Journal of proteome research

pubs.acs.org/jpr

Technical Note

NETGE-PLUS: Standard and Network-Based Gene Enrichment Analysis in Human and Model Organisms

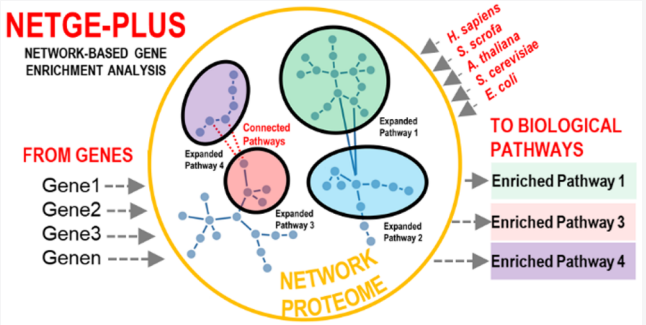
Samuele Bovo, Pier Luigi Martelli,* Pietro Di Lena, and Rita Casadio

Cite This: *J. Proteome Res.* 2020, 19, 2873−2878

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Omics techniques provide a spectrum of information at the genomic level, whose analysis can characterize complex traits at a molecular level. The relationship among genotype and phenotype implies that from genome information the molecular pathways and biological processes underlying a given phenotype are discovered. In dealing with this problem, gene enrichment analysis has become the most widely adopted strategy. Here we present NETGE-PLUS, a Web server for standard and network-based functional interpretation of gene sets of human and of model organisms, including *Sus scrofa*, *Saccharomyces cerevisiae*, *Escherichia coli*, and *Arabidopsis thaliana*. NETGE-PLUS enables the functional enrichment of both simple and ranked lists of genes, introducing also the possibility of exploring relationships among KEGG pathways. A Web interface makes data retrieval complete and user-friendly. NETGE-PLUS is publicly available at http://net-ge2.biocomp.unibo.it.



KEYWORDS: *network-based gene enrichment analysis, over-representation analysis, gene set enrichment analysis, functional association, model organisms, Web server*

INTRODUCTION

The analysis of large-scale genomic and proteomic data aims at characterizing complex traits at a molecular level. By this, lists of "relevant" genes/proteins can focus on putative specific phenotype(s). Over-representation analysis (ORA) is a widely adopted method to endow a gene set with functional annotations, with the aim of disentangling the complex relationships among genes, their functions, and the phenotype of interest. ORA tests whether some biological features are significantly more frequent in the given gene set than is expected by chance. Alternatively, when genes are endowed with some numerical score (e.g., the fold change of differential expression), gene set enrichment analysis (GSEA) allows the inclusion of the weighted contribution of a large ranked list of genes to perform the functional characterization.[1] Currently, few tools implement GSEA-related statistics in the analysis of protein sets. Examples include protein set enrichment analysis (PSEA)[2] and PSEA-Quant.[3] Like GSEA, PSEA is suited to analyze the results of differential experiments for the same set of genes. During input, PSEA endows each gene with a spectral index computed starting from mass spectrometry data.[2] Alternatively, PSEA-Quant takes as input the proteins highly expressed in different technical or biological replicates.[3] Both methods compute a statistical enrichment of the data by adopting background sets of previously clustered seed genes having the same functional annotation. In this respect, our method is novel. Indeed, NETGE-PLUS allows the computation of a GSEA statistical enrichment according to not only standard enrichment methods but also the adoption of a reference set of genes that are in subnetworks of the known interactome for each given species.

Previously, we described NET-GE, a tool for performing ORA on sets of human genes after their mapping on subnetworks of genes derived from the analysis of the human interactome.[4,5] Exploiting the information contained in biological networks,[6] NET-GE enriches the Gene Ontology (GO) terms,[7] KEGG pathway[8] and Reactome[9] annotations. NET-GE does not allow GSEA, and it is limited to human genes and their related interactions. Furthermore, NET-GE adopts the KEGG BRITE hierarchy (http://www.genome.jp/kegg/kegg3b.html), which provides a logical organization rather than a complete description of the proximity relationships among metabolic pathways.

Adopting network-based strategies stands on the notion that interactomes remain species-specific[10] and that their updates improve over the different releases. The advantage of such a procedure is that the enriched biological features have underlying sets of genes that are connected in the interactome

Special Issue: Methods for Omics Research

Received: November 6, 2019
Published: January 23, 2020

https://dx.doi.org/10.1021/acs.jproteome.9b00749
J. Proteome Res. 2020, 19, 2873−2878

Table 1. NETGE-PLUS Summary Statistics

	H. sapiens	*S. scrofa*	*A. thaliana*	*S. cerevisiae*	*E. coli*
annotated genes	17 060	16 406	19 144	5860	3790
annotation sets[a]	23 384	19 022	10 494	9707	5800
associations$_{std}$[b]	913 514	552 526	382 085	321 667	156 394
associations$_{net}$[c]	3 065 034	1 696 519	1 279 481	780 930	436 734

[a]Number of GO terms, KEGG, and Reactome pathways. [b]Number of gene-term associations used in the standard enrichment analysis. [c]Number of gene-term associations used in the network-based enrichment analysis.

in specific functional modules and not only genes that are clustered together based on identical functional annotations.

Here we introduce NETGE-PLUS, a new Web server that allows different enrichment procedures. First, the server can enrich the functional features of genes not only from *H. sapiens* but also from *S. scrofa*, *E. coli*, *S. cerevisiae*, and *A. thaliana*, organisms that are widely adopted in translational, biomedical, and biotechnological research. NETGE-PLUS can perform standard and network-based enrichment. In this version, as input, NETGE-PLUS takes both lists of genes/proteins and genes or proteins that are ranked according to some criteria (e.g., expression level, spectral index). The present version of NETGE-PLUS implements both the ORA and GSEA procedures for the analysis of differential expression data. Another interesting advancement is the possibility of linking different KEGG maps. With this, a given set of genes can retrieve a network of pathways (KEGG-NET), highlighting the underlying complexity of the phenotype.

MATERIALS AND METHODS

Databases

NETGE-PLUS includes the interactomes of *H. sapiens* (hsa; taxid: 9606), *S. scrofa* (ssc; taxid:9823), *S. cerevisiae* (sce; taxid: 4932), *E. coli* (eco; taxid: 51145), and *A. thaliana* (ath; taxid: 3702), as derived from STRING v.10.5.[11] The identifiers adopted in the procedure are (i) the ENSEMBL protein identifiers for *H. sapiens* and *S. scrofa* and (ii) the systematic locus identifiers, adopted for *S. cerevisiae*, *E. coli*, and *A. thaliana*. All of the links with a combined STRING score ≥0.4 (medium confidence) were retained irrespective of the supporting evidence. Term-specific modules of interacting genes have been computed for the Gene Ontology terms (v.169; https://www.ebi.ac.uk/GOA), for KEGG (v.83.2; http://www.kegg.jp/), and for Reactome (v.61; https://reactome.org/) pathways.

Processing the Interactomes

NETGE-PLUS relies on modules of functionally related genes, precomputed as described by Di Lena et al.[4] In brief, each module is built starting from (i) a set of genes sharing a specific functional term (seed set) and (ii) an interactome. Each seed set is extended into a compact and connected module of interacting proteins by computing all of the shortest paths among the seed genes. By applying measures based on graph and information theory, modules are reduced to minimal connecting networks while preserving the distances among seeds. The algorithmic details are included in the Supporting Information and online (http://net-ge2.biocomp.unibo.it/enrich/default/help).

The resulting modules, containing seed nodes and some of their interacting partners (connecting nodes), are at the basis of the functional enrichment. The number of seed genes, annotation sets, and gene-term associations is reported in Table 1. Statistics on the sizes of the annotation sets are presented in Table S1 (Supporting Information). The network-based modules are larger than the sets of seeds from which they are derived. Depending on the species and the annotation type, the mean sizes increase by a factor ranging between 1.4 and 4 upon the module construction.

Enrichment Procedures

NETGE-PLUS implements both a standard and a network-based gene enrichment analysis. Given an input gene/protein set, genes are mapped into the subnetworks of each annotation database. NETGE-PLUS allows us to perform: (i) ORA, through the Fisher's exact test and (ii) GSEA, through a Kolmogorov−Smirnov-like statistic, as described by Subramanian et al.[1] and reimplemented using the Python 2.7 package GSEAPY (v0.9.4; https://pypi.python.org/pypi/gseapy). For GSEA, we use a weighted enrichment-scoring statistic and 100 permutations. For the multiple testing correction, the user can select either the Bonferroni or the Benjamini−Hochberg (false discovery rate (FDR)) procedures.[12]

Implementation of the Web Server

The Web server runs on a web2py engine (http://www.web2py.com/), and it is optimized to work with all common Web browsers. The analysis runs asynchronously: Upon request submission, the server displays a bookmarkable page that is periodically updated until job completion. A link to the results page is given to the user as soon as the job is completed.

The final visualization of the results exploits the Graphviz library (http://www.graphviz.org/) for laying out the directed acyclic graphs for Gene Ontology, KEGG, and REACTOME. KEGG-NET results are rendered via the JavaScript library Cytoscape.js (http://js.cytoscape.org/). Enriched terms from these annotation systems are highlighted. In addition, the Web server shows dynamic network renderings based on the JavaScript library d3.js (http://d3js.org/) for the visualization of the underlying interaction networks involving a specific term and the network of pathways.

For multiple submissions, each request is queued, and it runs as soon as there is available computing power. The run time depends on the size of the input set (from 2 up to 200; higher numbers may require a long time) and ranges from 1 to 10 min.

The user can also provide an e-mail address to receive results.

Case Studies

Three gene sets, related to human and porcine diseases, were investigated to qualitatively evaluate the performance of NETGE-PLUS. Here we focus on only one of them, the nonalcoholic fatty liver disease (NAFLD) study case. The other two study cases are presented in the Supporting Information and online (http://net-ge2.biocomp.unibo.it/enrich/default/tutorial).

We retrieved from Phenopedia[13] a list of genes that possibly contribute to the development of NAFLD. Among the 408 NAFLD-related genes, we selected the ones supported by at least five publications. As a result, we obtained a list of 28 genes. The set was analyzed by running ORA over the KEGG-NET

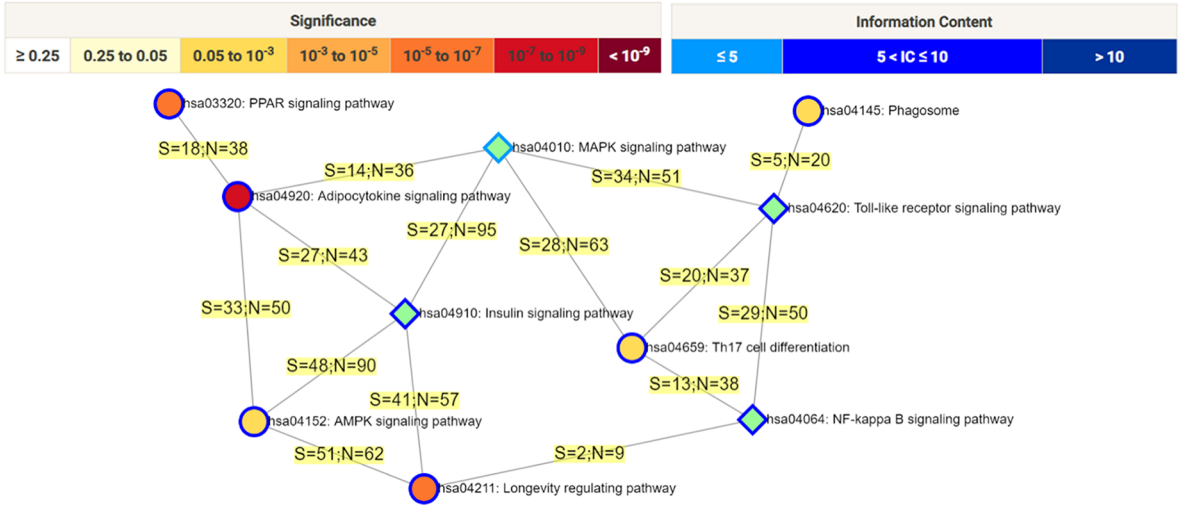
https://dx.doi.org/10.1021/acs.jproteome.9b00749
J. Proteome Res. 2020, 19, 2873−2878

A

Enrichment	TERM	Description	IC	N1	N2	Background	p-value	p-value (Bonferroni)
N+S	hsa04920	Adipocytokine signaling pathway	6.65	9	115	7921	1.15e-10	2.73e-08
N+S	hsa04211	Longevity regulating pathway	6.27	8	119	7921	4.95e-09	1.18e-06
N	hsa03320	PPAR signaling pathway	6.59	8	133	7921	1.20e-08	2.85e-06
N**	hsa04145	Phagosome	5.6	9	426	7921	9.56e-06	2.28e-03
N	hsa04659	Th17 cell differentiation	6.09	6	175	7921	2.68e-05	6.39e-03

S: term enriched with the standard method.
N: term enriched with the network-based method.
N+S: term enriched with both the standard and the network-based method.
N**: term enriched with the network-based method and not included in the annotations of the input set.
IC: Information Content; **N1**: The number of input genes/proteins associated to the term;
N2: The whole set of genes/proteins annotated with the term.

B



C

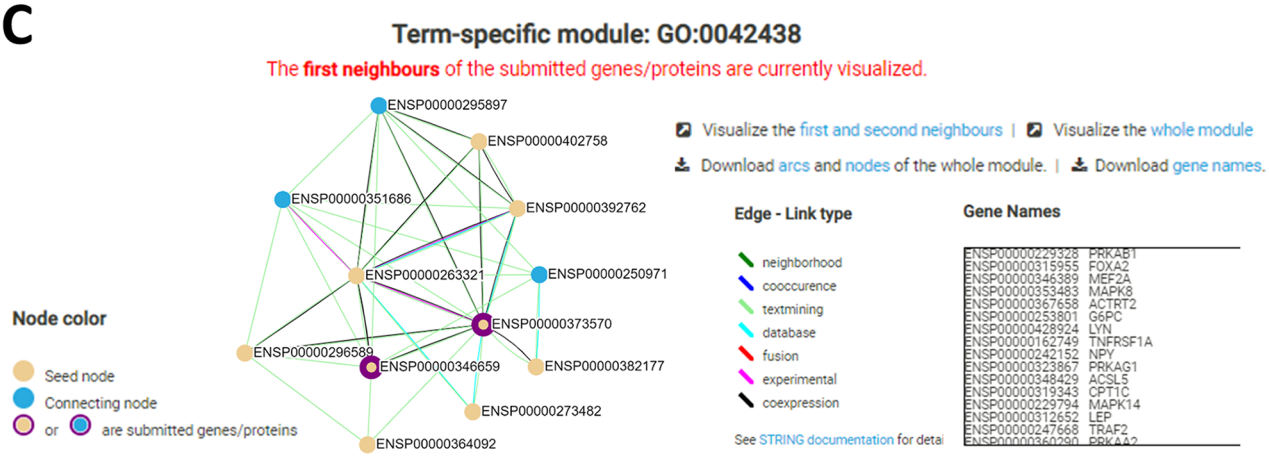


Figure 1. NETGE-PLUS results. (A) Table of network-based over-represented KEGG-NET pathways. (B) Network of pathways. Circles represent enriched pathways, and diamonds represent the connecting pathways. The circle color represents the magnitude of enrichment, and the green diamonds highlight the presence of at least one input gene associated with them. Each edge is labeled with the number of genes shared between pathways in the standard gene enrichment procedure (S) and in the network-based procedure (N). (C) Example of a functional module of the over-represented GO term GO:0042438. Seed nodes (yellow) represent genes directly annotated with the term, and connecting nodes (blue) represent connecting genes. Nodes presenting a purple border identify the part of the submitted genes. The link type is highlighted based on the seven different channels of STRING (https://string-db.org/).

resource. We considered statistically enriched terms with a p value <0.01 after the correction with the Bonferroni procedure.

RESULTS AND DISCUSSION

NETGE-PLUS Web Server

NETGE-PLUS accepts different gene names and identifiers (UniProtKB, ENSEMBL gene and protein, official gene name, and the systematic locus identifier) plus the related scores in the case of GSEA. As output, NETGE-PLUS returns two tables (one for the standard method and one for the network-based method; Figure 1A) listing the over-represented terms. The tables report basic information (e.g., term name, p value(s), input genes) plus other statistics that allow a better understanding of the specificity of the enriched terms. These include the information content (IC), a measure of function specificity,[14] and the ontology hierarchy level. The results page also provides graphs

https://dx.doi.org/10.1021/acs.jproteome.9b00749
J. Proteome Res. 2020, 19, 2873−2878

Table 2. Nonalcoholic Fatty Liver Disease Case Study[a]

enrichment[b]	term[c]	N1[d]	N2[e]	background[f]	Bonferroni[g]	description[h]
S	hsa04920	6	69	6516	8.44×10^{-5}	adipocytokine signaling pathway
S	hsa04211	6	90	6516	4.12×10^{-4}	longevity regulating pathway
S	hsa04152	6	121	6516	2.32×10^{-3}	AMPK signaling pathway
N	hsa03320	8	133	7921	2.85×10^{-6}	PPAR signaling pathway
N**	hsa04145	9	426	7921	2.28×10^{-3}	phagosome
N	hsa04659	6	175	7921	6.39×10^{-3}	Th17 cell differentiation

[a]Gene enrichment analysis over the KEGG-NET resource. [b]Standard (S) and network-based (N) procedures. N** indicates a new enriched term not directly associated with the input gene/proteins. [c]Functional annotation identifier. [d]Input genes/proteins belonging to the term. [e]Genes associated with the functional term. [f]Number of genes used as background at the Fisher's exact test. [g]p value corrected by using the Bonferroni procedure. [h]Brief explanation of the term.

depicting (i) the functional connection among enriched terms (the hierarchy or the network of pathways; Figure 1B) and (ii) the organization of functional modules (subnetworks; Figure 1C). Tables and graphs can be downloaded and locally managed.

Input and output details are presented in the Supporting Information. Additional details are also available online (http://net-ge2.biocomp.unibo.it/enrich/default/help).

Understanding the Functional Relationships among Over-Represented KEGG Pathways

When enriching functional pathways, it is useful to represent them in a larger context that comprises the most related pathways. Whereas for REACTOME a full hierarchy is provided, in the case of KEGG, the BRITE hierarchy gives only a categorization of the different KEGG maps. For a better understanding of the overall organization of the KEGG pathways, we exploit the information contained in the links among different KEGG maps by defining KEGG-NET. When a set of genes enriches different KEGG pathways, the most related connecting pathways are determined by computing the pairwise shortest paths through the links. Because the pathway network is highly connected, we retain only the paths with a maximum length equal to two (no. of edges). KEGG map01100 (whole metabolism) and the disease-related maps are not considered in this procedure because of their dense connectivity. By this, users have the possibility of obtaining putative biological dependences among the enriched pathways.

Nonalcoholic Fatty Liver Disease: A Case Study Involving the KEGG-NET Resource

In the following, we deal with a specific case study, described also in the online tutorial. We make use of the KEGG-NET resource to dissect the biological complexity of NAFLD. Defined as a genetic−environmental−metabolic stress-related disease, NAFLD is a pathology characterized by excessive fat accumulation in the liver, even in the absence of alcohol consumption. NAFLD encompasses a spectrum of diseases, from simple steatosis to nonalcoholic steatohepatitis (NASH), which can progress to cirrhosis and hepatocellular carcinoma. There is also increasing evidence that NAFLD represents the hepatic component of a metabolic syndrome characterized by obesity, hyperinsulinemia, peripheral insulin resistance, diabetes, hypertriglyceridemia, and hypertension.[15] Moreover, several studies have identified many genetic variations that may be associated with the development of NAFLD.[16]

The analysis of the 28 NAFLD-related genes highlighted 6 over-represented pathways (Table 2); 3 were detected via the standard enrichment analysis, whereas the other 3 were added by the network-based procedure.

Pathways over-represented by the standard method are the adipocytokine signaling pathway, the longevity-regulating pathway, and the AMPK signaling pathway. Considering the graph generated by linking the whole set of enriched pathways (Figure 1B), the three pathways are connected within a chain. All of them are also linked to the insulin-signaling pathway (not included in the enriched set). Insulin resistance plays an important role in NAFLD, and it is caused by adipocytokines, a specific kind of cytokine secreted by the adipose tissue. Adiponectin is an anti-inflammatory and antidiabetic adipocytokine, which exerts its actions by the activation of adenosine monophosphate (AMP)-activated kinase (AMPK) and PPARα.[17] Interestingly enough, the PPAR signaling pathway is one of the terms enriched with the network-based procedure.

Another important cytokine within the adipocytokine pathway is leptin. It binds the leptin receptor (LEP-R) and triggers a phosphorylation chain, resulting in the activation of the mitogen-activated protein kinase (MAPK) pathway.[18] This is another connecting pathway in the network. One of the members of the MAPK pathway, namely, the protein kinase c-Jun N-terminal kinase (JNK), is closely related to insulin resistance. Moreover, rat models with activated JNK present phenotypes related to NAFLD, such as hepatocyte fat accumulation and cell injury.[15]

In Figure 1B, the MAPK pathway links the adipocytokine and insulin signaling pathways with the Th17 cell differentiation pathway (enriched with the network-based procedure). Interestingly, Th17 cells have been associated with hepatocellular steatosis and inflammatory processes via the production of IL-17, which is also implicated in insulin resistance. Moreover, the secretion of IL-17 is triggered and perpetuated through the nuclear factor-κB (NF-κB),[19] and the NF-κB pathway is indeed a connecting node.

The last term enriched with the network procedure is "phagosome", which is linked to the other nodes through the "Toll-like receptor signaling pathway". Both of these terms computationally derived with NETGE-PLUS suggest the involvement of macrophages in NAFLD, in particular, in relation to the reprogramming induced by cytokines. The role of macrophages in NAFLD from initial steatosis to advanced fibrosis was previously reviewed.[20] By highlighting links among different metabolisms, KEGG-NET can help to highlight the complex interactions among different biological pathways underlying the phenotype of interest.

CONCLUSIONS

We describe a Web server, NETGE-PLUS, which, besides standard enrichment methods, computes online interesting alternatives for characterizing the molecular complexity of

https://dx.doi.org/10.1021/acs.jproteome.9b00749
J. Proteome Res. 2020, 19, 2873−2878

emergent phenotypes in organisms encompassing the different kingdoms of life. NETGE-PLUS, with a user-friendly Web page, offers: (i) the possibility of performing ORA and GSEA, both according for standard and network-based procedures; (ii) the possibility of performing network-based enrichment over sets of connected genes with updated annotations for *H. sapiens* and four other well-studied model organisms (*S. scrofa*, *S. cerevisiae*, *E. coli*, and *A. thaliana*); and (iii) the possibility of understanding the functional relationships among over-represented KEGG pathways via the KEGG-NET resource.

NETGE-PLUS is therefore a useful tool for disentangling biological complexity and, via the subnetworks (network-based enrichment analysis), for retrieving genes that, besides the ones given in the input, may play a fundamental role in the genotype−phenotype relationship.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jproteome.9b00749.

Workflow of the gene (set) enrichment analysis. Figure S1. Outline of the module generation in NETGE-PLUS. Table S1. Statistics of the seed sets and network-based modules adopted for NETGE-PLUS enrichment procedure. Table S2. Examples of identifiers for the asparagine synthase (glutamine-hydrolyzing) [EC:6.3.5.4]. Case study 1: Ichthyosis, cyclic, with epidermolytic hyperkeratosis (CIEHK). Table S3. CIEHK case study: over-represented GO-BP. Table S4. CIEHK case study: over-represented GO−CC. Table S5. CIEHK case study: over-represented Reactome pathways. Case study 2: Effect of the Pseudorabies virus (PRV) infection on the swine transcriptome. Table S6. PRV case study: GSEA over the GO-PB (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Pier Luigi Martelli − *Biocomputing Group, Department of Pharmacy and Biotechnology (FABIT), University of Bologna, 40126 Bologna, Italy*; Email: gigi@biocomp.unibo.it

Authors

Samuele Bovo − *Biocomputing Group, Department of Pharmacy and Biotechnology (FABIT) and Department of Agricultural and Food Sciences (DISTAL), Division of Animal Sciences, University of Bologna, 40126 Bologna, Italy*; orcid.org/0000-0002-5712-8211

Pietro Di Lena − *Department of Computer Science and Engineering (DISI), University of Bologna, 40126 Bologna, Italy*

Rita Casadio − *Biocomputing Group, Department of Pharmacy and Biotechnology (FABIT), University of Bologna, 40126 Bologna, Italy; Institute of Biomembranes, Bioenergetics and Molecular Biotechnologies (IBIOM), Italian National Research Council (CNR), 70126 Bari, Italy*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jproteome.9b00749

Author Contributions

S.B. implemented the Web application, performed the enrichment analyses, and drafted the manuscript. P.L.M. and P.D.L. supervised the module generation. R.C. and P.L.M. supervised the project and drafted the manuscript. All authors read and approved the final manuscript.

Notes

The authors declare no competing financial interest.

NETGE-PLUS is publicly available at http://net-ge2.biocomp.unibo.it.

■ REFERENCES

(1) Subramanian, A.; Tamayo, P.; Mootha, V. K.; Mukherjee, S.; Ebert, B. L.; Gillette, M. A.; Paulovich, A.; Pomeroy, S. L.; Golub, T. R.; Lander, E. S.; et al. Gene Set Enrichment Analysis: A Knowledge-Based Approach for Interpreting Genome-Wide Expression Profiles. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (43), 15545−15550.

(2) Cha, S.; Imielinski, M. B.; Rejtar, T.; Richardson, E. A.; Thakur, D.; Sgroi, D. C.; Karger, B. L. In situ proteomic analysis of human breast cancer epithelial cells using laser capture microdissection: annotation by protein set enrichment analysis and gene ontology. *Mol. Cell. Proteomics* **2010**, *9*, 2529−2544.

(3) Lavallée-Adam, M.; Rauniyar, N.; McClatchy, D. B.; Yates, J. R., 3rd. PSEA-Quant: a protein set enrichment analysis on label-free and label-based protein quantification data. *J. Proteome Res.* **2014**, *13* (12), 5496−5509.

(4) Di Lena, P.; Martelli, P. L.; Fariselli, P.; Casadio, R. NET-GE: A Novel NETwork-Based Gene Enrichment for Detecting Biological Processes Associated to Mendelian Diseases. *BMC Genomics* **2015**, *16* (S8), S6.

(5) Bovo, S.; Di Lena, P.; Martelli, P. L.; Fariselli, P.; Casadio, R. NET-GE: A Web-Server for NETwork-Based Human Gene Enrichment. *Bioinformatics* **2016**, *32* (22), btw508.

(6) Laukens, K.; Naulaerts, S.; Berghe, W. V. Bioinformatics Approaches for the Functional Interpretation of Protein Lists: From Ontology Term Enrichment to Network Analysis. *Proteomics* **2015**, *15* (5−6), 981−996.

(7) Gene Ontology Consortium. Gene Ontology Consortium: Going Forward. *Nucleic Acids Res.* **2015**, *43* (D1), D1049−1056.

(8) Kanehisa, M.; Sato, Y.; Kawashima, M.; Furumichi, M.; Tanabe, M. KEGG as a Reference Resource for Gene and Protein Annotation. *Nucleic Acids Res.* **2016**, *44* (D1), D457−462.

(9) Fabregat, A.; Sidiropoulos, K.; Garapati, P.; Gillespie, M.; Hausmann, K.; Haw, R.; Jassal, B.; Jupe, S.; Korninger, F.; McKay, S.; et al. The REACTOME Pathway Knowledgebase. *Nucleic Acids Res.* **2016**, *44* (D1), D481−487.

(10) Sharan, R.; Suthram, S.; Kelley, R. M.; Kuhn, T.; McCuine, S.; Uetz, P.; Sittler, T.; Karp, R. M.; Ideker, T. Conserved Patterns of Protein Interaction in Multiple Species. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (6), 1974−1979.

(11) Szklarczyk, D.; Morris, J. H.; Cook, H.; Kuhn, M.; Wyder, S.; Simonovic, M.; Santos, A.; Doncheva, N. T.; Roth, A.; Bork, P.; et al. The STRING Database in 2017: Quality-Controlled Protein-Protein Association Networks, Made Broadly Accessible. *Nucleic Acids Res.* **2017**, *45* (D1), D362−D368.

(12) Noble, W. S. How Does Multiple Testing Correction Work? *Nat. Biotechnol.* **2009**, *27* (12), 1135−1137.

(13) Yu, W.; Clyne, M.; Khoury, M. J.; Gwinn, M. Phenopedia and Genopedia: Disease-Centered and Gene-Centered Views of the Evolving Knowledge of Human Genetic Associations. *Bioinformatics* **2010**, *26* (1), 145−146.

(14) Louie, B.; Bergen, S.; Higdon, R.; Kolker, E. Quantifying Protein Function Specificity in the Gene Ontology. *Stand. Genomic Sci.* **2010**, *2* (2), 238−244.

(15) Zeng, L.; Tang, W. J.; Yin, J. J.; Zhou, B. J. Signal Transductions and Nonalcoholic Fatty Liver: A Mini-Review. *Int. J. Clin Exp Med.* **2014**, *7* (7), 1624−1631.

(16) Sookoian, S.; Pirola, C. J. Genetic Predisposition in Nonalcoholic Fatty Liver Disease. *Clin Mol. Hepatol* **2017**, *23* (1), 1−12.

(17) Berlanga, A.; Guiu-Jurado, E.; Porras, J. A.; Auguet, T. Molecular Pathways in Non-Alcoholic Fatty Liver Disease. *Clin. Exp. Gastroenterol.* **2014**, *7*, 221−239.

(18) Polyzos, S. A.; Kountouras, J.; Mantzoros, C. S. Leptin in Nonalcoholic Fatty Liver Disease: A Narrative Review. *Metab., Clin. Exp.* **2015**, *64* (1), 60−78.
(19) Procaccini, C.; De Rosa, V.; Galgani, M.; Carbone, F.; La Rocca, C.; Formisano, L.; Matarese, G. Role of Adipokines Signaling in the Modulation of T Cells Function. *Front. Immunol.* **2013**, *4*, 332.
(20) Krenkel, O.; Tacke, F. Macrophages in Nonalcoholic Fatty Liver Disease: A Role Model of Pathogenic Immunometabolism. *Semin. Liver Dis.* **2017**, *37* (3), 189−197.

https://dx.doi.org/10.1021/acs.jproteome.9b00749
J. Proteome Res. 2020, 19, 2873−2878