Santos Fernando (Orcid ID: 0000-0002-6281-7310)
Antón Josefa (Orcid ID: 0000-0002-5823-493X)

A manuscript re-submitted to *Environmental Microbiology*

Environmental dissolved DNA harbors meaningful biological information on microbial community structure

Borja Aldeguer-Riquelme[1], María Dolores Ramos-Barbero[1], Fernando Santos*[1], Josefa Antón[1,2].

[1] Department of Physiology, Genetics, and Microbiology, University of Alicante, 03080 Alicante, Spain. [2] Multidisciplinary Institute of Environmental Studies Ramón Margalef, University of Alicante, 03080 Alicante, Spain

Correspondence to: Fernando Santos

E-mail: fernando.santos@ua.es, Tel. +34 965 90 38 53, Fax. +34 965 90 95 69

Originality-Significance Statement. Extracellular DNA (eDNA) comprises all the DNA molecules outside cells and may serve as a source of nutrients and genetic information. Here, we have optimized a protocol for dissolved eDNA (dDNA) extraction from a hypersaline system, showing that centrifugation might induce bias in dDNA quantification and composition. In-depth metagenomic analysis of dDNA provided insight into its distinctive composition, as well as into virus-host interactions and microbial dynamics in hypersaline environments, unveiling the role of nanohaloarchaeal viruses as putative relevant players in the system.

Running title: Dissolved DNA from hypersaline environments

Keywords: eDNA, dDNA, hypersaline, *Nanohaloarchaeota*, nanohalovirus, *Haloquadratum*, halovirus

Summary

Extracellular DNA (eDNA) comprises all the DNA molecules outside cells. This component of microbial ecosystems may serve as a source of nutrients and genetic information. Hypersaline environments harbor one of the highest concentrations of eDNA reported for natural systems, which has been attributed to the physicochemical preservative effect of salts and to high viral abundance. Here, we compared centrifugation and filtration protocols for the extraction of dissolved DNA (dDNA, as opposed to eDNA that also includes DNA from free viral particles) from a solar saltern crystallizer pond (CR30) water sample. The crystallizer dDNA fraction has been characterized, for the first time, and compared with cellular and viral metagenomes from the same location. High-speed centrifugation affected CR30 dDNA concentration and composition due to cell lysis, highlighting that protocol optimization should be the first step in dDNA studies. Crystallizer dDNA, which accounted for lower concentrations than those previously reported for hypersaline anoxic sediments, had a mixed viral and cellular origin, was enriched in archaeal DNA and had a distinctive taxonomic composition compared to that from the cellular assemblage of the same sample. Bioinformatic analyses indicated that nanohaloarchaeal viruses could be a cause for these differences.

Introduction

The extracellular DNA (eDNA) of a natural sample comprises all the DNA molecules outside the cells (Torti *et al.*, 2015), and thus includes the DNA inside free viral particles and the dissolved DNA. Among different natural systems, hypersaline environments might harbor one of the highest concentrations of eDNA (Danovaro *et al.*, 2005). This fact could be attributed to the elevated concentration of salts, which preserves the dissolved DNA from thermodegradation or depurination (Marguet and Forterre, 1994; Tehei *et al.*, 2002; Corinaldesi *et al.*, 2008), and to the high abundance and activity of viruses, causing the release of DNA through capsid disruption and host lysis (Santos *et al*., 2011; Corinaldesi *et al.*, 2014). Although cellular and viral communities inhabiting hypersaline waters have been extensively described (Antón *et al.*, 1999, 2002; Benlloch *et al.*, 2001, 2002; Papke *et al.*, 2003, 2004; Bolhuis *et al.*, 2004; Maturrano *et al.*, 2006; Mutlu *et al.*, 2008; Santos *et al.*, 2010, 2011; Ghai *et al.*, 2011; Narasingarao *et al.*, 2012; Boujelben *et al.*, 2012; García-Heredia *et al.*, 2012; Podell *et al.*, 2013; Zhaxybayeva *et al.*, 2013; Fernández *et al.*, 2014; Gomariz *et al.*, 2015; Vavourakis *et al.*, 2016; Viver *et al.*, 2018), only three studies on eDNA in hypersaline environments or halophilic microbes have been

reported. However, they were focused on the quantification and dynamics of the extracellular genetic material in deep hypersaline anoxic sediments (Danovaro *et al.*, 2005; Corinaldesi *et al.*, 2014), and on the impact of the dissolved DNA in pure cultures of the haloarchaeon *Haloferax volcanii* (Chimileski *et al.*, 2014).

The microbiota of the multi-pond solar saltern Bras del Port (Santa Pola, Alicante, Spain) is one of the best-characterized hyperhalophilic communities so far, vastly studied for more than 30 years by both culture-dependent and -independent approaches (Antón *et al.*, 1999, 2002; Papke *et al.*, 2003, 2004; Legault *et al.*, 2006; Santos *et al.*, 2011; García-Heredia *et al.*, 2012; Zhaxybayeva *et al.*, 2013; Ventosa *et al.*, 2014; Gomariz *et al.*, 2015). In Bras del Port crystallizer ponds, where sodium chloride precipitates and is harvested, the extremely halophilic microbial community is mainly composed of dense populations of a few prokaryotic phyla (Guixa-Boixareu *et al.*, 1996; Benlloch *et al.*, 2002; Ventosa *et al.*, 2015). The "square" euryarchaeon *Haloquadratum walsbyi* dominates the community, comprising up to 80% of the total cells (Oh *et al.*, 2010), and *Salinibacter ruber* is the most abundant bacterium (Antón *et al.*, 1999; Ghai *et al.*, 2011; Gomariz *et al.*, 2015). In addition, uncultured representatives from the phylum *Nanohaloarchaeota* are also present in significant numbers, as well as variable proportions of readily culturable high-GC haloarchaea (Ghai *et al.*, 2011; Gomariz *et al.*, 2015). The dense viral assemblages in the crystallizers, reaching up to 10^9 viral-like particles (VLPs) per mL, are dominated by low-GC content viruses infecting *Hqr. walsbyi* (Santos *et al.*, 2010; García-Heredia *et al.*, 2012; Villamor *et al*., in prep.), although the high-GC viruses associated to *S. ruber* and high-GC haloarchaea seem to be more active in these ponds (Santos *et al.*, 2011). In summary, the high numbers of prokaryotes and viruses, which can contribute to maintain a high pool of dDNA, together with their relatively low diversity and the role of salt as a physicochemical preservative agent of dDNA, make crystallizers suitable models to delve into the role of dDNA in nature.

In the present work we have optimized the protocol for the extraction of dissolved DNA from a CR30 crystallizer brine. The obtained dDNA fraction has been characterized and compared with cellular and viral metagenomes from the same location. Our results indicate that filtration is the most suitable strategy to study the CR30 dDNA, in agreement with the recent work by Linney *et al.* (2021) in marine dDNA, avoiding cell damage induced by other techniques, such as high-speed centrifugation. We have also quantified, for the first time, the concentration of the CR30 dDNA, which was lower than that previously reported for hypersaline deep and anoxic sediments (Danovaro *et al.*, 2005; Corinaldesi *et al.*, 2014). The crystallizer dDNA, with a size range from 0.5 to 1.5 kb, had a mixed cellular and viral origin, with a taxonomic composition

which was different from that in cellular metagenome. Bioinformatic analyses indicated that nanohaloarchaeal viruses could be a cause for these observed differences.

Results and discussion

Centrifugation induces artifacts in crystallizer dDNA studies

The dDNA fraction from the CR30 water subsamples were obtained with two different protocols used to remove cells: high-speed centrifugation plus filtration ("C" protocol) or consecutive filtration steps without centrifuging ("F" protocol) (Fig. 1). The amount of dDNA recovered was higher for protocol C (hereinafter: CdDNA, with 19.04 µg/L) than for protocol F (hereinafter: FdDNA, with 10.3 µg/L), suggesting, as discussed below, that high-speed centrifugation might be introducing a bias in the results (Fig. 2A). These dDNA values were lower than those previously reported for deep-hypersaline anoxic sediments (338 ng/g – 343 µg/g) (Danovaro *et al.*, 2005; Corinaldesi *et al.*, 2014), and were in the range of those from marine environments (from 0.03 µg/L, for oligotrophic oceans, to 44 µg/L, for subtropical estuaries; Deflaun *et al.*, 1986; DeFlaun *et al.*, 1987; Paul *et al.*, 1991; Dell'Anno and Corinaldesi, 2004; Nielsen *et al.*, 2007; Ibáñez de Aldecoa *et al.*, 2017) and freshwater systems (0.5-25.6 µg/L; Deflaun *et al.*, 1986; Paul *et al.*, 1989; Torti *et al.*, 2015; Ibáñez de Aldecoa *et al.*, 2017). Nevertheless, these published data must be carefully considered since different techniques were used for DNA quantifications and, in some studies, DNA from free viral particles was not separated from the dDNA pool, and thus DNA inside viral capsids was wrongly considered as a part of the dDNA fraction.

The size range of the CR30 CdDNA and FdDNA was also evaluated by conventional electrophoresis, which predominantly showed nucleic acids between 0.5 and 1.5 Kb. Larger dDNA was also observed, which could correspond to viral genomes, plasmids or chromosomal fragments (Fig. 2B). Similar results were obtained in preliminary studies of CR30 crystallizer samples taken in December 2016 (data not shown). These dDNA sizes were more similar to those observed in marine waters (ranging from 0.12 to 35.2 Kb, with a high predominance of small fragments; DeFlaun *et al.*, 1987) than to those from freshwaters (from 0.1 to 0.5 Kb; Beebee, 1993). CdDNA contained fragments that were slightly smaller than those in the FdDNA fraction, which displayed a higher amount of larger DNA (see arrows in figure 2B). This difference might be partially due to the cell damage caused by centrifugation and the consequent release of intracellular nucleases, which could actively degrade dissolved DNA.

To explore in more depth the effect of the applied protocols on dDNA retrieval, the origin of DNA in the CdDNA and FdDNA pools was determined by 16S rRNA gene DGGE profiling and compared with that of the CR30 cellular fraction (hereinafter: iDNA, from intracellular DNA). The archaeal DGGE iDNA pattern was identical to those found in the CdDNA and FdDNA (Suppl. Fig. 1). However, bacterial primers failed to amplify 16S rRNA genes in both dDNA fractions. This result indicated that, if present, dissolved DNA from bacterial origin was below the detection limits of the standard PCR, and suggested that bacterial cells were more resistant to high-speed centrifugation-induced damages than their archaeal counterparts. It is noteworthy that the compositions of the cellular fractions collected by centrifugation (CiDNA) and filtration (FiDNA) were almost identical, as estimated by 16S metabarcoding (Suppl. Fig. 2).

In order to get a better description of FdDNA and CdDNA, metagenomics was performed. For this purpose, the CR30-derived FdDNA, CdDNA, iDNA and vDNA (from viral DNA) fractions were sequenced (Table 1). A Nonpareil analysis indicated that the sequencing depths (Table 1) were above the recommended threshold of 60% to establish comparisons among different data-sets and to obtain meaningful and unbiased biological conclusions (Rodriguez-R and Konstantinidis, 2014a; 2014b). Then, the CdDNA and FdDNA reads were compared with those from the iDNA and vDNA metagenomes. BLASTN-derived PCAs highlighted that iDNA was more similar to the CdDNA than to the FdDNA (Fig. 3A-3B). This clustering was also in good agreement with the taxonomic composition of the different metagenomes (except for vDNA), either considering all metagenomic reads or only the 16S rRNA gene reads (Fig. 4). Indeed, according to the 16S rRNA gene metagenomic profiles, the most abundant genus in the three fractions was *Haloquadratum*, whose proportion in the CdDNA (57%) was closer to the proportion in the iDNA (63%) than in the FdDNA fraction (42%) (Fig. 4A). In addition, the most abundant CR30 bacterium, *Salinibacter ruber*, was detected in the iDNA (2%), but it showed a lower abundance in the CdDNA fraction (0.6%) and it was absent in the FdDNA. This result could explain the absence of bacterial PCR products mentioned above and were also in good agreement with our previous analyses (data not shown). Similar results were obtained for the taxonomic profiles based on non-16S metagenomic reads (Fig. 4B).

All these data indicated that high-speed centrifugation might be biasing not only the amount but also the composition of the CR30 dDNA, most likely because of the partial lysis of some members of the archaeal hyperhalophilic community. Centrifugation-induced changes in cell membrane properties and, subsequently, in cell viability, had been previously observed in pure cultures (Allan and Pearce, 1979; Gilbert *et al.*, 1991; Pembrey *et al.*, 1999; Peterson *et al.*, 2012; Nagler *et al.*, 2018). Our data showed that the effect of the 30,000 *xg* centrifugation

applied to remove cells impacted the *Haloquadratum* CR30 assemblage, whose abundance was 15% higher in the CdDNA than in the FdDNA. This phenomenon is probably due to specific cell wall properties and agrees with previous experimental observations which showed that brine dilution, and the consequent osmotic shock, caused a dramatic decrease in the numbers of the *Haloquadratum*-dominated CR30 archaeal community (Santos *et al.*, 2011).

To verify if the centrifugation-induced cell damage was restricted to 30,000 *xg* or if it was independent from the centrifugal forces applied, different centrifugation speeds were tested with a CR30 water sample: 30,000 *xg*, 20,000 *xg*, 9,000 *xg* and 5,000 *xg*. The obtained supernatants were processed as in protocol C, and compared to the F protocol. The amounts of recovered dDNA, as well as the percentages of cells remaining in the supernatants and the viral concentration in the eDNA fraction, were measured. As expected, the highest recovery of dDNA was obtained using 30,000 *xg* and 20,000 *xg* centrifugal forces (Suppl. Fig. 3A). The amount of dDNA using lower centrifugation speeds (9,000 and 5,000 *xg*) was more similar to that obtained with the F protocol, although the replicates' variability was higher (Suppl. Fig. 3A), most likely as a consequence of the cells remaining in their corresponding supernatants (Suppl. Fig. 3B). In addition, the number of virus-like particles (VLPs) in the supernatants from centrifuged samples were higher than in the F protocol (Suppl. Fig. 3C). Thus, centrifugation forces ≥ 20,000 *xg* increased the amount of dDNA due to cell damage, and centrifugations ≤ 9,000 *xg* yielded supernatants with high cell and viral concentrations, whose effects on downstream steps of the protocol cannot be foreseen. Therefore, we suggest the sequential filtration protocol as the most suitable methodology to study the CR30 dDNA, as recently suggested by Linney *et al.* (2021) for marine environments. Nevertheless, since our results have only been tested with water samples from the CR30 crystallizer, it is not possible to asses that the centrifugation effect is the same with different environmental samples harboring different microbial assemblages. Thus, based in our observations, it is recommendable an *ad hoc* optimization protocol for the study of the environmental dissolved DNA.

Characterization of the CR30 dissolved DNA (FdDNA)

As previously mentioned, the concentration of dissolved DNA in the CR30 crystallizer was around 10.3 µg per liter of brine and showed a low molecular weight. The typical bimodal curve of iDNA for the CR30 crystallizer displays a predominant GC peak around 50%, related to *Haloquadratum*, and a lower high-GC peak associated with *S. ruber* and some haloarchaea such as *Halonotius* or *Halorubrum* (Legault *et al.*, 2006; Ghai *et al.*, 2011). However, metagenomic

FdDNA reads presented a GC content distribution with an intermediate profile between this typical bimodal curve and the vDNA profile, which showed the opposite trend (Fig. 5A). This fact initially suggested that the FdDNA should contain DNA fragments from both cellular and viral origin, as confirmed by the taxonomic assignment of metagenomic FdDNA reads (Fig. 4B).

Although most FdDNA reads were assigned to *Haloquadratum*, a significant fraction (20.4% of the 16S rRNA gene reads and 2.5% of total reads) (Suppl. Table 1-2) was affiliated with *Nanohaloarchaeota*, a phylum of uncultured small microbes with small genomes and low-to-intermediate GC values (Ghai *et al.*, 2011; Narasingarao *et al.*, 2012; Martínez García *et al.*, 2014), whose proportion was higher in the FdDNA than in the iDNA fraction (2.8% and 0.7%, respectively; Suppl. Table 1-2). To rule out the possibility that nanohaloarchaeal cells, due to their estimated small size, could have gone through the filters used to remove cells, their absence in the CR30 filtered concentrate used for the extraction of the FdDNA was checked by TEM. Since no cells were observed, the higher proportion of nanohaloarchaeal sequences in the FdDNA compared to the cellular metagenome was confirmed to be natural and not an artifact.

To date, the main functions of dDNA in the aquatic environment are to serve as a source of N and P (Chimileski *et al.*, 2014), and as a vehicle of horizontal gene transfer (HGT) (Ibáñez de Aldecoa *et al.*, 2017). In Chimileski *et al*., 2014, the nutritional role of dDNA was tested in *Hfx. volcanii* using concentrations from 50 to 250 mg/L, which are three to four orders of magnitude higher than the estimated concentration in the CR30 crystallizer (10.3 μg/L) and in the range of those estimated for deep-hypersaline anoxic sediments (338 ng/g – 343 μg/g) (Danovaro *et al.*, 2005; Corinaldesi *et al.*, 2014). In the case of HGT, short DNA fragments were able to transform the non-halophilic bacterium *Acinetobacter baylyi* even at concentrations as low as 10 ng/L (Overballe-Petersen *et al.*, 2013). The fact that most CR30 dissolved DNA originated from *Archaea* implies that bacterial dDNA might be less available for transformation, even though bacteria-to-archaea HGT events clearly occur in nature as previously documented (Méheust *et al.*, 2018). Our observation agrees with the same arguments described by Fuchsman and co-workers, which pointed that horizontal gene transfer from archaea to bacteria are more predominant than bacteria-to-archaea HGT events in extreme systems (Fuchsman *et al.*, 2017). Our analysis also agrees with previous studies showing that archaea-to-archaea HGT has been frequent in the haloarchaeal evolution (Papke *et al.*, 2004; Cuadros-Orellana *et al.*, 2007; Méheust *et al.*, 2018). On the other hand, in solar salterns, where the microbial community is exposed to high levels of UV radiation, dDNA could also act as a UV protector absorbing this harmful light, as we experimentally tested by using dDNA concentrations in the range of those estimated for the CR30 crystallizer (Suppl. Fig. 4). Nevertheless, other molecules different from

the dissolved DNA (i. e., dissolved pigments), could be also contributing to minimize the effects of the UV exposure.

To ascertain if the dissolved DNA was enriched in genes coding for specific functions, compared to the iDNA, a search for specifically overrepresented open reading frames was performed with all the FdDNA ORFs longer than 50 amino acids. They were quantified in the FdDNA, iDNA and vDNA metagenomes and were considered as overrepresented in FdDNA if their abundances were 5 times higher than in the iDNA, and if they were absent in the vDNA pool. With these criteria, 10,467 FdDNA ORFs (0.4%) were found to be overrepresented and were further analyzed. Taxonomically, they were mainly associated to *Candidatus* Nanosalina (17%; Narasingarao *et al.*, 2012), the uncultured archaeon J07HB67 (16%; Podell *et al.*, 2013), *Halorubrum* (10%) and *Haloquadratum* (7%) (Fig. 5B). Functionally, 34.2% were hypothetical proteins, 26.5% were "conserved hypothetical proteins", and 29.8% included low abundant functions (<0.2% each). The remaining functionally assigned ORFs were involved in DNA metabolism (3.8%), cell surface proteins (2.9%) and plasmids (2.2%) (Fig. 5C). To explore the origin of these plasmid-related ORFs, 96 plasmids from sequenced halophilic prokaryotes were recruited against all iDNA and FdDNA metagenomic reads. Only the PL47 plasmid from *Hqr. walsbyi* DSM 16790, the strain isolated from the CR30 pond (Bolhuis *et al.*, 2004, 2006), was fully covered by the metagenomics reads (Suppl. Fig. 5). Then, to ascertain if the plasmid/chromosome ratio was higher in the FdDNA than in the iDNA, the abundances of the whole *Hqr. walsbyi* DSM16790 genome and 5 of its housekeeping genes were calculated and compared to those from the PL47 plasmid (Suppl. Fig. 5). The calculated *Hqr. walsbyi* plasmid/chromosome ratios were, approximately, two times higher in the dissolved DNA pool than in the cellular metagenome. We might hypothesize that this abundance of plasmids in the extracellular environment could be caused by the inherent resistance of plasmids to exonucleases or by an active release as free molecules (Matsui *et al.*, 2003) or through vesicles (Erdmann *et al.*, 2017). Although metagenome-assembled plasmids (MAPs) were not obtained, most likely due to their genetic microdiversity and plasticity (Pfeifer and Ghahraman, 1993; DasSarma and Arora, 1997; Ng *et al.*, 1998; DasSarma *et al.*, 2009), the electrophoresed dDNA revealed the presence of low mobility DNA bands that could correspond to these molecules (Fig. 2B).

Finally, the taxonomic affiliation of the viral reads in the FdDNA (27% of total FdDNA reads, Suppl. Table 3) indicated that a group of fosmid-derived haloviral genomes related to *Hqr. walsbyi* (EHqrV-1 genus; Villamor *et al*., in prep.) was the most abundant among the already described haloviruses, followed by other "environmental halophages" (eHPs; García-Heredia *et*

al., 2012) with unknown hosts, and the nanohaloarchaeal virus 1 (NHV-1), discovered in the CR30 crystallizer by combining single-cell genomics with microarrays (Martínez García *et al.*, 2014). To retrieve information on the stability of the viral capsids, the vDNA/FdDNA normalized ratios were obtained for every published haloviral genome (Table 2). All the EHqrV-1 and eHP viruses were more than twice as abundant in the viral fraction than in the dissolved DNA pool, indicating either that their capsids were very stable in natural conditions or their production from infected cells exceeded their decay. Conversely, the nanohaloarchaeal NHV-1, with a lower vDNA/FdDNA ratio, could have a less stable capsid (Table 2). However, it is important to highlight that other factors could explain the differences in the vDNA/FdDNA ratios, such as the viral infectivity rates or the degradation of viral DNA.

The impact of viruses on the dissolved DNA pool

The higher proportion of nanohaloarchaeal sequences in the FdDNA, compared to the iDNA, could be due to an active release of DNA to the extracellular medium, a high rate of cell lysis (induced by viruses or antimicrobial compounds such as halocins), or a combination of both. To ascertain whether nanohaloarchaeal viruses could act shaping the pool of dissolved DNA, the activity of NHV-1 viruses in the CR30 sample was estimated by retrieving their associated reads from the CR30 metagenome (iDNA). These sequences, representatives from the NHV-1 viruses which were actively replicating inside their hosts, accounted for 0.042% of the iDNA reads (Suppl. Table 4). When this value was normalized against the corresponding genome size, NHV-1 was the most abundant viral population in the iDNA (Table 2), among all previously described haloviruses. Other abundant viruses included viruses eHP-6, eHP-12, eHP-16 and eHP-36 (García-Heredia *et al.*, 2012).

The NHV-1 virus was not only abundant in the iDNA sample, but also in another metagenome from the same pond (obtained in November 2014; Ramos-Barbero et al., in prep.) and in its corresponding metatranscriptome, whose analyses indicated that NHV-1 viruses were actively expressed at a higher proportion than any other reported halovirus (Table 2). Indeed, a viral metatranscriptome obtained from the CR30 crystallizer some years ago (Santos *et al.*, 2011), before the discovery of the *Nanohaloarchaeota*, showed that a virus closely related to NHV-1 was one of the most active viruses in the crystallizer microbial community. It was unexpected that the population represented by the nanohaloarchaeal virus NHV-1 was more active than the viruses associated to *Hqr. walsbyi*, given the low abundance of nanohaloarchaea compared to the “square” archaeon. This fact is even more intriguing if we consider the host

dependency of some *Nanohaloarchaeota* (Hamm *et al.*, 2019) since, if free nanohaloarchaea are not able to divide by themselves, viral replication could also be compromised.

In summary, all these data supported the hypothesis that the activity of nanohaloarchaeal viruses might contribute to shape the taxonomic composition of the CR30 dissolved DNA, and highlights the study of dDNA as a powerful analytic tool to better understand biological processes occurring in the environment.

Concluding remarks

Intracellular DNA and viral DNA have received the most attention by microbial molecular ecologists. Nevertheless, the study of the dissolved DNA fraction can help to shed light on the dynamics and evolution of ecosystems and should be considered as relevant in ecological studies. The isolation of dissolved DNA, however, must be carried out with as much care as possible, starting with the optimization of the protocol for each kind of sample. For hypersaline water samples, filtration has been the most suitable strategy to remove cells, avoiding cell damage and irreproducible outputs. Here, the study of dDNA from a hypersaline environment (the CR30 crystallizer pond in Santa Pola salterns) has revealed that it is composed of both cellular and viral DNA, with proportions which are different from those found in the intracellular or viral DNA pool, and it has provided clues on the activity of *Nanohaloarchaea* viruses. In addition, experimental evidence has supported a new function of dDNA as UV-protector.

Experimental procedures

Study site and sampling

Ten liters of hypersaline water sample were collected from the CR30 crystallizer, located in “Bras del Port” solar salterns (Santa Pola, Alicante, Spain) in May 2019. The salinity was 37%, measured *in situ* with a hand refractometer (Sper Scientific, 300006), and the temperature and pH at the time of sampling were 26.5 ^{o}C and 7.11, respectively. The sample was stored at 4 ^{o}C until being processed.

Dissolved DNA purification

Two different protocols were used to purify the CR30 eDNA (Fig. 1), namely centrifugation followed by filtration ("C" protocol, yielding CdDNA) or solely filtration ("F") (FdDNA). The C protocol, used to remove cells and to access eDNA (Corinaldesi *et al.*, 2005, 2008, 2011; Steinberger and Holden, 2005; Ascher *et al.*, 2009; Lever *et al.*, 2015; Torti *et al.*, 2018), consisted of the centrifugation of 3.5 liters of hypersaline water sample (30000 *xg*, 1 hour, 20 ^{o}C, in a Beckman Avanti J-30I centrifuge with a JA-14 rotor), and the subsequent filtration of the supernatant through 0.22 μm Durapore filters (Millipore®) to remove the remaining cells. The second protocol consisted solely of the sequential filtration of 5 liters of hypersaline water through 2, 0.22 and 0.1 μm filters (Durapore, Millipore®) to eliminate all cells and minimize the possible cell damage occurring during centrifugation.

The water samples harboring the eDNA (containing dissolved and viral DNA) obtained with either protocol were further concentrated to a final volume of 100 ml by tangential filtration (using a Vivaflow 200-PES 30 kDa, Sartorius) and then ultracentrifuged (185000 *xg*, 8 hours, 20 ^{o}C, in a Beckman Coulter® optima^{TM} Max-XP ultracentrifuge with a SW Ti 41 rotor) to separate the viruses (see below) from the supernatants containing the dDNA. Both CdDNA and FdDNA samples were finally desalted and concentrated to a final volume of 15 ml using Amicon Ultra 100 kDa devices (Millipore).

Nucleic acid concentrations in dDNA fractions were quantified using the Qubit 2.0 fluorometer (Invitrogen) and the dsDNA High Sensitivity kit. CdDNA and FdDNA size ranges were determined by gel electrophoresis (1% agarose, TAE 1X, 6 V/cm).

Microbial and viral DNA extraction

A 144 ml sample of the same water used for eDNA extraction was centrifuged (17,000 *xg*, 15 min, room temperature, in a Sorvall Legend Micro 17 centrifuge, Thermo Scientific) and the cell pellet used for the extraction of the intracellular DNA (iDNA, Fig. 1), using the protocol described in Mutlu *et al.*, 2008. To compare the iDNA obtained by centrifugation and filtration through 0.2 μm, 200 mL were filtered and DNA extracted using the DNeasy blood and tissue kit (Qiagen). 16S rRNA gene from both iDNA were partially amplified by PCR using 341F and 805R primers (Herlemann *et al.*, 2011), products were sequenced by Illumina Nextera 2 x 250 bp and analyzed using Qiime (Caporaso *et al.*, 2010).

Viral DNA (vDNA) was extracted from the viral pellets obtained after the ultracentrifugation step (see below) using the protocol described in Santos *et al.* (2010), which included embedding of the viral particles in agarose plugs. Prior to viral DNA purification from agarose plugs, the size range of vDNA was checked by pulsed field gel electrophoresis. The lack

of cellular contamination in the viral pellet was checked by PCR with archaeal 16S rRNA gene primers as previously reported (Ramos-Barbero *et al.*, 2019). All the extracted DNAs were also quantified by Qubit 2.0 (Invitrogen).

Denaturing gradient gel electrophoresis (DGGE)

16S rRNA gene profiles in iDNA, CdDNA and FdDNA were analyzed and compared by DGGE. All DNAs were PCR-amplified with primer pairs 341fGC-907r (for *Bacteria*) and 344fGC-907r (for *Archaea*) (Muyzer *et al.*, 1996; Schäfer *et al.*, 2001) in 50 μl reactions containing: PCR buffer 1X, $MgCl_2$ 1.5 mM, 0.2 mM of each dNTP, 0.2 mM of each primer and 1 U of recombinant *Taq* DNA polymerase (Invitrogen). PCRs were carried out as previously described (Gomariz *et al.*, 2015) including the extension and reconditioning steps.

PCR products were purified (GeneJET PCR Purification Kit, Thermo Scientific) and quantified using a NanoDrop ND-1000 spectrophotometer, and 500 ng from each sample were loaded into a 6% polyacrylamide gel (acrylamide: bis-acrylamide gel stock solution 37.5:1, Bio-Rad), with a 40-70% denaturing gradient (where 100% of denaturants contains 7 M urea and 40% formamide) in 1X TAE buffer (40 mM Tris-HCl, pH 8.0; 20 mM acetic acid; 1 mM EDTA). The electrophoresis was carried out for 16 hours at 60 ^{o}C and 70 V in a D-Code System (Bio-Rad). Finally, the gel was stained with 10X SybrGreen for 15 minutes, washed with milli-Q water (2 x 15 minutes) and visualized in a Typhoon 9410 image analyzer (Amersham Biosciences).

DNA sequencing and metagenomic analyses

Total DNA from CdDNA, FdDNA, iDNA and vDNA samples was sequenced in FISABIO (Valencia, Spain) with the Illumina Nextera Flex kit on a MiSeq 2 x 250 bp run. Metagenomic reads were then filtered by quality, Illumina adapters were removed using Trimmomatic 0.36 (Bolger *et al.*, 2014) (parameters: java -jar trimmomatic-0.36.jar PE -phred33 ILLUMINACLIP:NexteraPE-PE.fa:2:30:10 LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:36), and paired reads were joined and converted to fasta with fq2fa (option: "--merge"). In order to estimate the sequencing depth (or coverage) and the diversity of each metagenome, cleaned reads were analyzed with Nonpareil (Rodríguez-R and Konstantinidis, 2014a; 2016).

The similarity among the obtained metagenomes was estimated using a self-BLASTN comparison. Subsets of 10^6 reads from each metagenome were reciprocally compared and two different data were obtained: i) the averaged percentage of similarity of all reads aligning with > 70% coverage and ii) the percentage of identical sequences shared considering as hits those

reads aligning with 100 % of identity and 100% coverage. A Principal Component Analysis (PCA) was performed using ClustVis (Metsalu and Vilo, 2015) to visualize the differences among metagenomes.

The GC content distribution of reads was obtained with custom Linux commands, and *RNAscan* (Lagesen *et al.*, 2007) was used to extract 16S rRNA gene reads for the analysis of each metagenome's taxonomic composition by Qiime (Caporaso *et al.*, 2010). 16S rRNA gene sequences were clustered into OTUs (operational taxonomic units) using pick_open_reference_otus.py and the aligner UCLUST (Edgar, 2010). The longest sequence of each OTU was selected as the reference and taxonomy was assigned by BLAST against the SILVA_132 database (https://www.arb-silva.de/no_cache/download/archive/qiime/) (Quast *et al.*, 2013). Samples were normalized to the same number of reads for abundance comparison. In addition, Diamond's BLASTx of reads against the nr-NCBI and Megan 6.0 databases (Huson *et al.*, 2007; Sayers *et al.*, 2019) were used to determine the taxonomic affiliation of raw reads using permissive parameters (30% identity and 70% coverage).

Viral sequences in iDNA and FdDNA were identified by BLASTN against the vDNA reads. Sequences aligning with > 70% coverage were selected, extracted, and identified using Diamond's BLASTx of reads against the nr-NCBI database. The abundances of the viral genomes, in recruited metagenomic nucleotides normalized by reference genome and metagenome sizes, were calculated for each metagenome using reads aligning above 95% identity with, at least, 70% coverage, which are the thresholds recommended to quantify viral "species" (Adriaenssens and Rodney Brister, 2017; Gregory *et al.*, 2019). The percentage of genome covered by metagenomic reads was estimated using custom Linux commands.

To determine if the estimated viral activity and abundance in the CR30 sample taken in May 2019 (see below) were common in the crystallizer, the metagenome, metatranscriptome (free of 16S rRNA gene reads) and the metavirome obtained from this pond in November 2014 (Ramos-Barbero *et al*., in prep.) were analyzed by the fragment recruitment strategy as described below.

In order to identify the genes overrepresented in dDNA compared to iDNA, reads were translated using Prodigal (Hyatt *et al.*, 2010) and every open reading frame (ORF) longer than 50 amino acids extracted. The number of reads which hit each ORF in each metagenome (3 million reads per metagenome) was obtained using Diamond BLASTx. The output table was filtered by best hit (allowing multiple hits if identity was identical), coverage > 70%, identity 30 % and e-value < 10^{-5}. ORFs overrepresented more than 5-fold in dDNA than in iDNA and absent from

vDNA were selected. These ORFs were annotated by BLASTp against nr (coverage > 70 % and identity > 30 %).

A custom-made plasmid database from halophilic microorganisms was constructed to identify the more abundant plasmids in dDNA. BLASTn (cut-off parameters: identity > 95 %, coverage > 70 %) was employed to calculate the abundance in bp, normalized by molecule size and metagenome size, of plasmid and host.

Nucleotide sequence accession number

The raw sequences datasets have been deposited within the NCBI under the Bioproject accession number PRJNA713327.

Acknowledgements

The help of the owners and staff of the salterns Bras del Port is greatly appreciated. We thank Karen Neller for her professional English editing. This research was supported by the Spanish Ministry of Science, Innovation and Universities grant MICROMATES (PGC2018-096956-B-C44), which was also supported with European Regional Development Fund (FEDER) funds, and by the Generalitat Valenciana grant PROMETEO/2017/129. B.A.-R. is a FPU-Generalitat Valenciana fellow. The authors state that there is no conflict of interest that could be perceived to bias their work.

References

Adriaenssens, E.M. and Rodney Brister, J. (2017) How to name and classify your phage: An informal guide. *Viruses* **9**: 70.

Allan, I. and Pearce, J.H. (1979) Modulation by centrifugation of cell susceptibility to chlamydial infection. *J Gen Microbiol* **111**: 87–92.

Antón, J., Llobet-Brossa, E., Rodríguez-Valera, F., and Amann, R. (1999) Fluorescence *in situ* hybridization analysis of the prokaryotic community inhabiting crystallizer ponds. *Environ Microbiol* **1**: 517–523.

Antón, J., Oren, A., Benlloch, S., Rodríguez-Valera, F., Amann, R., and Rosselló-Móra, R. (2002) *Salinibacter ruber* gen. nov., sp. nov., a novel, extremely halophilic member of the Bacteria

from saltern crystallizer ponds. *Int J Syst Evol Microbiol* **52**: 485–491.

Ascher, J., Ceccherini, M.T., Pantani, O.L., Agnelli, A., Borgogni, F., Guerri, G., et al. (2009) Sequential extraction and genetic fingerprinting of a forest soil metagenome. *Appl Soil Ecol* **42**: 176–181.

Beebee, T.J.C. (1993) Identification and analysis of nucleic acids in natural freshwaters. *Sci Total Environ* **135**: 123–129.

Benlloch, S., Acinas, S.G., Antón, J., López-López, A., Luz, S.P., and Rodríguez-Valera, F. (2001) Archaeal biodiversity in crystallizer ponds from a solar saltern: Culture versus PCR. *Microb Ecol* **41**: 12–19.

Benlloch, S., López-López, A., Casamayor, E.O., Ovreas, L., Goddard, V., Daae, F.L., et al. (2002) Prokaryotic genetic diversity throughout the salinity gradient of a coastal solar saltern. *Environ Microbiol* **4**: 349–360.

Bolger, A.M., Lohse, M., and Usadel, B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**: 2114–2120.

Bolhuis, H., Palm, P., Wende, A., Falb, M., Rampp, M., Rodriguez-Valera, F., et al. (2006) The genome of the square archaeon *Haloquadratum walsbyi*: life at the limits of water activity. *BMC Genomics* **7**: 169.

Bolhuis, H., Poele, E.M. Te, and Rodríguez-Valera, F. (2004) Isolation and cultivation of Walsby's square archaeon. *Environ Microbiol* **6**: 1287–1291.

Boujelben, I., Yarza, P., Almansa, C., Villamor, J., Maalej, S., Antón, J., and Santos, F. (2012) Virioplankton community structure in Tunisian solar salterns. *Appl Environ Microbiol* **78**: 7429–7437.

Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., et al. (2010) QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* **7**: 335–336.

Chimileski, S, Dolas, K., Naor, A., Gophna, U., and Papke, R.T. (2014) Extracellular DNA metabolism in *Haloferax volcanii*. *Front Microbiol* **5**: 57.

Corinaldesi, C., Barucca, M., Luna, G.M., and Dell'Anno, A. (2011) Preservation, origin and genetic imprint of extracellular DNA in permanently anoxic deep-sea sediments. *Mol Ecol* **20**: 642–654.

Corinaldesi, C., Beolchini, F., and Dell'Anno, A. (2008) Damage and degradation rates of extracellular DNA in marine sediments: Implications for the preservation of gene sequences. *Mol Ecol* **17**: 3939–3951.

Corinaldesi, C., Danovaro, R., and Dell'Anno, A. (2005) Simultaneous recovery of extracellular and intracellular DNA suitable for molecular studies from marine sediments. *Appl Environ Microbiol* **71**: 46–50.

Corinaldesi, C., Tangherlini, M., Luna, G.M., and Dell'anno, A. (2014) Extracellular DNA can preserve the genetic signatures of present and past viral infection events in deep hypersaline anoxic basins. *Proceedings Biol Sci* **281**: 20133299.

Cuadros-Orellana, S., Martín-Cuadrado, A.B., Legault, B., D'Auria, G., Zhaxybayeva, O., Papke, R.T., and Rodríguez-Valera, F. (2007) Genomic plasticity in prokaryotes: The case of the square haloarchaeon. *ISME J* **1**: 235–245.

Danovaro, R., Corinaldesi, C., Dell'Anno, A., Fabiano, M., and Corselli, C. (2005) Viruses, prokaryotes and DNA in the sediments of a deep-hypersaline anoxic basin (DHAB) of the Mediterranean Sea. *Environ Microbiol* **7**: 586–592.

DasSarma, S. and Arora, P. (1997) Genetic analysis of the gas vesicle gene cluster in haloarchaea. *FEMS Microbiol Lett* **153**: 1–10.

DasSarma, S., Capes, M., and DasSarma, P. (2009) Haloarchaeal Megaplasmids. In, Schwartz,E. (ed), *Microbial Megaplasmids. Microbiology Monographs, vol 11*. Berlin: Springer, pp. 3–30.

Deflaun, M.F., Paul, J.H., and Davis, D. (1986) Simplified method for dissolved DNA determination in aquatic environments. *Appl Environ Microbiol* **52**: 654–659.

DeFlaun, M.F., Paul, J.H., and Jeffrey, W.H. (1987) Distribution and molecular weight of dissolved DNA in subtropical estuarine and oceanic environments. *Mar Ecol Prog Ser* **38**: 65–73.

Dell'Anno, A. and Corinaldesi, C. (2004) Degradation and turnover of extracellular DNA in marine sediments : Ecological and methodological considerations. *Appl Environ Microbiol* **70**: 4384–4386.

Edgar, R.C. (2010) Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**: 2460–2461.

Erdmann, S., Tschitschko, B., Zhong, L., Raftery, M.J., and Cavicchioli, R. (2017) A plasmid from

an Antarctic haloarchaeon uses specialized membrane vesicles to disseminate and infect plasmid-free cells. *Nat Microbiol* **2**: 1446–1455.

Fernández, A.B., Vera-Gargallo, B., Sánchez-Porro, C., Ghai, R., Papke, R.T., Rodríguez-Valera, F., and Ventosa, A. (2014) Comparison of prokaryotic community structure from Mediterranean and Atlantic saltern concentrator ponds by a metagenomic approach. *Front Microbiol* **5**: 196.

Fuchsman, C.A., Collins, R.E., Rocap, G., and Brazelton, W.J. (2017) Effect of the environment on horizontal gene transfer between bacteria and archaea. *PeerJ* **5**: e3865.

García-Heredia, I., Martín-Cuadrado, A.-B., Mojica, F.J.M., Santos, F., Mira, A., Antón, J., and Rodríguez-Valera, F. (2012) Reconstructing viral genomes from the environment using fosmid clones: the case of haloviruses. *PLoS One* **7**: e33802.

Ghai, R., Pasic, L., Fernández, A.B., Martín-Cuadrado, A.-B., Mizuno, C.M., McMahon, K.D., et al. (2011) New abundant microbial groups in aquatic hypersaline environments. *Sci Rep* **1**: 135.

Gilbert, P., Caplan, F., and Brown, M.R.W. (1991) Centrifugation injury of Gram-negative bacteria. *J Antimicrob Chemother* **27**: 550–551.

Gomariz, M., Martínez-García, M., Santos, F., Rodríguez, F., Capella-Gutiérrez, S., Gabaldón, T., et al. (2015) From community approaches to single-cell genomics: the discovery of ubiquitous hyperhalophilic Bacteroidetes generalists. *ISME J* **9**: 16–31.

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.

Guixa-Boixareu, N., Calderón-Paz, J.I., Heldal, M., Bratbak, G., and Pedrós-Alió, C. (1996) Viral lysis and bacterivory as prokaryotic loss factors along a salinity gradient. *Aquat Microb Ecol* **11**: 215–227.

Hamm, J.N., Erdmann, S., Eloe-Fadrosh, E.A., Angeloni, A., Zhong, L., Brownlee, C., et al. (2019) Unexpected host dependency of Antarctic Nanohaloarchaeota. *Proc Natl Acad Sci* **116**: 14661–14670.

Herlemann, D.P.R., Labrenz, M., Jürgens, K., Bertilsson, S., Waniek, J.J., and Andersson, A.F. (2011) Transitions in bacterial communities along the 2000 km salinity gradient of the Baltic Sea. *ISME J* **5**: 1571–1579.

Huson, D.H., Auch, A.F., Qi, J., and Schuster, S.C. (2007) MEGAN analysis of metagenomic data. *Genome Res* **17**: 377–386.

Hyatt, D., Chen, G.-L., LoCascio, P.F., Land, M.L., Larimer, F.W., and Hauser, L.J. (2010) Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* **11**: 119.

Ibáñez de Aldecoa, A.L., Zafra, O., and González-Pastor, J.E. (2017) Mechanisms and regulation of extracellular DNA release and its biological roles in microbial communities. *Front Microbiol* **8**: 1390.

Lagesen, K., Hallin, P., Rødland, E.A., Stærfeldt, H.-H., Rognes, T., and Ussery, D.W. (2007) RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res* **35**: 3100–3108.

Legault, B. a, López-López, A., Alba-Casado, J.C., Doolittle, W.F., Bolhuis, H., Rodríguez-Valera, F., and Papke, R.T. (2006) Environmental genomics of "Haloquadratum walsbyi" in a saltern crystallizer indicates a large pool of accessory genes in an otherwise coherent species. *BMC Genomics* **7**: 171.

Lever, M.A., Torti, A., Eickenbusch, P., Michaud, A.B., Santl-Temkiv, T., and Jorgensen, B.B. (2015) A modular method for the extraction of DNA and RNA, and the separation of DNA pools from diverse environmental sample types. *Front Microbiol* **6**: UNSP 476.

Linney, M.D., Schvarcz, C.R., Steward, G.F., DeLong, E.F., and Karl, D.M. (2021) A method for characterizing dissolved DNA and its application to the North Pacific Subtropical Gyre. *Limnol Oceanogr Methods*.

Marguet, E. and Forterre, P. (1994) DNA stability at temperatures typical for hyperthermophiles. *Nucleic Acids Res* **22**: 1681–1686.

Martínez García, M., Santos Sánchez, F., Moreno-Paz, M., Parro, V., and Antón Botella, J. (2014) Unveiling viral–host interactions within the 'microbial dark matter.' *Nat Commun* **5**: 4542.

Matsui, K., Ishii, N., and Kawabata, Z. (2003) Release of extracellular transformable plasmid DNA from *Escherichia coli* cocultivated with algae. *Appl Environ Microbiol* **69**: 2399–2404.

Maturrano, L., Santos, F., Rosselló-Mora, R., and Antón, J. (2006) Microbial diversity in Maras salterns, a hypersaline environment in the Peruvian Andes. *Appl Environ Microbiol* **72**: 3887–3895.

Méheust, R., Watson, A.K., Lapointe, F.J., Papke, R.T., Lopez, P., and Bapteste, E. (2018)

Hundreds of novel composite genes and chimeric genes with bacterial origins contributed to haloarchaeal evolution. *Genome Biol* **19**: 75.

Metsalu, T. and Vilo, J. (2015) ClustVis: a web tool for visualizing clustering of multivariate data using Principal Component Analysis and heatmap. *Nucleic Acids Res* **43**: W566–W570.

Mongodin, E.F., Nelson, K.E., Daugherty, S., Deboy, R.T., Wister, J., Khouri, H., et al. (2005) The genome of *Salinibacter ruber*: convergence and gene exchange among hyperhalophilic bacteria and archaea. *Proc Natl Acad Sci U S A* **102**: 18147–18152.

Mutlu, M.B., Martínez-García, M., Santos, F., Peña, A., Guven, K., and Antón, J. (2008) Prokaryotic diversity in Tuz Lake, a hypersaline environment in Inland Turkey. *FEMS Microbiol Ecol* **65**: 474–483.

Muyzer, G., Hottenträger, S., and Teske A, W.C. (1996) Denaturing gradient gel electrophoresis of PCR-amplified 16S rDNA-a new molecular approach to analyse the genetic diversity of mixed microbial communities. In, Akkermans,A., van Elsas,J., and de Bruijn,F. (eds), *Molecular Microbial Ecology Manual*. Dordrecht, Netherland: Kluwer, pp. 1–23.

Nagler, M., Podmirseg, S.M., Griffith, G.W., Insam, H., and Ascher-Jenull, J. (2018) The use of extracellular DNA as a proxy for specific microbial activity. *Appl Microbiol Biotechnol* **102**: 2885–2898.

Narasingarao, P., Podell, S., Ugalde, J.A., Brochier-Armanet, C., Emerson, J.B., Brocks, J.J., et al. (2012) De novo metagenomic assembly reveals abundant novel major lineage of Archaea in hypersaline microbial communities. *ISME J* **6**: 81–93.

Ng, W. V., Ciufo, S.A., Smith, T.M., Bumgarner, R.E., Dale Baskin, J.F., Hall, B., et al. (1998) Snapshot of a large dynamic replicon in a halophilic archaeon: megaplasmid or minichromosome? *Genome Res* **8**: 1131–1141.

Nielsen, K.M., Johnsen, P.J., Bensasson, D., and Daffonchio, D. (2007) Release and persistence of extracellular DNA in the environment. *Environ Biosafety Res* **6**: 37–53.

Oh, D., Porter, K., Russ, B., Burns, D.G., and Dyall-Smith, M. (2010) Diversity of *Haloquadratum* and other haloarchaea in three, geographically distant, Australian saltern crystallizer ponds. *Extremophiles* **14**: 161–169.

Overballe-Petersen, S., Harms, K., Orlando, L.A.A., Mayar, J.V.M., Rasmussen, S., Dahl, T.W., et al. (2013) Bacterial natural transformation by highly fragmented and damaged DNA. *Proc Natl Acad Sci* **110**: 19860–19865.

Papke, R.T., Douady, C.J., Doolittle, W.F., and Rodríguez-Valera, F. (2003) Diversity of bacteriorhodopsins in different hypersaline waters from a single Spanish saltern. *Environ Microbiol* **5**: 1039–1045.

Papke, R.T., Koenig, J.E., Rodríguez-Valera, F., and Doolittle, W.F. (2004) Frequent recombination in a saltern population of *Halorubrum*. *Science (80)* **306**: 1928–1929.

Paul, J.H., Jeffrey, W.H., David, A.W., Deflaun, M.F., and Cazares, L.H. (1989) Turnover of extracellular DNA in eutrophic and oligotrophic freshwater environments of southwest Florida. *Appl Environ Microbiol* **55**: 1823–1828.

Paul, J.H., Jiang, S.C., and Rose, J.B. (1991) Concentration of viruses and dissolved DNA from aquatic environments by vortex flow filtration. *Appl Environ Microbiol* **57**: 2197–204.

Pembrey, R.S., Marshall, K.C., and Schneider, R.P. (1999) Cell surface analysis techniques: What do cell preparation protocols do to cell surface properties? *Appl Environ Microbiol* **65**: 2877–2894.

Peña, A., Teeling, H., Huerta-Cepas, J., Santos, F., Yarza, P., Brito-Echeverría, J., et al. (2010) Fine-scale evolution: genomic, phenotypic and ecological differentiation in two coexisting *Salinibacter ruber* strains. *ISME J* **4**: 882–895.

Peterson, B.W., Sharma, P.K., van der Mei, H.C., and Busscher, H.J. (2012) Bacterial cell surface damage due to centrifugal compaction. *Appl Environ Microbiol* **1**: 120–125.

Pfeifer, F. and Ghahraman, P. (1993) Plasmid pHH1 of *Halobacterium salinarium*: characterization of the replicon region, the gas vesicle gene cluster and insertion elements. *Mol Gen Genet* **238**: 193–200.

Podell, S., Ugalde, J.A., Narasingarao, P., Banfield, J.F., Heidelberg, K.B., and Allen, E.E. (2013) Assembly-driven community genomics of a hypersaline microbial ecosystem. *PLoS One* **8**: e61692.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., et al. (2013) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* **41**: D590-6.

Ramos-Barbero, M.D., Martínez, J.M., Almansa, C., Rodríguez, N., Villamor, J., Gomariz, M., et al. (2019) Prokaryotic and viral community structure in the singular chaotropic salt lake Salar de Uyuni. *Environ Microbiol* **21**: 2029–2042.

Rodriguez-R, Luis M and Konstantinidis, K.T. (2014) Estimating coverage in metagenomic data sets and why it matters. *ISME J* **8**: 2349–2351.

Rodriguez-R, Luis M. and Konstantinidis, K.T. (2014) Nonpareil: a redundancy-based approach to assess the level of coverage in metagenomic datasets. *Bioinformatics* **30**: 629–635.

Rodríguez-R, L.M. and Konstantinidis, K.T. (2016) The enveomics collection : a toolbox for specialized analyses of microbial genomes and metagenomes. *Peer J Prepr* **4:** e1900v1**.**

Santos, F., Moreno-Paz, M., Meseguer, I., López, C., Rosselló-Móra, R., Parro, V., and Antón, J. (2011) Metatranscriptomic analysis of extremely halophilic viral communities. *ISME J* **5**: 1621–1633.

Santos, F., Yarza, P., Parro, V., Briones, C., and Antón, J. (2010) The metavirome of a hypersaline environment. *Environ Microbiol* **12**: 2965–2976.

Sayers, E.W., Agarwala, R., Bolton, E.E., Brister, J.R., Canese, K., Clark, K., et al. (2019) Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res* **47**: D23–D28.

Schäfer, H., Bernard, L., Courties, C., Lebaron, P., Servais, P., Pukall, R., et al. (2001) Microbial community dynamics in Mediterranean nutrient-enriched seawater mesocosms: changes in the genetic diversity of bacterial populations. *FEMS Microbiol Ecol* **34**: 243–253.

Steinberger, R.E. and Holden, P.A. (2005) Extracellular DNA in single- and multiple-species unsaturated biofilms. *Appl Environ Microbiol* **71**: 5404–5410.

Tehei, M., Franzetti, B., Maurel, M.-C., Vergne, J., Hountondji, C., and Zaccai, G. (2002) The search for traces of life: the protective effect of salt on biological macromolecules. *Extremophiles* **6**: 427–430.

Torti, A., Jørgensen, B.B., and Lever, M.A. (2018) Preservation of microbial DNA in marine sediments: insights from extracellular DNA pools. *Environ Microbiol* **20**: 4526–4542.

Torti, A., Lever, M.A., and Jorgensen, B.B. (2015) Origin, dynamics, and implications of extracellular DNA pools in marine sediments. *Mar Genomics* **24 Pt 3**: 185–196.

Vavourakis, C.D., Ghai, R., Rodriguez-Valera, F., Sorokin, D.Y., Tringe, S.G., Hugenholtz, P., and Muyzer, G. (2016) Metagenomic insights into the uncultured diversity and physiology of microbes in four hypersaline soda lake brines. *Front Microbiol* **7**: 211.

Ventosa, A., Fernández, A.B., León, M.J., Sánchez-Porro, C., and Rodríguez-Valera, F. (2014) The

Santa Pola saltern as a model for studying the microbiota of hypersaline environments. *Extremophiles* **18**: 811–824.

Ventosa, A., de la Haba, R.R., Sánchez-Porro, C., and Papke, R.T. (2015) Microbial diversity of hypersaline environments: A metagenomic approach. *Curr Opin Microbiol* **25**: 80–87.

Villamor, J., Ramos-Barbero, M.D., González-Torres, P., Gabaldón, T., Rosselló-Móra, R., Meseguer, I., et al. (2017) Characterization of ecologically diverse viruses infecting co-occurring strains of cosmopolitan hyperhalophilic *Bacteroidetes*. *ISME J* **12**: 424–437.

Viver, T., Orellana, L., González-Torres, P., Díaz, S., Urdiain, M., Farías, M.E., et al. (2018) Genomic comparison between members of the *Salinibacteraceae* family, and description of a new species of *Salinibacter* (*Salinibacter altiplanensis* sp. nov.) isolated from high altitude hypersaline environments of the Argentinian Altiplano. *Syst Appl Microbiol* **41**: 198–212.

Zhaxybayeva, O., Stepanauskas, R., Mohan, N.R., and Papke, R.T. (2013) Cell sorting analysis of geographically separated hypersaline environments. *Extremophiles* **17**: 265–275.

FIGURE LEGENDS

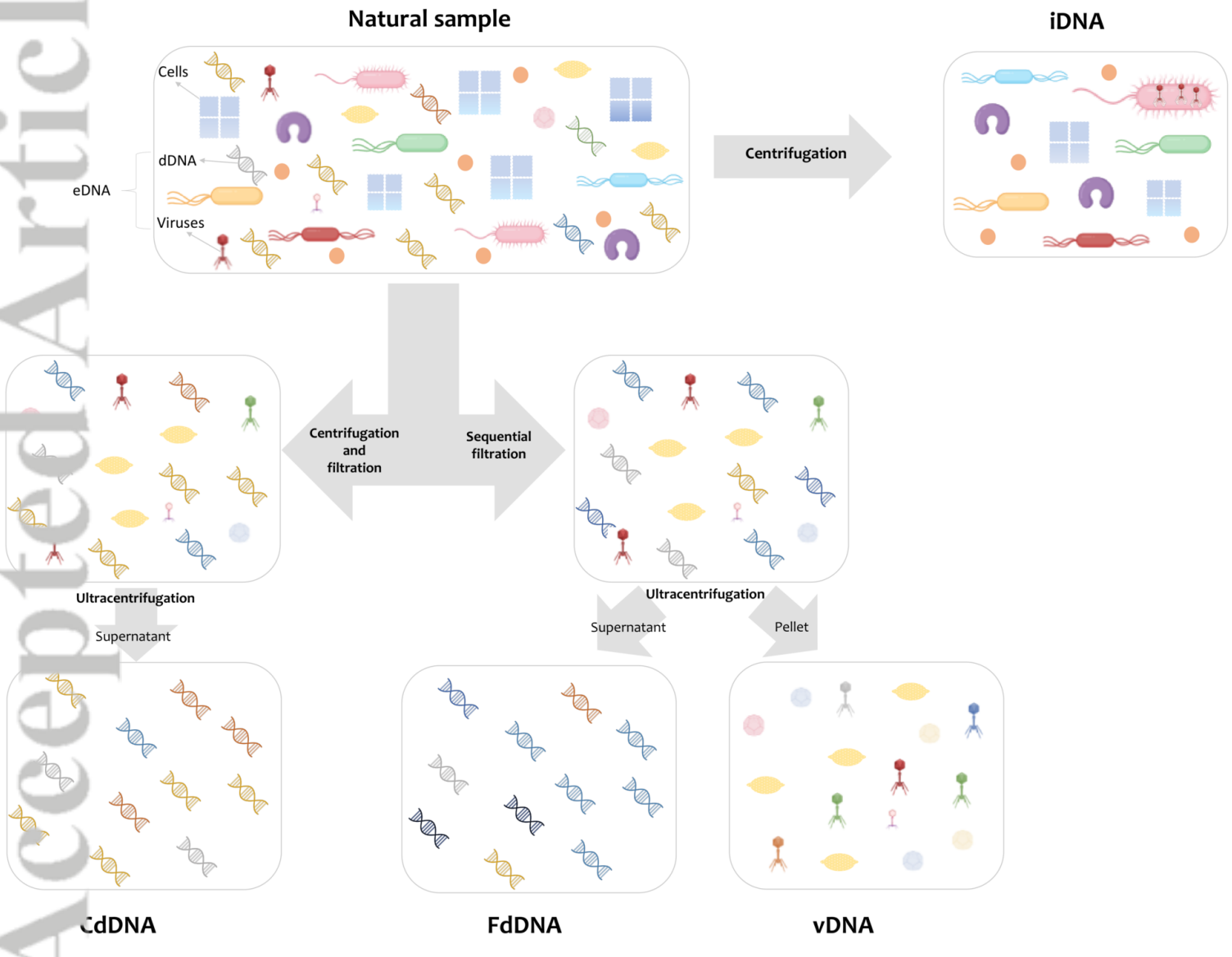
Figure 1. Schematic representation of the workflow used to purify the different DNA fractions from the CR30 crystallizer. eDNA: extracellular DNA; dDNA: dissolved DNA; iDNA: intracellular DNA (from a cell pellet); CdDNA: dissolved DNA obtained using the centrifugation protocol; FdDNA: dissolved DNA obtained using the filtration protocol; vDNA: viral DNA.

Figure 2. Main features of the CR30 dissolved DNA (dDNA). A) Concentrations of CR30 dDNAs DNAs according to the extraction protocol used: sequential filtration (F sample) and centrifugation + filtration (C sample). B) Electrophoresis of the dissolved DNA fractions. Arrows indicate the presence of high molecular weight (>20 Kb) DNA bands.

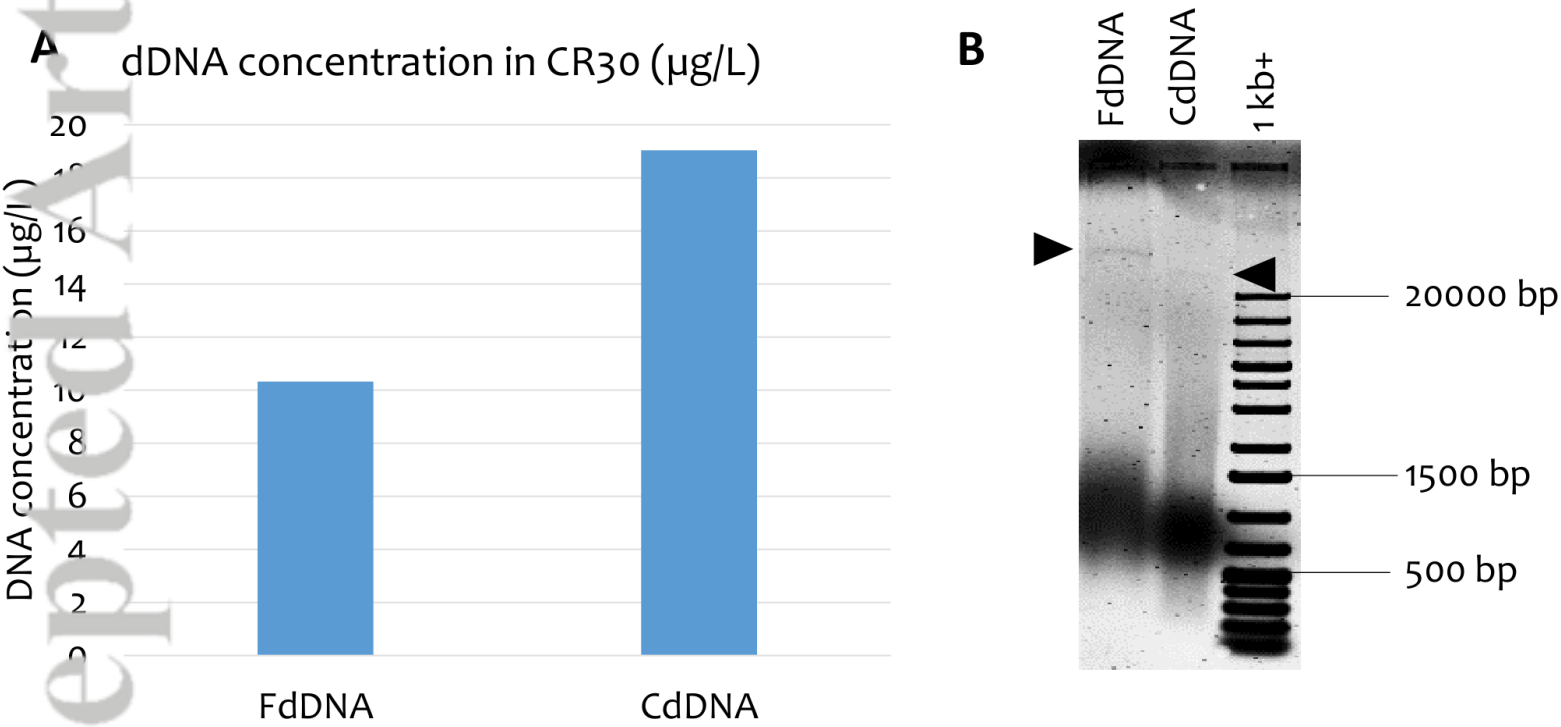
Figure 3. BLASTN-based comparison of all metagenomes obtained in this study. A) PCA of shared sequences (alignment coverage ≥ 70%) among the metagenomes. The percentage of shared reads is indicated in bold, while parentheses indicate the averaged identity of shared sequences. B) PCA of identical sequences (alignment coverage 100%, identity 100%) among metagenomes. Numbers indicate the percentage of identical reads.

Figure 4. Detected genera in the cellular metagenome (iDNA) and dissolved DNA fractions (FdDNA and CdDNA) based on 16S rRNA gene reads (A) or on the affiliation of total reads (B). In pink, *Halobacterota*; *Halobacterales*; *Haloferacaceae*. In red, *Halobacterota*; *Halobacterales*, *Halomicrobiaceae*. In green, *Nanohaloarchaeota*. 1: *Haloquadratum*; 2: *Halonotius*; 3: *Halorubrum*; 4: *Haloparvum*; 5: *Halopiger*; 6: *Halobellus*; 7: Uncultured *Haloferacaceae*; 8: *Halomicrobium*; 9: *Halomicroarcula*; 10: *Salinibacter*; 11: *Candidatus* Nanosalina; 12: *Candidatus* Nanosalinarum; 13: Nanohaloarchaeon SG9; 14: Unclassified *Nanohaloarchaeota*; 15: Others (< 0.5% each); 16: Unclassified; 17: *Haloferax*; 18: *Haloarcula*; 19: Unclassified *Halobacterales*; 20: Viruses

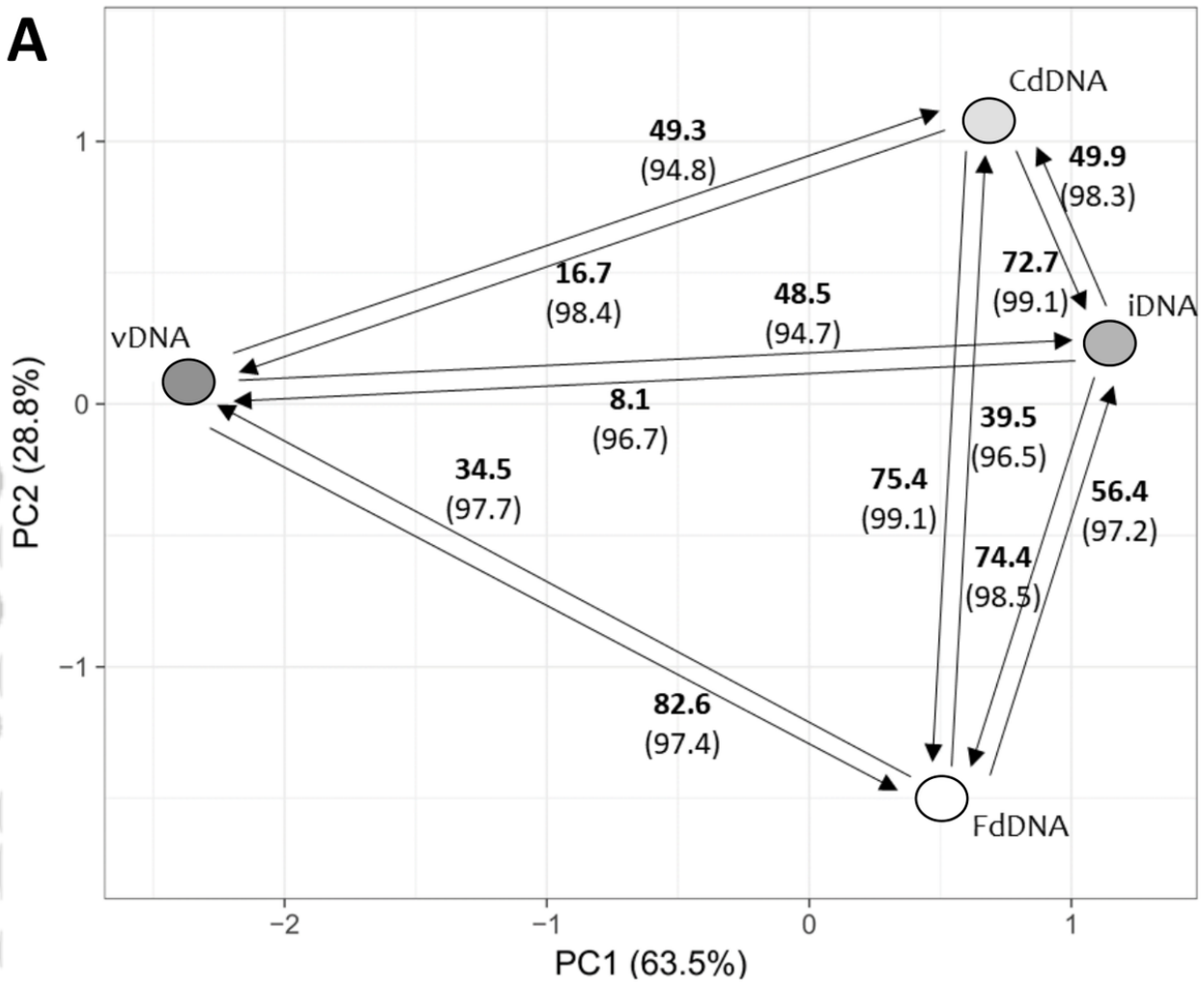
Figure 5. Characterization of FdDNA. A) GC content distribution of reads in the CR30 cellular metagenome (iDNA), metavirome (vDNA) and dissolved DNA fraction (FdDNA). B) Taxonomic affiliation of overrepresented FdDNA ORFs. C) Abundance of ORFs overrepresented in FdDNA compared to iDNA.

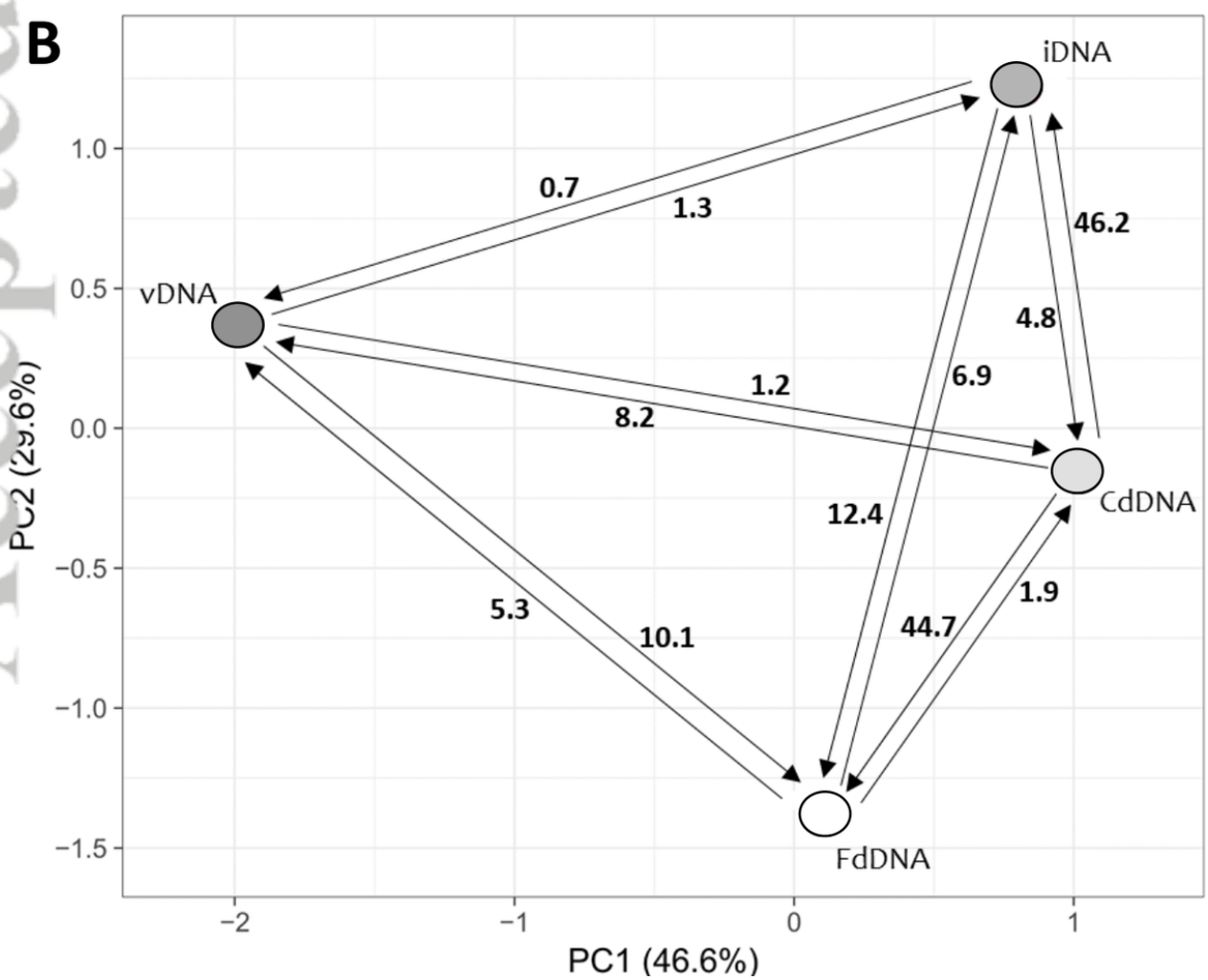


EMI_15510_Aldeguer-Riquelme_et_al._March2021_Figure1.tiff

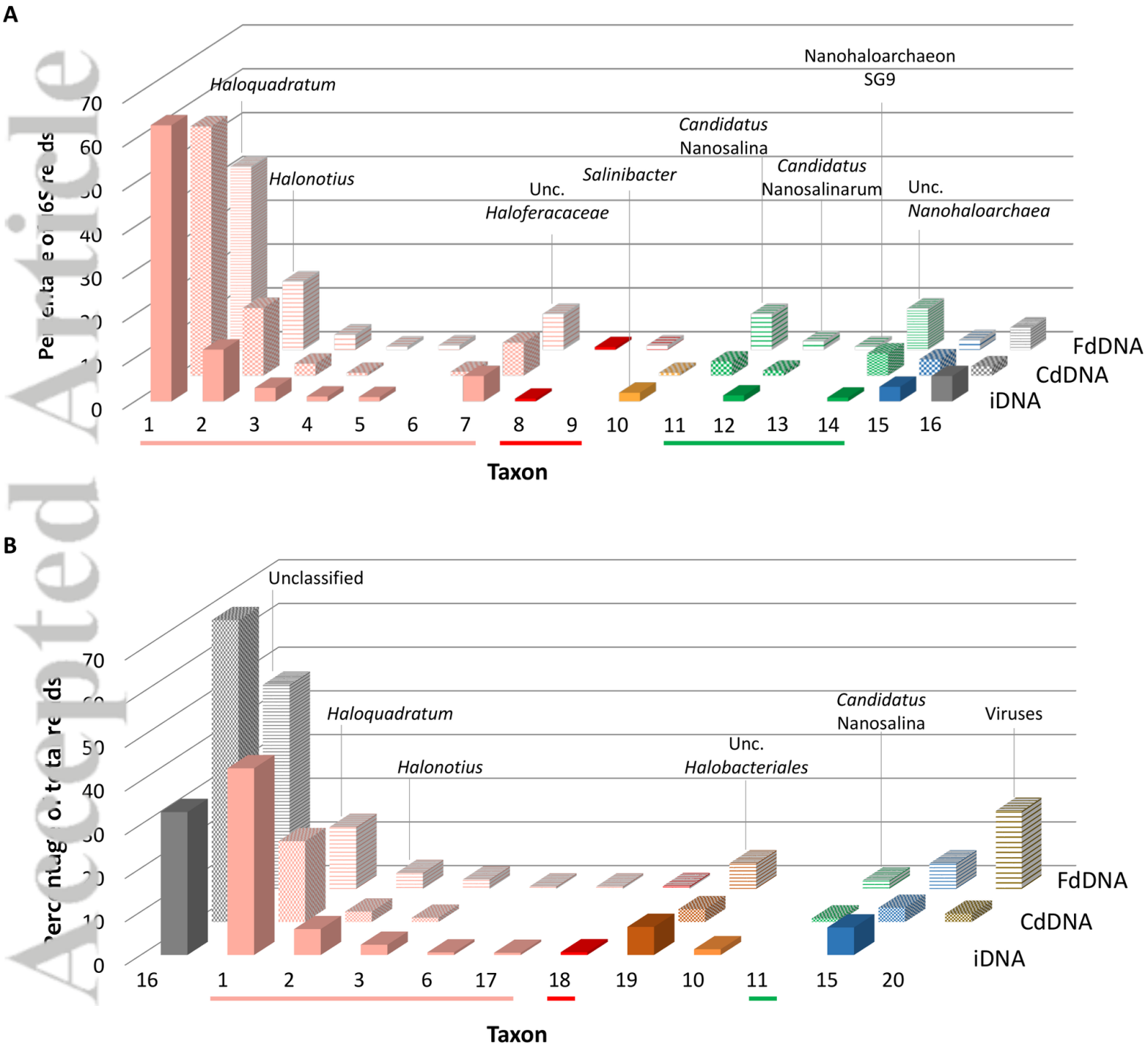


EMI_15510_Aldeguer-Riquelme_et_al._March2021_Figure2.tiff

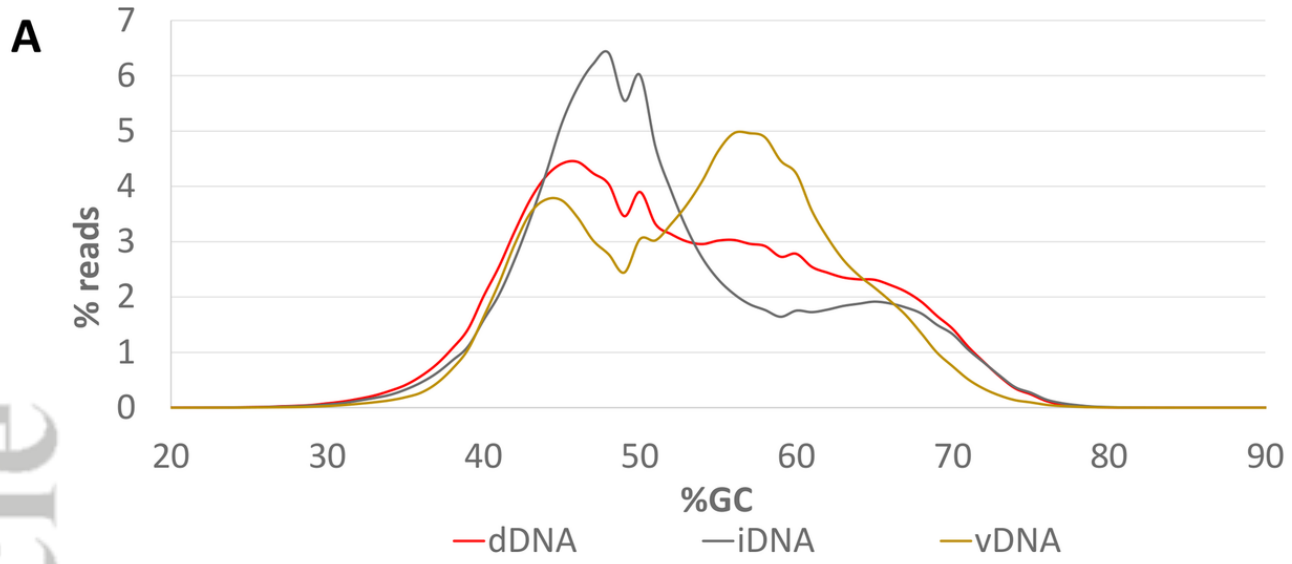


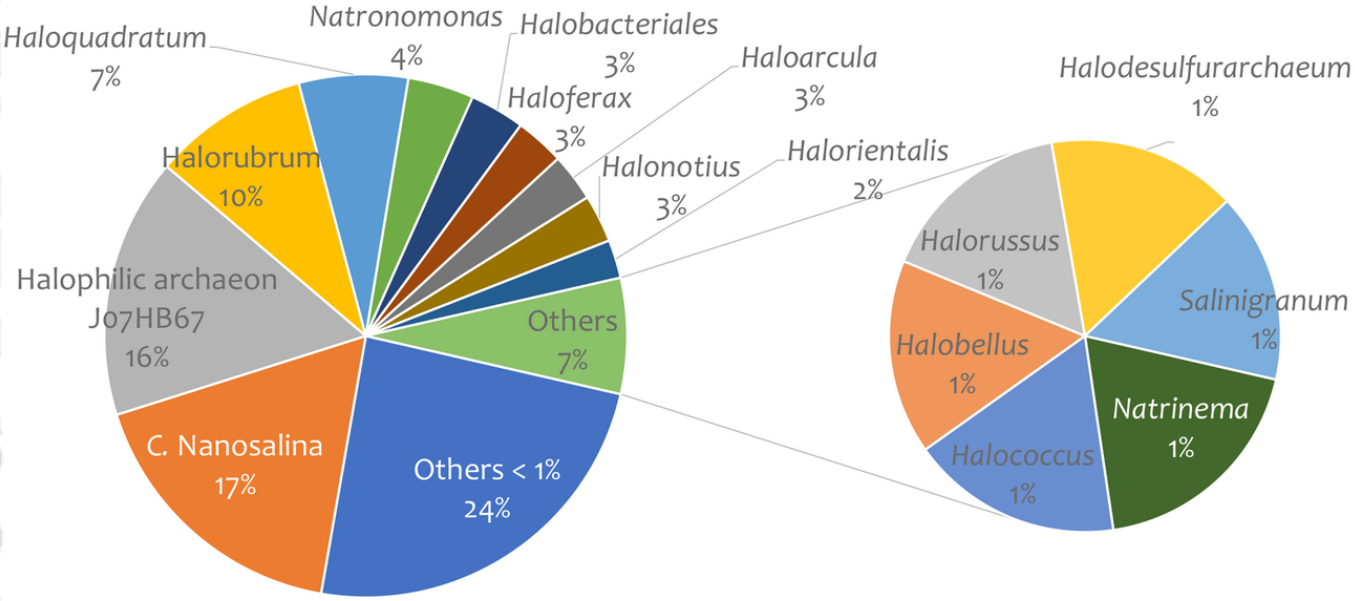


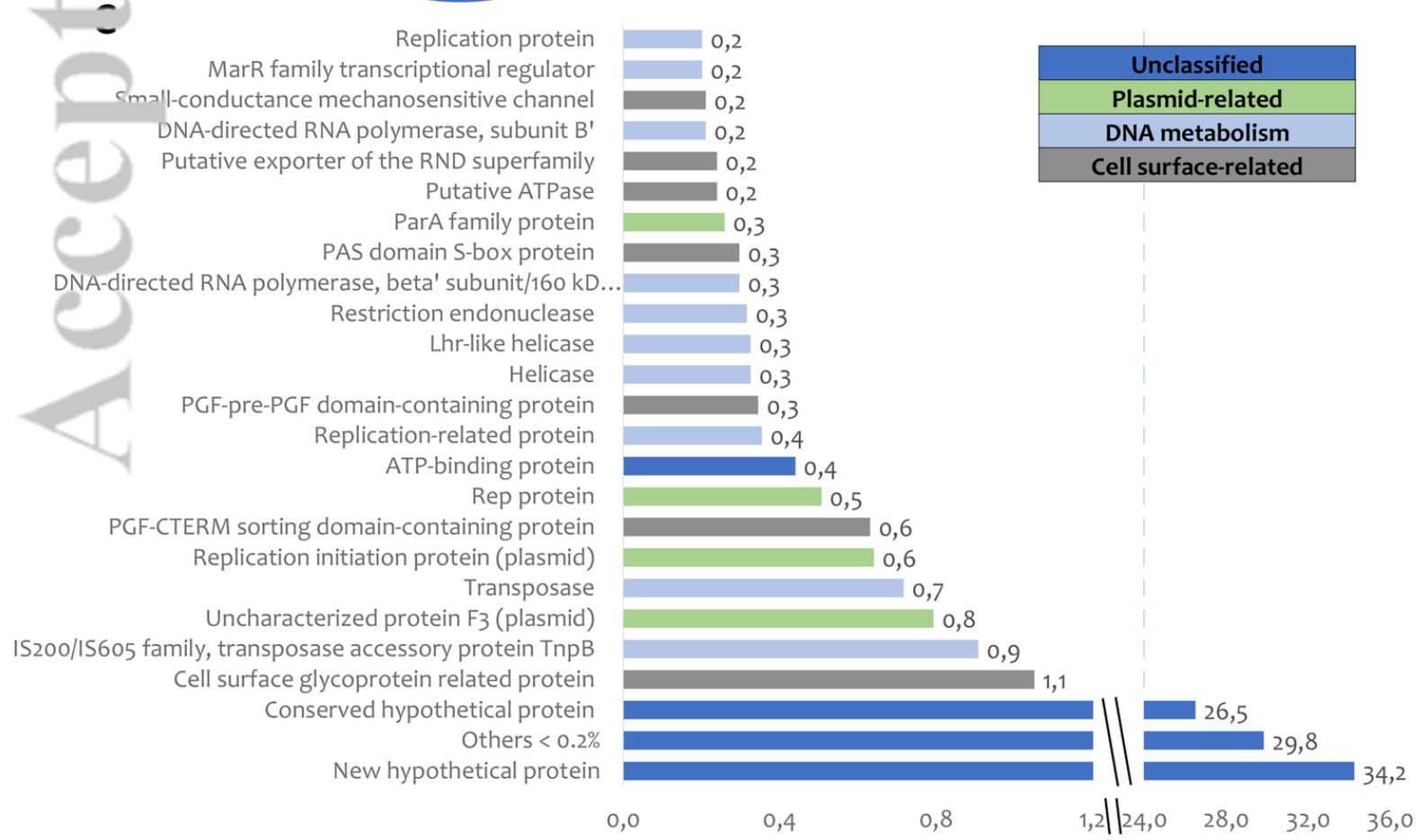
EMI_15510_Aldeguer-Riquelme_et_al._March2021_Figure3.tiff



EMI_15510_Aldeguer-Riquelme_et_al._March2021_Figure4.tiff







EMI_15510_Aldeguer-Riquelme_et_al._March2021_Figure5.tiff

Table 1. General characteristics of the metagenomes analysed in this study.

Metagenome ID	Sample	bp	Reads	%GC	Nonpareil
iDNA	CR30 metagenome May 2019	2.12E+09	8776324	52.33	95.00
FdDNA	CR30 dDNAFilt metagenome May 2019	8.49E+08	3461294	53.11	88.70
CdDNA	CR30 dDNACent metagenome May 2019	7.06E+08	6745542	52.15	92.98
vDNA	CR30 metavirome May 2019	1.90E+09	7664182	53.93	95.68
iDNA_CR30Nov14	CR30 metagenome Nov. 2014	4.36E+09	14993022	55.2	61.56
MT_CR30Nov14	CR30 metatranscriptome Nov. 2014	2.52E+08	1856815	56.38	99.68
vDNA_CR30Nov14	CR30 metavirome Nov.2014	6.54E+08	3084906	51.65	94.88

Table 2. Quantification of the most abundant described haloviruses in different metagenomes from two CR30 samples (May 2019 and November 2014). Metatranscriptomic data from November 2014 are also shown. Abundances are represented in bold, as recruited nucleotides normalized by both metagenome and reference genome sizes. The percentage of each recruited genome (coverage) is provided in parentheses.

	CR30 May 2019				CR30 November 2014				
Genome	**iDNA**	**FdDNA**	**vDNA**	**ratio vDNA/FdDNA**	**iDNA**	**MT**	**vDNA**	**Host**	**Reference**
Hqr. walsbyi DSM 16790	**71.4** (98.8)	**24.7** (96.9)	**0.3** (39.4)	-	**51.1** (97.5)	**179** (13.6)	**0.7** (26.2)	-	Bolhuis et al., 2006
eHP-2	**6.9** (89.9)	**297.4** (96.9)	**670.4** (97)	2,3	**8.7** (95.4)	**0.7** (6)	**2382.2** (97)	*Hqr. walsby*	García-Heredia et al., 2012
eHP-5	**7.4** (94.1)	**314.6** (99.3)	**728.5** (100)	2,3	**9.4** (95.3)	**1** (6.6)	**2375.1** (100)	*Hqr. walsby*	García-Heredia et al., 2012
eHP-22	**6.5** (91.8)	**276.9** (99.8)	**633.9** (100)	2,3	**8.9** (98.1)	**1** (5.5)	**2268.5** (100)	*Hqr. walsby*	García-Heredia et al., 2012
eHP-37	**8** (97.6)	**342.1** (99.7)	**817.9** (99.9)	2,4	**11.4** (98.8)	**1** (7.3)	**2622.8** (99.8)	*Hqr. walsby*	García-Heredia et al., 2012
Nanohaloarchaea AB578	**3.7** (39)	**2.2** (31.1)	**1.3** (9.6)	-	**3.7** (56.9)	**3.3** (1.8)	**6.7** (17.6)	-	Martínez-García et al., 2014
VNH-1	**14.2** (100)	**52.6** (100)	**94.8** (100)	1,8	**15.3** (100)	**2.1** (4.1)	**132.2** (100)	Nanohaloarchaea AB578	Martínez-García et al., 2014
eHP-6	**1.9** (68.6)	**70.9** (97)	**210.8** (99.8)	3,0	**1.6** (82.2)	**0.2** (1.3)	**275.9** (100)	Unknown	García-Heredia et al., 2012
eHP-12	**1.7** (70.6)	**62.3** (97.5)	**185.3** (99)	3,0	**1.6** (77.9)	**0.4** (1.9)	**249.2** (98.7)	Unknown	García-Heredia et al., 2012
eHP-16	**1.4** (64.7)	**52.6** (95.5)	**159.3** (97.5)	3,0	**1.4** (75)	**0.3** (1.1)	**214.6** (98.4)	Unknown	García-Heredia et al., 2012
eHP-36	**1.4** (70)	**57.8** (97.5)	**188.9** (99.9)	3,3	**1.5** (84.1)	**1.9** (0.7)	**247.4** (99.5)	Unknown	García-Heredia et al., 2012