

Patterns and ecological drivers of ocean viral communities

Jennifer Brum, J Cesar Ignacio-Espinoza, Simon Roux, Guilhem Doulcier, Silvia G. Acinas, Adriana A. Alberti, Samuel Chaffron, Corinne Cruaud, Colomban de Vargas, Josep Gasol, et al.

▶ To cite this version:

Jennifer Brum, J Cesar Ignacio-Espinoza, Simon Roux, Guilhem Doulcier, Silvia G. Acinas, et al.. Patterns and ecological drivers of ocean viral communities. Science, 2015, 348 (6237), pp.1261498. 10.1126/science.1261498 . hal-01768331

HAL Id: hal-01768331 https://amu.hal.science/hal-01768331

Submitted on 8 Jan 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés. 1

Title: Patterns and ecological drivers of ocean viral communities

2

Authors: Jennifer R. Brum^{§,1}, J. Cesar Ignacio-Espinoza^{§,2}, Simon Roux^{§,1}, Guilhem Doulcier^{1,3}, Silvia G. Acinas⁴, Adriana Alberti⁵, Samuel Chaffron^{6,7,8}, Corinne Cruaud⁵, Colomban de Vargas^{9,10}, Josep M. Gasol⁴, Gabriel Gorsky^{11,12}, Ann C. Gregory¹³, Lionel Guidi^{11,12}, Pascal Hingamp¹⁴, Daniele Iudicone¹⁵, Fabrice Not^{9,10}, Hiroyuki Ogata¹⁶, Stephane Pesant^{17,18}, Bonnie T. Poulos¹, Sarah M. Schwenck¹, Sabrina Speich^{19,†}, Celine Dimier^{9,10,20}, Stefanie Kandels-Lewis^{21,22}, Marc Picheral^{11,12}, Sarah Searson^{11,12}, *Tara* Oceans Coordinators[‡], Peer Bork^{21,23}, Chris Bowler²⁰, Shinichi Sunagawa²¹, Patrick Wincker^{5,24,25}, Eric Karsenti^{20,22,*}, & Matthew B. Sullivan^{1,2,13,*} 3 4 5 6 7 8 9 10 11 [§]co-first authors 12

13

14 Affiliations:

- 15 ¹ Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, Arizona,
- 85721. USA 16
- ² Department of Molecular and Cellular Biology, University of Arizona, Tucson, Arizona, 85721, 17 18 USA
- ³ Environmental and Evolutionary Genomics Section, Institut de Biologie de l'Ecole Normale 19
- Supérieure (IBENS), CNRS, UMR8197, INSERM U1024, 75230 Paris, France 20
- ⁴ Department of Marine Biology and Oceanography, Institute of Marine Science (ICM)-CSIC, 21
- Pg. Marítim de la Barceloneta 37-49, Barcelona, E08003, Spain 22
- ⁵ CEA Institut de Génomique, GENOSCOPE, 2 rue Gaston Crémieux, 91057, Evry, France 23
- 24 ⁶ Department of Microbiology and Immunology, Rega Institute KU Leuven, Herestraat 49, 3000 Leuven, Belgium 25
- ⁷ Center for the Biology of Disease, VIB, Herestraat 49, 3000 Leuven, Belgium 26
- ⁸ Department of Applied Biological Sciences, Vrije Universiteit Brussel, Pleinlaan 2, 1050 27 Brussels, Belgium 28
- ⁹ CNRS, UMR 7144, Station Biologique de Roscoff, Place Georges Teissier, 29680 Roscoff, 29 30 France
- ¹⁰ Sorbonne Universités, UPMC Univ Paris 06, UMR 7144, Station Biologique de Roscoff, Place 31 Georges Teissier, 29680 Roscoff, France 32
- ¹¹ CNRS, UMR 7093, LOV, Observatoire Océanologique, F-06230 Villefranche-sur-mer, France 33
- ¹² Sorbonne Universités, UPMC Univ Paris 06, UMR 7093, LOV, Observatoire Océanologique, 34
- F-06230, Villefranche-sur-mer, France 35
- ¹³ Soil, Water, and Environmental Science, University of Arizona, Tucson, Arizona, 85721, USA 36

- 37 ¹⁴ Aix Marseille Université CNRS IGS UMR 7256 13288, Marseille, France
- 38 ¹⁵ Stazione Zoologica Anton Dohrn, Villa Comunale, 80121, Naples, Italy
- ¹⁶ Institute for Chemical Research, Kyoto University, Gokasho, Uji, Kyoto, 611-001, Japan
- 40 ¹⁷ PANGAEA, Data Publisher for Earth and Environmental Science, University of Bremen,
- 41 Bremen, Germany
- ¹⁸ MARUM, Center for Marine Environmental Sciences, University of Bremen, Bremen,
 Germany
- ¹⁹ Laboratoire de Physique des Océan UBO-IUEM Palce Copernic 29820 Polouzané, France
- ²⁰ Ecole Normale Supérieure, Institut de Biologie de l'ENS (IBENS), and Inserm U1024, and
 CNRS UMR 8197, Paris, F-75005, France
- 47 ²¹ Structural and Computational Biology, European Molecular Biology Laboratory, Meyerhofstr.
 48 1, 69117 Heidelberg, Germany
- ²² Directors' Research European Molecular Biology Laboratory Meyerhofstr. 1 69117
 Heidelberg, Germany
- 51 ²³ Max-Delbr ckCentre for Molecular Medicine, 13092 Berlin, Germany
- 52 ²⁴ CNRS, UMR 8030, CP5706, Evry, France
- 53 ²⁵ Université d'Evry, UMR 8030, CP5706, Evry, France
- [†]Current address: Department of Geosciences, Laboratoire de Météorologie Dynamique (LMD),
- 55 Ecole Normale Supérieure, 24 rue Lhomond 75231 Paris, Cedex 05, France
- 56 *Tara* Oceans coordinators and affiliations are listed at following the Acknowlegments.
- 57 *Correspondence to: mbsulli@gmail.com, karsenti@embl.de

58 **Abstract:** Viruses influence ecosystems by modulating microbial population size, diversity, 59 metabolic outputs, and gene flow. Here we use quantitative double-stranded DNA (dsDNA) viral-fraction metagenomes (viromes) and whole viral community morphological datasets from 60 61 43 Tara Oceans expedition samples to assess viral community patterns and structure in the upper ocean. Protein cluster cataloging defined pelagic upper-ocean viral community pan and core 62 63 gene sets and suggested this sequence space is well-sampled. Analyses of viral protein clusters, 64 populations, and morphology revealed biogeographic patterns whereby viral communities were 65 passively transported on oceanic currents and locally structured by environmental conditions that impact host community structure. Together these investigations establish a global ocean dsDNA 66 67 viromic dataset with analyses supporting the seed-bank hypothesis to explain how oceanic viral 68 communities maintain high local diversity.

69

One Sentence Summary: Global patterns that emerge from the *Tara* Oceans Virome dataset
 support a seed-bank structure underlying observed biogeography in ocean viral communities.

Main Text: Ocean microbes produce half of the oxygen we breathe (1) and drive much of the
substrate and redox transformations that fuel Earth's ecosystems (2). However, they do so in a

constantly evolving network of chemical, physical and biotic constraints – interactions which are

only beginning to be explored. Marine viruses are presumably key players in these interactions

- (3, 4) as they affect microbial populations through lysis, reprogramming of host metabolism, and
 horizontal gene transfer. Here we strive to develop an overview of ocean viral community
- 79 patterns and ecological drivers.

80 The *Tara* Oceans expedition provided a platform for sampling ocean biota from viruses to 81 fish larvae with comprehensive environmental context (5). Prior virus-focused work from this expedition has helped optimize the dsDNA viromic sample-to-sequence workflow (6), evaluate 82 ecological drivers of viral community structure as inferred from morphology (7), and map 83 ecological patterns in the large dsDNA nucleo-cytoplasmic viruses using marker genes (8). Here 84 85 we explore global patterns and structure of ocean viral communities using 43 samples from 26 86 stations in the Tara Oceans expedition (Supplementary File S1) to establish dsDNA viromes 87 from viral-fraction ($<0.22 \mu m$) concentrates and quantitative whole viral community 88 morphological datasets from unfiltered seawater. Viruses lack shared genes that can be used for 89 investigation of community patterns. Therefore, we used three levels of information to study 90 such patterns: (i) protein clusters (PCs, 9) as a means to organize virome sequence space 91 commonly dominated by unknown sequences (63–93%, 10), (ii) populations, using established

92 metrics for viral contig recruitment (11), and (*iii*) morphology, using quantitative transmission

- electron microscopy (qTEM, 7).
- 94

95 The Tara Oceans Viromes (TOV) dataset

96 The 43 *Tara* Oceans Viromes (TOV) dataset is comprised of 2.16 billion ~101-bp paired-end

- 97 Illumina reads (Supplementary File 1), largely representing epipelagic ocean viral communities
- 98 (only 1 of 43 viromes are from mesopelagic waters, Environment Ontology feature
- 99 ENVO:00000213) from the surface (ENVO:00002042) and deep chlorophyll maximum (DCM;

100 ENVO:01000326) throughout seven oceans and seas (Supplementary File S1). The TOV dataset 101 offers deeper sampling of surface ocean viral communities, but under-represents the deep ocean 102 relative to the Pacific Ocean Viromes dataset (POV, 10) which includes 16 viromes from aphotic 103 zone waters. In all viromes, sampling and processing affects what viruses are represented (6, 12-104 14). We filtered TOV seawater samples through $0.22 \,\mu m$ pore-sized filters and then concentrated 105 viruses in the filtrate using iron chloride flocculation (15). These steps would have removed most cells, but also excluded any viruses larger than 0.22 µm. We then purified the resulting TOV 106 107 viral concentrates using DNase treatment, which is as effective as density gradients for purifying 108 ocean viral concentrates (14). This DNAse-only step is unlikely to impact viral representation in 109 the viromes, but reduces non-viral DNA contamination. Finally, we extracted DNA from the 110 samples and prepared sequence libraries using linker amplification (13). These steps preserve 111 quantitative representation of dsDNA viruses in the resulting viromes (12, 13), but the ligation 112 step excludes RNA viruses, and is biased against single-stranded DNA (ssDNA) viruses (12). 113 We additionally applied qTEM (7) to paired whole seawater samples to evaluate patterns in 114 whole viral communities. This method simultaneously considers ssDNA, dsDNA, and RNA 115 viruses, though without knowledge of their relative abundances since particle morphology does 116 not identify nucleic acid type. In the oceans, total virus abundance estimates based on TEM 117 analyses, which include all viral particles, are similar to estimates based on fluorescent staining, 118 which inefficiently stains ssDNA and RNA viruses (16-24). This suggests that most ocean 119 viruses are dsDNA viruses. However, one study quantifying nucleic acids at a single marine 120 location suggests RNA viruses may constitute as much as half of the viral community there (16). 121 It remains unknown what the relative contribution of these viral types is to the whole viral 122 community, but our analyses suggest small dsDNA viruses likely dominate as follows. The 123 viromes capture the <0.22 µm dsDNA viruses of bacteria and archaea that are thought to 124 dominate marine viral communities, whereas qTEM analysis includes all viruses regardless of size, nucleic acid type, or host (7). In these whole seawater samples used for qTEM, we found 125 that viral capsid diameters ranged from 26 to 129 nm, with the per-sample average capsid 126 127 diameter constrained at 46-66 nm (Fig. 1). We detected no viral particles larger than 0.22 µm 128 among 100 randomly counted particles from each of 41 gTEM samples. These findings are 129 similar to those from a subset of these Tara Oceans stations (14 of the 26 stations; 7), and 130 indicate that size fractionation using 0.22 µm filtration to prepare viromes did not substantially 131 bias the TOV dataset.

132

TOV Protein Clusters for Comparison of Local and Global Genetic Richness and Diversity Across the 43 viromes, a total of 1,075,763 PCs were observed, with samples beyond the 20th

Across the 43 viromes, a total of 1,075,763 PCs were observed, with samples beyond the 20th virome adding few PCs (Fig. 2A). When combining TOV with 16 photic-zone viromes from the POV dataset (*10*), the number of PCs increased to 1,323,921, but again approached a plateau (Fig. 2B). These results suggest that, while impossible to sample completely, the sequence space corresponding to dsDNA viruses from the epipelagic ocean is now relatively well sampled. This

139 contrasts results from marine microbial metagenomic surveys using older sequencing

140 technologies (9), but is consistent with those from this expedition (25), as well as findings from

viral sequence datasets which suggest a limited range of functional diversity derived from

142 bacterial and archaeal viral isolates (26) and the POV dataset (27).

- PCs were next used to establish the core genes shared across the TOV dataset (Fig. 2A).
- Broadly, there were 220, 710 and 424 core PCs shared across all surface and DCM viromes,
- surface viromes only, and DCM viromes only, respectively. The number of core PCs in the
- 146 upper-ocean TOV samples (220 PCs) was thus less than the number of photic-zone core PCs in
- 147 POV (565 PCs; 28), likely because the POV dataset includes only the Pacific Ocean while TOV
- includes samples from seven oceans and seas. However, the number of core PCs in the upper-ocean TOV samples exceeded the total number of core PCs observed in POV (180 PCs; 28),
- 150 likely because of deep-ocean representation in POV (half of the samples in POV are from the
- 151 aphotic zone). Consistent with the latter, the addition of the sole deep-ocean TOV sample,
- 152 TARA 70 MESO, decreased the number of core PCs shared by all TOV samples from 220 to
- 153 65, which suggests that deep-ocean viral genetic repertoires are different from those in the upper
- 154 oceans. Indeed, niche-differentiation has been observed in viromes sampled across these oceanic
- zones in the POV dataset (28), and similar findings were observed in the microbial metagenomic
 counterparts from the *Tara* Oceans Expedition (25). Thus viral communities from the deep ocean
- remain poorly explored and appear to hold different gene sets from those in the epipelagicoceans.
- 159 Beyond core and pan metagenomic analyses, PCs also provide a metric for viral community
- 160 diversity comparisons (Fig. 3A; Supplementary File S1) from which three trends emerge in the
- 161 TOV dataset. First, high-latitude viromes (82_DCM and 85_DCM) were least diverse
- 162 (Shannon's H' of 8.93 and 9.22 nats), consistent with patterns in marine macroorganisms (29)
- and epipelagic ocean bacteria (25, 30). Second, the remaining viromes had similar diversity
- 164 (Shannon's H' between 9.47 and 10.55 nats) and evenness (Pielou's *J* from 0.85 to 0.91)
- 165 indicating low dominance of any particular PCs (*31*). Third, local diversity was relatively similar
- to global diversity (local:global ratios of H' from 0.73 to 0.87), suggesting high dispersal of viral
 genes (32) across the sampled ocean viral communities.
- 168

169 TOV Viral Populations for Assessing Global Viral Community Structure

- 170 We next estimated abundances of the 5,476 dominant viral populations in TOV, which
- represented up to 14.5% of aligned reads in a sample and were defined by applying empirically-
- derived recruitment cut-offs from naturally-occurring T4-like cyanophages (11) to high-
- 173 confidence contigs from bacterial and archaeal viruses (see Methods). Assigning viral
- 174 populations based on virome data remains challenging (11, 33), but here assembly of large
- 175 contigs (up to 100 kb) aided our ability to accomplish not only analyses at the gene-level using
- 176 PCs, but also the genome-level using viral populations. Viral populations were rarely endemic to
- 177 one station (15%), and instead were commonly observed across >4 stations (47%), and up to 24
- 178 of the 26 stations (Fig. 4 and Fig. 5A). Exceptional samples include those from the Benguela
- upwelling region (TARA_67_SUR) and high-latitude samples from the Antarctic Circumpolar

and Falklands currents (TARA 82 DCM and TARA 85 DCM, respectively). These samples 180 181 were also divergent when assessing microbial communities (TARA 82 DCM and 182 TARA 85 DCM displayed lower microbial genetic richness; (25)) and eukaryotic communities 183 (TARA 67 SUR had specific and unique eukaryotic communities in all size fractions; 34). While many viral populations were broadly distributed, they were much more abundant at the 184 185 original location (origin inferred from longest contig assembled; see Methods) compared to 186 alternate stations (Fig. 5B). Thus most populations were relatively widespread, but with variable 187 sample-to-sample abundances. As was observed with PCs, diversity and evenness estimates 188 based on viral populations were similar across all samples except for high-latitude samples 189 (TARA 82 DCM and TARA 85 DCM) and one sample in the Red Sea (TARA 32 DCM) that 190 displayed lower diversity (Fig. 3B; Supplementary File S1). Finally, local diversity was 191 relatively similar to global diversity (local:global ratios of H' from 0.23 to 0.86, average 0.74, 192 Supplementary File S1), reflecting the high dispersal of viruses as highlighted by PC analysis. 193 Only 39 of the 5,476 populations we identified could be affiliated to cultured viruses, 194 reflecting the dearth of reference viral genomes in databases. These cultured viruses include those infecting the abundant and widespread hosts SAR11, SAR116, Roseobacter, 195 196 Prochlorococcus and Synechococcus (Fig. 6). The most abundant and widespread viral populations observed in TOV lack cultured representatives (Fig. 6), which suggests that most 197 198 upper ocean viruses remain to be characterized even though viruses from known dominant 199 microbial hosts (35-39) have been cultured. Methods independent of cultivation, including viral 200 tagging (11) and mining of microbial genomic datasets (40, 41), show promise to expand the 201 number of available viral reference genomes (33).

202

203 Drivers of Global Viral Community Composition and Distribution

204 We next leveraged this global dataset to evaluate ecological drivers (including environmental variables, sample location, and microbial abundances; Supplementary File S1) of viral 205 206 community structure using all three data types – morphology, populations, and PCs. These 207 metrics revealed increasing resolution, respectively, and showed that viral community structure 208 was influenced by region and/or environmental conditions (Table 1). We conducted the analysis 209 of ecological drivers using all samples in this study as well as a sample subset that omitted 210 samples with exceptional environmental conditions and divergent viral communities observed 211 using PC and population analyses (see above; TARA 67 SUR, TARA 82 DCM, 212 TARA 85 DCM, and TARA 70 MESO). Within the sample subset, oceanic viral communities 213 varied significantly with Longhurst province, biome, latitude, temperature, oxygen 214 concentration, and microbial concentrations (including total bacteria, Synechococcus, and 215 Prochlorococcus). Viral communities were not structured by depth (surface vs DCM) except when considering PCs, likely reflecting minimal variation between samples in the epipelagic 216 217 zone compared to that of globally-sourced samples, and higher resolution provided by PCs. 218 Nutrients influenced viral community structure when considering the whole dataset, but were 219 much less explanatory when the few high-nutrient samples were removed, except for the

influence of phosphate concentration on viral populations. Thus nutrient concentrations mayinfluence viral community structure, but testing this hypothesis would require analysis of

samples across a more continuous nutrient gradient.

Global-scale analyses of oceanic macro- (29) and micro-organisms (30) have been 223 224 conducted, including a concurrent Tara Oceans study showing that temperature and oxygen 225 influence microbial community structure (25). Environmental conditions have also been shown 226 to affect global viral community morphological traits (7). Our TOV study is consistent with these 227 earlier findings in that viral communities are influenced by temperature and oxygen 228 concentration, but not chlorophyll concentration (Table 1). Biogeographic structuring of TOV 229 viral communities based on the significant influence of latitude and Longhurst provinces is also 230 consistent with the conclusion that geographic region influences community structure in Pacific 231 Ocean viruses (42). While only PC analysis showed depth-based divergence, this likely reflects 232 poor (*n*=1) deep sample representation in the TOV dataset as discussed above. Prior POV viral 233 investigation and concurrent Tara Oceans microbial analysis, both of which have better deep-234 water representation, show stronger depth patterns whereby photic and aphotic zone 235 communities diverge (25, 28, 42). Thus our results suggest biogeography of upper-ocean viral 236 communities is structured by environmental conditions. 237 Since viruses require host organisms to replicate, viral community structure follows from

238 environmental conditions shaping the host community, as observed in paired Tara Oceans 239 microbial samples (25), which would then indirectly affect viral community composition. 240 However, global distribution of viruses can also be directly influenced by environmental 241 conditions, such as salinity, that affect their ability to infect their hosts (43). Additionally, the 242 variable decay rates of cultivated viruses and whole viral communities (44) could also influence 243 their distribution as viruses with lower inherent decay rates will persist for longer in the 244 environment, and environments with more favorable conditions (such as fewer extracellular enzymes) will also contribute to increased viral persistence. Until methods to link viruses to their 245 host cells in natural communities mature to the point of investigating this issue at larger scales 246 247 (emerging possible methods reviewed by 33, 45), analyses such as ours remain the only means to 248 assess ecological drivers of viral community structure.

249 To further investigate how ocean viral communities are distributed throughout the oceans, we 250 compared population abundances between neighboring samples to assess the net direction and 251 magnitude of population exchange (Fig. 7, see Methods). These genomic signals revealed that 252 population exchange between dsDNA viral communities was largely directed along major oceanic current systems (46). For example, the Agulhas current and subsequent ring formation 253 254 (47) connects viral communities between the Indian and Atlantic Oceans, as also observed in 255 planktonic communities from the Tara Oceans expedition (48), while increased connection between the high-latitude stations (TARA 82 and TARA 85) reflects their common origin at the 256 divergence of the Falklands and Antarctic Circumpolar currents. Further, current strength (46) 257 258 was generally related to the magnitude of inter-sample population exchange, as higher and lower 259 exchange was observed, respectively, in stronger currents such as the Agulhas current, and

260 within the open ocean gyres or between land-restricted basins such as the Mediterranean and Red 261 Seas. These findings suggest that the intensity of water mass movement, in addition to 262 environmental conditions, may explain the degree to which viral populations cluster globally 263 (Fig. 4). Beyond such current-driven biogeographic evidence, vertical viral transport from surface to DCM samples was also observed (Fig. 4). This is consistent with POV observations 264 265 wherein deep-sea viromes include a modest influx of genetic material derived from surface-266 ocean viruses that are presumably transported on sinking particles (28). Exceptions include areas such as the Arabian Sea upwelling region, where increased mixing and upwelling likely exceed 267 268 sinking within the upper ocean.

269 Our TOV results enabled evaluation of a hypothesis describing the structure of viral 270 communities in the environment. Gene-marker-based studies targeting subsets of ocean viruses 271 previously found high local and low global diversity (49), a pattern also recently observed 272 genome-wide in natural cyanophage populations (11). To explain this, a seed-bank viral 273 community structure has been invoked whereby high local genetic diversity can exist by drawing 274 variation from a common and relatively limited global gene pool (49). Our results support this 275 hypothesis regarding viral community structure. Ecological driver analyses suggests that such 276 local 'seed' communities are influenced by environmental conditions, which directly impact their 277 microbial hosts and then indirectly restructure viral communities. These seed communities then 278 form the 'bank' in neighboring samples, presumably when passively transported by ocean 279 currents as shown here through the population-level analyses of net viral movement between 280 samples. This systematically-sampled, global dataset suggests large- and small-scale processes 281 play roles in structuring viral communities and offers empirical grounding for the seed-bank 282 hypothesis with regards to viral community distribution and structure.

283284 Conclusions

285 Our large-scale dataset provides a picture of global upper-ocean viral communities in which we assessed patterns using multiple parameters including morphology, populations and PCs. Our 286 287 data provide advanced and complementary views on viral community structure including nonmarker-gene-based diversity estimates and broad application of population-based viral ecology. 288 We affirm the seed-bank model for viruses, hypothesized nearly a decade ago (49), which 289 290 explains how high local viral diversity can be consistent with limited global diversity (11, 27). 291 The mechanism underlying this seed-bank population structure appears to be a local production 292 of viruses under small-scale environmental constraints and passive dispersal with oceanic 293 currents. Improving sequencing, assembly and experimental methods are transforming the 294 investigation of viruses in nature (33, 45), and pave the way towards assessment of viral 295 community structure and analysis of virus-host co-occurrence networks (50) without requiring 296 marker genes (51, 52). Such experimental and analytical progress, coupled to sampling 297 opportunities from the Tara Oceans expedition, are advancing viral ecology towards the 298 quantitative science needed to model the nano- (viruses) and micro- (microbes) scale entities 299 driving Earth's ecosystems.

300

301302 Materials and Methods

303

304 *Sample Collection*

305 Forty-three samples were collected between November 2, 2009, and May 13, 2011, at 26 306 locations throughout the world's oceans (Supplementary File S1) through the Tara Oceans 307 Expedition (5). These included samples from a range of depths (surface, deep chlorophyll 308 maximum, and one mesopelagic sample) located in 7 oceans and seas, 4 different biomes and 11 309 Longhurst oceanographic provinces (Supplementary File S1). Longhurst provinces and biomes 310 are defined based on Longhurst (53) and environmental features are defined based on 311 Environment Ontology (http://environmentontology.org/). Sampling strategy and methodology 312 for the Tara Oceans Expedition is fully described by Pesant et al. (54).

313

314 Environmental Parameters

315 Temperature, salinity, and oxygen data were collected from each station using a CTD 316 (Sea-Bird Electronics, Bellevue, WA, USA; SBE 911plus with Searam recorder) and dissolved 317 oxygen sensor (Sea-Bird Electronics; SBE 43). Nutrient concentrations were determined using 318 segmented flow analysis (55) and included nitrite, phosphate, nitrite plus nitrate, and silica. 319 Nutrient concentrations below the detection limit (0.02 μ mol kg⁻¹) are reported as 0.02 μ mol kg⁻¹ 320 ¹. Chlorophyll concentrations were measured using HPLC (56, 57). These environmental 321 parameters are available in Pangaea (www.pangaea.de) using the accession numbers in 322 Supplementary File S1.

323

324 *Microbial Abundances*

Flow-cytometry was used to determine the concentration of *Synechococcus*, *Prochlorococcus*, total bacteria, low-DNA bacteria, high-DNA bacteria, and the percent of
bacteria with high DNA in each sample (58).

328

329 Morphological Analysis of Viral Communities

330 Quantitative transmission electron microscopy (qTEM) was used to evaluate the capsid 331 diameter distributions of viral communities as previously described (7). Briefly, preserved 332 unfiltered samples (EM-grade glutaraldehyde; Sigma-Aldrich, St. Louis, MO, USA; 2% final 333 concentration) were flash-frozen and stored at -80°C until analysis. Viruses were deposited onto 334 TEM grids using an air-driven ultracentrifuge (Airfuge CLS, Beckman Coulter, Brea, CA, USA), 335 followed by positive staining of the deposited material with 2% uranyl acetate (Ted Pella, 336 Redding, CA, USA). Samples were then examined using a transmission electron microscope 337 (Philips CM12, FEI, Hilsboro, OR, USA) with 100 kV accelerating voltage. Micrographs of 100 338 viruses were collected per sample using a Macrofire Monochrome CCD camera (Optronics, 339 Goleta, CA, USA) and analyzed using ImageJ software (US National Institutes of Health, 340 Bethesda, MD, USA; 59) to measure the capsid diameter. A subset (21) of the 41 samples

341 presented here had previously been analyzed in a different study (7).

342

343 Virome Construction

344 For each sample, 20 L of seawater were 0.22 µm-filtered and viruses were concentrated 345 from the filtrate using iron chloride flocculation (15) followed by storage at 4°C. After resuspension in ascorbic-EDTA buffer (0.1 M EDTA, 0.2 M Mg, 0.2 M ascorbic acid, pH 6.0), 346 347 viral particles were concentrated using Amicon Ultra 100 kDa centrifugal devices (Millipore), 348 treated with DNase I (100U/mL) followed by the addition of 0.1 M EDTA and 0.1 M EGTA to 349 halt enzyme activity, and extracted as previously described (14). Briefly, viral particle 350 suspensions were treated with Wizard PCR Preps DNA Purification Resin (Promega, WI, USA) 351 at a ratio of 0.5 mL sample to 1 mL resin, and eluted with TE buffer (10 mM Tris, pH 7.5, 1 mM 352 EDTA) using Wizard Minicolumns. Extracted DNA was Covaris-sheared and size selected to 353 160–180 bp, followed by amplification and ligation per the standard Illumina protocol. 354 Sequencing was done on a HiSeq 2000 system at the Genoscope facilities (Paris, France).

355

356 *Quality Control of Reads and Assembly*

357 Individual reads of 43 metagenomes were quality controlled using a combination of 358 trimming and filtering as previously described (60). Briefly, bases were trimmed at the 5' end if the number of base calls for any base (A, T, G, C) diverged by more than two standard 359 360 deviations from the average across all cycles. Conversely, bases were trimmed at the 3' end of 361 reads if the quality score was <20. Finally, reads that were shorter than 95 bp or reads with a 362 median quality score <20 were removed from further analyses. Assembly of reads was done 363 using SOAPdenovo (61) where insert and k-mer size are calculated at run-time and are specific 364 to each virome as implemented in the MOCAT pipeline (62). On average, 34.2% of the virome reads were included in the assembled contigs (min: 21.08%, max: 48.52%). Virome reads were 365 366 deposited in the European Nucleotide Archive (http://www.ebi.ac.uk/ena/) under accession 367 numbers reported in Supplementary File S1.

368

369 *Protein Clustering*

370 Open Reading Frames (ORFs) were predicted from all quality-controlled contigs using 371 Prodigal (63) with default settings. Predicted ORFs were clustered based on sequence similarity 372 as described previously (9, 10). Briefly, ORFs were initially mapped to existing clusters (POV, 373 GOS and phage genomes), using cd-hit-2d ('-g 1 -n 4 -d 0 -T 24 -M 45000'; 60% percent 374 identity and 80% coverage). Then the remaining, unmapped ORFs were self-clustered, using cd-375 hit with the same options as above. Only protein clusters (PCs) with more than two ORFs were 376 considered *bona fide* and were used for subsequent analyses. To develop read counts per PC for 377 statistical analyses, reads were mapped back to predicted ORFs in the contigs dataset using 378 Mosaik with the following settings: "-a all -m all -hs 15 –minp 0.95 –mmp 0.05 -mhp 100 -act 379 20" (version 1.1.0021; http://bioinformatics.bc.edu/marthlab/Mosaik). Read counts to PCs were 380 normalized by sequencing depth of each virome. Shannon diversity (H') was calculated from PC 381 read counts using only PCs with more than two predicted ORFs. Observed richness is reported as

the total number of reads in each PC. Pielou's evenness (*J*) was calculated as the ratio of H'/ H_{max} , where $H_{max} = \ln N$, and N = total number of observed PCs in a sample.

384

385 Analysis of Viral Populations

386 Considering the size of the entire dataset (3,821,756 assembled contigs), we decided to 387 focus the analysis of viral populations using contigs most likely originating from bacterial or 388 archaeal viruses. For this, we mined only the 22,912 contigs with more than 10 predicted genes 389 (corresponding to an average of 6.41% of the assembled reads per sample, min: 1.29%, max: 390 14.52%), as the origin of contigs with only a few predicted genes can be spurious. First, we 391 removed 6,706 contigs suspected of having originated from cellular genomes (64), whether due 392 to free genomic DNA contamination or viral-encapsidation of cellular DNA (for example, in 393 gene transfer agents or generalized transducing phages). These suspect cellular contigs were 394 those containing no typical viral genes (such as virion-related genes including major capsid 395 proteins and large subunits of the terminase) and displaying as many 'characterized genes' (such 396 as genes with a significant similarity to a PFAM domain through Hmmsearch, 65) as a typical 397 cellular genome, whereas phage genomes are typically enriched in 'uncharacterized genes' (40). 398 We also removed all contigs posited to originate from eukaryotic viruses. These were contigs that contained at least three predicted proteins with best BLAST hits to a eukaryotic virus, and 399 400 more than half of the affiliated proteins were not associated to bacteriophages or archaeal 401 viruses. Not surprisingly, given that eukaryotes are outnumbered by bacteria and archaea in the 402 marine environment, this step removed only 142 contigs associated with eukaryotic viruses. From the remaining 16,124 contigs most likely to have originated from bacterial or archaeal 403 404 viruses, the population study only used those longer than 10kb in size – a total of 6,322 contigs, 405 which corresponded to an average of 4.04% of the assembled reads per sample, min: 0.98%, 406 max: 9.97%).

These 6,322 contigs were then clustered into populations if they shared more than 80% of their genes at >95% nucleotide identity; a threshold derived from naturally-occurring T4-like cyanophages (11). This resulted in 5,476 'populations' from the 6,322 contigs, where as many as 12 contigs (average 1.15 contigs) were included per population. For each population, the longest contig was chosen as the 'seed' sequence.

412 The relative abundance of each population was computed by mapping all quality-413 controlled reads to the set of 5,476 non-redundant populations (considering only mapping quality 414 scores greater than 1) with Bowtie 2 (66). For each sample-sequence pair, if more than 75% of the reference sequence was covered by virome reads, the relative abundance was computed as 415 416 the number of base pairs recruited to the contig normalized to the total number of base pairs 417 available in the virome and the contig length. Shannon diversity index (H') and Pielou's evenness 418 (J) were calculated as done for PCs using the relative abundance of viral populations. 419 The sample containing the seed sequence (the longest contig in a population) was also

420 considered the best estimate of that population's origin. We reasoned this was because the421 longest contig in a population would derive most often from the sample with the highest

422 coverage (a metric for population abundance) and likely corresponded to the location with the
423 greatest viral abundance for this population. This assumption was supported by the results
424 showing that populations were most abundant in their original samples (Fig. 4, Fig. 5B). Even
425 though some individual cases could diverge from this rule, we expected to correctly identify
426 most of these original locations, hence getting an accurate global signal.

427 The seed sequence was also used to assess taxonomic affiliation of the viral population. 428 Cases where >50% of the genes were affiliated to a specific reference genome from RefSeq 429 (based on a BLASTp comparison with thresholds of 50 for bit score and 10^{-5} for e-value) with an 430 identity percentage of at least 75% (at the protein sequence level) were considered as confident 431 affiliations to the corresponding reference virus.

432 Finally, estimations of net viral population movement between samples were made based 433 on the relative abundance of populations in one sample compared to that of its neighboring 434 samples (Fig. 4). For each neighboring sample pair, the average relative abundance of 435 populations originating from sample A in sample B was compared with the relative abundance of 436 populations originating from sample B in sample A. The origin of each population was defined 437 as the sample in which the longest contig of the population was assembled. The magnitude of 438 these differences was carried through the analysis to estimate the level of transport between each pair of samples (depicted as line width in Fig. 7) and the difference between these values was 439 440 used to estimate the directionality of the transfer. For example, if sample B contains many 441 populations from sample A, but very few populations from sample B are detected in sample A, we calculate that the net movement is from sample A to sample B. Again, while the sampling of 442 443 some populations may not be strong, the net movement was calculated as the average of all 444 shared populations between neighboring sample pairs, which corresponded to 105 different 445 populations on average (ranging from 2 to 412).

446

447 Statistical Ordination of Samples

448 Viral community composition based on capsid diameter distributions (from qTEM; using 449 7-nm histogram bin sizes), population abundances, and normalized PC read counts (using only 450 protein clusters with more than 20 representatives) were compared using non-metric 451 multidimensional scaling (NMDS) performed using the 'metaMDS' function (default 452 parameters) of the vegan package (67) in R version 2.15.2 (68). The influence of metadata on 453 sample ordination was evaluated using the functions 'envfit' for factor variables including depth category, Longhurst province, and biome, and 'ordisurf' for all linear variables, in the vegan 454 package (67, 69). Several samples had exceptional environmental conditions (TARA 67 SUR, 455 TARA 70 MESO, TARA 82 DCM, and TARA 85 DCM), thus all statistical ordination 456 457 analyses were conducted with and without these samples (referred to as the 'sample subset') to evaluate their influence. 458

459

460 References

461 1. C. B. Field, M. J. Behrenfeld, J. T. Randerson, P. Falkowski, Primary production of the
462 biosphere: Integrating terrestrial and oceanic components. *Science* 281, 237-240 (1998).

- 463 2. P. G. Falkowski, T. Fenchel, E. F. Delong, The microbial engines that drive Earth's
- 464 biogeochemical cycles. *Science* **320**, 1034-1039 (2008).
- 465 3. M. Breitbart, Marine viruses: Truth or dare. *Ann. Rev. Mar. Sci.* **4**, 425-448 (2012).
- 466 4. C. A. Suttle, Marine viruses major players in the global ecosystem. *Nat. Rev. Microbiol.* 5,
 467 801-812 (2007).
- 5. E. Karsenti, *et al.*, A holistic approach to marine eco-systems biology. *PLoS Biol.* 9, e1001177
 (2011).
- 470 6. S. Solonenko, *et al.*, Sequencing platform and library preparation choices impact viral
- 471 metagenomes. *BMC Genomics* **14**, 320 (2013).
- 472 7. J. R. Brum, R. O. Schenck, M. B. Sullivan, Global morphological analysis of marine viruses
 473 shows minimal regional variation and dominance of non-tailed viruses. *ISME J.* 7, 1738-1751
 474 (2013).
- 475 8. P. Hingamp, *et al.*, Exploring nucleo-cytoplasmic large DNA viruses in Tara Oceans microbial
 476 metagenomes. *ISME J.* 7, 1678-1695 (2013).
- 9. S. Yooseph, *et al.*, The Sorcerer II global ocean sampling expedition: expanding the universe
 of protein families. *PLoS Biol.* 5, 0432-0466 (2007).
- 479 10. B. L. Hurwitz, M. B. Sullivan, The Pacific Ocean Virome (POV): A marine viral
- 480 metagenomic dataset and associated protein clusters for quantitative viral ecology. *PLoS ONE* 8,
 481 e57355 (2013).
- 482 11. L. Deng, *et al.*, Viral tagging reveals discrete populations in *Synechococcus* viral genome
 483 sequence space. *Nature* 513, 242 (2014).
- 484 12. M. B. Duhaime, M. B. Sullivan, Ocean viruses: Rigorously evaluating the metagenomic
 485 sample-to-sequence pipeline. *Virology* 434, 181-186 (2012).
- 486 13. M. B. D. Duhaime, L. Deng, B. T. Poulos, M. B. Sullivan, Towards quantitative
- 487 metagenomics of wild viruses and other ultra-low concentration DNA samples: a rigorous
- assessment and optimization of the linker amplification method. *Environ. Microbiol.* 14, 25262537 (2012).
- 490 14. B. L. Hurwitz, L. Deng, B. T. Poulos, M. B. Sullivan, Evaluation of methods to concentrate
- and purify ocean virus communities through comparative, replicated metagenomics. *Environ. Microbiol.* 15, 1428-1440 (2013).
- 493 15. S. G. John, *et al.*, A simple and efficient method for concentration of ocean viruses by
- 494 chemical flocculation. *Environ. Microbiol. Rep.* **3**, 195-202 (2011).
- 495 16. G. F. Steward, *et al.*, Are we missing half of the viruses in the ocean? *ISME J.* 7, 672-679
 496 (2013).
- 497 17. K. Holmfeldt, D. Odic, M. B. Sullivan, M. Middelboe, L. Riemann, Cultivated single
- 498 stranded DNA phages that infect marine Bacteroidetes prove difficult to detect with DNA
 499 binding stains. *Appl. Environ. Microbiol.* 78, 892-894 (2012).
- 500 18. Y. Tomaru, K. Nagasaki, Flow cytometric detection and enumeration of DNA and RNA 501 viruses infecting marine eukaryotic microalgae. *J. Oceanogr.* **63**, 215-221 (2007).
- 502 19. C. P. D. Brussaard, D. Marie, G. Bratbak, Flow cytometric detection of viruses. J. Virol.
 503 Methods 85, 175-182 (2000).
- 504 20. Y. Bettarel, T. Sime-Ngando, C. Amblard, H. Laveran, A comparison of methods for
- 505 counting viruses in aquatic systems. *Appl. Environ. Microbiol.* **66**, 2283-2289 (2000).
- 506 21. K. P. Hennes, C. A. Suttle, Direct counts of viruses in natural waters and laboratory cultures
- 507 by epifluorescence microscopy. *Limnol. Oceanogr.* **40**, 1050-1055 (1995).

- 508 22. M. G. Weinbauer, C. A. Suttle, Comparison of epifluorescence and transmission electron
- microscopy for counting viruses in natural marine waters. *Aquat. Microb. Ecol.* 13, 225-232
 (1997).
- 511 23. R. T. Noble, J. A. Fuhrman, Use of SYBR Green I for rapid epifluorescence counts of marine 512 viruses and bacteria. *Aquat. Microb. Ecol.* **14**, 113-118 (1998).
- 513 24. D. Marie, C. P. D. Brussaard, R. Thyrhaug, G. Bratbak, D. Vaulot, Enumeration of marine
- viruses in culture and natural samples by flow cytometry. *Appl. Environ. Microbiol.* 65, 45-52 (1999).
- 516 25. S. Sunagawa, Structure and function of the global ocean microbiome. (in review).
- 517 26. D. M. Kristensen, *et al.*, Orthologous gene clusters and taxon signature genes for viruses of
 518 prokaryotes. *J. Bacteriol.* 195, 941-950 (2013).
- 519 27. C. J. Ignacio-Espinoza, S. A. Solonenko, M. B. Sullivan, The global virome: not as big as we
 520 thought? *Curr. Opin. Virol.* 3, 566-571 (2013).
- 521 28. B. L. Hurwitz, J. R. Brum, M. B. Sullivan, Depth-stratified functional and taxonomic niche
- 522 specialization in the 'core' and 'flexible' Pacific Ocean Virome. *ISME J.* 9, 472-484 (2015).
- 523 29. D. P. Tittensor, *et al.*, Global patterns and predictors of marine biodiversity across taxa.
 524 *Nature* 466, 1098-1101 (2010).
- 525 30. W. J. Sul, T. A. Oliver, H. W. Ducklow, L. A. Amaral-Zettler, M. L. Sogin, Marine bacteria 526 exhibit a bipolar distribution. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 2342-2347 (2013).
- 527 31. D. I. Jarvis, et al., A global perspective of the richness and evenness of traditional crop-
- variety diversity maintained by farming communities. *Proc. Natl. Acad. Sci. U.S.A.* 105, 53265331 (2008).
- 530 32. S. P. Hubbell, *The Unified Neutral Theory of Biodiversity and Biogeography* (Princeton
 531 University Press, Princeton, NJ, 2001).
- 33. J. R. Brum, M. B. Sullivan, Rising to the challenge: accelerated pace of discovery transforms
 marine virology. *Nat. Rev. Microbiol.* doi:10.1038/nrmicro3404 (2015).
- 534 34. C. de Vargas, *et al.*, Global oceans eukaryotic plankton diversity. (in review).
- 535 35. I. Kang, H.-M. Oh, D. Kang, J.-C. Cho, Genome of a SAR116 bacteriophage shows the
- prevalence of this phage type in the oceans. *Proc. Natl. Acad. Sci. U.S.A.* 110, 12343-12348
 (2013).
- 36. S. J. Labrie, *et al.*, Genomes of marine cyanopodoviruses reveal multiple origins of diversity. *Environ. Microbiol.* 15, 1356-1376 (2013).
- 540 37. M. B. Sullivan, et al., Genomic analysis of oceanic cyanobacterial myoviruses compared
- 541 with T4-like myoviruses from diverse hosts and environments. *Environ. Microbiol.* 12, 3035-
- 542 3056 (2010).
- 543 38. Y. Zhao, et al., Abundant SAR11 viruses in the ocean. Nature 494, 357-360 (2013).
- 544 39. F. Rohwer, et al., The complete genomic sequence of the marine phage Roseaphage SIO1
- shares homology with nonmarine phages. *Limnol. Oceanogr.* **45**, 408-418 (2000).
- 546 40. S. Roux, *et al.*, Ecology and evolution of viruses infecting uncultivated SUP05 bacteria as
- 547 revealed by single-cell- and meta- genomics. *eLife* 10.7554/eLife.03125 (2014).
- 548 41. C. M. Mizuno, F. Rodriguez-Valera, N. E. Kimes, R. Ghai, Expanding the marine virosphere 549 using metagenomics. *PLoS Genet.* 9, e1003987 (2013).
- 42. B. Hurwitz, A. Westvald, J. Brum, M. Sullivan, Modeling ecological drivers in marine viral
- 551 communities using comparative metagenomics and network analyses. Proc. Natl. Acad. Sci.
- **552** U.S.A. **111**, 10714-10719 (2014).

- 43. P. Kukkaro, D. H. Bamford, Virus-host interactions in environments with a wide range of ionic strengths. *Environ. Microbiol. Rep.* **1**, 71-77 (2009).
- 44. K. E. Wommack, R. R. Colwell, Virioplankton: viruses in aquatic ecosystems. *Microbiol. Mol. Biol. Rev.* 64, 69-114 (2000).
- 557 45. V. Dang, M. B. Sullivan, Emerging methods to study bacteriophage infection at the single-558 cell level. *Front. Microbiol.* (in press).
- 46. L. D. Talley, G. L. Pickard, W. J. Emery, J. H. Swift, *Descriptive Physical Oceanography: An Introduction (Sixth Edition)* (Elsevier, Boston, 2011).
- 561 47. D. B. Olson, R. H. Evans, Rings of the Agulhas current. *Deep Sea Res. A* 33, 27-42 (1986).
- 562 48. E. Villar, Dispersal and remodeling of plankton communities by Agulhas rings. (in review).
- 49. M. Breitbart, F. Rohwer, Here a virus, there a virus, everywhere the same virus? *Trends Microbiol.* 13, 278-284 (2005).
- 565 50. G. Lima-Mendez, *et al.*, Top-down determinants of ocean microbial community structure. (in review).
- 567 51. D. M. Needham, *et al.*, Short-term observations of marine bacterial and viral communities:
- 568 patterns, connections and resilience. *ISME J.* 7, 1274-1285 (2013).
- 569 52. C.-E. T. Chow, D. Y. Kim, R. Sachdeva, D. A. Caron, J. A. Fuhrman, Top-down controls on
- bacterial community structure: microbial network analysis of bacteria, T4-like viruses and
 protists. *ISME J.* 8, 816-829 (2014).
- 572 53. A. Longhurst, *Ecological Geography of the Sea* (Elsevier, Inc., London, 2007).
- 573 54. S. Pesant, *et al.*, *Tara* Oceans Data: A sampling strategy and methodology for the study of 574 marine plankton in their environmental context. (in review).
- 575 55. A. Aminot, R. Kerouel, S. C. Coverly, in *Practical Guidelines for the Analysis of Seawater*,
- 576 O. Wurl, Eds (CRC Press, Boca Raton, 2009), vol. pp. 143-178.
- 577 56. J. Ras, H. Claustre, J. Uitz, Spatial variability of phytoplankton pigment distributions in the
- 578 Subtropical South Pacific Ocean: comparison between *in situ* and predicted data. *Biogeosciences* 579 5 353 369 (2008)
- **579 5**, 353-369 (2008).
- 580 57. L. Van Heukelem, C. S. Thomas, Computer-assisted high-performance liquid
- 581 chromatography method development with applications to the isolation and analysis of
- 582 phytoplankton pigments. J. Chromatogr. A 910, 31-49 (2001).
- 583 58. J. M. Gasol, P. A. del Giorgio, Using flow cytometry for counting natural planktonic bacteria
- and understanding the structure of planktonic bacterial communities. *Sci. Mar.* 64, 197-224
 (2000).
- 586 59. M. D. Abramoff, P. J. Magalhaes, S. J. Ram, Image processing with ImageJ. *Biophotonics*587 *International* 11, 36-42 (2004).
- 588 60. S. Schloissnig, *et al.*, Genomic variation landscape of the human gut microbiome. *Nature*589 493, 45-50 (2013).
- 61. R. Luo, *et al.*, SOAPdenovo2: an empirically improved memory-efficient short-read de novo
 assembler. *GigaScience* 1, 18 (2012).
- 592 62. J. R. Kultima, *et al.*, MOCAT: A metagenomics assembly and gene prediction toolkit. *PLoS*593 *ONE* 7, e47656 (2012).
- 594 63. D. Hyatt, *et al.*, Prodigal: prokaryotic gene recognition and translation initiation site
- identification. *BMC Bioinformatics* **11**, 119 (2010).
- 596 64. S. Roux, M. Krupovic, D. Debroas, P. Forterre, F. Enault, Assessment of viral community
- 597 functional potential from viral metagenomes may be hampered by contamination with cellular
- **598** sequences. *Open Biol.* **3**, 130160 (2013).

599 65. R. D. Finn, J. Clements, S. R. Eddy, HMMER web server: interactive sequence similarity 600 searching. Nucleic Acids Res. 39, W29-W37 (2011). 66. B. Langmead, S. L. Salzberg, Fast gapped-read alignment with Bowtie 2. Nat. Methods 9, 601 602 357-359 (2012). 603 67. J. Oksanen, et al., vegan: Community Ecology Package, R package version 2.1-27/r2451 604 (2013).605 68. R Core Team, R: A language and environment for statistical computing. v. 2.15.2 (2012). 606 69. S. N. Wood, Fast stable restricted maximum likelihood estimation of semiparametric 607 generalized linear models. J. R. Stat. Soc. Ser. A 73, 3-36 (2011). 608 609 Acknowledgements. We thank Jesse Czekanski-Moir for advice on statistics and Laurent 610 611 Coppola for assistance with validating nutrient data. We thank the commitment of the following 612 people and sponsors: CNRS (in particular Groupement de Recherche GDR3280), European 613 Molecular Biology Laboratory (EMBL), Genoscope/CEA, VIB, Stazione Zoologica Anton 614 Dohrn, UNIMIB, Fund for Scientific Research – Flanders, Rega Institute, KU Leuven, The 615 French Ministry of Research, the French Government 'Investissements d'Avenir' programmes OCEANOMICS (ANR-11-BTBR-0008), FRANCE GENOMIQUE (ANR-10-INBS-09-08), 616 MEMO LIFE (ANR-10-LABX-54), PSL* Research University (ANR-11-IDEX-0001-02), ANR 617 (projects POSEIDON/ANR-09-BLAN-0348, PHYTBACK/ANR-2010-1709-01, 618 PROMETHEUS/ANR-09-PCS-GENM-217, TARA-GIRUS/ANR-09-PCS-GENM-218), 619 620 European Union FP7 (MicroB3/No.287589, IHMS/HEALTH-F4-2010-261376), ERC Advanced 621 Grant Award to CB (Diatomite: 294823), Gordon and Betty Moore Foundation grant (#3790) to 622 MBS, Spanish Ministry of Science and Innovation grant CGL2011-26848/BOS MicroOcean 623 PANGENOMICS to SGA, TANIT (CONES 2010-0036) from the Agència de Gestió d'Ajusts 624 Universitaris i Reserca to SGA, JSPS KAKENHI Grant Number 26430184 to HO, and FWO, 625 BIO5, Biosphere 2 to MBS. We also thank the support and commitment of Agnès b. and Etienne 626 Bourgois, the Veolia Environment Foundation, Region Bretagne, Lorient Agglomeration, World 627 Courier, Illumina, the EDF Foundation, FRB, the Prince Albert II de Monaco Foundation, the 628 Tara schooner and its captains and crew. We thank MERCATOR-CORIOLIS and ACRI-ST for 629 providing daily satellite data during the expedition. We are also grateful to the French Ministry 630 of Foreign Affairs for supporting the expedition and to the countries who graciously granted 631 sampling permissions. Tara Oceans would not exist without continuous support from 23 632 institutes (http://oceans.taraexpeditions.org). We also acknowledge excellent assistance from the 633 European Bioinformatics Institute (EBI), in particular Guy Cochrane and Petra ten Hoopen, as 634 well as the EMBL Advanced Light Microscopy Facility (ALMF), in particular Rainer 635 Pepperkok. The authors further declare that all data reported herein are fully and freely available 636 from the date of publication, with no restrictions, and that all of the samples, analyses, 637 publications, and ownership of data are free from legal entanglement or restriction of any sort by the various nations whose waters the Tara Oceans expedition sampled in. Data described herein 638 is available at EBI (project identifiers PRJEB402 and PRJEB7988) and Pangaea (see 639 Supplementary Table S1), and the data release policy regarding future public release of *Tara* 640 641 Oceans data is described in Pesant et al (54). We also acknowledge the EMBL Advanced Light 642 Microscopy Facility (ALMF), and in particular Rainer Pepperkok. This research is funded in part 643 by the Gordon and Betty Moore Foundation through grants GBMF2631 and GBMF3790 to 644 MBS. We also acknowledge support from UA high-performance computing; the foundation for 645 France-American Cultural Exchange, Partner University Fund program, awarded to Ecole

- Normale Supérieure and the University of Arizona; and a grant to the University of Arizona 646
- Ecosystem Genomics Institute through the UA Technology and Research Initiative Fund and the 647
- 648 Water, Environmental and Energy Solutions Initiative. All authors approved the final
- 649 manuscript. This article is contribution number XXX of the Tara Oceans Expedition.
- Supplement contains additional data. 650

Tara Oceans Coordinators 651

- Silvia G. Acinas¹, Peer Bork^{2,3}, Emmanuel Boss⁴, Chris Bowler⁵, Colomban de Vargas^{6,7}, Michael Follows⁸, Gabriel Gorsky^{9,31}, Nigel Grimsley^{10,11}, Pascal Hingamp¹², Daniele Iudicone¹³, Olivier Jaillon^{14,15,16}, Stefanie Kandels-Lewis^{2,17}, Lee Karp-Boss¹⁸, Eric Karsenti^{5,17}, Uros Krzic¹⁹, Fabrice Not^{6,7}, Hiroyuki Ogata²⁰, Stephane Pesant^{21,22}, Jeroen Raes^{23,24,25}, Emmanuel G. Reynaud²⁶, Christian Sardet^{27,28}, Mike Sieracki^{29,†}, Sabrina Speich^{30,‡}, Lars Stemmann^{9,31}, Matthew B. Sullivan³², Shinichi Sunagawa², Didier Velayoudon³³, Jean Weissenbach^{14,15,16}, Patrick Wincker^{14,15,16} 652
- 653
- 654
- 655 656
- 657
- 658
- ¹ Department of Marine Biology and Oceanography, Institute of Marine Science (ICM)-CSIC, 659 Pg. Marítim de la Barceloneta 37-49, Barcelona, E08003, Spain 660
- ² Structural and Computational Biology, European Molecular Biology Laboratory, Meyerhofstr. 661 662 1, 69117 Heidelberg, Germany
- ³ Max-Delbr ckCentre for Molecular Medicine, 13092 Berlin, Germany 663
- ⁴ School of Marine Sciences, University of Maine, Orono, Maine, USA 664
- ⁵ Ecole Normale Supérieure, Institut de Biologie de l'ENS (IBENS), and Inserm U1024, and 665 CNRS UMR 8197, Paris, F-75005 France 666
- ⁶ CNRS, UMR 7144, Station Biologique de Roscoff, Place Georges Teissier, 29680 Roscoff, 667 668 France
- ⁷ Sorbonne Universités, UPMC Univ Paris 06, UMR 7144, Station Biologique de Roscoff, Place 669 Georges Teissier, 29680 Roscoff, France 670
- ⁸ Dept of Earth. Atmospheric and Planetary Sciences. Massachusetts Institute of Technology. 671 672 Cambridge, Massachusetts, USA
- ⁹ CNRS, UMR 7093, LOV, Observatoire Océanologique, F-06230 Villefranche-sur-mer, France 673
- ¹⁰ CNRS UMR 7232, BIOM, Avenue du Fontaulé, 66650 Banyuls-sur-Mer, France 674
- ¹¹ Sorbonne Universités Paris 06, OOB UPMC, Avenue du Fontaulé, 66650 Banyuls-sur-Mer, 675 676 France
- ¹² Aix Marseille Université CNRS IGS UMR 7256 13288 Marseille, France 677
- ¹³ Stazione Zoologica Anton Dohrn, Villa Comunale, 80121, Naples, Italy 678

- 679 ¹⁴ CEA Institut de Génomique, GENOSCOPE, 2 rue Gaston Crémieux, 91057 Evry, France
- 680 ¹⁵ CNRS, UMR 8030, CP5706, Evry, France
- 681 ¹⁶ Université d'Evry, UMR 8030, CP5706, Evry, France
- ¹⁷ Directors' Research European Molecular Biology Laboratory Meyerhofstr. 1 69117
 Heidelberg, Germany
- 684 ¹⁸ School of Marine Sciences, University of Maine, Orono, USA
- ¹⁹ Cell Biology and Biophysics, European Molecular Biology Laboratory, Meyerhofstr. 1, 69117
 Heidelberg, Germany
- ²⁰ Institute for Chemical Research, Kyoto University, Gokasho, Uji, Kyoto, 611-001, Japan
- ⁶⁸⁸ ²¹ PANGAEA, Data Publisher for Earth and Environmental Science, University of Bremen,
 ⁶⁸⁹ Bremen, Germany
- ⁶⁹⁰ ²² MARUM, Center for Marine Environmental Sciences, University of Bremen, Bremen,
 Germany
- ²³ Department of Microbiology and Immunology, Rega Institute, KU Leuven, Herestraat 49,
 3000 Leuven, Belgium
- ²⁴ Center for the Biology of Disease, VIB, Herestraat 49, 3000 Leuven, Belgium
- ²⁵ Department of Applied Biological Sciences, Vrije Universiteit Brussel, Pleinlaan 2, 1050
 Brussels, Belgium
- ²⁶ Earth Institute, University College Dublin, Dublin, Ireland
- ²⁷ CNRS, UMR 7009 Biodev, Observatoire Océanologique, F-06230 Villefranche-sur-mer,
 France
- ²⁸ Sorbonne Universités, UPMC Univ Paris 06, UMR 7009 Biodev, F-06230 Observatoire
 Océanologique, Villefranche-sur-mer, France
- 702 ²⁹ Bigelow Laboratory for Ocean Science, East Boothbay, Maine, USA
- ³⁰ Laboratoire de Physique des Océan UBO-IUEM Palce Copernic 29820 Polouzané, France
- ³¹ Sorbonne Universités, UPMC Univ Paris 06, UMR 7093, LOV, Observatoire Océanologique,
 F-06230, Villefranche-sur-mer, France
- ³² Department of Ecology and Evolutionary Biology, Depts Molecular and Cellular Biology and
 Soil, Water and Environmental Science, University of Arizona, Tucson, Arizona, 85721, USA
- 708 ³³ DVIP Consulting, 92310, Sèvres, France

- 709 [†] Current address: National Science Foundation, Arlington, Virginia, USA
- [†]Current address: Department of Geosciences, Laboratoire de Météorologie Dynamique (LMD),
- 711 Ecole Normale Supérieure, 24 rue Lhomond 75231 Paris, Cedex 05, France

712 Supplementary Information

713

Supplementary File S1. Description of samples and relevant virome data. Metadata is
presented for each *Tara* Oceans sample in this study including the PANGAEA accession
numbers, sample location and environmental conditions, and the abundances of selected
microorganisms. Detailed information is also presented for the viromes in this study including
ENA accession numbers, the total number of reads and PCs for each virome, and diversity and
evenness data for each virome based on PCs and viral populations.

- 720 721
- 722 Figure Legends
- 723

Fig. 1. Distribution of viral capsid diameters in each sample (n = 100 viruses per sample). 724 725 Data are not available for samples TARA 18 DCM and TARA 70 MESO. Boxplots are constructed with the upper and lower lines corresponding to the 25th and 75th percentiles, while 726 727 outliers are displayed as points. Longhurst provinces are indicated below samples (MEDI, 728 Mediterranean Sea; REDS, Red Sea; ARAB, NW Arabian Upwelling; MONS, Indian Monsoon Gyres; ISSG, Indian S. Subtropical Gyre; EAFR, E. Africa Coastal; BENG, Benguela Current 729 730 Coastal; SATL, S. Atlantic Gyre; FKLD, SW Atlantic Shelves; APLR, Austral Polar; PNEC, N. 731 Pacific Equatorial Countercurrent).

732

Fig. 2. Protein cluster (PC) richness in core and pan viromes from the TOV and POV

datasets. A) Accumulation curves of core and pan PCs in the TOV dataset. Vertical axis shows
the number of shared (core virome) and total (pan virome) PCs when *n* viromes are compared (*n*

- 736 = 1 to 43; from 3 to 41 only 1000 combinations are shown). Lines: i) total number of PCs
- 737 (1,075,763 PCs), ii) core surface virome (710 PCs), iii) core DCM virome (424 PCs), iv) core
- surface and DCM virome (220 PCs), v) all samples (including the deep-ocean sample
- TARA_70_MESO; 65 PCs). B) Core and pan PCs in all TOV and photic-zone POV samples
- combined. Vertical axis shows the number of shared (core virome) and total (pan-virome) PCs
- 741 when *n* virones are compared (n = 1 to 57; from 3 to 57 only 1,000 combinations are shown).
- 742 Overall, 1,323,921 PCs were identified in all viromes combined.
- 743

Fig. 3. Alpha diversity measurements in TOV dataset. A) Shannon's richness H' and Pielou's

evenness J calculated from protein clusters counts for each sample and a pool of all samples,

- normalized to 5 million reads. B) Shannon's richness H' and Pielou's evenness J calculated from
- relative abundances of viral populations for each sample and a pool of all samples, with
- subsamples of 100,000 reads. Outliers corresponding to values outside of the average value plus
- or minus two standard deviations are colored in green and red, respectively. Values calculated

- from the pool of all samples are colored in blue. Longhurst provinces are indicated belowsamples using the same abbreviations as in Fig. 1.
- 752

Fig. 4. Relative abundance of viral populations in TOV by sample. This heatmap displays the
relative abundance of each population (sorted according to its original sample; y-axis) in each
sample (x-axis). Relative abundance of one population in a sample is based on recruitment of
reads to the population reference contig, and only considered if more than 75% of the reference
contig is covered. Longhurst provinces are indicated below samples (using the same
abbreviations as in Fig. 1) and outlined in black on the heatmap.

759

Fig. 5. Relative abundance of viral populations in TOV by station. A) Evaluation of viral
 population distribution showing the number of stations (y-axis) in which each population (sorted

population abstroation bile wing the number of stations (y and) in which each population (served
 by their original station, x-axis) is distributed. Populations are grouped by station, merging
 surface and DCM samples from the same station. B) Relative abundance of populations at the
 original stations where the contigs were assembled compared to their abundance at other stations.

- 765 Boxplots are constructed as in Fig. 1.
- 766

767 Fig. 6. Taxonomic affiliation of TOV viral populations sorted by distribution and average

768 **abundance.** A population was considered as similar to a known virus when less than half of its

reference contig genes were uncharacterized, and all characterized genes had taxonomic

affiliations to the same reference genome. As in Fig. 4, the relative abundance (y-axis) is

computed for each sample as the number of bp mapped to a contig per kb of contig per Mb of

772 metagenome sequenced. Here, the relative abundance of a population is defined as the average773 abundance of its reference contig across all samples.

774

775 Fig. 7. Net movement of viral populations throughout the oceans. Calculations are based on 776 reciprocal comparison of viral population abundances between neighboring samples (see Fig. 3 777 and Methods). For each sample pair, the average relative population abundances in one sample 778 originating from a neighboring sample were calculated and compared (for example, relative 779 abundance of populations from sample A found in sample B are compared with relative 780 abundance of populations from sample B found in sample A). The sign of the relative abundance 781 difference between neighboring samples was used to estimate the movement direction 782 (arrowhead), and the absolute value of the difference was interpreted as reflecting the movement 783 magnitude (line width). Stations are labeled with station number. 'Down' and 'up' refer to net 784 vertical movement of viral populations between the surface and DCM samples at the same 785 station.



Fig. 1. Distribution of viral capsid diameters in each sample (n = 100 viruses per sample). Data are not available for samples TARA_18_DCM and TARA_70_MESO. Boxplots are constructed with the upper and lower lines corresponding to the 25th and 75th percentiles, while outliers are displayed as points. Longhurst provinces are indicated below samples (MEDI, Mediterranean Sea; REDS, Red Sea; ARAB, NW Arabian Upwelling; MONS, Indian Monsoon Gyres; ISSG, Indian S. Subtropical Gyre; EAFR, E. Africa Coastal; BENG, Benguela Current Coastal; SATL, S. Atlantic Gyre; FKLD, SW Atlantic Shelves; APLR, Austral Polar; PNEC, N. Pacific Equatorial Countercurrent).



Fig. 2. Protein cluster (PC) richness in core and pan viromes from the TOV and POV datasets. A) Accumulation curves of core and pan PCs in the TOV dataset. Vertical axis shows the number of shared (core virome) and total (pan virome) PCs when *n* viromes are compared (n = 1 to 43; from 3 to 41 only 1000 combinations are shown). Lines: i) total number of PCs (1,075,763 PCs), ii) core surface virome (710 PCs), iii) core DCM virome (424 PCs), iv) core surface and DCM virome (220), v) all samples (including the deep-ocean sample TARA_70_MESO; 65 PCs). B) Core and pan PCs in all TOV and photic-zone POV samples combined. Vertical axis shows the number of shared (core virome) and total (pan-virome) PCs when *n* viromes are compared (n = 1 to 57; from 3 to 57 only 1,000 combinations are shown). Overall, 1,323,921 PCs were identified in all viromes combined.



Fig. 3. Alpha diversity measurements in TOV dataset. A) Shannon's richness H' and Pielou's evenness J calculated from protein clusters counts for each sample and a pool of all samples, normalized to 5 million reads. B) Shannon's richness H' and Pielou's evenness J calculated from relative abundances of viral populations for each sample and a pool of all samples, with subsamples of 100,000 reads. Outliers corresponding to values outside of the average value plus or minus two standard deviations are colored in green and red, respectively. Values calculated from the pool of all samples are colored in blue. Longhurst provinces are indicated below samples using the same abbreviations as in Fig. 1.



Fig. 4. Relative abundance of viral populations in TOV by sample. This heatmap displays the relative abundance of each population (sorted according to its original sample; y-axis) in each sample (x-axis). Relative abundance of one population in a sample is based on recruitment of reads to the population reference contig, and only considered if more than 75% of the reference contig is covered. Longhurst provinces are indicated below samples (using the same abbreviations as in Fig. 1) and outlined in black on the heatmap.



Fig. 5. Relative abundance of viral populations in TOV by station. A) Evaluation of viral population distribution showing the number of stations (y-axis) in which each population (sorted by their original station, x-axis) is distributed. Populations are grouped by station, merging surface and DCM samples from the same station. B) Relative abundance of populations at the original stations where the contigs were assembled compared to their abundance at other stations. Boxplots are constructed as in Fig. 1.



Fig. 6. Taxonomic affiliation of TOV viral populations sorted by distribution and average abundance. A population was considered as similar to a known virus when less than half of its reference contig genes were uncharacterized, and all characterized genes had taxonomic affiliations to the same reference genome. As in Fig. 4, the relative abundance (y-axis) is computed for each sample as the number of bp mapped to a contig per kb of contig per Mb of metagenome sequenced. Here, the relative abundance of a population is defined as the average abundance of its reference contig across all samples.



Fig. 7. Net movement of viral populations throughout the oceans. Calculations are based on reciprocal comparison of viral population abundances between neighboring samples (see Fig. 3 and Methods). For each sample pair, the average relative population abundances in one sample originating from a neighboring sample were calculated and compared (for example, relative abundance of populations from sample A found in sample B are compared with relative abundance of populations from sample B found in sample A). The sign of the relative abundance difference between neighboring samples was used to estimate the movement direction (arrowhead), and the absolute value of the difference was interpreted as reflecting the movement magnitude (line width). Stations are labeled with station number. 'Down' and 'up' refer to net vertical movement of viral populations between the surface and DCM samples at the same station.

Table 1. Relationships between viral community structure (based on viral morphology, populations, and PCs) and metadata using NMDS analysis of all samples and the sample subset (all samples except for TARA_67_SRF, TARA_70_MESO, TARA_82_DCM, and TARA_85_DCM due to exceptional environmental conditions at these locations). Significant relationships are italicized and in bold.

		Viral Morphology	Populations	Protein Clusters
		(qTEM)	(contigs)	(PCs)
Depth Category	all samples	p = 0.354 (n = 41)	p = 0.362 (n = 43)	p = 0.033 (n = 43)
	sample subset	p = 0.228 (n = 38)	p = 0.105 (n = 39)	p = 0.011 (n = 39)
Province	all samples	p = 0.098 (n = 41)	p < 0.001 (n = 43)	p = 0.014 (n = 43)
	sample subset	p = 0.029 (n = 38)	p < 0.001 (n = 39)	p = 0.008 (n = 39)
Biome	all samples	p = 0.099 (n = 41)	p < 0.001 (n = 43)	p = 0.097 (n = 43)
	sample subset	p = 0.120 (n = 38)	p < 0.001 (n = 39)	p = 0.543 (n = 39)
Latitude	all samples	p = 0.003 (n = 41)	p < 0.001 (n = 43)	p = 0.002 (n = 43)
	sample subset	p = 0.014 (n = 38)	p < 0.001 (n = 39)	p = 0.010 (n = 39)
Temperature	all samples	p = 0.001 (n = 41)	p < 0.001 (n = 43)	p < 0.001 (n = 43)
	sample subset	p = 0.001 (n = 38)	p < 0.001 (n = 39)	p = 0.015 (n = 39)
Salinity	all samples	p = 0.118 (n = 39)	p = 0.035 (n = 41)	p = 0.029 (n = 41)
	sample subset	p = 0.138 (n = 36)	p = 0.075 (n = 37)	p = 0.001 (n = 37)
Oxygen	all samples	p = 0.001 (n = 41)	p < 0.001 (n = 43)	p < 0.001 (n = 43)
	sample subset	p = 0.005 (n = 38)	p < 0.001 (n = 39)	p < 0.001 (n = 39)
Chlorophyll	all samples	p = 0.711 (n = 41)	p < 0.001 (n = 43)	p = 0.001 (n = 39)
	sample subset	p = 0.738 (n = 38)	p = 0.412 (n = 39)	p = 0.059 (n = 39)
Nitrite	all samples	p = 0.951 (n = 39)	p = 0.648 (n = 41)	p = 0.828 (n = 41)
	sample subset	p = 0.851 (n = 36)	p = 0.509 (n = 37)	p = 0.999 (n = 37)
Phosphate	all samples	p = 0.275 (n = 39)	p < 0.001 (n = 41)	p < 0.001 (n = 41)
	sample subset	p = 0.411 (n = 36)	<i>p</i> < 0.001 (<i>n</i> = 37)	p = 0.583 (n = 37)
Nitrite+Nitrate	all samples	p = 0.046 (n = 39)	p < 0.001 (n = 41)	p < 0.001 (n = 41)
	sample subset	p = 0.290 (n = 36)	p = 0.052 (n = 37)	p = 0.643 (n = 37)
Silica	all samples	p = 0.008 (n = 39)	p = 0.002 (n = 41)	p = 0.008 (n = 41)
	sample subset	p = 0.255 (n = 36)	p = 0.285 (n = 37)	p = 0.191 (n = 37)
Bacteria	all samples	p = 0.579 (n = 39)	p < 0.001 (n = 40)	p = 0.119 (n = 40)
	sample subset	p = 0.329 (n = 36)	p = 0.003 (n = 36)	p = 0.007 (n = 36)
Low DNA bacteria	all samples	p = 0.227 (n = 39)	p = 0.090 (n = 40)	p = 0.123 (n = 40)
	sample subset	p = 0.468 (n = 36)	p = 0.018 (n = 36)	p = 0.005 (n = 36)
High DNA bacteria	all samples	p = 0.967 (p = 39)	p < 0.001 (n = 40)	p = 0.273 (n = 40)
	sample subset	p = 0.174 (n = 36)	p = 0.027 (n = 36)	p = 0.024 (n = 36)
% high DNA bacteria	all samples	p = 0.007 (n = 39)	p = 0.078 (n = 40)	p = 0.009 (n = 40)
	sample subset	p = 0.017 (n = 36)	p = 0.059 (n = 36)	p < 0.001 (n = 36)
Synechococcus	all samples	p = 0.143 (n = 39)	p = 0.094 (n = 40)	p = 0.041 (n = 40)
	sample subset	p = 0.142 (n = 36)	p = 0.023 (n = 36)	p = 0.013 (n = 36)
Prochlorococcus	all samples	p = 0.118 (n = 39)	p = 0.076 (n = 40)	p = 0.123 (n = 40)
	sample subset	p = 0.249 (n = 37)	p = 0.161 (n = 37)	p = 0.140 (n = 37)

(YAXA, ristinal)	[TARA stations service meetal feature]	Environmental Factors	Mediller of	wrony-addre contactual data yabilided of PANGAEA	na-Militani	[dagroos Herth] [dogram	East) (m	J (Longbarst 2007)	Areas (883) (M9040 registered at menuscripter option.com)	washinger and	(110)	than (see	u'i Ine	an') an	14.5 10	and lag") (pres	Na') Na) (aso'e	ar's (n)	10 ed) b	(30° ml°) (#20°	mi') (r.10'm	5 -
IAFA PH	TARA ON DOM	(DOM) deep charapter werknam lever (DMO:01006558)	TARA, A100000172	http://www.pargs.st.de/search143dap1143tA_A190000112	1009-11-0211403	26,7828	14,2765	#1 Wastarles Ginns	(MS) Medimenten Sea (MRGR:1996)	(JBDD) Medberrenkes San, Black See Prevines (MRGID:21445)	384	375	254	A13	0.05	0.02	8 42	4.62	16A	144	RA.	MA	MA
TARA OIL	TARA ON SRI	(SRF) wathin weight layer (LMVII)00002042)	TARA, A180939171	http://www.pangasa.de/wareh?kHap=TARA_A110000171	1008-11-01186-13	35.487	14,14,14	\$ Westarles Diams	(MS) Mediamanesa Sea (MROID:1996)	(MEDD Maddamarata Sas, Black Sas Prevince (MROD)214467	214	3/9	204	0.05	0.62	0.03		456	1.0	213	634	4.32	2.62
IARA DEE	TARA OLL SRF	(389) surface writer layer (LNV/IECOCCOM)	TARA SIMULICI	http://www.pasagesa.de/search7AHLgFTARA_3290002705	1009-11-10100-10	200000	12,4185	E Wastarfier Binne	(MS) Mediterrumen Sea (MRGJD:1998)	(MAD) Modernanan Sas, Black Sas Previews (MRGD-21468)	v	376	221	A.38	667	0.02	9 21	243	619	0.64	5.94	9.09	225
TARA 012	TARA 911.0CM	(DOM) does alterated markenin layer (MMVD.01006328)	TARA XIMOROI 1912	http://www.paegasa.de/eservirtAlle=TARA_X200001312	2001-11-1011220	46_173F	17,7288	65 Washinfor Binnis	(M3) Mediumanean San (MRGE:1994)	(MRDD Medbersmen Sea, Black See Prevince (MRGD):21446)	157	39.4	224	8.24	0.62	8.62	8 6 5	1.65	104	134	\$.95	935	26
IARA, INF	TARA OF DOM	(DOM) deep attangeter mentionen layer (EMVO \$1000222)	TARA LEUXIDUU06	http://www.pangene.do/exarch?AALq=TARA_EB00000005	2009-11-23413:52	10,0001	11,3997	D) Westarles Bloms	(MS) Nodimeranesa Sos (MRGID:1996)	(MEDD Madhimmake San, Black See Prevince (MRGID:21446)	15 2	305	256	÷2	667	0.67	415	122	MA	-		m	
IARA DES	1ARA 026 SRI	(SRF) surface writer layer (EMV/0.00002042)	TARA, FEROMOTIN	http://www.pangean.du/search?ABLav?A.R.A_b290000218	2009-11-23-099-12	202400	19,3996	6 Wartarfer Sinns	(MS) Hedlaruman See (HRGD:1996)	(MED) Modernana San, Bind, Son Previous (MEGD)(1008)	183	382	23.8	6 12	0.62	0.02	4.03	1.45	254	F 14	264	159	• 73
TARA 000	TARA GIO DOM	(DCM) daup alforeging? maximum layor (DWO)(1000224)	TARA, X000001318	http://www.pangasa.ds/search?ABLq+TABLA_X000001322	1001-12-1611628	ar 33,46%	32,0112	76 Westarles Blome	(WS) Hadrenson San (MROD; 1998)	(MEDD Medhamatan Sur, Black San Prevince (MINGO:21494)	70	395	210	\$14	0.42	0.02	4 42	4.67	142	3/4	\$ 20	1.94	2.65
TARA OLI	TARA,001,784	(SRF) warnes were layer (EHV0:00000012)	TARA A100001211	http://www.pangaua.do/search1ABCar1ARA_A180801311	2810-01-0818246	27,18	34,891	5 Cuestal Ficane	(REI) Red See (MRCID:A2M)	(REDS) Red Sea, Parsing Call Province (MRGID 21474)	21	-	209	0.05		0.02		4.12	• 63	24.3	to.	42	23
TARA 612	TARA 632 DCM	IDC60 deep altempty! markets layer (UWO.01006324)	1ARA, A100001510	http://www.pangasa.de/search?A36e+1A3A_A100001510	2010-01-11114-07	23,4483	31741	60 Courtal Binare	(As) Hed See (MRGID:4244)	(RLDS) Red Jun. Paralas Call Province (MRCBD\$1474)	261	40.2	196	0.22	0.02	0.03	4 82	B 85	•11	30	3.63	12	171
TARA 012	TARA 491, SRI	(SRF) earlies writer layer RMV0.000020421	TARA AIBOODISIS	http://www.pangasa.sla/eserols?ABCq=TASUA_A180001515	2010-01-010020	23.30	17,6111	5 Coortal Biome	(Ast Red See (MRG/DA244)	(Ht05) Red Sus, Paraias Guil Province (HRGID \$1174)	25 8	397	100	•	0 4 <u>2</u>	0.02	442	4.37	• 67	19.4	\$ 25	3 16	2.09
TARA OH	TARA ON DOM	(DC6U deep alteraphyl merkman layer (SMYO.01000024)	TARA RISCOUGOS	http://www.pargs-auto/search?AbbarTASIA_R110000000	En10-01-10111-00	14,4417	11.6557	M Coortal Finna	Citil Red See (MIXCID:4244)	(RCD3) Red Sea, Paralas Gulf Province (MR08021474)	274	399	175	018	629	049		135	• 63	1.42	**	MA .	HA.
TARA CH	1ARA, 534, 3RF	(SNV0.00002042)	TARA RI0000006	http://www.pangava.du/avaroh1436.gr1ARA_R10000006	2810-01-201042/	1000	39,915	5 Coortal Dieme	(RSI Red See (MRGID:R2H)	(RLSS) Red Sec. Pareira Dell' Province (MROID £16/4)	2/1	34.6	384	019	0.02	619		191	695	33	2.07	3.93	194
TARA, OH	TARA, NY DOM	IDC64 dawp obligations in a rinner layor (UMV0.01006324)	TARA RINKORDO	http://www.pargesa.de/warsh?A36g=TA55A_R180000830	1810-08-12111211	10,0221	63,6133	17 Courtal Home	(10) Index Ocean (MRGID: 104)	(ASUAD) Herthwood Acabian Son Upwalling Previous (MEDID:21475)	23.4	345	212	0.35	6.51	851	2.08	10	124		ш,	131	48
TARA, 004	TARA, GOH, SRP	(SRF) waters water layer IENV0.000000121	TARA RI 00000027	http://www.pargaua.do/asaroh?ABCartTABCAR100000017	1010-03-12190:00	EG A HES	41.0047	5 Coortal Piezo	DOI Swien Ocean (MIROID:1104)	(ARAB) Northword Arabian Say Upwalling Province (MROID (1476)	86	34.5	211	012	0.65	÷37	•10	12	11.7	151	11.4	121	636
1ARA,038	TARA, 104, DCM	(DCM) daug alteraphyl marinen layar (191/0.01004304)	TARA R/0000004	http://www.pangasaule/search7Aldor*TARA_R10000004	2010-00-05311;40	12,0244	44.812R	25 Trades Classe	UDI Sudan Ocean (MRGID;1904)	(MONA) Indian Vianasan Oyree Province (MRGID (1117)	23.5	HA.	105	••	•/	•42	1.86	165	57	20	111		1.09
TARA 031	TARA 436 SRF	(ERF) warture weter layer [EHV0,00002042]	TARA RI 0000001	http://www.pangana.de/search?A36e=TARA_R100000061	2010-03-06102,26	10,0393	464413	5 Erndes Dires	CIOI Sedan Ocean (MICOR) I POU	(SIGHS) Indian Shanasan Oyree Prevines (MRGID 21471)	76.2	346	200	616	8.02	• 31		114	14/	29.4	10.3	230	435
LARA 019	IARA, GOP DOM	(DOM deep shorehol marken lover (DNV0,01009221)	TARA, RIBOROOGIE	http://www.pangata.de/waareb?AMapr1ARA_R160000412	2010-02-01111/23	16,6430	66,612/	16 Frades Witnes	(IOI Julian Ocean (MRC20;1204)	(J60H1) Indian Manusch Optim Province UKRGAU (14/1)	26.4	54.3	193	0.38	0.02	\$26	0 0 9	1.38	754	22.4	9.62	7.0	33
1454,033	TARA 029, SRF	(SRF) earlies water layer (EMV/0;00002042)	TARA RIBORODATE	http://www.purgesaude/waarthTAMageTASIA_R10000011	2010-03-1010621	14,9914	et.egg	5 Trades Blome	IIOI Index Coven (MRGID:1994)	(MONS) befor, Manager, Dynes Previore (MRCaD (1471)	26.4	34.5	393	• •	6.62	623	443	144	14.7		6.95	40	4.52
TARAON	INRA ONI DOM	(DCM) dauge africange by a new larger (RHVD).010005.549	1ARA,R10000438	http://www.pangasa.do/search?AblantTANA_R10000464	2810-02-30110;91	14,9556	100124	Of Trades Chanes	(IDI bulan Casan, (MRCM);1994)	GROWSS Indian Managem Oyree Prevince (MRG2D (1473)	21	34.5	144		0.07	034	0.49	1.54			105	310	1.0
TARA OHI	TARA 044,950	(SNV0:00000001)	TARA RIDORCOUTS	http://www.pangawa.dk/search?Alk.ap 1ASUA_R100000455	2810-00-301024/	14,0059	41.8776	3 Trades Dimes	(ICI Index Doesn (MICOD) IPOI)	(MORN) Falles Minasim Oyres Previous (MiCaD (1111)	B1		20.	8.64	6.02	•1•	0.49	147	1.10	14.	7.09	4.0	1.75
TARA,042	TARA OILDCM	LDCM deep chierephyl mertenen leyer (MMYO,01000524)	TARA R100000152	http://www.pungenaute/www.httlawTARA.R10000152	1010-01-01130-50	23991	21,0042	bi Trades Diese	(IOI Indust Down (MRCID;1904)	CHORESP Indian Manason Oyres Province UNRCED \$1471	u)	351	50	637	672	634	139	32	144	39	322	1 59	163
TARA, OHE	TARA ONL SHIP	(SHI) surface water layer IEHVD:00000000	TANKA, R180800149	http://www.pangane.de/march1All.g=1A36A_R110000149	2010-04-04105.47	E dite i	12,6965	5 Trudes Planes	(ID) Sulan Down (MRCID; (IN)	(MONS) Indian Memory Oyres Province UNITOD 21471		346	159	0.01	0.02	0.04	045	221	137	24.6	709	433	10
TARA OH	TARA ON SRI	(SRF) markee welter lever (ENVD;60002042)	1A/0A/R10000000	http://www.parganal.de/sugrek?Allig=TANA_R110000408	2010-04-1310201	-64125	71.161	5 Frades Dises	(ID) Indian Count (MRGID; IPH)	(BONS) Falles Minasian Oyres Prevince (MICOD 21471)	201	351		012	0.02			194		5	10	542	241
FARA, DOZ	TARA, WIZ, DOM	(DOM) deep shiprophyl martmen layer (CMV0.01000024)	TARA R 00000134	Migc//www.pangasa.de/search1AMar=1ASA_R180007254	2010-05-17111-02	-169934	571001	/S Trades Diams	(ICI Instan Court (MRCAD) POIL)	(USIC) Indian Statis Exportation Cyty Province (MPCED21477)	24.5	347	133	0.34	0.00	016		112	4.24	100		30	1.91
LANA, 084	TARA, MOLDON	EDCMI deep objected markeen layer (EMVO 01000024)	TARA, RI 0000015	http://www.pangeau.de/march?Ald.g=TA30A_R190000015	2010-07-0010021	-18.8535	47,0112	ES Courtui Bhrmo	(IDI Indian Dynam (MRCAD; I (CA)	(EAPIU Eastern Ables Courtal Province (MRCED)21473)	23	353	20/	621	0.02	0.04	442	1/5	5 75		56	3/4	3/1
TARA CHI	TARA OF SRU	(ERF) surface mater layer (EXV/D,00002042)	1ARA A100000322	http://www.pangenaude/www.ch?Aling=1AttA_R180000012	2010-07-07194-0	-18.9010	27.5665	5 Goertal Blows	(ID) Sullah Gepah (MHGID;104U	(EO IO Caller Alice Geeta Previce (Alica)(1173)	at	13.1	2.	014			6 V Z	111		21			
TARA INT	TARA, OBS, DOM	ID CALL dauge shierophyd manimum laywr (LMVO.01000024)	TARA RIDOODIIN	http://www.pangeaa.de/search?Allar=TARA.R100001114	2010-05-12111-03	21 -352421	26 2041	30 Coortal Blone	UCI Index Cosen (MRC20:1994)	UATIO Lagine Ales Caula Prerves (MICD 21473)									14	201			•15
TANA, ONS	TARA, 045, ERF	(SR*) and so swith by w (LMVD;00002042)	TARA RIDODOTESO	MBp://www.panga.sa.de/aparth1A34apt1AStA_R1900011218	2010-01-1210525	-43,1725	34,3441	5 Coordial Blocket	COLEMEN Down (MICOD) POU	(DV IO Califier View Carlos builded and a 211/3)	A1	17.4		۰u								300	
1474,098	TARA 266 DOM	(DOMI does alterated maximum layer (EMVD.01000222)	TARA RI0000000	MBg://www.tpanga.au.de/search?AMapiTARA_R100000066	1010-01-1511538	10 -34,6301	14,0451	38 Coortal Rismo	(EAD) South Atlantic Ocean (MRGID:1914)	LECHOL Bargenia Current Desetal Province (MIGUD 21470)	13	33.3	240	0.42	0.54	• 37	323	24/	1.54	142	103	435	4/1
IARA,064	TARA, ONL, SRI	(UN7) surface writer layer (UNVO;00000040)	TARA RI 00000000	http://www.parearearearearearearearearearearearearea	2010-01-1511201	-31344	124145	5 Coastal Blowe	(EAD) Seeth Adaptic Cosan (MPGID:1914)	LEXICI Surgaria Current Deastal Province UNICID:114/01	13	33.3	200	625	• • •		3.94	LIA	135	1.19	14	4.1	10
TARA 047	TARA OF SHE	(SRF) warbers weber leyer (ERVD:00002042)	TARA, R10000051	http://www.parguea.de/wearchTALLapTANA_R/10000051	2010-00-01100:11	-51.5401	12,2140	5 Geertal Bioare	(SAD) Seeth Allerthe Deven (MPGID:1914)	(BEHC) Berganta Current Deantal Prevince (MICLO 31410)			249	135	617	1.02		13.00	18.3		151	2.44	
TAPA DH	TARA ON DOM	LDCSII deep and see in the lower layer (LMYO,C1000224)	TARA RIDOODDIS	http://www.pangs.ve.ds/weards1All.enTASA_R18000915	2010-00-14113:30	~31,007	41892	50 Trades Shows	EAD) South Allerdio Death (MHOLD;1914)	(SATE) SAMA AMARINE OVID HIGHNALE CAROLINE THE		<u></u>				12	1.0				2.03	10	
3454,061	1684,981,597	(SRF) suffree werter layer (ENVD)00000043	TARA, R.(00000018	http:// energy sector/ meanshil/Addam TARA, RI (COCCOLD	1010-00-1410431	-31,0204	4.895	3 fradra titma	(SKO) Seets Applies Onen (Brand) (and	CALC SHER AGAINS ON TOWARD CHIMAD 214141		35.7	154	••	0.5							12	
1454,070	TARA GIO UR SO	(MLS) mesopologie zone (MVV0.90000213)	LYNY BURGENESE	http://www.pangate.de/een/ef/14864/1494.je/fcco/rig/y	1010-00-111112/	54 *32.4975	-71041	BOD TREASE DISH.	ISKOD SERVI ANAMAS OFEN DATOD, IFTO	CALC SOUR AGAINS OF THE TOWNER CONCAVATION			194			2.14					10		
JARA D/D	TARA OTO SRI	(SRF) surface weter layer (LNV/D.00002048)	TARA RIDOODIDIS	http://www.pangaoa.de/wearch?AlderTAPA.RI10001815	1010-00-11100-51	-12.4041	-21121	5 Trader Sinte	ISADI Seeth Allartic Denni (MPCID)(\$14)	(SATO SHIR ADDIS OVE PROPER (BRODZINS)			216					1.04	1.2		10	171	
IARA, DII	TARA, 6/2, DCM	(DCtill dree alterupted masterias layer (DIVO.01006023)	TARA 8100001002	http://www.exesses.es.de/wearoh14.mag~14.PA.JR190001962	2010-10-0011930	-1.1310	-12.000	ICC Frades Street	(3YO) 2640 VOME OF THE DRIVES	SALD SHER AGAINS OVER HOWING CHRONIZING					0.42		0.02						
IARA,972	TARA 4/2, SRF	(SRF) auflice water layer (EHVO:00001942)	TARA RI00001079	http://www.paressia.de/www.chttAlda=TASA_SU60001819	2010-10-06100-00	-0.7760	-17,8099	5 Feeder Billion	ISADI Seeth Allentis Desan (MCMD;1914)	(SATU SHARADING OVER PROVINE GROUD THEFT			199			41	0.05			174		247	54
1484,070	TARA GRUDCH	EDCMI does objerce tell maximum layar (EMVO.01004522)	TARA RIPODOTIZO	http://www.pargate.ate/searchTALLer1ASA_R180801111	1010-10-10110-09	-210002	-11,3491	HC Treder States	(EAD) Sauth Allande Usean, (MHOLD)(1114)	ISATO SAUK AGANGE OTHE PROVIDE DIROUCTION					1.42		0.02				103		
1484,0/0	TARA, O'N, SRF	(205) autors weter leyer (UNVD:00402042)	TARA, R100001124	http://www.burgesaude/abareh1AALer=TARA_R180001118	1010-10-14109,64	-10,5354	-11.1003	8 tredes Bhitte	(pr. or prep, Aname Cesan (MRGD) 1914	CAN ID SHARE ADAVID VYTE ITEVING UNROID 11400	23	541		= 03	1.42		390			10.0	20		2.07
TARA, DEZ	TARA, NR, DOM	IDCMI does chlorighyd maximum lwywr (DAVO 01000524)	TARA RI 80800344	http://www.batapasa.ate/wwatch?Alder/TARA.Al100000644	2010-25-0011194	-47,3067		so central Bloose	134.07 SHOR ADAMON CHAIN (MRCCO):1914)	WANT SAUDER AND AND FURTHER UNDER 1				2.02		-4	30						
TAPA DI	TARA DAL DOM	ID050 data chloradoll martenin layar RHV0.01000328)	TARA, R100001317	http://www.pargesaude/www.chTAMup/TAMA_R100001317	1011-01-00113-34	21 -912231	-41,2135	16 Paler Blains	CPU) Sauthern Cassa (MRGID:198/)	USINA AND THE PROPERTY OF CONTRACT		24.5	345	- 2		431		P. 12			1.94		-14
TARA IM	TARA, 101, DCM	IDDMI deep alterapted martenan layer (CHVO.01000028)	TARA RIBOROIIII	http://www.pargessi.de/wearch?Alitup?TASIA_R100001810	2011-09-11122/1	2.0255	-141144	ar usstil films	CAPUT PROFES PRODUCE COMMA (MRCAD) 19081	Control Camponers Contraint Control Preventes (1988) (2022) 1439-1	14.5									254			
TARA IN	TARA, 109, 38J	(\$95) auftace water layer (LNV0;00002042)	TARA RIPOROIDO	http://www.parapara.de/waarch1Allag=TARA_R100001509	1011-05-1111440	13128	-041701	S Genetal Blooks	(NPO) Neyth Pscille Ocean (MIGRO;1908)	(DBU Date Para Generit Grantal Prevince (URGICCT1976)	<i>v</i> .	35.4	197	. (F	0.04			1.13	1.34		10.5	1.11	

Station Identitie (TARA_station#)	(TARA_station#_onvicus montal-feature]	Sample label [TARA_station9_environmental- feature_size-traction]	Size fraction inner throth [micrometre]	s Bins fract ald apper the (micromet	na shuld 18300 ma seassion nanànris) ra)	Corresponding mateotices data published at BNA	Paired cook	reads /	Aligned A ments P	Reade in the	nd Am	adı in desi poletinan Ta	nalifica Uniqu	n PG Int	nanos's X' for PCs arms land to SM reads)	Piología Evenenes (A Por PC (normalizad Va Silt rands)	i Sheenen's X' fer 72s is siran cont (secondizat to 50th conds)	ign Local-Global diversity ratio for f (normalized to Shi reads)	(normalized to 100	nal y ag-alation k younds)	e Pisieu's Luceness (A for sits) populations (normalized to JBM reads)	Local:Global dipensity patte (jacarmatized to 1200 reeds)
TABA.014	TARA DIR DOM	TARA DIA DOM 4-022		<u>ر</u>	0.22 ERRSHITT	http://www.ebiac.uk/www/inter/www/ER#574152	41/5635	10512712	25174900	233570/1	1793334	1024291	702/4	49925	24		•/	2.54	0.01	,	м	3.87
NRA DIA	TABA QIA SRP	TARA DIA SRF (-0.12		<	0.22 ERR294350	http://www.ekianak/arm/data/view/ERRE14358	4750850	5 15017812	45106031	44220183 3	(04/711	2179284	\$4270	56767	10		.06	7.39	0 41	,	63	1.41
BA 022	TARA DEL SRF	TARA DIZ SRF (-0.22		<	0.22 ERR394379ERR394409	http://www.ablan.uk/wwa/data/view/ER/R514378J/RR514406	\$757324	1 115146412	50519251	45060623 4	011110	1623367	149546	78791	19.43			6.62	0.005		46	1.73
ARA 073	TARA 023 DOM	TARA 023 DCM 4-022		<	0.12 ERR394404	http://www.ablac.uk/ene/data/view/ERUIS76400	4138/02	5 12/74050	30778047	20011014 3	10/0031	1454814	120126	66984	15.02		.07	8.16	0.01		23	1.92
ARA 023	TARA 021 DOM	TARA DES DOM 4-022		۲.	0.22 CRA694978	http://www.ablacub/ens/data/view/DUR514375	\$255475	105112/30	46691361	434/9717 2	100507	1382904	149410	92250	10.47			1.89	0.06	6	20	190
TARA 025	TARA 025 SRF	TARA DES SRF. 5-922		۲.	0.22 ERR594399	http://www.ubjac.ub/wew/dela/view/ERR614314	\$1347	104229666	46737761	41739880 3	609305	2205600	134834	642.59	19.44			\$21	0.56	6	31	1.91
TARA 020	TARA 010 DCM	TARA 530 DCM 4-0.22		۲.	0.22 ERREPANE	http://www.sklac.sk/wws/data/siew/ERR\$14495	3909431	7 70100734	301195/3	20438464 2	91994	1686322	05995	53433	99		187	7.65	041	5	/6	190
TARA 021	TARA 011 SRF	TARA 011.587. (+0.12		۲.	0.22 ERRST4401 ERR834410	http://www.sklasuk/www/data/view/ERE534401_ERE534410	66796324	133976492	64232168	38034883 3	3746026	2644591	162635	100396	10.19		196	6.14	0 65	6	.40	1.91
TARA 032	TARA 012 DCH	TARA OF DCH. CO22		< .	0.22 ERRSPACED	http://www.eblacub/www/deta/view/ERR514390	6464374	5 129187490	45773312	44945740	588548	449423	\$6432	15564	\$49			631	0.78		22	10
TARA 012	TARA 032 38F	TARA DJI SRF (-Q.1)		<	0.32 ERR04010	http://www.ebian.uk/www/deta/view/ERURS74310	5075653	5 101513150	3934/181	33569647 2	1043755	17/14/3	136282	/4633	10.4			6.33	0.85		2.	1.91
TABA 024	TABA ON DOM	TARA DIS DOM C-022			O.EZ ERRITAIN	http://www.shian.uk/ora/deta/view/EURST4390	4064875	\$1697510	27/61/99	23512131 2	475336	1441845	91/05	56060	10.42		1.50 E	7.93	0.62	\$.86	1.89
TANA 034	TABLE DOLL SRF	TARA 016 SRF (-0.22		÷.	0.22 EXPORTEM NEWSTAT	http://www.abiaouk/aca/deta/view/ERUE314368 URUE314370	613950	122/50156	3089/504	19425495 1	731130	3116554	116162	99615	19.36		ez	1.66	0.45		45	122
ARA DOF	TARA ON DOM	TABA ON DOM C-0.22		÷	9.22 ERR814402	http://www.ablacutt/arm/data/view/EU0014402	3052590	0 7/251800	20199575	20000195 1	/92668	854077	00058	31517	9.9/			1.11	0.42	,	44	4.87
TAGA AND	TADA DIE SOC	TAGA 018 SEF (-0.17		è	0.72 FAREN47AD	http://www.sblac.uk/ana/data/dow/ERR614368	3715154	7 74323174	26670063	23298945 1	1340632	79266	62308	31090	97			125	0.00	,	59	116
T404 618	TARA BIR OCH	TABA DIE DONI C-018		÷	0.77 FERSIAL MERSIAN	http://come.shiac.ab/ana/data/view/ERR\$14310 ERR\$14301	4712143	17 M242878	26/69138	24913425 2	143/42	162,8648	108957	50428	10 27			715	0.85	5	29	190
T404 018	TABA DIE SOF	TABA AND SRF C-0.32		÷	0.72 FAR2942200788.114400	http://www.eblacus/ana/data/view/ERR\$14374_ERR\$14400	5702443	114042560	34124947	31008543 1	1873066	8906/5	120496	61967	10.14		47	2.55	0.83	5	46	1.90
TARA ON	TABA BY DUN	TABA 010 DOM 6-017		è	OIL ERRITANISTIAN?	http://www.etian.uk/www/data/view/URE914303_ERE534307	4541611	3112226	11953764	20308973 1	B06449	633684	111796	38162	20.22			74	0.64	5	A5	3.99
100.000	TABLE DES COLO	TABLE 018 SEE (-011			0 17 E0 01 41 10 E0 01 4145	http://www.ablac.uk/aca/data/dau/FRRTA158 FRRTA345	3/24546	114493726	33219196	30500500 2	000000	1249016	\$42067	60413	20.55		- 80	1.09	0.87	6	4)	4.91
484 041	TABLE DAL DOW	TABA MI DOM COTI		2	OT I EPOTA MERDY MANE TO DE TANT 7 1158 PTRATT 2 158 PTRATT	http://www.ablanuk/are/data/view/IR/0714301/578574307/5785514271/EUU314972/EU8914273	1000244	3 200048945	642 82 190	57894189 2	755513	1169127	143206	9/446	10.26	i		2.0	0.04	6	.01	4.89
	7484 041 085	TARA ONI SEE CORIS			A 22 PEDauábuá	http://www.shi.aou/k/ara/data/data/ERRIN314	426588		1/983356	16942944	776079	334610	90032	50839	10.41		191	644	0.86	5		3,90
	TABLE INT DOW	TARA MI DOW COTT			Att Epplatent	http://www.ablac.uk/wwa/data/views/ER/014411	6000430	2 12003601	27960507	27764917 1	142362	2014	120704	82624	10.25		187	7.45	0.84	6	36	3,50
				2	Att Constantions fram)	http://www.shine.sh/www/fdata/shur/FDRSHAMI FRUITIANI	4915621	B 99112476	30237421	20033037 1	12 14 07	1146194	100198	60985	18		195	614 .	0.62	,	4 .	20
TADOLOGIC		7474 04 07 402		2	All Constants	http://www.shime.ck/www/data/data/FUD914378	3071202	1 7/415642	25554107	23934644	490412	223822	10427	\$7973	2.77		186	7.26	0.00	,	45	4.90
1000,044	TABLE DRIVE SHO	7484 08 000 000		2		http://www.eblac.uk/www/data/view/FUIP1143364	464364	0 92993760	29985215	25854768	1901209	1114077	115731	74634	10 26			7.86	0.84		.02	4.91
1000,002	THE STORE	7404 012 000 000		2		http://www.eblacc.it/acm/data/dam/FRIPHIA107	4782564		200/3137	10/07557	75/0/7	406180	103357	67393	10.10		. 24	6.00	0 34		19	4.91
				2		http://www.shime.ch/ante/view/ERD162307	100 710	0 96542150	12185492	10175459	1648065	1296512	12/452	69/64	10 16			749	0.04		4	a.92
10000044	induction and					http://www.shines.duine.duine.duine.dup.the.wr.Constants	5610041	1111006014	110647.14	12000000 3	161160	141115	199111	71854	10.2		15	1.09	0.05	6	1	4.92
ARA OID	TARA DIS DOM	CARCENS DON CHILL		2		http://www.ekias.co/acce/acce/acce/acce/acce/acce/acce/ac	COLUMN	0 1101/004	20000774	2/041979 1	1053670	611645	163053	\$5/96	19.37		35	5.90	0.04		62	0.52
TADGA DED	TADA, DID, SHO	1000,000,500,0022				Mary / www.eliantic and Jan (days (days / DOTA/14)	460004	10 10 10 10 10 10 10 10 10 10 10 10 10 1	1010117	11246311	11111	1027385	109741	44925	10		35	154	9.82	ŝ	24	0 92
ARA, OFF	TASA DELOCH	TARGE OF DEAL CORE		2		http://www.ekias.ch/acc/data/dam/TDPCN111	410486	16 100000	1644/744	ME27911 3	1914505	2054397	120464	\$4621	10.21		27	195	9.86		28	9.92
ASCA, OHI	TARA DIRESIO	1000,000,587,0-022		2		http://www.elias.ch/ang/dag/dag/dag/dag/dag/dag/dag/	419770		1/7 10150	11001146	401501	1708018	00544	/0819	10		35	1.63	081	ś	90	0.90
AXA,047	1420-017-510	1000,007,507,0022		2		http://www.eliante/ang/data/data/data/data	4003044		229912296	21106306	1177597	8/9745	91497	44117	1.9			7.4	0 82			4.87
	TABLE OF COM	1000 011 000 0021		2		http://www.ski.ac.ck/acac/data//200194901	6481577	-	1314 9492	11695053	1418283	1197573	122997	/561/	10 04			7.48	0.83		.22	4.92
ACC UNIT	12102040.000	1000,010,002,022					411630		3 16 /14/20	ascession 1	116 5016	1044434	61961	11004	*47			6.60	0.78		.90	4.60
ARA UTO	TAGE OF MESO	1000 010 1005 - 022		2		http://www.shieu.ch/accidente/interior	4731944		10001144	20100246	192671	1010751	133412	88178	10.3		11/	7.55	0.85	6	.02	0.00
TARA 070	14364_070_510	TARA 070, SHP. C-0.22			012 EXCOLUTS	http://www.ecca.do.en/ web/ toro/coverses/	4001331		11111111	Termsed 1	thrus (a)	150265	994/78	ED1 m				7.62	0.82		64	0.00
TARA 072	TARA 0/2 DCM	TAKA 072 DEN (*022		- 2	0.22 EXISTAT	http://www.education.com/and/org/and/and/a	hadarit		20549134	11406.000 1	1400000	MINO	1046	540.58	10.1			7.61	0 23		04	0.00
TARA 0/2	TANA DIZ SHF	1AHA 0/2 3HF (-022		1	U		599,2014		4574710	414412/10 1	INGISAR	2100040	110/0	80248	10 12			*11	0.87		37	0.91
1484,014	TARA DIT DON	TANA OF DOM CORE		1	ALL EXAMINES		Salara 1		41254500	104/07148	243533	1755070	121463	41741	10.4		59	7.69	0.85		39	0.89
TANA_0/0	TANA DIE SKE	TARA 0/1, 3HP, C-0.22		2	N.L. EUGITAIN	http://www.states.com/data/com/filleration	4187771	In In Italia	19062414	17965081	1000214	1215124	11/01	24062			35	6.01	0.73			0.33
TARA OIZ	TARA OR DOM	TARA OF DOM COT			0.22 239354409		443 8984		14444111	111111111	11464016	600670	(541)	41835				6.93	0./6		64	0.84
TARA CES	TARA DIS DOM	TARA OUT DOM (-021		1	037 510(3143)1		46/09		10100	STREET, STREET, ST		16413.05	1000	11/100	1016			7.87	0.04			0.05
TARA, 109	TARA_101, DOM	TARA_100_DCM_C-0.22		<u>.</u>	0.22 2.0033435/JE20134434242424045443811[EAV/5943836CRR394387	WWW/WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW	6215473		334 37622		1000272	1110070	10001		1026				0.05		19	0.91
TARA, 109	TARA, 199, SRF	TARA_109_SRP_<-0.22		۰.	0.22 EJ04094412	Durke, v Miller dotrettift, etter, anne, stean staat it	4341494	14 37431477	31,236400	********		******	******	******	1027						-	

.