

Absolute quantification of infecting viral particles by chip-based digital PCR

Africa McMullen¹, Francisco Martinez-Hernandez¹, and Manuel Martinez-Garcia^{1*}

¹Department of Physiology, Genetics, and Microbiology, University of Alicante. Alicante, Spain.

* Correspondence:

Manuel Martinez-Garcia

m.martinez@ua.es

Originality-Significance Statement

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/1758-2229.12804

In this study, we employ a novel chip-based digital PCR to unveil the temporal variations of infecting viruses of the uncultured Pelagibacter virus 37-F6 discovered by single-virus genomics. This virus is likely the most abundant viral population in marine ecosystems having a major role and impact on biogeochemical cycles and the 'viral shunt'.

Summary

In silico and empirical quantification of viruses is paramount for obtaining information on viral populations that have a major impact on biogeochemical cycles. The uncultured Pelagibacter virus vSAG 37-F6 discovered via single-virus genomics is one of the most abundant and cosmopolitan marine viruses, however little is understood about its temporal variation. Here, we estimated the absolute number of infecting 37-F6 viruses in coastal bacterioplankton from the Mediterranean Sea by using a novel, feasible SYBR Green I chip-based digital PCR (SYBR dPCR) technique, not implemented before for enumerating (uncultured)-microbes. Quantitative SYBR dPCR estimated 450-3,480 genome copies of virus 37-F6 in cells/mL (i.e. infecting viruses) and a total of \approx 10-400 putative infected cells/mL with a potential C release of 0.12-4.9 pg/mL in the analyzed samples. Considering that virus 37-F6 is ubiquitous and abundant in all *Tara* samples, an enormous amount of C could be transformed by this virus through the 'viral shunt'. Thus, this SYBR dPCR technique has enabled the absolute quantification of an ecologically relevant -uncultured- virus in nature and the estimation of its potential contribution on biogeochemical cycles. Overall, our study also shows that this approach has a broad applicability for quantifying any other target loci in Microbiology and Virology.

Introduction

Nowadays, the ecological role of viruses in nature is unquestionable (Danovaro *et al.*, 2011). In the case of marine ecosystems, viral metagenomics (i.e viromics), and more recently metatranscriptomics, have been undertaken to characterize the biogeography, diversity, *in silico* abundance, activity and transcriptional dynamics of important and abundant viruses (Brum and Sullivan, 2015), which has greatly expanded our current knowledge on marine viruses (Mizuno *et al.*, 2013; Zhao *et al.*, 2013; Brum and Sullivan, 2015; Paez-Espino *et al.*, 2016; Roux *et al.*, 2016; Luo *et al.*, 2017; Martinez-Hernandez *et al.*, 2017; Gregory *et al.*, 2019). However, recently, the novel single-virus genomic (SVG) approach has unveiled that some of the putative most abundant and cosmopolitan viruses in the ocean surface, such as the single virus vSAG 37-F6, had been overlooked by these techniques (Martinez-Hernandez *et al.*, 2017). This 'hidden' uncultured and microdiverse virus that also populates the deep ocean was uncovered by combining cutting-edge fluorescence-activated virus sorting, whole-genome amplification and the sequencing of one virus at a time (Martinez-Hernandez *et al.*, 2017). Later, single-cell genomics showed that 37-F6 viruses and relatives were present in sorted single *Pelagibacter* cells (Martinez-

Hernandez *et al.*, 2019a), and transcriptomic data provided first evidence of high in situ activity in coastal temperate waters of the NE Atlantic (Alonso-Sáez *et al.*, 2018). The 37-F6 viral population appears to have a major contribution to marine C cycles by controlling one of the most abundant bacteria on Earth (Martinez-Hernandez *et al.*, 2017, 2019a), but nothing is known about temporal dynamics of this recently discovered and ubiquitous virus.

Undoubtedly, metagenomics and viromics are powerful approaches to shed light on different biological aspects of the ecogenomics of microbes and their viruses (Roux *et al.*, 2016; Aylward *et al.*, 2017; Martinez-Hernandez *et al.*, 2017; Yoshida *et al.*, 2018). However, often, the *-in silico-* units (e.g. RPKM or FKPM) obtained to measure relative viral abundances from these techniques are unable to provide direct information about total number of viruses or infecting viral particles. Over the last years, different culture-independent tools have been used to estimate the abundance of specific viruses, such as phage-FISH (Allers *et al.*, 2013), quantitative PCR (qPCR) (Eggleston and Hewson, 2016) or the more recent PCR polony method (Baran *et al.*, 2018). Recently, digital PCR (dPCR) is becoming a powerful technique for estimating the absolute abundances of target loci in microbiology. The vast potential of dPCR in different fields (Ottesen *et al.*, 2006) is enormous, from Environmental Microbiology to Clinical Virology (Ottesen *et al.*, 2006; Vynck *et al.*, 2016). In dPCR, the sample is separated into a large number of partitions and the reaction is carried out in each partition individually, either in droplets (i.e. droplet digital PCR (ddPCR)) (Hindson *et al.*, 2011) or in pico-volume chip wells (i.e. chip-based dPCR) (Baker, 2012). In the case of chip-based dPCR and ddPCR, TaqMan chemistry is the preferred approach. The chip-based SYBR Green I dPCR quantification strategy (hereinafter

SYBR dPCR) has been rarely employed in Biology, with only a couple of examples reported within the human/rat framework (Burnham *et al.*, 2016; Mahood *et al.*, 2016), but has never been implemented and optimized for enumerating target (uncultured)-microorganisms.

Furthermore, a side-by-side comparison with the standard TaqMan chemistry chip-based dPCR is lacking. Given the simplicity of this SYBR dPCR which only requires two specific primers targeting the studied microorganism and the SYBR Green dye, this approach has an enormous potential not only in Microbial or Viral Ecology but also in other disciplines.

Thus, in this study we used this novel quantitative method based on chip-based digital PCR (dPCR) technology to investigate the number of infecting viral particles (i.e. active infecting viruses in bacterioplankton cell fraction) and applied this novel technology for enumerating the cosmopolitan abundant uncultured virus 37-F6, which is paramount overall for estimating the impact of this and other reference uncultured viruses in nutrient cycles.

Results and discussion

For our pilot SYBR dPCR study, we collected surface marine samples (1L) from two different coastal locations of the Mediterranean Sea (Fig. S1) near our dPCR laboratory at daytime and at night-time (before dawn): Cape Huertas (Alicante city, Spain; coordinates 38.353681, -0.426162; daytime 14:10 pm; night-time 5:45 am) and Campello Beach (coordinates 38.419096, -0.388555; daytime 14:40 pm; night-time 6:15 am). Sunrise and sunset time was 7:09 am and 19:11 pm, respectively. The Campello beach sampling site is environmentally impacted by the discharge of a nearby wastewater treatment plant and shows higher microbial abundance values (Fig. S2), in contrast to Cape Huertas, which is a

relatively well-conserved area with minimal human impact (Fig. S1). DNA extractions were performed from cell pellets in order to estimate the abundance of vSAG 37-F6 inside or attached to *Pelagibacter* spp. host cells (most likely infected cells). For that, seawater samples (600-800 mL) were centrifuged with an Ultra Centrifuge Beckman (model Avanti J-30I and rotor JA-14) during 90 minutes at 30,000 g in order to collect the bacterioplankton fraction and separated viruses (supernatant) from cells (pellet). This protocol ensures the collection of all ultramicrobacteria *Pelagibacter* cells present in marine samples (as per recommendations of Stephen Giovannoni, personal communication). Our bacterioplankton cell sampling and purification ensured that in principle no free viruses were collected, although we cannot rule out that some of the putative enumerated infecting viral particles in this study could indeed be some viruses attached to the cell and not actively replicating. However, considering that virus attachment is mainly host-specific, this would not negatively impact nor bias our dPCR quantification data of infecting virus 37-F6. DNA extractions from cellular pellets were carried out with MasterPure Complete DNA and RNA Purification Kit (Epicentre) following manufacture's protocol. DNA was resuspended in 35 μ L of buffer TE and quantified using the Invitrogen QubitTM dsDNA Assay Kit (Thermo Fisher Scientific). Extracted DNA was further purified to remove PCR inhibitors with DNeasy Blood and Tissue kit (Qiagen) according to manufacture's protocol.

The chip-based dPCR assays were run on a QuantStudio 3D Digital PCR System (ThermoFisher) and a detailed guide with tips for implementing SYBR dPCR in other laboratories for enumerating target microorganisms is available in Supplementary Materials. In dPCR, the DNA target template should be diluted according to Poisson distribution so that

each individual partition (e.g. each chip well in dPCR) contains approximately $0.6-1.6$ copies of the target DNA (Hindson *et al.*, 2011). First, for SYBR Green I chip-based dPCR, a specific primer set (named Seq6; amplicon size ≈ 520 bp; Fig. S3A; see Supplementary information for more details) targeting the virus vSAG 37-F6 was designed as previously described (Martinez-Hernandez *et al.*, 2019a) and further validated with Illumina amplicon sequencing demonstrating that $\approx 95\%$ of sequenced reads were indeed from the target virus at the species level ($\geq 95\%$ nucleotide identity, Fig. S3B) (Roux *et al.*, 2016; Martinez-Hernandez *et al.*, 2017). For the TaqMan chip-based dPCR assays, a previous validated probe and primer set named ddSeq4 (size of PCR amplicon band ≈ 100 ; Fig. S3A) was employed (Martinez-Hernandez, *et al.*, 2019b). First, we ensured by real time PCR that primer sets Seq6 and ddSeq4 used for TaqMan and SYBR chemistry assays had similar efficiencies and reproducibility for monitoring abundances of virus 37-F6 genome (Fig. S4). Then, TaqMan and SYBR dPCR assays were applied in parallel to estimate the number of virus 37-F6 genome copies in an aliquot containing an unknown amount of DNA template (Fig. 1) obtained in a previous single-virus genomics study from the Blanes Bay Microbial Observatory (Western Mediterranean Sea) (Martinez-Hernandez *et al.*, 2017). Nearly identical copy numbers were obtained by both dPCR assays with a high overall precision ($\approx 3\%$; Fig. 1). Negative (no template) and positive (one template per well) PCR-amplified wells in the dPCR chips were clearly distinguishable according to the fluorescence signal (Fig. 1). In both assays, the DNA template was properly diluted with <1 copy of the viral genome per well, as per dPCR recommendations. Next, we demonstrated from a serial dilution experiment with the above mentioned aliquot of pure DNA template of virus 37-F6,

that SYBR dPCR quantification data correlated with consecutive dilutions as expected, showing a robust and accurate estimation of the number of virus genome copies (Fig. 2). As a logical further step, TaqMan and SYBR dPCR technologies were applied to one of the collected bacterioplankton environmental samples (Cape Huertas, non-impacted). From both dPCR chemistries, nearly identical abundance data were obtained for numbers of infecting virus 37-F6 (≈ 450 viral copies per mL of seawater; Fig. 3). Remarkably, for “ground truthing” our dPCR data and corroborating that we were indeed enumerating our virus 37-F6 and not any other undesired non-specific target genes (yielding unexpected PCR amplicon sizes), chips were unsealed after fluorescence reading, and PCR amplicons were recovered (Fig. S5). PCR amplicon size in dPCR chip wells (one chip contains 20,000 wells) was as expected for virus 37-F6. A consistent amplicon size, together with Illumina sequencing data from the obtained PCR amplicons, confirmed that dPCR enumeration of 37-F6 virus was fully specific (Figs. S3B and S5).

Once proving the accuracy, robustness, and specificity of this novel SYBR dPCR assay targeting 37-F6, it was then used to estimate the number of infecting viral particles of 37-F6 in the collected coastal bacterioplankton fraction. Data overall ranged from 450 to 3,480 infecting viruses/mL (Fig. 4). At daytime (14:00), no differences regarding 37-F6 abundance were observed between impacted and non-impacted samples, despite having contrasting local conditions, although higher overall values were obtained for the impacted sample. Significant diel variations were only observed for the non-impacted marine sample (Cape Huertas; 10-fold higher at daytime), while minor variations were observed between

day-night for the impacted-sample (Campello Beach) likely because the natural diel cycles of bacterioplankton is altered by the constant input of organic matter in this location.

The estimated mean viral burst size of heterotrophic prokaryotes in marine systems has been reported to be 24 (Parada *et al.*, 2006). In the case of isolated pelagiphages, this value ranged between 9-49 (Zhao *et al.*, 2013). Although the burst size of the uncultured pelagiphage vSAG 37-F6 is unknown, it seems reasonable that, although speculative, it could somehow be close to the range of that of isolated pelagiphages, which in any case is near to the mean reported burst size for heterotrophic marine prokaryotes (Parada *et al.*, 2006). Thus, based on these burst size values, we estimated that in our samples at a given time, the total number of infected cells per mL ranged from 10 to ≈ 400 , which means a total potential C release from 124 fg to 4.9 pg (assuming total C cell content in oceanic bacterial assemblages of 12.4 fg as described in (Fukuda *et al.*, 1998). These theoretical estimations for virus 37-F6 have to be further corrected with empirical burst size data. However, regardless of the deviation of the 37-F6 burst size value from the reported mean value (Parada *et al.*, 2006) and the consideration of virus 37-F6 being ubiquitous and abundant in all *Tara* samples collected from tropical and subtropical oceans (Martinez-Hernandez *et al.*, 2017), an enormous amount of C could be transformed by this virus through the viral shunt. Recently, using ddPCR assay, which requires a more sophisticated and costly equipment, we estimated that the absolute number of free viral particles of virus vSAG 37-F6 (standing stock) in seawater ranged from 360 to 8,510 per mL (Martinez-Hernandez *et al.*, 2019b). In the case of the isolated pelagiphage HTVC010P (Zhao *et al.*, 2013), qPCR data in different samples collected from

the open Atlantic Ocean (Eggleston and Hewson, 2016) showed abundance values within the order of $10^3 - 10^4$ viruses/mL.

Here, with a limited number of samples and resources, we have demonstrated the utility of SYBR Green I chip-based dPCR for enumerating uncultured target microorganisms. Using this approach, we estimated the number of infecting viral particles of 37-F6 (i.e. actively replicating in cells) and the total infected cells in the bacterioplankton fraction of the Mediterranean Sea. In contrast to TaqMan dPCR chemistry, this novel SYBR Green I chip-based dPCR is more simple, feasible, and affordable, since probes are not required, which reduces the assay setup and running costs. Quantification of specific ecologically relevant viruses with readily available technology, such as the SYBR dPCR used, can broaden our understanding of viral assemblage structure, and the contribution of specific groups to the 'viral shunt' and biogeochemical cycles (Baran *et al.*, 2018). The use of these and complementary technologies applied in high-resolution, comprehensive time-series studies of marine ecosystems should help advance understanding of microbial dynamics in the ocean.

Acknowledgments

We greatly appreciate Professor Andres Moya and Dr. Raquel Sarmiento for sharing their dPCR equipment for performing the dPCR assays. This work has been supported by Spanish Ministry of Economy and Competitiveness to MMG (ref. RTI2018-094248-B-100), Generalitat Valenciana to FMH (ACIF/2015/332), and Gordon and Betty Moore Foundation to MMG (grant 5334). Gordon and Betty Moore Foundation to EFD (3777) and Simons Foundation Grant #329108 (to EFD).

Conflict of Interest Statement

The authors declare that they have no conflict of interest.

References

- Allers, E., Moraru, C., Duhaime, M.B., Beneze, E., Solonenko, N., Barrero-Canosa, J., et al. (2013) Single-cell and population level viral infection dynamics revealed by phageFISH, a method to visualize intracellular and free viruses. *Environ Microbiol* **15**: 2306–2318.
- Alonso-Sáez, L., Anxelu Morán, X.G., Clokie, M.R., and Alonso-Sáez Laura, L. (2018) Low activity of lytic pelagiphages in coastal marine waters. *ISME J* **12**: 2100–2102.
- Aylward, F.O., Boeuf, D., Mende, D.R., Wood-Charlson, E.M., Vislova, A., Eppley, J.M., et al. (2017) Diel cycling and long-term persistence of viruses in the ocean's euphotic zone. *Proc Natl Acad Sci U S A* **114**: 11446–11451.
- Baker, M. (2012) Digital PCR hits its stride. *Nat Methods* **2012** 96.
- Baran, N., Goldin, S., Maidanik, I., and Lindell, D. (2018) Quantification of diverse virus populations in the environment using the polony method. *Nat Microbiol* **3**: 62–72.
- Brum, J.R. and Sullivan, M.B. (2015) Rising to the challenge: accelerated pace of discovery transforms marine virology. *Nat Rev Microbiol* **13**: 147–159.
- Burnham, P., Kim, M.S., Agbor-Enoh, S., Luikart, H., Valentine, H.A., Khush, K.K., and De Vlaminc, I. (2016) Single-stranded DNA library preparation uncovers the origin and diversity of ultrashort cell-free DNA in plasma. *Sci Rep* **6**: 27859.

Danovaro, R., Corinaldesi, C., Dell'anno, A., Fuhrman, J.A., Middelburg, J.J., Noble, R.T., and Suttle, C.A. (2011) Marine viruses and global climate change. *FEMS Microbiol Rev* **35**: 993–1034.

Eggleston, E.M. and Hewson, I. (2016) Abundance of Two Pelagibacter ubique Bacteriophage Genotypes along a Latitudinal Transect in the North and South Atlantic Oceans. *Front Microbiol* **7**: 1534.

Fukuda, R., Ogawa, H., Nagata, T., and Koike, I. (1998) Direct determination of carbon and nitrogen contents of natural bacterial assemblages in marine environments. *Appl Environ Microbiol* **64**: 3352–8.

Gregory, A.C., Zayed, A.A., Conceição-Neto, N., Temperton, B., Bolduc, B., Alberti, A., et al. (2019) Marine DNA Viral Macro- and Microdiversity from Pole to Pole. *Cell* **177**: 1109-1123.e14.

Hindson, B.J., Ness, K.D., Masquelier, D.A., Belgrader, P., Heredia, N.J., Makarewicz, A.J., et al. (2011) High-Throughput Droplet Digital PCR System for Absolute Quantitation of DNA Copy Number. *Anal Chem* **83**: 8604–8610.

Mahood, T.H., Johar, D.R., Iwasiow, B.M., Xu, W., and Keijzer, R. (2016) The transcriptome of nitrofen-induced pulmonary hypoplasia in the rat model of congenital diaphragmatic hernia. *Pediatr Res* **79**: 766–775.

Martinez-Hernandez, F., Fornas, O., Lluesma Gomez, M., Bolduc, B., de la Cruz Peña, M.J., Martínez, J.M., et al. (2017) Single-virus genomics reveals hidden cosmopolitan and

abundant viruses. *Nat Commun* **8**: 15892.

Martinez-Hernandez, F., Fornas, Ò., Lluesma Gomez, M., Garcia-Heredia, I., Maestre-Carballa, L., López-Pérez, M., et al. (2019a) Single-cell genomics uncover Pelagibacter as the putative host of the extremely abundant uncultured 37-F6 viral population in the ocean. *ISME J* **13**: 232–236.

Martinez-Hernandez, F., Garcia-Heredia, I., Lluesma Gomez, M., Maestre-Carballa, L., Martínez Martínez, J., and Martinez-Garcia, M. (2019b) Droplet Digital PCR for Estimating Absolute Abundances of Widespread Pelagibacter Viruses. *Front Microbiol* **10**: 1226.

Ottesen, E.A., Hong, J.W., Quake, S.R., and Leadbetter, J.R. (2006) Microfluidic digital PCR enables multigene analysis of individual environmental bacteria. *Science* **314**: 1464–7.

Paez-Espino, D., Eloë-Fadrosh, E.A., Pavlopoulos, G.A., Thomas, A.D., Huntemann, M., Mikhailova, N., et al. (2016) Uncovering Earth's virome. *Nature* **536**: 425–430.

Parada, V., Herndl, G.J., and Weinbauer, M.G. (2006) Viral burst size of heterotrophic prokaryotes in aquatic systems. *J Mar Biol Assoc United Kingdom* **86**: 613–621.

Roux, S., Brum, J.R., Dutilh, B.E., Sunagawa, S., Duhaime, M.B., Loy, A., et al. (2016) Ecogenomics and potential biogeochemical impacts of globally abundant ocean viruses. *Nature* **537**: 689–693.

Vynck, M., Trypsteen, W., Thas, O., Vandekerckhove, L., and De Spiegelaere, W. (2016) The Future of Digital Polymerase Chain Reaction in Virology. *Mol Diagn Ther* **20**: 437–

447.

Yoshida, T., Nishimura, Y., Watai, H., Haruki, N., Morimoto, D., Kaneko, H., et al. (2018)

Locality and diel cycling of viral production revealed by a 24 h time course cross-omics analysis in a coastal region of Japan. *ISME J* **12**: 1287–1295.

Zhao, Y., Temperton, B., Thrash, J.C., Schwalbach, M.S., Vergin, K.L., Landry, Z.C., et al.

(2013) Abundant SAR11 viruses in the ocean. *Nature* **494**: 357–360.

Accepted Article

Figure and Figure legends

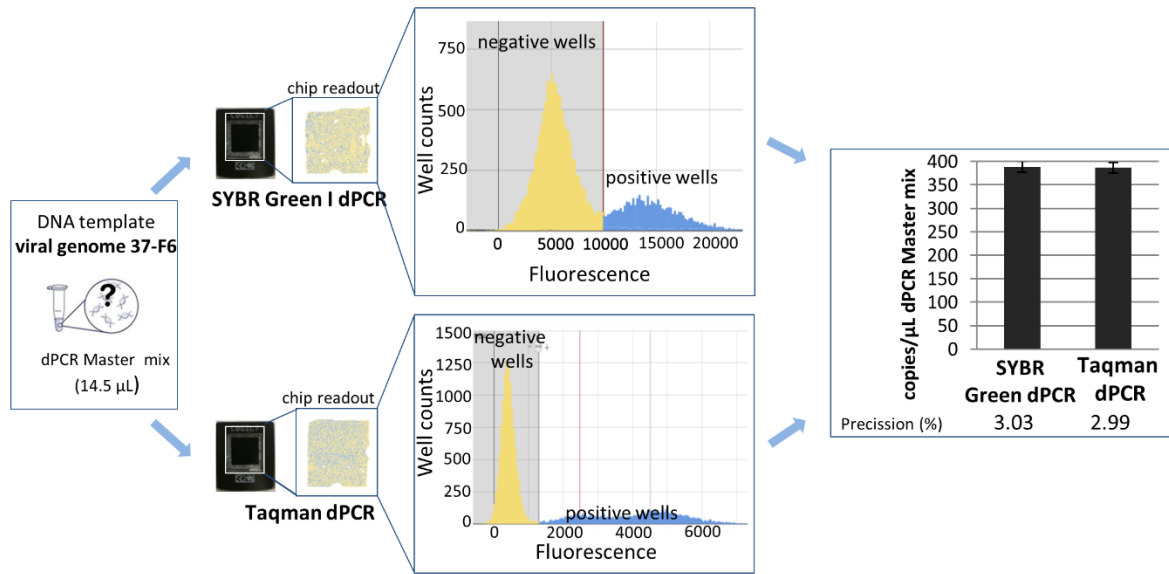
Figure 1. dPCR optimization assays with virus 37-F6 DNA template. Comparison of SYBR Green I chip-based dPCR chemistry vs TaqMan chip-based dPCR with virus 37-F6 DNA template. dPCR chip readouts are shown for both assays and frequency histograms of negative and positive chip wells are depicted. According to general guidelines of dPCR, a bimodal peak of negative and positive wells should be obtained, being positive-well peak always much lower than positive-wells peak. In dPCR, it is paramount that DNA target template is diluted according to Poisson distribution so that each individual chip well contains approximately 0.6-1.6 copies of the target DNA (Hindson *et al.*, 2011). Bar chart indicates copies of virus 37-F6 in the dPCR master mix (precision of the technique is indicated below bar chart; $\approx 3\%$).

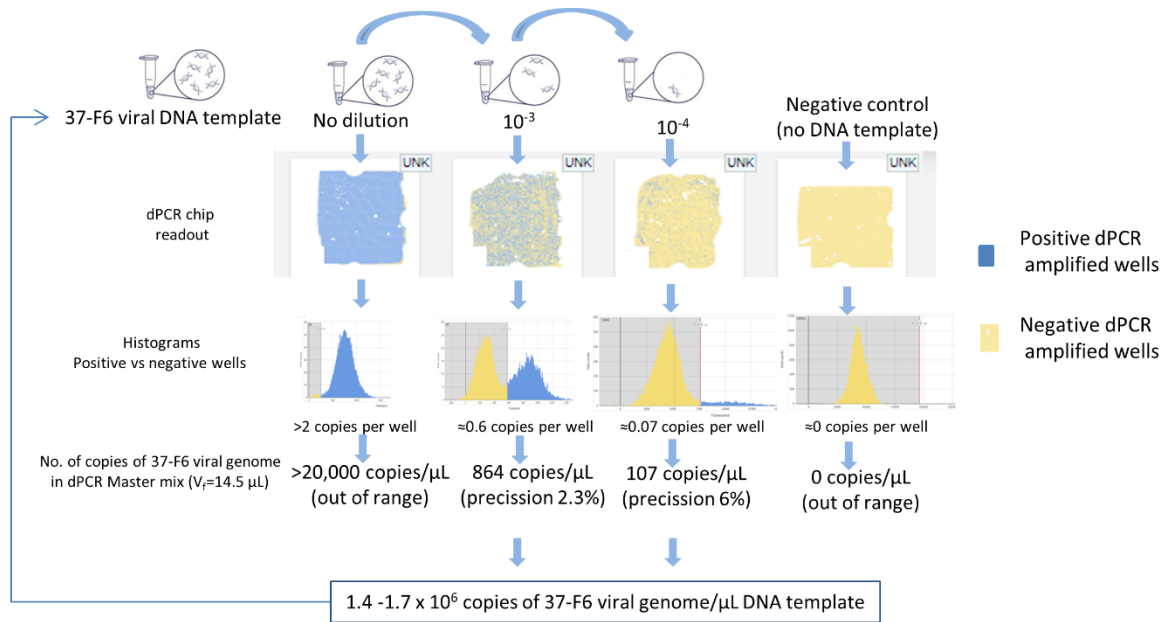
Figure 2. SYBR Green I chip-based dPCR experiments with serial dilution of 37-F6 DNA viral template. An aliquot containing an unknown number of viral copies of 37-F6 obtained in a previous study by single-virus genomics (Martinez-Hernandez *et al.*, 2017) was used as a template for the dilution dPCR experiment. When dPCR chips were loaded with undiluted DNA template or diluted up to 10^{-2} , the number of DNA template copies exceeded that of recommended value for dPCR of 1.6 per well, and thus all wells contained more than 1 copy per well. This, according to Poisson distribution, leads to an inaccurate measurement and thus out of range. Same happens when template is too diluted (out of range). In this case, dilution 10^{-3} and 10^{-4} of the original aliquot led to an accurate measurement of total copies of 37-F6 (1.4-1.7 copies/uL). dPCR quantification data from these consecutive 1-fold diluted aliquots yielded a quantitative data near to that expected by theory (only a very minor deviation was observed, likely because of the inherent error of manual micropipetting or stochastic dilution process of a DNA template)

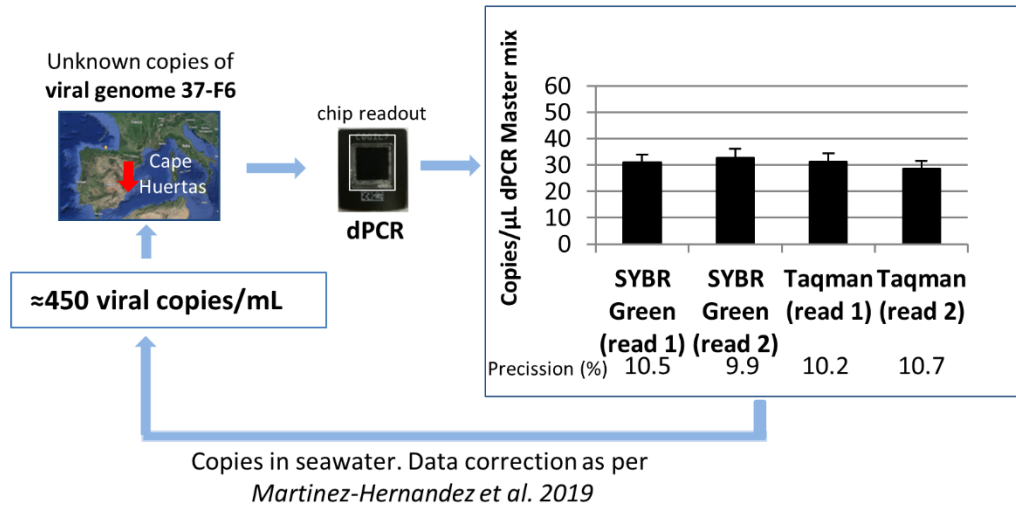
Figure 3. dPCR comparison of SYBR Green I chip-based dPCR chemistry vs TaqMan chip-based dPCR with an environmental DNA template (bacterioplankton cell fraction) collected at Cape Huertas (Mediterranean Sea). Total number of infecting viruses per mL of seawater is indicated according to dPCR data correction by (Martinez-Hernandez *et al.*, 2019b).

Figure 4. Chip-based SYBR Green I dPCR data of infecting viral particles in the bacterioplankton cell fraction collected from the Mediterranean Sea at two times (night and day) and locations (Cape Huertas and Campello Beach, the latter environmentally impacted by the discharging of the effluent of a wastewater treatment plant, see Fig. S1). **(A)** dPCR data of total copies of virus 37-F6 per mL of seawater. **(B)** Estimation of total infected cells per mL of seawater. Burst size data of pelagiphage was taken from (Zhao *et al.*, 2013). It is worth noting that empirical burst size of virus 37-F6 is unknown, and thus this bar chart only shows a theoretical possibly burst size according to literature.

Accepted Article







Accepted Article

