tanzania), offshore Paje lagoon	2	0.1-0.8	107741	1,86E-05
GS000a	Sargasso Station 11	1	0.1-0.8	644551	1,55E-06
GS020	Lake Gatun	1	0.1-0.8	296355	3,37E-06
GS036	Cabo Marshall, Isabella Island	1	0.1-0.8	77538	1,29E-05
GS122b	International waters between Madagascar and South Africa	1	0.8 - 3.0	50095	2,00E-05
GS149	West coast Zanzibar (Tanzania), harbour region	1	0.1-0.8	110984	9,01E-06

Annex 15: OlV1 Reads distribution.

Sample	Localitation	num Reads	Filter size	Seq. effort	Relative virus importance
GS012	Chesapeake Bay, MD	26	0.1-0.8	126162	2,06E-04
GS013	Off Nags Head, NC	26	0.1-0.8	138033	1,88E-04
GS008	Newport Harbor, RI	20	0.1-0.8	129655	1,54E-04
GS011	Delaware Bay, NJ	20	0.1-0.8	124435	1,61E-04
GS007	Northern Gulf of Maine	19	0.1-0.8	50980	3,73E-04
GS002	Gulf of Maine	16	0.1-0.8	121590	1,32E-04
GS014	South of Charleston, SC	15	0.1-0.8	128885	1,16E-04
GS010	Cape May, NJ	10	0.1-0.8	78304	1,28E-04
GS003	Browns Bank, Gulf of Maine	9	0.1-0.8	61605	1,46E-04
GS006	Bay of Fundy, Nova Scotia	9	0.1-0.8	59679	1,51E-04
GS005	Bedford Basin, Nova Scotia	6	0.1-0.8	61131	9,81E-05
GS122a	International waters between Madagascar and South Africa	4	0.1-0.8	112278	3,56E-05
GS117a	St. Anne Island, Seychelles	3	0.1-0.8	346952	8,65E-06
GS004	Outside Halifax, Nova Scotia	2	0.1-0.8	52959	3,78E-05
GS020	Lake Gatun	2	0.1-0.8	296355	6,75E-06
GS123	International water between Madagascar and South Africa	1	0.1-0.8	107966	9,26E-06

Annex 16: OtV1 Reads distribution

Sample	Localitation	num Reads	Filter size	Seq. effort	Relative virus importance
GS013	Off Nags Head, NC	74	0.1-0.8	138033	5,36E-04
GS012	Chesapeake Bay, MD	70	0.1-0.8	126162	5,55E-04
GS002	Gulf of Maine	56	0.1-0.8	121590	4,61E-04
GS007	Northern Gulf of Maine	46	0.1-0.8	50980	9,02E-04
GS011	Delaware Bay, NJ	41	0.1-0.8	124435	3,29E-04
GS008	Newport Harbor, RI	40	0.1-0.8	129655	3,09E-04
GS003	Browns Bank, Gulf of Maine	36	0.1-0.8	61605	5,84E-04
GS010	Cape May, NJ	19	0.1-0.8	78304	2,43E-04
GS014	South of Charleston, SC	16	0.1-0.8	128885	1,24E-04
GS009	Block Island, NY	15	0.1-0.8	79303	1,89E-04
GS004	Outside Halifax, Nova Scotia	11	0.1-0.8	52959	2,08E-04
GS006	Bay of Fundy, Nova Scotia	9	0.1-0.8	59679	1,51E-04
GS005	Bedford Basin, Nova Scotia	6	0.1-0.8	61131	9,81E-05
GS123	International water between Madagascar and South Africa	4	0.1-0.8	107966	3,70E-05
GS000a	Sargasso Station 11	1	0.1-0.8	644551	1,55E-06
GS015	Off Key West, FL	1	0.1-0.8	127362	7,85E-06
GS116	Outside Seychelles, Indian Ocean	1	0.1 - 0.8	60932	1,64E-05
GS121	International water between Madagascar and South Africa	1	0.1-0.8	110720	9,03E-06
GS122a	International waters between Madagascar and South Africa	1	0.1-0.8	112278	8,91E-06
GS148	East coast Zanzibar (Tanzania), offshore Paje lagoon	1	0.1-0.8	107741	9,28E-06
GS149	West coast Zanzibar (Tanzania), harbour region	1	0.1-0.8	110984	9,01E-06

Annex 17: OtV2 Reads distribution

Sample	Localitation	num Reads	Filter size	Seq. effort	Relative virus importance
GS012	Chesapeake Bay, MD	47	0.1-0.8	126162	3,73E-04
GS007	Northern Gulf of Maine	31	0.1-0.8	50980	6,08E-04
GS011	Delaware Bay, NJ	28	0.1-0.8	124435	2,25E-04
GS008	Newport Harbor, RI	26	0.1-0.8	129655	2,01E-04
GS002	Gulf of Maine	25	0.1-0.8	121590	2,06E-04
GS013	Off Nags Head, NC	24	0.1-0.8	138033	1,74E-04
GS003	Browns Bank, Gulf of Maine	20	0.1-0.8	61605	3,25E-04
GS014	South of Charleston, SC	15	0.1-0.8	128885	1,16E-04
GS010	Cape May, NJ	8	0.1-0.8	78304	1,02E-04
GS006	Bay of Fundy, Nova Scotia	7	0.1-0.8	59679	1,17E-04
GS004	Outside Halifax, Nova Scotia	6	0.1-0.8	52959	1,13E-04
GS005	Bedford Basin, Nova Scotia	6	0.1-0.8	61131	9,81E-05
GS123	International water between Madagascar and South Africa	5	0.1-0.8	107966	4,63E-05
GS036	Cabo Marshall, Isabella Island	4	0.1-0.8	77538	5,16E-05
GS122a	International waters between Madagascar and South Africa	4	0.1-0.8	112278	3,56E-05
GS117a	St. Anne Island, Seychelles	3	0.1-0.8	346952	8,65E-06
GS115	Indian Ocean	2	0.1 - 0.8	127362	1,57E-05
GS116	Outside Seychelles, Indian Ocean	2	0.1 - 0.8	60932	3,28E-05
GS122b	International waters between Madagascar and South Africa	2	0.8 - 3.0	50095	3,99E-05
GS148	East coast Zanzibar (Tanzania), offshore Paje lagoon	2	0.1-0.8	107741	1,86E-05
GS149	West coast Zanzibar (Tanzania), harbour region	2	0.1-0.8	110984	1,80E-05
GS000c	Sargasso Stations 3	1	0.22-0.8	368835	2,71E-06
GS020	Lake Gatun	1	0.1-0.8	296355	3,37E-06

Annex 18: OtV5 Reads distribution.

Acknowledgements

This work wouldn't be possible without the ICM infrastructure and the people who "live" in; so thanks to those people that gave me a helping hand at some occasion.

Thanks to Ramiro Logares for introduce me in the bioinformatics world and to Luis Quevedo for his enormous help with MatLab. Thanks so much to Elisabet Sa, for explain to me all the laboratory methods and techniques, for all of her time and, specially, for her patience.

Specially thanks to my thesis directors, Dolors Vaqué and Ramon Massana, for the opportunity, for their knowledge, advises, suggestions, challenges… and for give me up a place to work!

And finally, thanks to all my family and friends that asked me, listened to me and borne me, especially to Anna Salas (my particular English teacher) and Alex Fernandez.