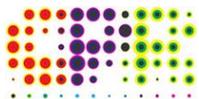


Standards Promoting and Uptake Activities at the Hellenic Centre for Marine Research

Evangelos Pafilis

Institute of Marine Biology, Biotechnology and Aquaculture (IMBBC)
Hellenic Centre for Marine Research (HCMR), Heraklion Crete, Greece

pafilis@hcmr.gr, <http://epafilis.info>





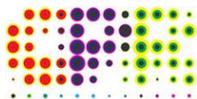
Database, 2016, 1–10
doi: 10.1093/database/bav126
Original article



Original article

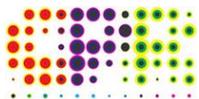
Value, but high costs in post-deposition data curation

Petra ten Hoopen^{1,*}, Clara Amid¹, Pier Luigi Buttigieg²,
Evangelos Pafilis³, Panos Bravakos³, Ana M. Cerdeño-Tárraga¹,
Richard Gibson¹, Tim Kahlke⁴, Aglaia Legaki³, Kada Narayana Murthy⁵,
Gabriella Papastefanou³, Emiliano Pereira⁶, Marc Rossello¹,
Ana Luisa Toribio¹ and Guy Cochrane¹

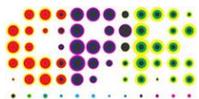


Environmental context metadata annotation

- **Interactive**
- **Lightweight**
- **Term look up assistant**
- **Standards-compliant term suggestions**
- **Metagenomics records**



EXTRACT





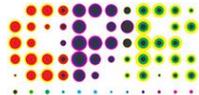
<http://www.environmentontology.org/>

Buttigieg PL, *et al.* 2013, J Biomed Semant.4:43.



<http://www.ncbi.nlm.nih.gov/Taxonomy>

Benson DA, *et al.* 2009, NAR



Bioinformatics, 2015, 1–3

doi: 10.1093/bioinformatics/btv045

Advance Access Publication Date: 24 January 2015

Applications Note

Data and text mining

ENVIRONMENTS and EOL: identification of Environment Ontology terms in text and the annotation of the Encyclopedia of Life

Evangelos Pafilis^{1,*}, Sune P. Frankild², Julia Schnetzer^{3,4}, Lucia Fanini¹, Sarah Faulwetter¹, Christina Pavloudi¹, Katerina Vasileiadou¹, Patrick Leary⁵, Jennifer Hammock⁶, Katja Schulz⁶, Cynthia Sims Parr^{6,†}, Christos Arvanitidis¹ and Lars Juhl Jensen^{2,*}

<http://environments.hcmr.gr>,

<http://environments.jensenlab.org>

OPEN ACCESS Freely available online

 PLOS ONE

The SPECIES and ORGANISMS Resources for Fast and Accurate Identification of Taxonomic Names in Text

Citation: Pafilis E, Frankild SP, Fanini L, Faulwetter S, Pavloudi C, *et al.* (2013) The SPECIES and ORGANISMS Resources for Fast and Accurate Identification of Taxonomic Names in Text. PLoS ONE 8(6): e65390. doi:10.1371/journal.pone.0065390

<http://species.hcmr.gr>, <http://species.jensenlab.org>

EXTRACT

Interactive Extraction of Metadata

extract@hcmr.gr

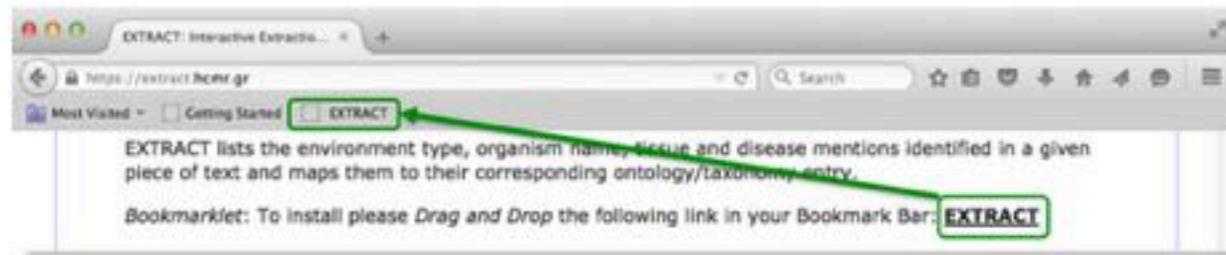
About

Demo

Help

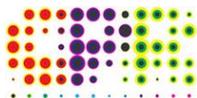
EXTRACT lists the environment type, organism name, tissue and disease mentions identified in a given piece of text and maps them to their corresponding ontology/taxonomy entries.

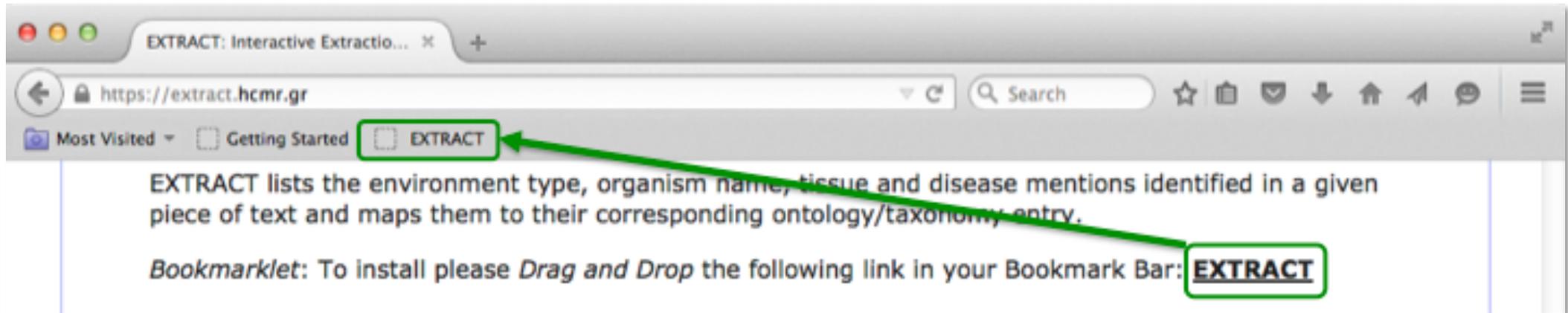
Bookmarklet: To install please *Drag and Drop* the following link in your Bookmark Bar: [EXTRACT](#)



Usage: a. select a piece of text of interest in a web page and then b. click on the bookmarklet. c. A pop-up such as the following will appear (supported browsers: [Chrome](#), [Firefox](#), [Safari](#)). By hovering the mouse cursor over the text tags or the table rows you can visually inspect which words have been identified as which entities.

EXTRACT: interactive extraction of environment metadata and term suggestion for metagenomic sample annotation Pafilis E, Buttigieg PL, Ferrell B, et al.. (2016). **Bioinformatics**, 2016, baw005. doi:10.1093/bioinformatics/btv04





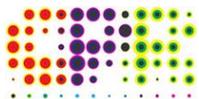
EXTRACT: Interactive Extractio... ✕ +

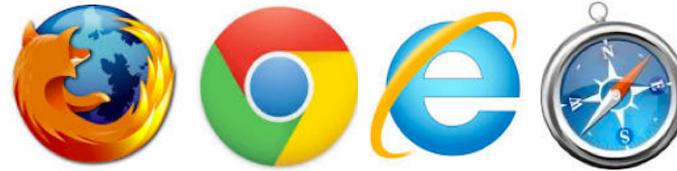
https://extract.hcmr.gr

Most Visited ▾ Getting Started EXTRACT

EXTRACT lists the environment type, organism name, tissue and disease mentions identified in a given piece of text and maps them to their corresponding ontology/taxonomy entry.

Bookmarklet: To install please *Drag and Drop* the following link in your Bookmark Bar: **EXTRACT**



A screenshot of a web browser window. The address bar shows 'https://extract.hcmr.gr'. The page content includes a description of EXTRACT and a bookmarklet installation instruction. A green arrow points from a boxed 'EXTRACT' link in the text to a boxed 'EXTRACT' bookmarklet in the browser's bookmark bar.

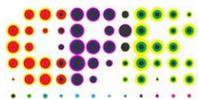
EXTRACT: Interactive Extractio... ✕ +

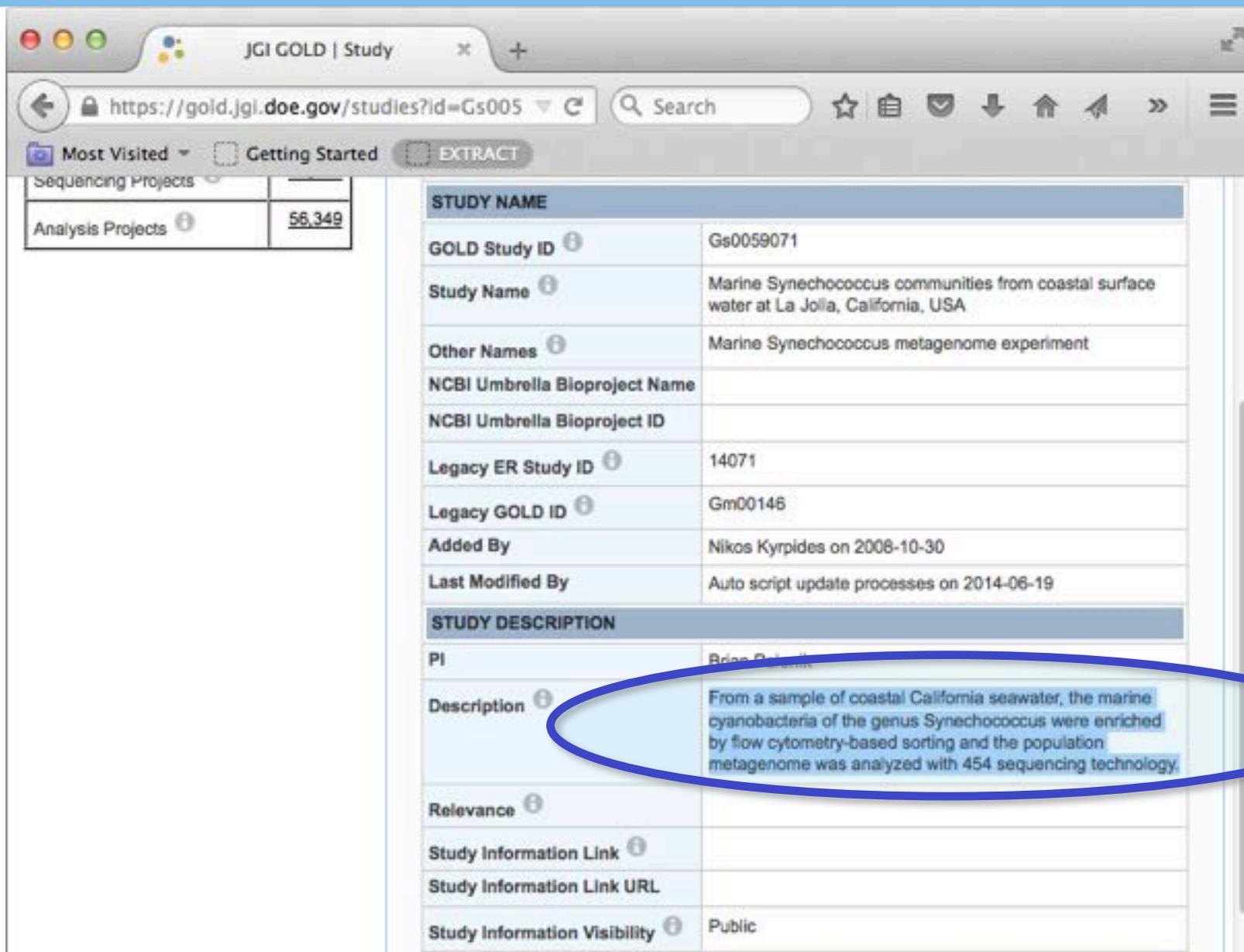
https://extract.hcmr.gr

Most Visited ▾ Getting Started **EXTRACT**

EXTRACT lists the environment type, organism name, tissue and disease mentions identified in a given piece of text and maps them to their corresponding ontology/taxonomy entry.

Bookmarklet: To install please *Drag and Drop* the following link in your Bookmark Bar: **EXTRACT**





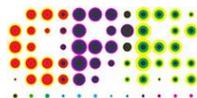
Sequencing Projects

Analysis Projects **56,349**

EXTRACT

| STUDY NAME | |
|-------------------------------|---|
| GOLD Study ID | Gs0059071 |
| Study Name | Marine Synechococcus communities from coastal surface water at La Jolla, California, USA |
| Other Names | Marine Synechococcus metagenome experiment |
| NCBI Umbrella Bioproject Name | |
| NCBI Umbrella Bioproject ID | |
| Legacy ER Study ID | 14071 |
| Legacy GOLD ID | Gm00146 |
| Added By | Nikos Kyrpides on 2008-10-30 |
| Last Modified By | Auto script update processes on 2014-06-19 |
| STUDY DESCRIPTION | |
| PI | Brian Palnitkar |
| Description | From a sample of coastal California seawater, the marine cyanobacteria of the genus <i>Synechococcus</i> were enriched by flow cytometry-based sorting and the population metagenome was analyzed with 454 sequencing technology. |
| Relevance | |
| Study Information Link | |
| Study Information Link URL | |
| Study Information Visibility | Public |

<https://gold.jgi.doe.gov/studies?id=Gs0059071>



[Show help page](#)
[Open popup in a new tab](#)
[Close](#)

EXTRACT ? ↑ ×

Selected text

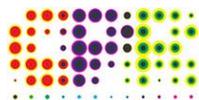
From a sample of **coastal** California **seawater**, the marine **cyanobacteria** of the genus **Synechococcus** were enriched by flow cytometry-based sorting and the population **metagenome** was analyzed with 454 sequencing technology.

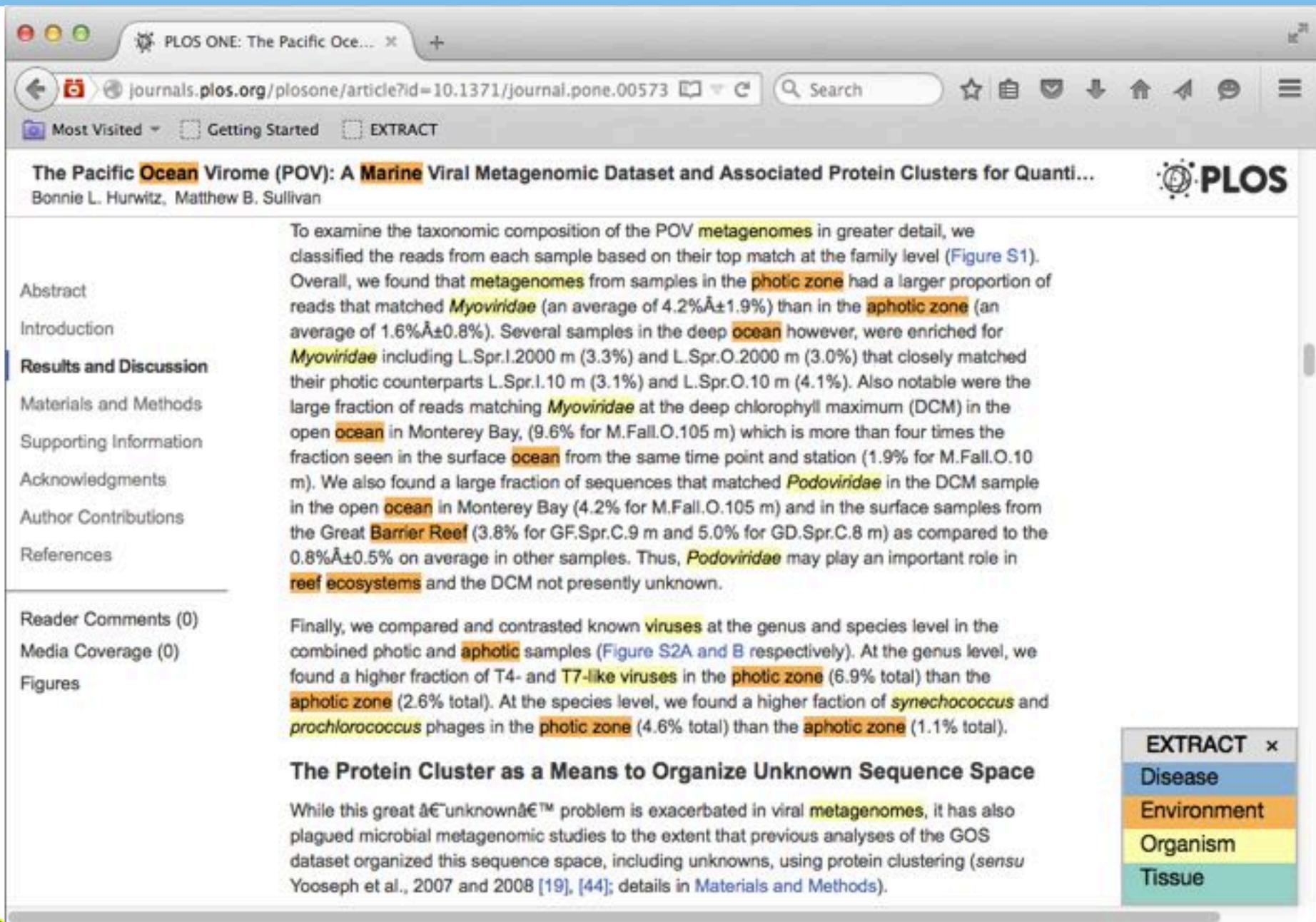
Identified terms

| Type | Name | Identifier |
|-------------|---------------|-------------------------------|
| Environment | Coast | ENVO:00000303 |
| Environment | Sea water | ENVO:00002149 |
| Organism | Cyanobacteria | 1117 |
| Organism | Metagenomes | 408169 |
| Organism | Synechococcus | 1129 |

Annotated user selected text
 Identified entity summary table
 Highlighted on mouse over: related tags and entities
 Additional information link

Copy to Clipboard and **Save** as tab separated values the list of extracted entities along with the selected text and the source page URL





The Pacific Ocean Virome (POV): A Marine Viral Metagenomic Dataset and Associated Protein Clusters for Quantification
Bonnie L. Hurwitz, Matthew B. Sullivan

Abstract

To examine the taxonomic composition of the POV metagenomes in greater detail, we classified the reads from each sample based on their top match at the family level (Figure S1). Overall, we found that metagenomes from samples in the photic zone had a larger proportion of reads that matched *Myoviridae* (an average of $4.2\% \pm 1.9\%$) than in the aphotic zone (an average of $1.6\% \pm 0.8\%$). Several samples in the deep ocean however, were enriched for *Myoviridae* including L.Spr.I.2000 m (3.3%) and L.Spr.O.2000 m (3.0%) that closely matched their photic counterparts L.Spr.I.10 m (3.1%) and L.Spr.O.10 m (4.1%). Also notable were the large fraction of reads matching *Myoviridae* at the deep chlorophyll maximum (DCM) in the open ocean in Monterey Bay, (9.6% for M.Fall.O.105 m) which is more than four times the fraction seen in the surface ocean from the same time point and station (1.9% for M.Fall.O.10 m). We also found a large fraction of sequences that matched *Podoviridae* in the DCM sample in the open ocean in Monterey Bay (4.2% for M.Fall.O.105 m) and in the surface samples from the Great Barrier Reef (3.8% for GF.Spr.C.9 m and 5.0% for GD.Spr.C.8 m) as compared to the $0.8\% \pm 0.5\%$ on average in other samples. Thus, *Podoviridae* may play an important role in reef ecosystems and the DCM not presently unknown.

Introduction

Finally, we compared and contrasted known viruses at the genus and species level in the combined photic and aphotic samples (Figure S2A and B respectively). At the genus level, we found a higher fraction of T4- and T7-like viruses in the photic zone (6.9% total) than the aphotic zone (2.6% total). At the species level, we found a higher fraction of *synechococcus* and *prochlorococcus* phages in the photic zone (4.6% total) than the aphotic zone (1.1% total).

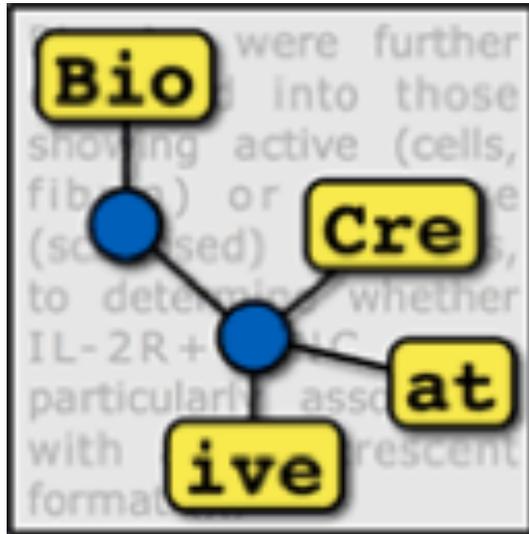
Results and Discussion

The Protein Cluster as a Means to Organize Unknown Sequence Space

While this great “unknown” problem is exacerbated in viral metagenomes, it has also plagued microbial metagenomic studies to the extent that previous analyses of the GOS dataset organized this sequence space, including unknowns, using protein clustering (*sensu* Yooseph et al., 2007 and 2008 [19], [44]; details in Materials and Methods).

EXTRACT ×

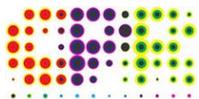
- Disease
- Environment
- Organism
- Tissue



<http://www.biocreative2015.org>

BioCreative V: Interactive Annotation Task (IAT) Dr. L. Hirschman, Dr. C. Arighi *et al.*
Challenge: **March – August 2015**
Presentation: **September 2015**, Sevilla, Spain
Metagenomics Record Annotation Session
(Department of Energy [DE-SC0010838])

- **Easy**: installation, tagging a web page, invoke the popup by processing selected text, saving results to a file
- **Adequate** NER accuracy (4/10 reported FN/FP, still two of them satisfied)
- **Speedup** in the range of 15–25%.
- Saving by time by avoiding looking up the ENVO identifier for every term
- **Manual** document inspection still needed
- EXTRACT helps **in finding terms** that would have been missed by the curators (e.g. due to non-familiarity with terminology)
- Average score 8.3 out of 10: they would **recommend EXTRACT**



EXTRACT

Interactive Extraction of Metadata

extract@hcmr.gr

About

Dem

Help

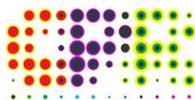
Please find below:

- Practical tips on how to use EXTRACT
- Curation assistance points
- Technical points and troubleshooting cases
- Using EXTRACT within other resources

Points in blue are a good starting point as they provide you with basic information about the EXTRACT bookmarklet, such as how to install and use the bookmarklet, and the EXTRACT popup description. Some points about record annotation with standardized metadata are listed afterwards (in green), followed by troubleshooting cases (in orange). Information on how to use EXTRACT within other resources can be found at the end (in purple).

[Show All](#) / [Hide All](#)

- ▶ How do I install the EXTRACT bookmarklet?
- ▶ How do I use EXTRACT?
- ▶ How can I use the EXTRACT popup for curation?
- ▶ Which types of entities can EXTRACT identify?
- ▶ How can I enlarge the EXTRACT summary popup?
- ▶ How can I use EXTRACT on documents that are not web pages?
- ▶ Why should I annotate samples with standards-compliant metadata?
- ▶ How should I annotate an outdoor sample with environment metadata?
- ▶ How can I annotate a host-associated/disease-related sample?
- ▶ Can EXTRACT suggest sections to study in a full-text article?

<https://extract.hcmr.gr>

EXTRACT

Interactive Extraction of Metadata

extract@hcmr.gr

About

Dem

Help

► How can I add the EXTRACT popup in my own web pages?

▼ Can I invoke the EXTRACT tagger programmatically?

In addition to the high-level `ExtractPopup` web method used in the previous section, EXTRACT offers a robust and fine-grained Application Programming Interface (API) to its named entity recognition engine. The core methods of this REST API are presented below:

GetEntities

`GetEntities` (<http://tagger.jensenlab.org/GetEntities>) returns the unique list of the entities identified in the document. The entities belong to the specified `entity_types` and the response follows the specified `format`.

Request:

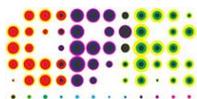
```
http://tagger.jensenlab.org/GetEntities?
document=Both+samples+were+dominated+by+Zetaproteobacteria+Fe+oxidizers.+This+gro
up+was+most+abundant+at+Volcano+1,+where+sediments+were+richer+in+Fe+and+containe
d+more+crystalline+forms+of+Fe+oxides.&entity_types=-2+-25+-26+-27&format=tsv
```

Response:

```
Zetaproteobacteria    -2    580370
sediments            -27    ENVO:00002007
Volcano -27          ENVO:00000247
```

| Parameter | Type | Content |
|-----------|----------|---|
| document | required | the plain or html-formatted text to be tagged |

<https://extract.hcmr.gr>, screenshot cropped for illustration purposes



EXTRACT

Interactive Extraction of Metadata

extract@hcmr.gr

Prospective record annotation

About

Dem

Help

► How can I add the EXTRACT popup in my own web pages?

▼ Can I invoke the EXTRACT tagger programmatically?

In addition to the high-level `extractPopup` web method used in the previous section, EXTRACT offers a robust and fine-grained Application Programming Interface (API) to its named entity recognition engine. The core methods of this REST API are presented below:

GetEntities

`GetEntities` (<http://tagger.jensenlab.org/GetEntities>) returns the unique list of the entities identified in the document. The entities belong to the specified `entity_types` and the response follows the specified `format`.

Request:

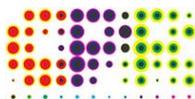
```
http://tagger.jensenlab.org/GetEntities?
document=Both+samples+were+dominated+by+Zetaproteobacteria+Fe+oxidizers.+This+gro
up+was+most+abundant+at+Volcano+1,+where+sediments+were+richer+in+Fe+and+containe
d+more+crystalline+forms+of+Fe+oxides.&entity_types=-2+-25+-26+-27&format=tsv
```

Response:

```
Zetaproteobacteria    -2    580370
sediments            -27    ENVO:00002007
Volcano -27          ENVO:00000247
```

| Parameter | Type | Content |
|-----------|----------|---|
| document | required | the plain or html-formatted text to be tagged |

<https://extract.hcmr.gr>, screenshot cropped for illustration purposes



EXTRACT

Interactive Extraction of Metadata

extract@hcmr.gr

About

Dem

Help

Prospective record annotation

► How can I add the EXTRACT popup in my own web pages?

▼ Can I invoke the EXTRACT tagger programmatically?

In addition to the high-level `extractPopup` web method used in the previous section, EXTRACT offers a robust and fine-grained Application Programming Interface (API) to its named entity recognition engine. The core methods of this REST API are presented below:

GetEntities

`GetEntities` (<http://tagger.jensenlab.org/GetEntities>) returns the unique list of the entities identified in the document. The entities belong to the specified `entity_types` and the response follows the specified `format`.

Request:

```
http://tagger.jensenlab.org/GetEntities?
document=Both+samples+were+dominated+by+Zetaproteobacteria+Fe+oxidizers.+This+gro
up+was+most+abundant+at+Volcano+1,+where+sediments+were+richer+in+Fe+and+containe
d+more+crystalline+forms+of+Fe+oxides.&entity_types=-2+-25+-26+-27&format=tsv
```

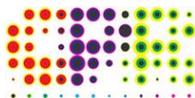
Response:

```
Zetaproteobacteria    -2    580370
sediments             -27    ENVO:00002007
Volcano -27          ENVO:00000247
```

| Parameter | Type | Content |
|-----------|----------|---|
| document | required | the plain or html-formatted text to be tagged |

JSON-LD, OpenAnnotation compliant
tagger API available
in collaboration with Dr. S. Pyssalo and
Prof. Lars Juhl Jensen

<https://extract.hcmr.gr>, screenshot cropped for illustration purposes



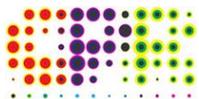
New version of SPECIES (rule based extension for improved bacteria strain name matching)

- Top place in BioNLP – Shared Task 2016 BB3 (April 2016)
Bacteria Biotope 3 - Event extraction of microorganisms and habitats with ontologies and their linking, in collaboration with Ms. Helen V. Cook and Prof. Lars Juhl Jensen (BB-cat-ner subtask)

<https://sites.google.com/site/bionlpst2016/tasks/bb2>

EXTRACT v2.0: More entity types

- Tissues
- Chemicals
- Biological Processes
- Molecular Functions
- Cellular Components
- talk coming up in ICBO/BioCreative 2016 by Dr. Lars Juhl Jensen





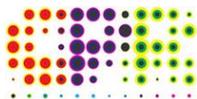
Database, 2016, 1–10
doi: 10.1093/database/bav126
Original article



Original article

Value, but high costs in post-deposition data curation

Petra ten Hoopen^{1,*}, Clara Amid¹, Pier Luigi Buttigieg²,
Evangelos Pafilis³, Panos Bravakos³, Ana M. Cerdeño-Tárraga¹,
Richard Gibson¹, Tim Kahike¹, Aglaia Legaki³, Kada Narayana Murthy⁵,
Gabriella Papastefanou³, Emiliano Pereira⁶, Marc Rossello¹,
Ana Luisa Toribio¹ and Guy Cochrane¹





LifeWatchGreece Research (e-)Infrastructure



<https://www.lifewatchgreece.eu/>

<https://portal.lifewatchgreece.eu/>



LifeWatchGreece - Data Management Team

Go to

Collaborators

Dr. Christos M. Tsipras
He has been involved in a PhD student in the University of Crete, Department of Biology, currently in the Marine Biology Station of Crete, Department of Biology. His main research interests focus on biodiversity and taxonomy of plants, and more specifically on the macroalgae of Crete. His scientific field also includes molecular phylogenetics and phylogeography of plants, as well as field research on Crete and South Aegean islands. He has studied tropical field areas in Belize and has conducted field research in several Mediterranean islands. He is currently working on the development of a Biodiversity Atlas.

Dr. Vasilios J. Zotos
He is currently a post doc researcher in the Department of Environmental and Natural Resources Management (University of Patras, Greece). He has been researching focus on plant taxonomy, more specifically, the vascular plants of the island of Crete. He has also conducted studies regarding the reproductive biology of the plants of the island.

Dr. Vasilios J. Zotos
He is currently a post doc researcher in the Department of Environmental and Natural Resources Management (University of Patras, Greece). He has been researching focus on plant taxonomy, more specifically, the vascular plants of the island of Crete. He has also conducted studies regarding the reproductive biology of the plants of the island.

Dr. Vasilios J. Zotos
He is currently a post doc researcher in the Department of Environmental and Natural Resources Management (University of Patras, Greece). He has been researching focus on plant taxonomy, more specifically, the vascular plants of the island of Crete. He has also conducted studies regarding the reproductive biology of the plants of the island.

Dr. Vasilios J. Zotos
He is currently a post doc researcher in the Department of Environmental and Natural Resources Management (University of Patras, Greece). He has been researching focus on plant taxonomy, more specifically, the vascular plants of the island of Crete. He has also conducted studies regarding the reproductive biology of the plants of the island.

Dr. Vasilios J. Zotos
He is currently a post doc researcher in the Department of Environmental and Natural Resources Management (University of Patras, Greece). He has been researching focus on plant taxonomy, more specifically, the vascular plants of the island of Crete. He has also conducted studies regarding the reproductive biology of the plants of the island.

Dr. Vasilios J. Zotos
He is currently a post doc researcher in the Department of Environmental and Natural Resources Management (University of Patras, Greece). He has been researching focus on plant taxonomy, more specifically, the vascular plants of the island of Crete. He has also conducted studies regarding the reproductive biology of the plants of the island.

Dr. Vasilios J. Zotos
He is currently a post doc researcher in the Department of Environmental and Natural Resources Management (University of Patras, Greece). He has been researching focus on plant taxonomy, more specifically, the vascular plants of the island of Crete. He has also conducted studies regarding the reproductive biology of the plants of the island.

Dr. Vasilios J. Zotos
He is currently a post doc researcher in the Department of Environmental and Natural Resources Management (University of Patras, Greece). He has been researching focus on plant taxonomy, more specifically, the vascular plants of the island of Crete. He has also conducted studies regarding the reproductive biology of the plants of the island.

Dr. Vasilios J. Zotos
He is currently a post doc researcher in the Department of Environmental and Natural Resources Management (University of Patras, Greece). He has been researching focus on plant taxonomy, more specifically, the vascular plants of the island of Crete. He has also conducted studies regarding the reproductive biology of the plants of the island.



<https://www.lifewatchgreece.eu/?q=content/collaborators>

✓ Coordination of a team of 20 scientists;

✓ Representatives of different disciplines, as well as different Institutes & Universities all over Greece;

✓ Acting as data collectors, data providers and data managers



LifeWatchGreece – **Data Management**

data management tasks can be summarized as:

- ✓ **Collecting & Receiving data** from national, regional and international programmes
- ✓ **Collecting & Digitizing Historical biogeographic data;**
- ✓ **Verifying the quality** of the data (using agreed upon standards);
- ✓ **Reporting the results** of quality control directly to data providers as part of the quality assurance process;
- ✓ **Making data available**, nationally and internationally through MedOBIS and GBIFgreece ipt and
- ✓ Ensuring the **long term preservation** of the data and associated information required for correct interpretation of the data



LifeWatchGreece – c. Managing Data

Taxonomic Quality Control

| | | |
|----------------------------------|---|---|
| Nemalion helmintoides | → | Nemalion helmintoides |
| Didemnids | → | Didemnidae |
| Littorina saxatilis | → | Littorina saxatilis |
| Gastropods | → | Gastropoda |
| Prorocentrum triestnum | → | Prorocentrum triestinum |
| Indetermined | → | Biota or Higher Classification if given |
| Corallina sp | → | Corallina |
| Corallina spp | → | Corallina |
| Jania rubens/Haliptilon virgatum | → | ??? |

- NEMERTINEA ----- NEMERTEA
- Polydontes frons ----- Polydontes frons
- Amaena tribolata ----- Amaena trilobata
- Asyhis biceps ----- Asyhis biceps
- Magelona papilicornis ----- Magelona papilicornis
- Ochnesoma steenstrupi ----- Ochnesoma steenstrupi
- Aplakophora ----- Aplacophora
- GASTEROPODA ----- GASTROPODA
- Strobiformis subulata ----- Strobiformis subulatus

- ✓ typical taxonomic errors are checked (misspelling, invalid names, inconsistencies, misidentifications)
- ✓ taxon names are linked to one of the following bases: World Register of Marine Species (WoRMS), Catalogue of Life (CoL), ITIS, IRMNG, Paleobiology Data, International Plant Name Index (IPNI) & Euro+Med Plantbase, Index Fungorum, Fishbase, Barcode of Life Database (BOLD) and Web of Science.
- ✓ If no link possible relevant papers and data provider(s) are consulted
- ✓ originally delivered name is always safeguarded!



LifeWatchGreece – c. Managing Data

Spatial Quality Control



Long/Lat expression

| | |
|---------------------------|----------------------------------|
| Decimal Degrees | lat: 35.241078° long: 25.153364° |
| Degrees, Minutes, Seconds | 35°14'27.88" N 25° 9'12.11" E |
| Degrees, Decimal Minutes | 35° 14.465' N 25° 7.202' E |
| UTM | Zone 35N |

CONVERSION

Degrees, Minutes, Seconds → Decimal Degrees

$$\text{Degrees} + ((\text{minutes}/60) + (\text{seconds}/3600)) = \text{dd}$$

35°14'27.88" N 25° 9'12.11" E

$$35 + (14/60) + (27.88/3600) = 35.241078$$

$$25 + (9/60) + (12.11/3600) = 25.153364$$



- Coordinates are always:
 - ✓ converted in WGS84 geodetic reference coordinate system (decimal degrees)
 - ✓ plotted on map and visual checks are performed

When not available:

- ✓ are derived from provided map, <http://www.geodata.gov.gr>, related paper etc
- ✓ Marine Regions webpage is checked to assign coordinates
- ✓ Data provider is contacted





LifeWatchGreece – c. Managing Data

Darwin Core:



<http://rs.tdwg.org/dwc/terms/index.htm#theterms>

✓ a body of standards intended to facilitate the sharing of information about biological diversity

✓ a glossary of terms meant to provide stable semantic definitions with the goal of being maximally reusable

| Region | Country | Institution or organisation | Data owner | Sampling area | Name of sampling station | Geographic coordinates of the station (WGS 84 - decimal format) |
|--|--|---|---|---|--|---|
| DwC term: <i>waterBody</i> ... <i>The name of the water body in which the location occurs</i> | DwC term: <i>country</i> ... <i>The name of the country or major administrative unit in which the location occurs</i> | DwC term: <i>institutionCode</i> ... <i>The name (or acronym) in use by the institution having custody of the object(s) or</i> | DwC term: <i>rightsHolder</i> ... <i>a person or organization owning or managing rights over the</i> | DwC term: <i>locality</i> ... <i>The specific description of the place</i> | DwC term: <i>locationRemarks</i> ... <i>Comments or notes about the Location.</i> | DwC term: <i>verbatimCoordinates</i> ... <i>The verbatim original spatial coordinates of the Location.</i> |

| Weather type (sunny, cloudy, rain, storm) | Sample track code (country-station-level-plot-replicate) | Species (number of individuals)/Stations | Number per sample-unit | WORMS AphiaID |
|--|--|--|--|--|
| DwC term: <i>fieldNotes</i> ... <i>the text of notes taken in the field about the Event</i> | DwC term: <i>fieldNumber</i> ... <i>an identifier given to the event in the field</i> | DwC term: <i>taxonRemarks</i> ... <i>the original name received</i> | DwC term: <i>individualCount</i> ... <i>The number of individuals represented present at the time of the Occurrence</i> | DwC term: <i>taxonID</i> ... <i>An identifier for the set of taxon information (data associated with the Taxon class)</i> |

[Home](#)[Manage Resources](#)[About](#)Resource Title [CINCS: Pelagic-Benthic Coupling In the oligotrophic Cretan Sea](#)

Basic Metadata

The resource title and description are required. The resource's three main contact's must also be described here: Resource contact, resource creator and metadata provider. For each contact you must supply at least a last name, a position or an organisation before you can make the resource public. The person(s) or organisation(s) responsible for the creation of the resource as it appears in the IPT and for effectively publishing the resource should add themselves as an associated party with role 'publisher'.

Title*

Description*

Metadata Language

Resource Language

Type

Subtype

Section

[Basic Metadata](#)

[Geographic Coverage](#)[Taxonomic Coverage](#)[Temporal Coverage](#)[Keywords](#)[Associated Parties](#)[Project Data](#)[Sampling Methods](#)[Citations](#)[Collection Data](#)[External links](#)[Additional Metadata](#)



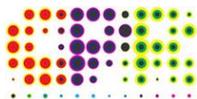
Database, 2016, 1–10
doi: 10.1093/database/bav126
Original article



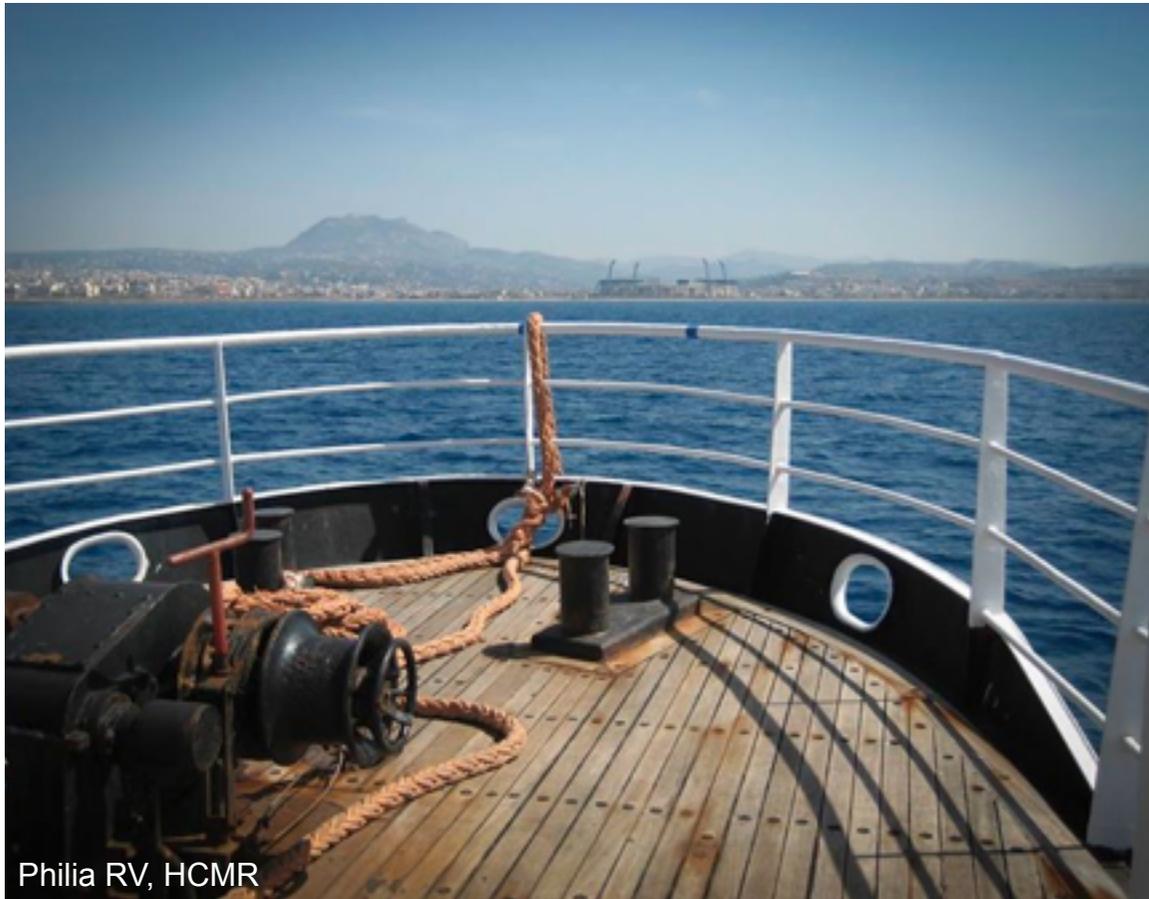
Original article

Value, but high costs in post-deposition data curation

Petra ten Hoopen^{1,*}, Clara Amid¹, Pier Luigi Buttigieg²,
Evangelos Pafilis³, Panos Bravakos³, Ana M. Cerdeño-Tárraga¹,
Richard Gibson¹, Tim Kahlke⁴, Aglaia Legaki³, Kada Narayana Murthy⁵,
Gabriella Papastefanou³, Emiliano Pereira⁶, Marc Rossello¹,
Ana Luisa Toribio¹ and Guy Cochrane¹







Philia RV, HCMR

Heraklion, May 2014, MicroB3 Summer School



Niskin bottle

www.osil.co.uk



CTD

<http://www.seabird.com/sbe25plus-sealogger-ctd>





OSD Main Event 2014 Pipeline

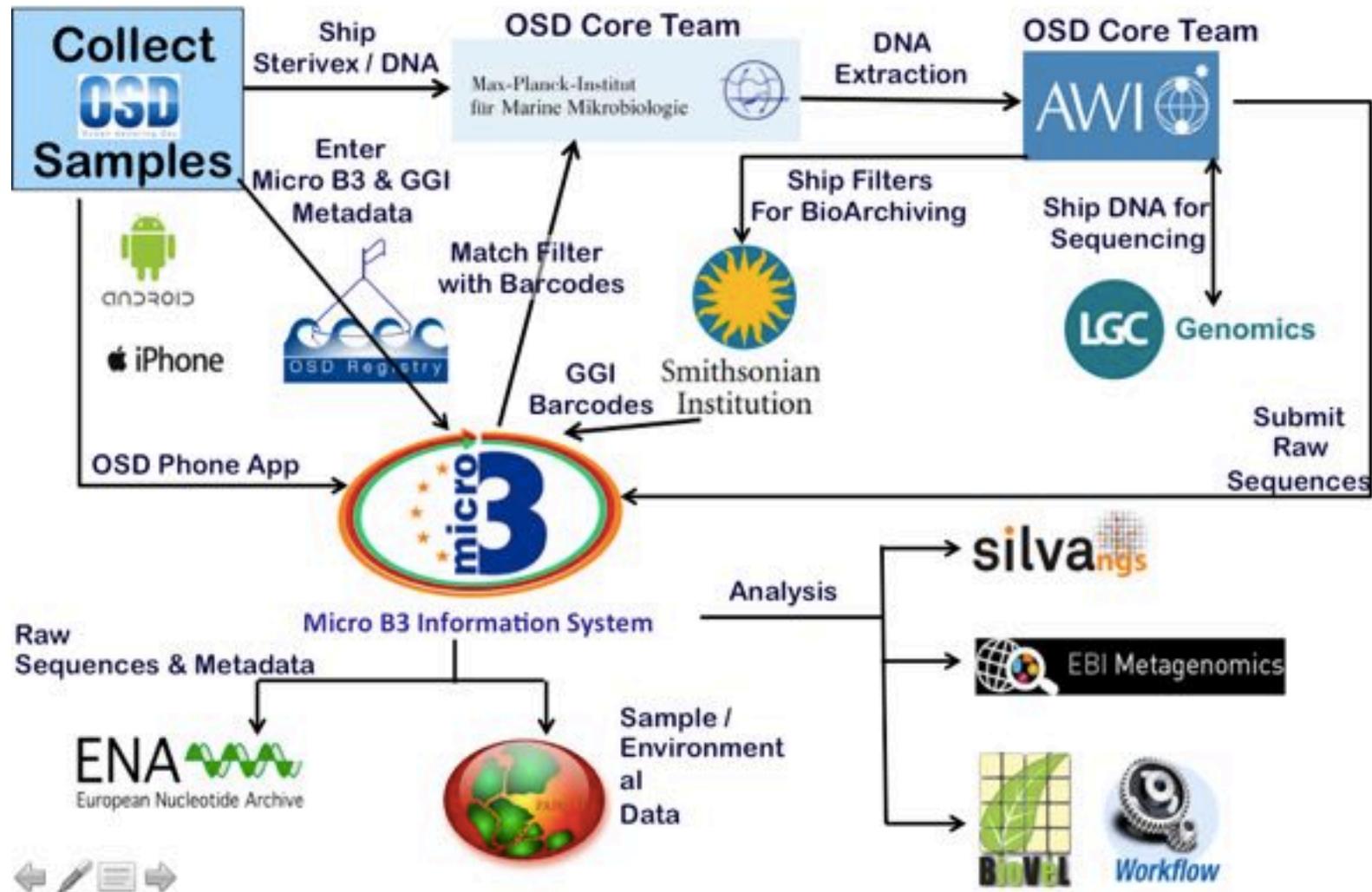


image: <http://oceansamplingday.blogspot.com>

Annex I OSD Logsheets

BEFORE

| | | | | |
|--------------------|---------------------------|---------------------------------|------------|--------------------|
| | SAMPLING_Site (ID, Name): | OSD-SUMMER-SCHOOL | | |
| | SAMPLING_Platform: | RV Philia | | |
| | SAMPLING_Campaign: | OSD-SUMMER-SCHOOL | | |
| | SAMPLING_Project: | MicroB3-OSD2014 - summer-school | | |
| SAMPLING | SAMPLING_Investigators: | Last name | First name | email |
| | | CHRISTAKIS | CHRISTOS | christakis@hcmr.gr |
| | | GROUP 2 | | |
| SAMPLING_Objective | summer school training | | | |

DURING

| | | | | | | | | |
|----------------|--------------------------------|-------|---------|----|----|-------|----------|-----|
| EVENT_DateTime | Start | yyyy | mm | dd | hh | mm | ss | UTC |
| | End | 2014 | 05 | 27 | 14 | 15 | 00 | +2 |
| EVENT_Lat/Long | Start | +N/-S | dd.0000 | | | +E/-W | ddd.0000 | |
| | End | | 35.2172 | | | + | 25.15833 | |
| EVENT_Device: | NISKIN BOTTLE | | | | | | | |
| EVENT_Method: | BOTTLES IN-LINE WITH MESSENGER | | | | | | | |
| EVENT_Comment: | | | | | | | | |

AFTER

| | | |
|---------------------------|---|------------------------------|
| ENVIRONMENT_Marine_Region | e.g. Adriatic Sea | HERAKLION BAY - SEA OF CRETE |
| ENVIRONMENT_Biome | e.g. ENVO:00000447 for "marine biome" | |
| ENVIRONMENT_Feature | e.g. ENVO:00000569 for "marine habitat" | |
| ENVIRONMENT_Material | e.g. ENVO:00002042 for "surface water" | |

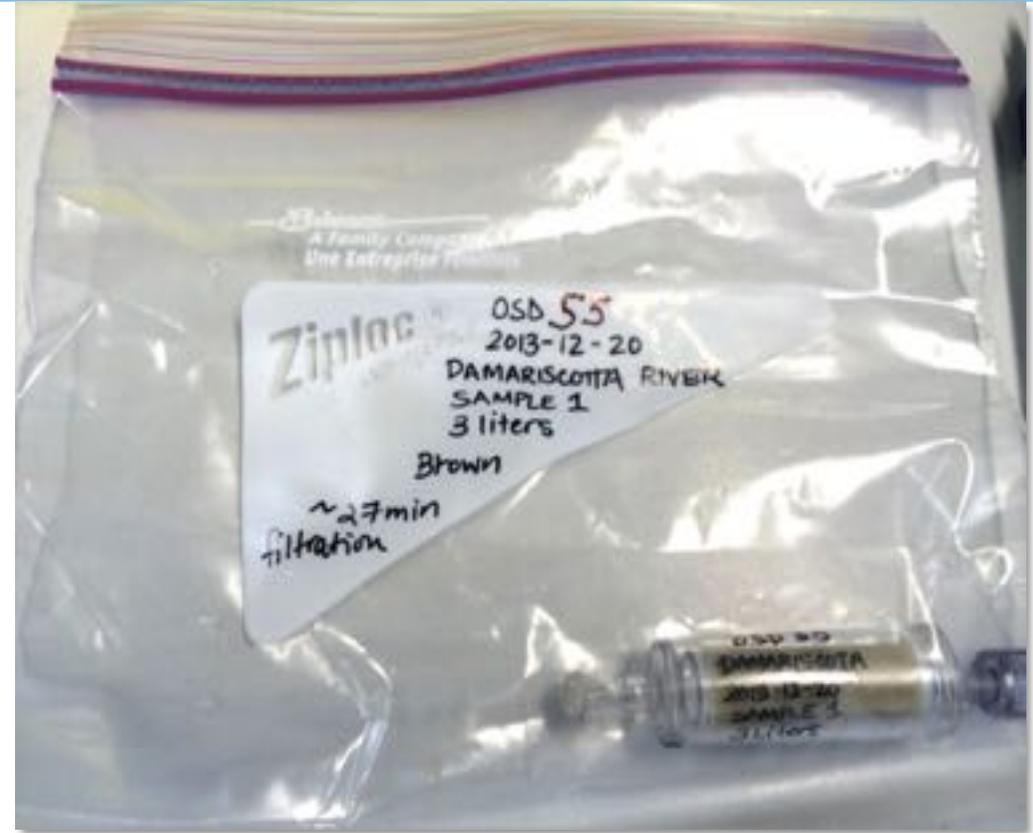
Stephane Pesante
Renzo Kottman
Christos Christakis
MicroB3 Summer
School Team

- One sampling event **per bag**
- Use **freezable** bags
- Label **filter and bag**
- Seal the label
 - Use **adhesive** tape
 - not parafilm
- International shipping
 - Pack filters on **dry ice**
 - Dispatch samples on a **Monday**
 - Specify: **no commercial value content**, fill-in: **custom forms**



Photo and slide input by Dr. A. Klindworth

Label filter and bag



<OSD_SiteID>_<Month>_<Year>_<SiteName>_<SampleNo>_<Depth>

e.g. OSD3_06_14_Helgoland_1_surface

e.g. OSD3_06_14_Helgoland_1_20m

e.g. OSD5_06_14_Crete_1_2m

Photo and slide input by Dr. A. Klindworth

Thank You!

EXTRACT is based on components from: ENVIRONMENTS, SPECIES, Reflect <http://reflect.ws>
Developed by the LifeWatchGreece Research Infrastructure and the group of Dr. Lars Juhl Jensen (Uni Copenhagen); with input from the groups of Genomes OnLine Database, Virome / Metagenomes Online, MegX.net and Dr. Pier L. Buttigieg (AWI). **BioCreative**: All evaluators, Dr. Cecilia Arighi and Dr. Lynette Hirschman (DoE Award No DE-SC0010838)

LifeWatchGreece/IMBBC: Christos Arvanitidis, Dimitra Mavraki, Stamatina Nikolopoulou, Sarah Faulwetter, Nikos Minadakis (ICS/FORTH) and many more (<https://www.lifewatchgreece.eu/>)

NNF CPR: Dr. Sune Frankild, Helen Cook, **MM-MPI**: Frank Oliver Glöckner, Dr. A. Klindworth, Renzo Kottman, E. Pereira, Dr. Julia Schnetzer *et al.* **Uni Delaware**: Dr. Barbra Ferrel, **EBI**: Dr. Guy Cochrane, Dr. Petra ten Hoopen *et al.* **Pangea**: Dr. Stephane Pesant, **HCMR/IMBBC**: Dr. P. Polymenakou, Dr. G. Kotoulas, Dr. Anastasios Oulas, Anastasia Tsiola, Christina Pavloudi, Gabriela Papastefanou, Aglaia Legaki, **UoA**: Panos Bravakos. Many more I might be omitting

