Global Alignment and Collaboration

Jo McEntyre

@jomcentyre
www.europepmc.org
www.ebi.ac.uk
About EMBL-EBI

- Part of the European Molecular Biology Laboratory
- International, non-profit research institute
- Europe’s hub for biological data services and research
Science is global
2. Discoverability through structured data

Structured data is one of the true enablers of life science

- Discovery of homology between genes across species
- Predicting function based on protein folds

• Structured data can be cross-analysed, compared by algorithm, and encourages development of new products and tools
Osteoarthritis
Data Ecosystem

Citations & Links
- ORCIDs
- DOIs, ROIs
- PIDs

Unstructured/semi-structured

Literature

Ontologies Metadata

Structured

Added Value

Algorithms
- Text and data mining
- Analytical tools
A stable, freely available, shared repository

Europe PMC

- Abstracts: 30 million
- Full-text articles: 3 million
- Article citation counts
- Grants
- ORCIDs
- Semantic annotation
- Data citations
- Data integration

Europe PMC is a member of the PMC International Collaboration.

Supports OA mandates of 26 European funders of life science research
Europe PMC content

3 million full-text articles

500,000 Agricola records

50,000 biomedical research grants

500,000 open access articles

30 million abstracts

4 million patents

UniProt

HMS guidelines

European Nucleotide Archive

Protein Data Bank in Europe

Institution networks
The scent of information

“retaining files and being prepared to share them” ≠ accessible data

It's healthy to remember that users are selfish, lazy, and ruthless in applying their cost-benefit analyses

Information Foraging: by JAKOB NIELSEN, June 30, 2003
http://www.nngroup.com/articles/information-scent/

• Data placement: where people will find it
• Scent alone is not scientific …
Data Citation Principles

- Data as legitimate, citable products of research
- Attribution and credit
- Cited as evidence for a claim
- Unique identification
- Access
- Persistence
- Specificity and verifiability (provenance)
- Interoperability and flexibility

https://www.force11.org/datacitation
BioStudies database for unstructured data
PIIDs build infrastructure for scientific discovery and credit systems
(1) Europe PMC and ORCID

Use of the claiming tool
- ~400K article claims
- ~21K unique ORCIDs

Incorporation into Europe PMC
- 1.2M unique articles in Europe PMC linked to ORCIDs
  (cf. PubMed 12K)

An integrated encyclopedia of DNA elements in the human genome. (PMID:22955616 PMCID:PMC3439153)
The putative tumor suppressor gene PTPN13/PTPL1 induces apoptosis through insulin receptor substrate-1 dephosphorylation.
(PMID: 17638892)

Abstract

INSERM U826, Contrôle de la Progression des Cancers Hormonodépendants, Génétique de Cancer, Université Montpellier I, CNRS UMR 5054 Val d'Aurelle Paul Levesque Cancer Research [2007, 67(14):6806-6813]

Type: Journal Article, Research Support, Non-U.S. Gov't
DOI: 10.1158/0008-5472.CAN-07-0513

The protein tyrosine phosphatase (PTP) PTPL1/PTPN13 is a cotransporter that is overexpressed in several tumors. Indeed, PTPL1 activity has been reported recently to be decreased, or even restored in some tumors, and we showed that PTPL1 is necessary for inhibition of Akt activation and induction of apoptosis. Implications of the phosphatidylinositol 3-kinase (PI3K) and progression are now well established, and our study was therefore sufficient to inhibit this pathway. We also showed that PTPL1 expression affects the dephosphorylation of insulin receptor substrate-1 (IRS-1) and RNA interference confirms this finding. Finally, we report that PTPL1 expression is sufficient to inhibit the insulin-like growth factor-1 effect on cell apoptosis. Altogether, these data provide the first evidence for a role of the tumor suppressor gene PTPL1/PTPN13 on apoptosis and identify it as a new signaling pathway.

Genes & Proteins

Found 5 UniProt record(s) citing this article

Insulin receptor substrate 1
(UniProt:P35568)

HAL Open Archive

HAL - author self-archived e-prints

- The Putative Tumor Suppressor Gene PTPN13/PTPL1 Induces Apoptosis through Insulin Receptor Substrate-1 Dephosphorylation.
  http://www.hal.inserm.fr/inserm-00165318

DEPOD

DEPOD - the human DEPhOosphorylation Database is a manually curated database collecting human active phosphatases.

- Substrates (established/proposed) of PTPN13 include "IRS1"
Cross Disciplinary Integration

Prokaryotic community structure and diversity in the sediments of an active submarine mud volcano (Kazan mud volcano, East Mediterranean Sea).

(PMID:20370830)

Abstract

We investigated 16S rRNA gene diversity at a high sediment depth (20 cm) in an active site of the Kazan mud volcano, East Mediterranean Sea. 374 bacterial clones were analysed, which were attributed to 38 and 32 families, respectively (> or = 90% similarity). Most of the bacterial phyotypes were dominated by phyla such as Proteobacteria, Bacteroidetes, and Planctomycetes. The diversity index diversity was calculated to be 2.0. The distribution of the bacterial phyla was different from other sites in the Mediterranean Sea. Our study revealed different ecological strategies for microorganisms living in the sediments of the Kazan mud volcano.
Elixir: An international distributed infrastructure for

- Data
- Standards
- Tools
- Compute
- Training
- Industry

Alignment with BD2K (NIH)