LifeWatch e-Science European Infrastructure for Biodiversity and Ecosystem Research



ATTIVITA' IN CORSO & PRIORITA' STRATEGICHE The Biomolecular Thematic Centre and MoBiLab

Graziano Pesole Università di Bari & IBBE-CNR, Bari, Italy

CONFERENZA ANNUALE LIFEWATCH ITALIA Biblioteca della Accademia Nazionale delle Scienze 16 Dicembre 2015 ROMA



Molecular Biodiversity



One of the emerging topics in the **post-genomic era** is represented by the study of **Molecular Biodiversity** aimed at understanding and classifying the **taxonomic and functional diversity** of life forms, whatever their size and their level of organization...







CONFERENZA ANNUALE LIFEWATCH ITALIA Biblioteca della Accademia Nazionale delle Scienze 16 Dicembre 2015 ROMA







Life

ch



In **Biomolecular Thematic Centre (CTB)** skills and advanced facilities for molecular and bioinformatics analyses are integrated to provide the scientific community with **services** and **counselling** for Molecular Biodiversity studies.





The activities of the Thematic Centre are supported by Molecular Biodiversity Laboratory (MoBiLab), with its fully operative platforms based on the most innovative NGS technologies and powerful resources of data storage and computational analysis. In addition to their support to LifeWatch, the services hosted by MoBiLab contribute to the Italian node of the European infrastructure ELIXIR.

About MoBiLab



Advanced molecular and bioinformatic infrastructure aimed at supporting genomic, metagenomic, phylogenetic, phylogeographic and DNA barcoding data analysis in biodiversity studies.

MoBiLab is currently equipped state of the art advance instrumentations including Illumina HTS platforms.



- Meta-barcoding and shotgun protocols based on high-throughput sequencing (HTS) technologies;
- tools for molecular and bioinformatic analyses of free-living/symbiotic prokaryotic and eukaryotic communities;
- 3. tools for molecular and bioinformatic analyses of genomes and their expression from prokaryotes and eukaryotes.

ON LINE AT:

http://www.servicecentrelifewatch.eu/web/lifewatch-italia/the-biomolecular-thematic-centre



ICT tools & services Genomic and Metagenomic Ready-to-use Tools & Services ICT tools & services Meta-barcoding Multiple Alignments Data Analysis **MSA-PAD BioMaS** Specialized Sequence and Database Metadata Retrieval **ITSoneDB** Shotgun Metagenomic DataRetrieval **HMAdb** Data Analysis: MetaShot

Meta-barcoding Data Analysis BioMaS



BioMaS: a modular pipeline for Bioinformatic analysis of Metagenomic AmpliconS

Bruno Fosso¹, Monica Santamaria¹, Marinella Marzano¹, Daniel Alonso-Alemany², Gabriel Valiente², Giacinto Donvito³, Alfonso Monaco³, Pasquale Notarangelo³ and Graziano Pesole^{1,4,5*}



Available online at https://recasgateway.ba.infn.it/







(CrossMark



BioMaS is an automated pipeline able to quickly convert Meta-barcoding NGS Illumina and Roche 454 raw data in taxonomic characterization of environmental microbial communities.

Specialized Database ITSoneDB

BRIEFINGS IN BIOINFORMATICS. VOL 13. NO 6. 682–695 Advance Access published on 10 July 2012



doi:10.1093/bib/bbs036

Reference databases for taxonomic assignment in metagenomics

Monica Santamaria*, Bruno Fosso*, Arianna Consiglio, Giorgio De Caro, Giorgio Grillo, Flavio Licciulli, Sabino Liuni, Marinella Marzano, Daniel Alonso-Alemany, Gabriel Valiente and Graziano Pesole

Submitted: I4th March 2012; Received (in revised form): 2nd June 2012

SIMPLE SEARCH ADVANCED SEARCH	
Fungal Ribosomal Internal Transcribed Spacer 1 Database	
ITSoneDB is a comprehensive collection of the fungal ribosomal RNA Internal Transcribed Spacer 1 (ITS1) sequences aime metagenomic surveys of fungal environmental communities. The sequences were extracted from GenBank (GB) and arrang taxonomy tree. ITS1 start and end boundaries were defined by GB annotations and/or designed by mapping Hidden Marko profiles of flanking 185 and 5.85 ribosomal RNA coding genes on each sequence. <i>Current GenBank release: 202 (June 2014)</i>	d at supporting jed on the NCBI v Model (HMM)
Number of the nucleotide sequences	491,919
Total number of species (according to the NCBI taxonomy)	40,855
Total number of species (according to the NCBI taxonomy) Number of ITS1 nucleotide sequences with start and end positions inferred by both GB information and HMM profiles	40,855 34,316
Total number of species (according to the NCBI taxonomy) Number of ITS1 nucleotide sequences with start and end positions inferred by both GB information and HMM profiles Number of ITS1 nucleotide sequences with start and end position inferred only by HMM profiles mapping	40,855 34,316 101,253

CNR Istituto di Tecnologie Biomediche, Bari - CNR Istituto di Biomembrane e Bioenergetica, Bari - Università degli studi di Bari

ITSONEDB is a curated collection of taxonomically annotated ITS1 sequences suitable for metagenomic studies of fungal communities.

Available online at http://itsonedb.ba.itb.cnr.it:8080/ITS1

Specialized Database

HMAdb



HMAdb: a specialized Reference database of 18S rRNAs from harmful microalgae



Shotgun Metagenomic Data Analysis:



MetaShot

MetaShot is a pipeline for the taxonomic characterization of shotgun NGS metagenomic data, particularly oriented towards the study of human and other host microbiome (Fosso et al., in preparation).



Multiple Alignments

MSA-PAD



Sequence Analysis

MSA-PAD: DNA Multiple Sequence Alignment Framework based on PFAM Accessed Domain Information

Bachir Balech¹, Saverio Vicario², Giacinto Donvito³, Alfonso Monaco³, Pasquale Notarangelo³, Graziano Pesole^{1,4*}.

MsaPAD workflow in Taverna Workbench Biodiversity 2.5

MSA-PAD aligns DNA sequences encoding either single or multiple protein domains by two alignment options: **Gene and Genome**





Sequence and Metadata Retrieval (in progress)

DataRetrieval



DataRetrieval is a RESTfull service able to query Public DNA barcode databases and extract the relevant data and metadata corresponding to a given taxon.



How to call DataRetrieval web service?

Required arguments:

- TaxonName
- Division: Animals, Fungi, Protists, Plants

Outputs:

- Summary data
- Full data+metadata

CTB CALL

for investigating alien and invasive species through NGS data production and analysis

AFFILIATION	PI	RESEARCH TITLE	APPLICATIONS
SZN, Napoli	Casotti	Isolation and genetic characterization of bacterial strains isolated in the presence of polyunsaturated aldehydes from sea water samples of diatoms bloom.	Genomics
CNR-ISMAR, Venezia	Quero	MALWARe: MetAgenomic study of microbiaL invasive species introduced by Ballast WateRs.	Metagenomics
University of Milan Bicocca	Casiraghi	Design of an innovative pipeline for DNA metabarcoding of eukaryotes to monitor potential invasive species in atmospheric particulate during Milan EXPO 2015.	Metagenomics
SZN, Napoli	Ferrante	An investigation of the genetic basis of toxin production in the diatom species Pseudo-nitzschia.	Genomics
University of Salento	Basset	Biotic Interactions in Detritus food Webs: metagenomics analysis of the microfungi- detritivores and detritivore-detritivore interactions in lagoon ecosystems (BID_Webs).	Metagenomics
CNR-ISE, Verbania	Eckert	Daphnia as Vectors for Potential Pathogens and Antibiotic Resistant Bacteria in Freshwater Ecosystems.	Metagenomics
CNR-ISE, Verbania	Fontaneto	MEIOfauna DIVErsity (MEIODIVE): Testing ecogenomics of meiofauna as a tool for biological monitoring of marine sediments.	Metagenomics
University of Trieste	Pallavicini	Detection of alien species from gut contents of commercially important fish in the Adriatic Sea, possible effect on feeding behaviour and food webs.	Metagenomics
CNR-ISMAR, Venezia	Bongiorni	Responses of benthic microbial diversity and functioning to the recent invasion of Gracilaria vermiculophylla in the lagoon of Venice.	Metagenomics
University of Milan	Gissi	Comparing fresh with formalin-preserved samples: application of the NGS technology to the molecular characterization of old museum specimens	Genomics

LifeWatch

e-Science European Infrastructure for Biodiversity and Ecosystem Research



Lif<mark>ewatch-Italy Call f</mark>or in<mark>terdisciplinary</mark> research projects at the distributed Laboratory of "<mark>Molecular Biodiversity</mark>" (deadline 18.04.2014)

Ten projects have been approved and currently under processing



RELATED PROJECTS







https://www.microb3.eu/osd

Publications in international journals

Biol Invasions (2015) 17:923–940 DOI 10.1007/s10530-014-0810-2

MOLECULAR TOOLS

The influence of invasive jellyfish blooms on the aquatic microbiome in a coastal lagoon (Varano, SE Italy) detected by an Illumina-based deep sequencing strategy

Caterina Manzari • Bruno Fosso • Marinella Marzano • Anita Annese • Rosa Caprioli • Anna Maria D'Erchia • Carmela Gissi • Marianna Intranuovo • Ernesto Picardi • Monica Santamaria • Simonetta Scorrano • Giuseppe Sgaramella • Loredana Stabili • Stefano Piraino • Graziano Pesole

Received: 14 November 2013 / Accepted: 31 October 2014 / Published online: 19 November 2014 © The Author(s) 2014. This article is published with open access at Springerlink.com

> Bioinformatics, 31(15), 2015, 2571–2573 doi: 10.1093/bioinformatics/btv141 Advance Access Publication Date: 26 March 2015 Applications Note

Sequence analysis

MSA-PAD: DNA multiple sequence alignment framework based on PFAM accessed domain information

Bachir Balech¹, Saverio Vicario², Giacinto Donvito³, Alfonso Monaco³, Pasquale Notarangelo³ and Graziano Pesole^{1,4,*}

¹Istituto di Biomembrane e Bioenergetica and ²Istituto Tecnologie Biomediche, Consiglio Nazionale delle Ricerche, 70126 Bari, Italy, ³Istituto Nazionale di Fisica Nucleare, 70126 Bari, Italy and ⁴Dipartimento Bioscienze, Biotecnologie e Biofarmaceutica, University of Bari, 70126 Bari, Italy

*To whom correspondence should be addressed. Associate Editor: David Posada

Received on August 8, 2014; revised on March 6, 2015; accepted on March 9, 2015



Draft genome sequence of *Sphingobium* sp. strain ba1, resistant to kanamycin and nickel ions

Caterina Manzari¹, Matteo Chiara², Alessandra Costanza¹, Claudia Leoni¹, Mariateresa Volpicella¹, Ernesto Picardi^{1,3,4}, Anna Maria D'Erchia¹, Antonio Placido³, Massimo Trotta⁵, David S. Horner², Graziano Pesole^{1,3,4,6} & Luigi R. Ceci³

Scientific Results



Scientific Results

BRIEFINGS IN BIOINFORMATICS. VOL 13. NO 6. 682–695 Advance Access published on 10 July 2012 doi:10.10

Reference databases for taxonomic assignment in metagenomics

Monica Santamaria*, Bruno Fosso*, Arianna Consiglio, Giorgio De Caro, Giorgio Grillo, Flavio Licciulli, Sabino Liuni, Marinella Marzano, Daniel Alonso-Alemany, Gabriel Valiente and Graziano Pesole

Submitted: 14th March 2012; Received (in revised form): 2nd June 2012



METHODOLOGY

Open Access

BioMaS: a modular pipeline for Bioinformatic analysis of Metagenomic AmpliconS

Bruno Fosso¹, Monica Santamaria¹, Marinella Marzano¹, Daniel Alonso-Alemany², Gabriel Valiente², Giacinto Donvito³, Alfonso Monaco³, Pasquale Notarangelo³ and Graziano Pesole^{1,4,5}^{*}



Staff





Personnel & contact details

Collaborations

INFN ISTITUTO NAZIONALE DI FISICA NUCLEARE



















