

ClueGO

Documentation



Gabriela Bindea and Bernhard Mlecnik
Laboratory of Integrative Cancer Immunology
UMRS1138 Cordeliers Research Center, Paris,
France

Contents

ClueGO Cytoscape App	2
Installation	2
License	2
Documentation	3
Annotation and ontology sources	3
ClueGO features	5
ClueGO overview	7
ClueGO analysis steps	8
ClueGO analysis with multiple lists	8
ClueGO visualization of preselected GO terms and pathways	10
ClueGO supported organisms	11
ClueGO predefined selection criteria	12
ClueGO result visualizations	14
ClueGO visual styles	15
Kappa score	17
Statistics	19
ClueGO automation	21
Bibliography	23

ClueGO Cytoscape App

Cytoscape is a major computational platform for the analysis and visualization of biological networks [1]. Apps with specific functionality extend Cytoscape and are available at the Cytoscape App Store [2].

ClueGO is a Cytoscape App that facilitates the biological interpretation of large lists of genes and proteins by selecting representative Gene Ontology terms and pathways from multiple ontologies and visualizes them into functionally organized networks. ClueGO is extended by CluePedia, that allows a detailed pathways analysis.

ClueGO supports many organisms. Human and mouse data are included by default in the App, and more than 200 organisms are available for download.

ClueGO and CluePedia should be downloaded from the Cytoscape App Store through the Cytoscape App Manager (Cytoscape menu, Apps → ClueGO).

ClueGO v2.0.0 (Cytoscape 3.0.0) was released in 13th of December 2012. The functionality of ClueGO earlier described is valid in versions ported to Cytoscape 3.

Installation

System Requirements:

- Windows, Linux, Unix or MacOS operating system.
- 16 GB RAM recommended. Hard disk with at least 100 MB free (for example files).
- Java 1.8+ needed.
- Cytoscape 3.7.+ installed: <https://cytoscape.org/>

At the first startup, the ClueGOConfiguration folder containing precompiled ClueGO files and sample files will be created in the user home folder. If this folder is removed or the content damaged, it will be recreated automatically at the next startup.

License

The authors wish to make ClueGO available on a nonexclusive basis to interested parties for noncommercial internal research purposes.

ClueGO is available free of charge only to academic, government, and other nonprofit institutions for noncommercial, nonprofit internal research purposes. Note that the license terms specifically limit its use to such purposes. License applications are reviewed manually and an email including the license key will be sent to you as soon as the evaluation process has been completed.

To obtain a license key please visit: <http://www.ici.upmc.fr/cluego/cluegoLicense.shtml>.

Documentation

The biological interpretation of large gene clusters derived from high-throughput experiments is a real challenge. Many ontology sources exist in order to capture biological information in a meaningful way.

Annotation and ontology sources

The Gene Ontology (GO) project [3] aims to capture the increasing knowledge on gene function in a controlled vocabulary applicable to all the organisms. GO describes gene products in terms of their associated biological processes, cellular components and molecular functions, terms that have a hierarchical relationship (parent-child) (Figure 1).

Due to the complexity of hierarchy structure (directed acyclic graph), the terms can be found in several different levels. The specificity of the terms varies along the tree: from very general terms (in first levels of GO) to very specific ones.

Kyoto Encyclopedia of Genes and Genomes (KEGG) [4] is a database of biological systems that integrates genomic, chemical and systemic functional information in pathways.

Reactome is a free, open-source, curated and peer-reviewed pathway database that provides intuitive bioinformatics tools for the visualization, interpretation and analysis of pathway

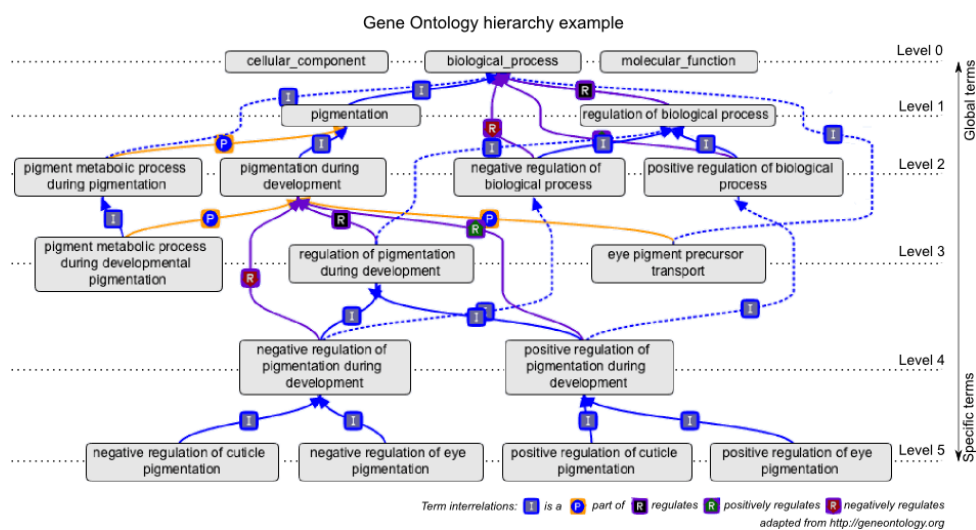


Figure 1: GO hierarchy example

knowledge to support basic research, genome analysis, modeling, systems biology and education [5].

WikiPathways is an open, collaborative platform dedicated to the curation of biological pathways [6].

The Planteome is a resource for common reference ontologies and applications for plant biology [7].

Ensembl [8] is a genome browser for vertebrate genomes that supports research in comparative genomics, evolution, sequence variation and transcriptional regulation. Ensembl annotate genes, computes multiple alignments, predicts regulatory function and collects disease data. Ensembl tools include BLAST, BLAT, BioMart and the Variant Effect Predictor (VEP) for all supported species.

The National Center for Biotechnology Information [9] advances science and health by providing access to biomedical and genomic information.

InterPro [10] provides functional analysis of proteins by classifying them into families and predicting domains and important sites. Protein signatures from a number of member databases are combined into a single searchable resource, capitalising on their individual strengths to produce a powerful integrated database and diagnostic tool.

STRING [11] is a database of known and predicted protein-protein interactions. The inter-

actions include direct (physical) and indirect (functional) associations; they stem from computational prediction, from knowledge transfer between organisms, and from interactions aggregated from other (primary) databases.

The CORUM [12] database is a collection of experimentally verified mammalian protein complexes, key molecular entities that integrate multiple gene products to perform cellular functions.

IntAct [13] provides a freely available, open source database system and analysis tools for molecular interaction data. All interactions are derived from literature curation or direct user submissions and are freely available.

BioCyc [14] is a collection of 14560 Pathway/Genome Databases (PGDBs), plus software tools for exploring them.

For a complete view on the studied process, several ontology sources should be consulted in order to integrate their complementary information. For each gene there is a large amount of information in each of these ontology sources. This makes the analysis of the relationship between genes and between terms very difficult to represent and comprehend. Also, for close related terms, a high degree of redundancy of their associated genes exists.

ClueGO, an open-source Java tool was design to extract the non-redundant biological information for large clusters of genes, using multiple ontologies. ClueGO is a Cytoscape [1] App and is taking advantage of its complex visualization environment.

ClueGO features

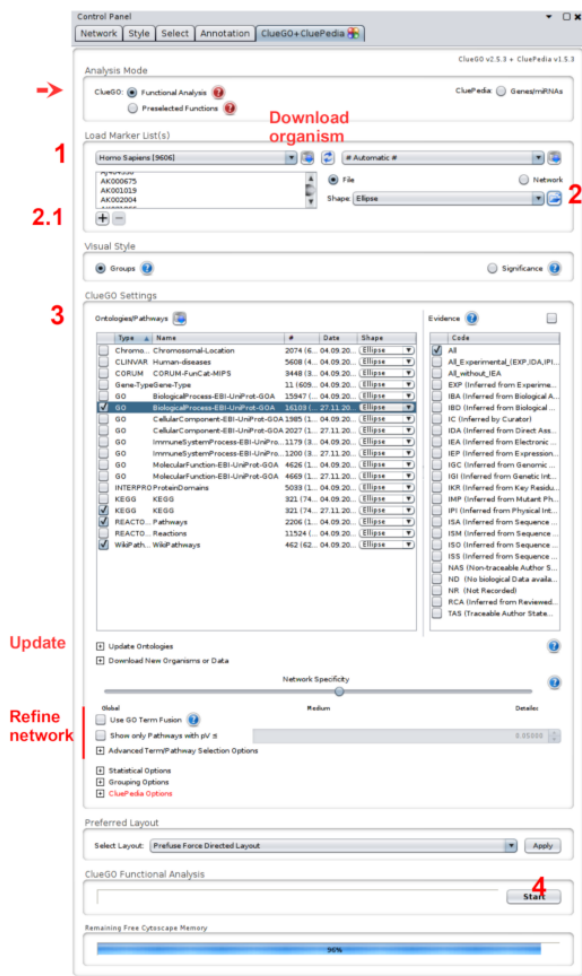
1. ClueGO allows the analysis of a single marker list (cluster) or compares multiple lists with markers.
2. ClueGO supports many organisms. Human and mouse are included in ClueGO by default and more than 200 other organisms are available for download.
3. ClueGO recognizes automatically multiple gene and protein identifiers.
4. Identifiers can be pasted, uploaded from lists or from existing networks.

5. ClueGO allows the simultaneous analysis of multiple annotation and ontology sources.
6. Users of ClueGO can automatically update the ontology sources.
7. Predefined criteria for the selection of terms/pathways are provided. These criteria were defined using lists of 200 genes and can be customized.
8. Multiple statistical methods for the enrichment calculation are provided.
9. Results are illustrated as a functionally grouped network, as plots and tables.
10. Results are automatically mapped on the network in different visual styles.
11. The results and the entire project can be saved.
12. Selected terms and pathways are included in the network and are linked based on the kappa score that shows how similar are their associated genes.
13. Multiple visualization styles of the network are available.
14. To reduce the redundancy of GO terms can be applied the fusion of related terms with similar associated genes.
15. ClueGO can visualize also preselected GO terms and pathways.
16. ClueGO functionality is REST enabled, and can be accessed from R, Python, etc.
17. ClueGO can be used in combination with CluePedia to get detailed information of pathways.
18. Easy extendable.

The selection panel includes ClueGO features for enrichment analyses. The results of the enrichment analysis are shown as a functionally grouped network, as tables and graphs. A summary of uploaded markers can be found under the network.



ClueGO analysis steps



ClueGO: Functional Analysis

Representative GO terms and pathways for one or multiple lists of markers are selected and visualized as a functionally grouped network.

- >200 organisms
- one marker list (cluster) or multiple clusters loaded from files or network
- identifiers: genes, proteins, miRNAs (CluePedia step: find target genes)
- stats: hypergeometric test, Benjamini step down correction
- results as network of terms, stats table, graphs
- project can be saved & reloaded

Requirements:

Download the organism/ask for ClueGO files for additional organisms

Update ontologies

Analysis steps:

1. Select organism
2. Upload markers
 - 2.1 add additional lists/clusters
3. Select ontologies
4. Start

Refine the network (optional features):

If too many terms:

- Use GO Term Fusion
- Show only Pathways with $pV < 0.05$
- Advanced Term/Pathway Selection Options
 - increase #/% of genes/term
 - restrict the GO interval
- Use EXP evidence code

If no terms found:

- Advanced selection options
 - decrease #/% of genes/term

Figure 3: ClueGO analysis steps

Requirements:

Download the organism or ask for additional organisms. If needed, update ontologies.

Analysis steps:

1. Select the organism
2. Upload markers. Optional (2.1) Add additional lists
3. Select ontologies
4. Start the analysis

ClueGO analysis with multiple lists

After starting Cytoscape, ClueGO can be found in the Cytoscape menu/Apps. Once selected, it will display the ClueGO selection panel options (Figure 4). ClueGO provides short explanations for each feature (click on help icons).

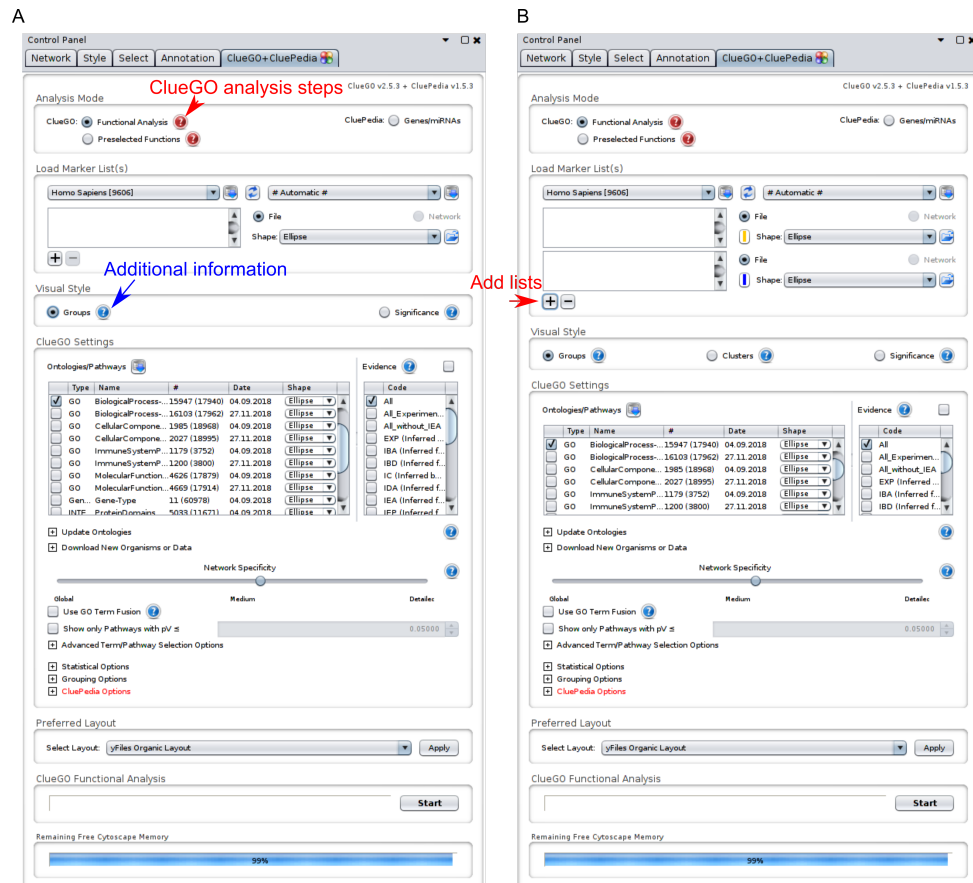


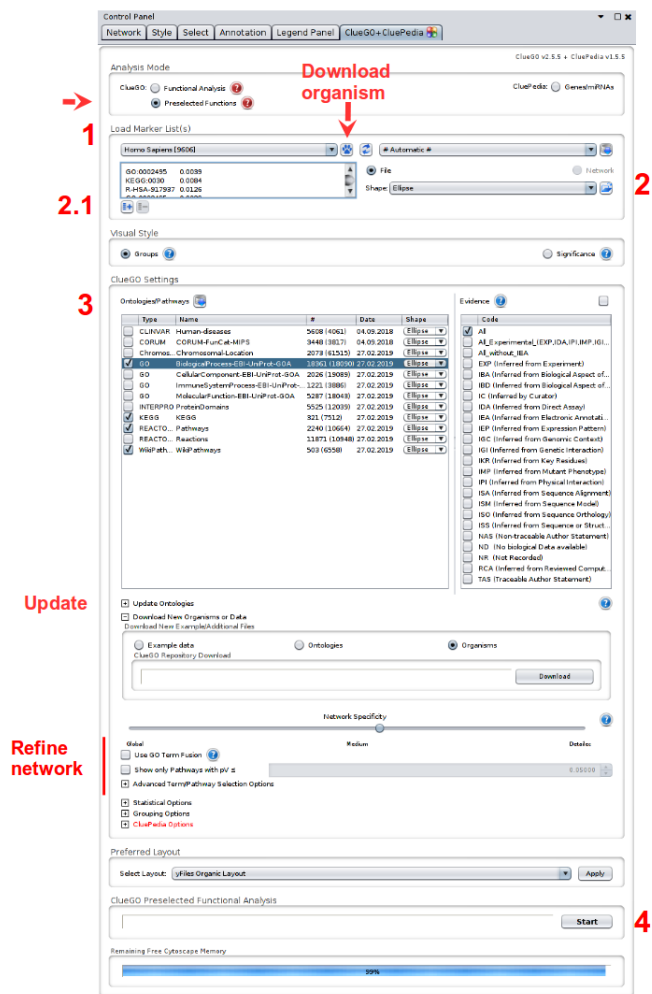
Figure 4: *ClueGO selection Panel with one (A) or multiple (B) lists*

The user can choose between a ClueGO functional analysis, a ClueGO like visualization of preselected terms or to run CluePedia for a detailed analysis of pathways (Analysis Mode section). ClueGO analysis steps are summarized (click on the help icon).

Additional lists can be added by selecting on "+" in Load Marker Lists(s) section. A color is automatically attributed to each uploaded list.

ClueGO visualization of preselected GO terms and pathways

GO terms and pathways resulted from enrichments can be visualized with ClueGO. Upload your file with pathways and Pvalues obtained from the enrichment analysis, mapp the terms on ontologies and visualized them into a ClueGO funtionally grouped network.



ClueGO: Preselected Functions

GO terms and pathways resulting from enrichment analyses are visualized as a ClueGO functionally grouped network.

Requirements:

Download the organism/ask for ClueGO files for additional organisms
Update ontologies

Identifiers for GO terms and/or pathways (KEGG, Reactome, WikiPathways) and corresponding p-values (optional) as a tab delimited file. Identifiers can be also directly pasted into the Load Marker List text field.

Goid.PathwayID	P value
GO:0008089	0.0208
GO:0015721	0.0039
GO:0019369	0.0157
GO:0002495	0.0039
KEGG:00330	0.0084
R-HSA-917937	0.0126
WP:2586	0.0075

Analysis steps:

1. Select organism
2. Upload pathways and p-values (optional)
 - 2.1 add additional lists/clusters
3. Select ontologies
4. Start

Figure 5: ClueGO preselected pathways

ClueGO supported organisms

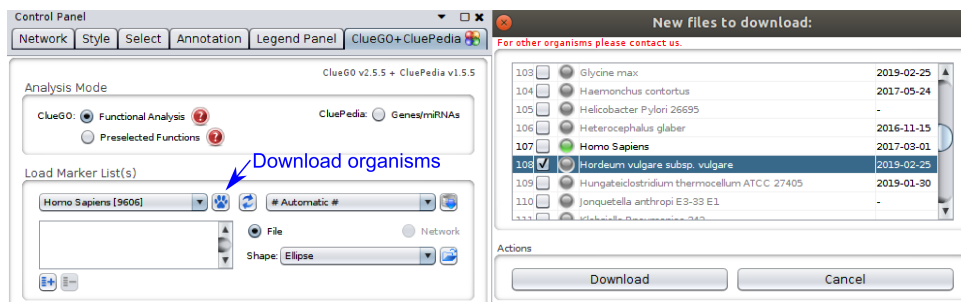


Figure 6: *ClueGO supported organisms*

Human and mouse data are included in ClueGO. More than 200 other organisms can be downloaded (Figure 6). Downloaded organisms are shown in green. Select a new organism and click Download to install it.

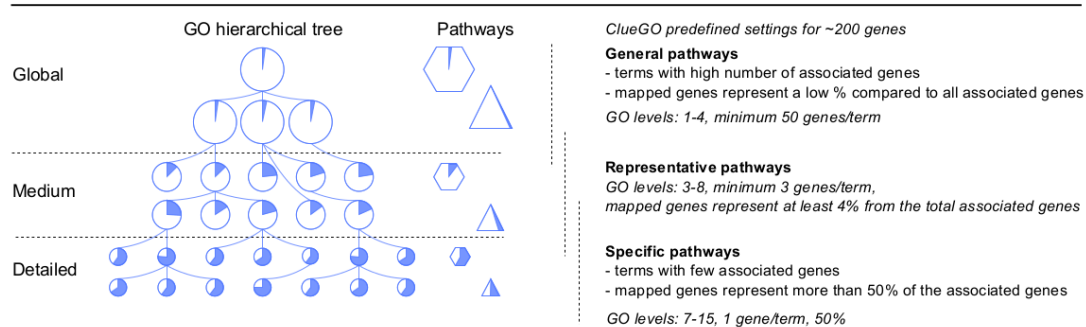
Additional organisms can be added upon request (contact the authors). The taxonomy identifier and functional annotations: genes/proteins and their corresponding GO terms are needed. This data should be GAF or tab delimited.

Functional annotations for non-model organism sequences can be obtained using annotation software like: EggNOG mapper [15].

ClueGO predefined selection criteria

ClueGO provides predefined selection criteria of representative pathways defined using lists of around 200 genes. These criteria include the 3-8 GO tree interval, a minimum of 3 genes from the uploaded list found to be associated to a term, and that these genes represent at least 4% from the total number of associated genes.

Network specificity



Example:

ClueGO NK cell top 200 upregulated genes, GO, KEGG, Reactome (09.09.2016), Percentage of mapped genes/term is shown. Node size: term significance

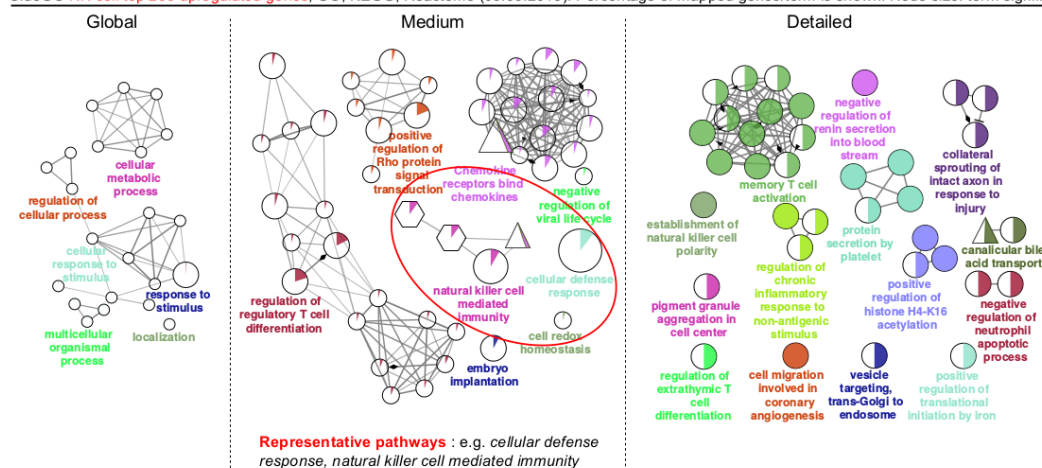


Figure 7: ClueGO predefined selection criteria

ClueGO Term selection

ClueGO: GO Term/Pathway Selection (#/% Genes)

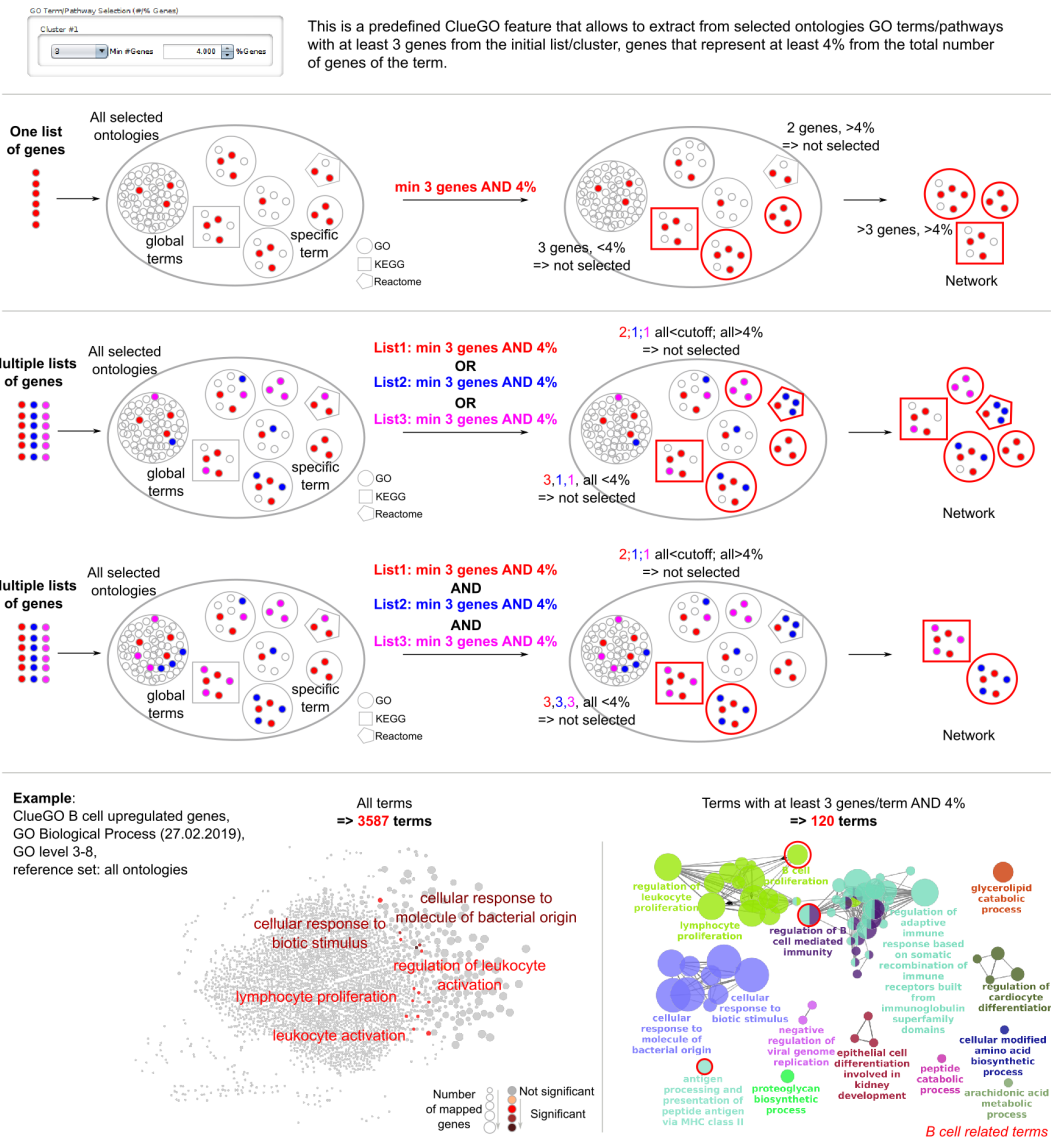


Figure 8: Term selection

ClueGO result visualizations

ClueGO Graphical representations

ClueGO results are illustrated as a functionally grouped network of terms/pathways. The most significant term of a group is considered to be the leading term and it is highlighted on the network and in the statistical summary. If groups overlap, the next most significant term is highlighted. In addition, selected terms and functional groups are illustrated as bar or pie charts. A term included in several groups is multiple times listed/ counted in the charts. The multiple occurrence is marked in the term name with "-" and the number of additional occurrences.

All statistical results are included in the ClueGO table and it can be saved as project, graphical representations (svg, png) or table (xls).

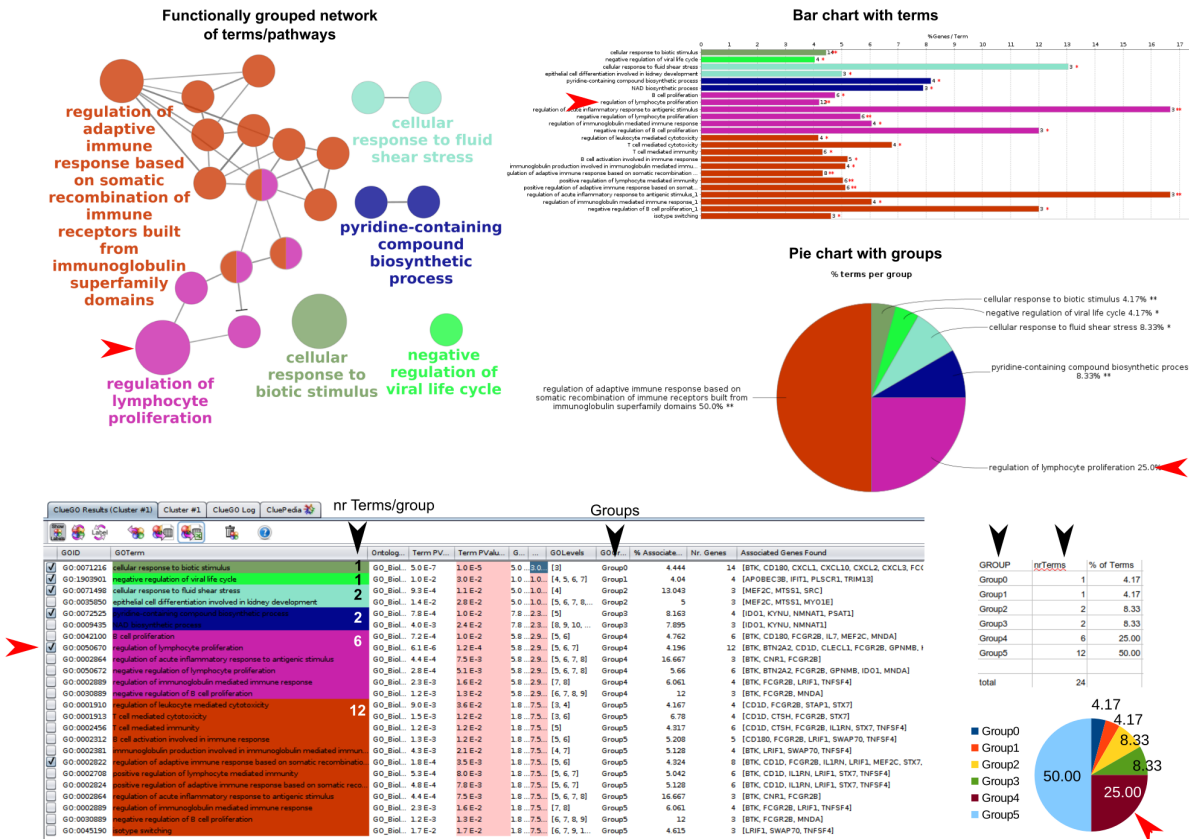


Figure 9: ClueGO visualizations

ClueGO visual styles

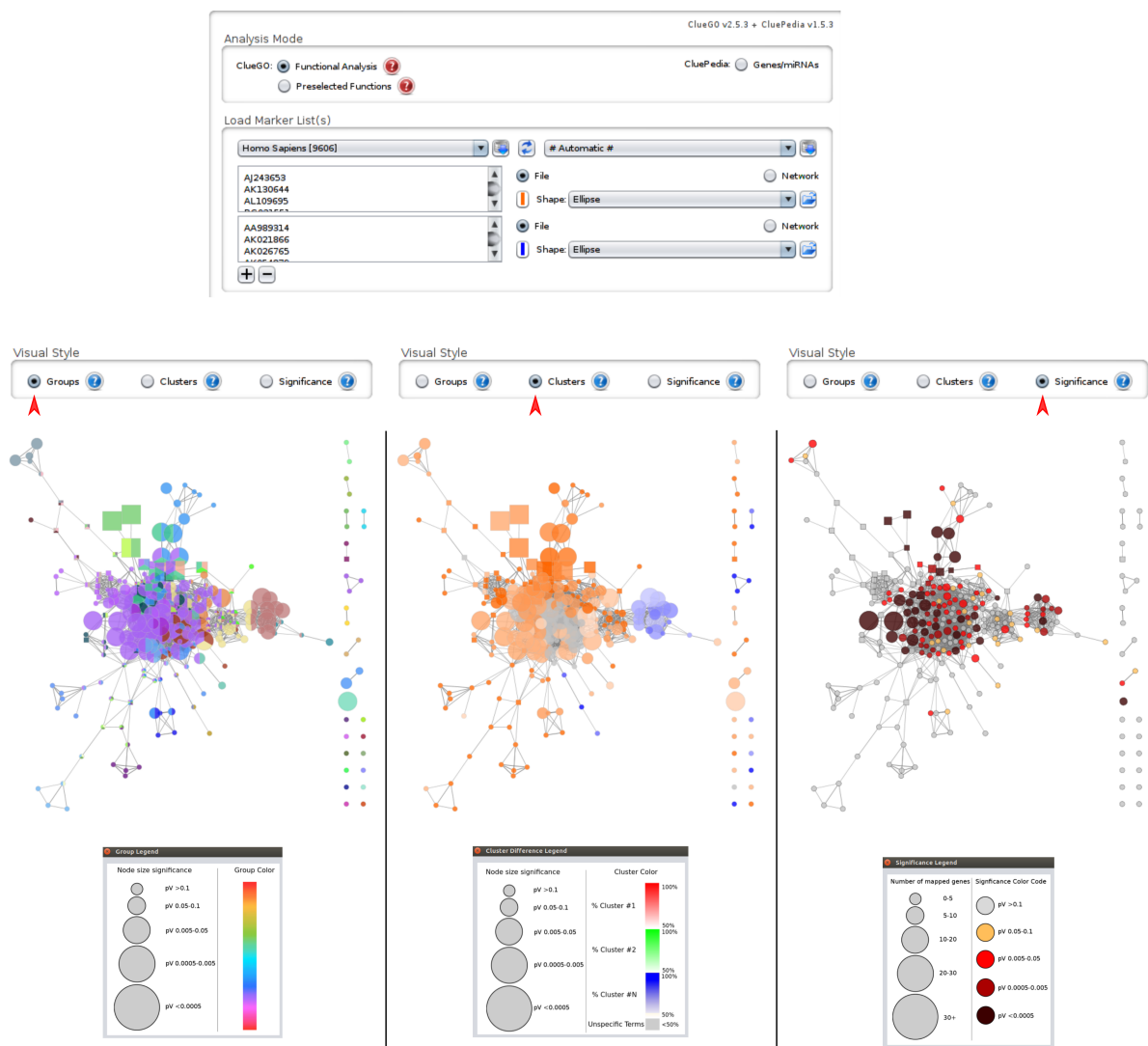


Figure 10: *ClueGO* visual styles

Kappa score

In ClueGO the Kappa score is used to define term-term interactions shown as edges on the network.

Kappa score

Cohen's kappa coefficient measures inter-rater agreement for categorical items.

$$\kappa = \frac{p_o - p_e}{1 - p_e} = 1 - \frac{1 - p_o}{1 - p_e}$$

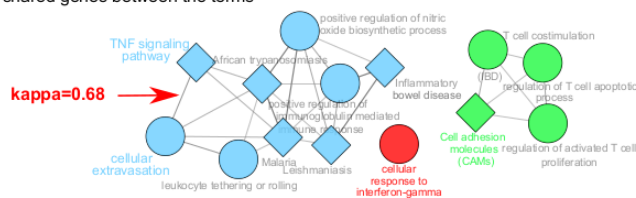
po: relative observed agreement among raters
pe: hypothetical probability of chance agreement

		GO Term 2	
		yes	no
GO Term1	yes	a	b
	no	c	d

$po = (a + d) / (a+b+c+d)$
 Rater/Class GOTerm1: $marginal_a = ((a + b) * (a+c)) / (a+b+c+d)$
 Rater/Class GOTerm2: $marginal_b = ((c + d) * (b+d)) / (a+b+c+d)$
 $pe = (marginal_a + marginal_b) / (a+b+c+d)$

https://en.wikipedia.org/wiki/Coherent%27s_kappa

In ClueGO kappa score is used to define term-term interrelations (edges) and functional groups based on shared genes between the terms



Example: Functional analysis of ClueGO Th1 related genes. GO, KEGG (2016.09.09), Standard selection, Fusion

Step 1: A binary matrix with GO terms/pathways and genes is created

[illegible]

Step 2: Kappa score is calculated between all the terms. A term-term similarity matrix is calculated.

		TNF signaling pathway	
		yes	no
cellular extravasation	yes	4	1
	no	1	8

$$p_o = (4 + 8) / (4 + 1 + 1 + 8) = 0.857$$

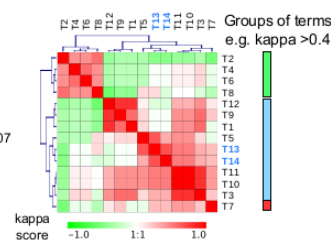
Class GOTerm1: $\text{marginal}_a = ((4 + 1) * (4 + 1)) / (4 + 1 + 1 + 8) = 1.785$

Class GOTerm2: $\text{marginal}_b = ((1 + 8) * (1 + 8)) / (4 + 1 + 1 + 8) = 5.785$

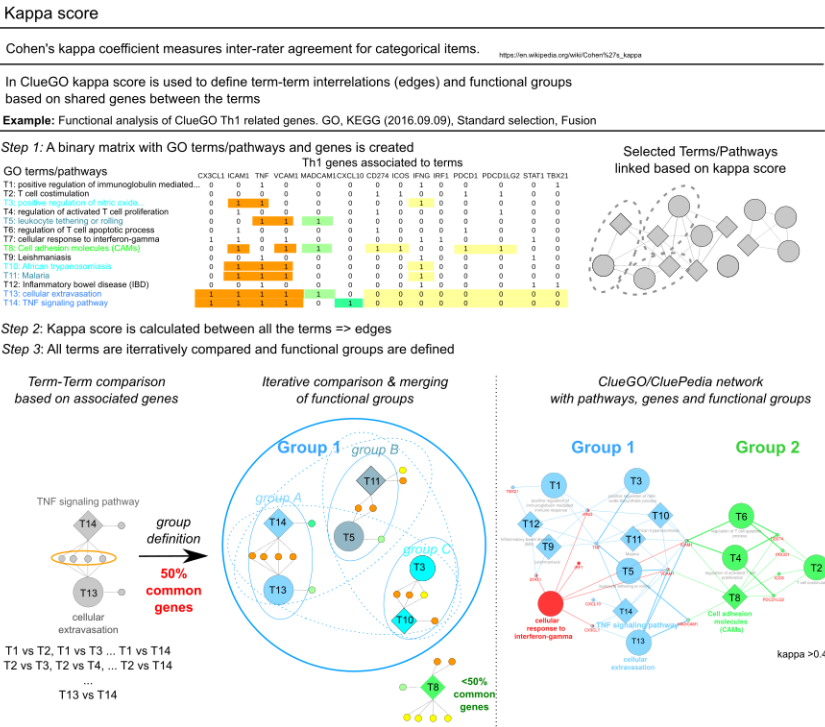
$$pe = (\text{marginal}_a + \text{marginal}_b) / (a+b+c+d) = (1.785+5.785)/(4+1+1+8) = 0.5407$$

$$k = 1 - (1 - 0.857)(1 - 0.5407) = 0.685$$

Step 3: All terms are iteratively compared and functional groups are defined

Figure 12: *Kappa score*

Kappa score is used also to associate terms and pathways into functional groups based on shared genes.



Statistics

P value calculation

ClueGO term P-value calculation example

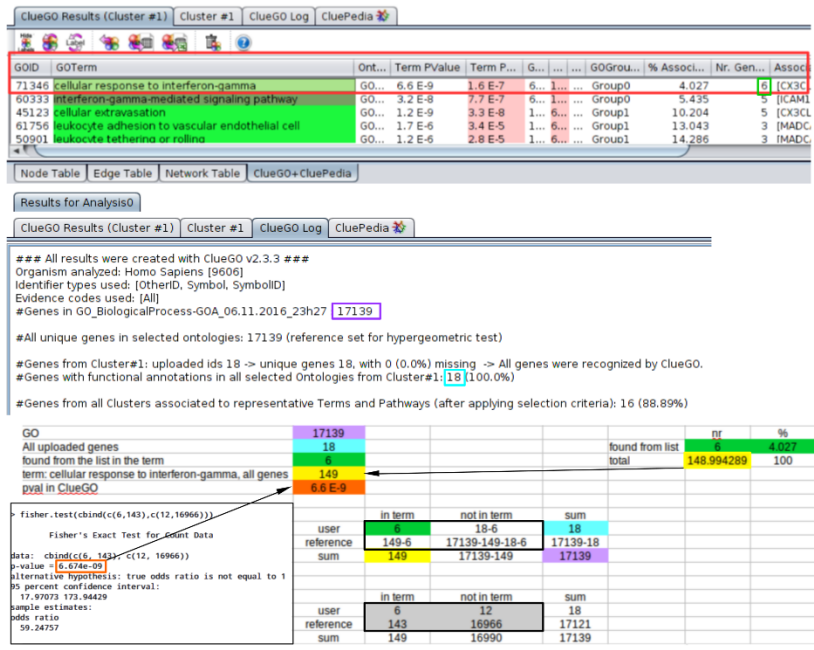


Figure 14: Stats example

Reference Set

Reference Set

The uploaded list is iteratively compared to all terms/pathways from selected ontologies to find terms having associated at least one uploaded gene. The significance is calculated using as reference set all selected Ontologies. Additionally, predefined reference sets or custom reference lists can be used.

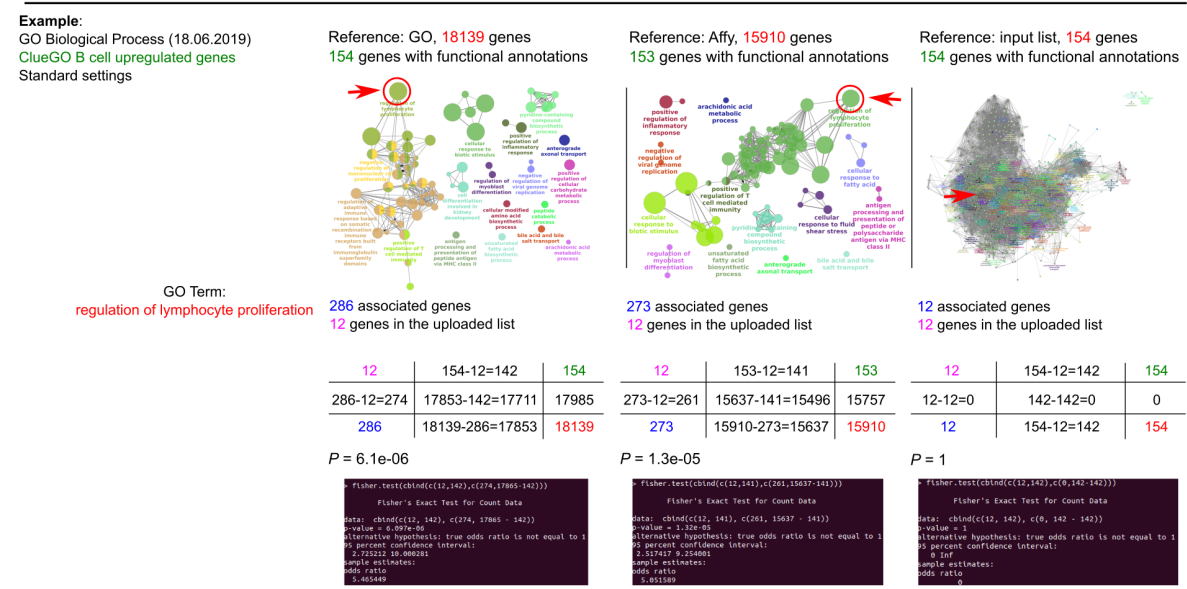
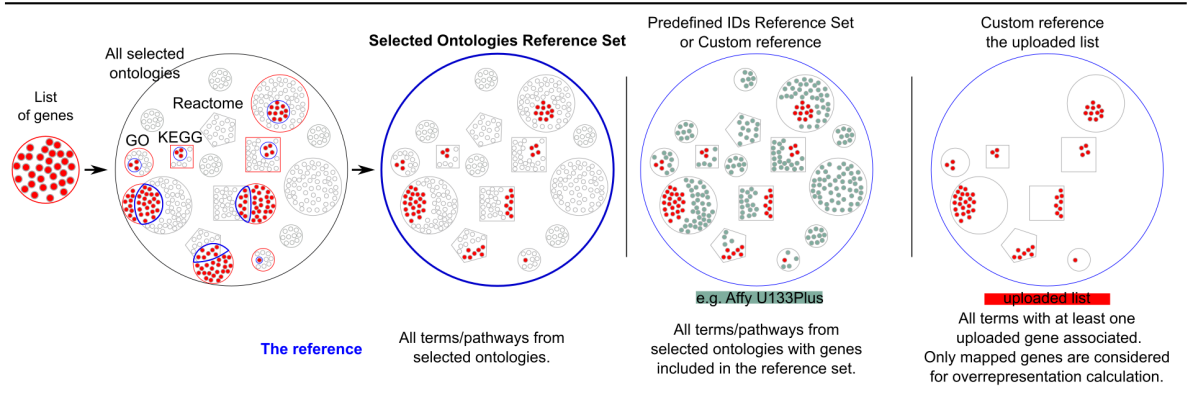


Figure 15: Reference Set

ClueGO automation

Cytoscape Automation enables scientific workflows written in many languages through the CyREST [16]. ClueGO (v2.5.0+) implements the cyREST core plugin API and provides programmatical access to its functionality [17]. ClueGO features REST enabled can be explored in the cyREST API Swagger (Figure 16) (Cytoscape menu, Help/Automation/CyREST API).

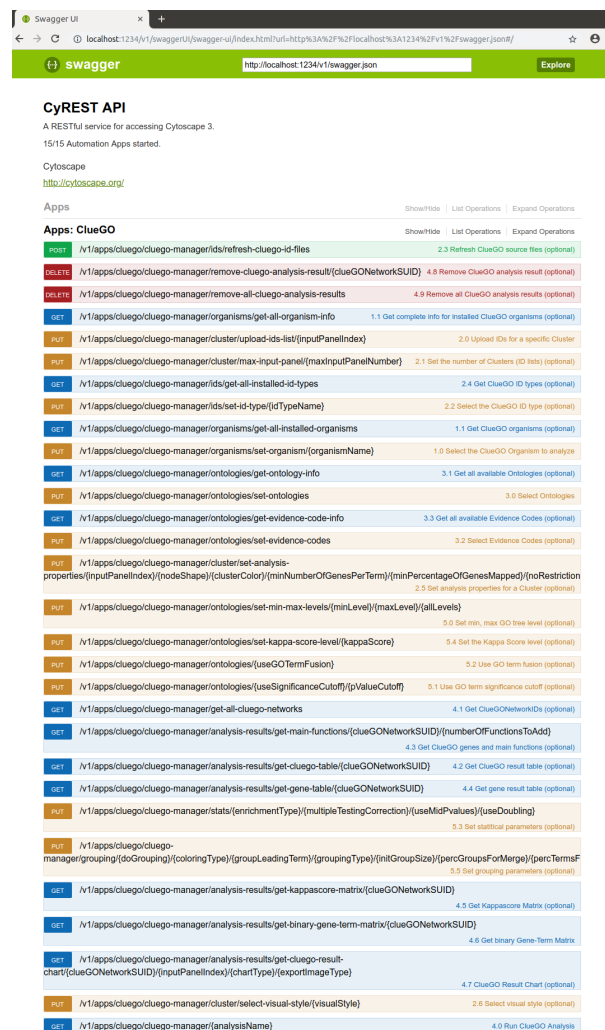


Figure 16: ClueGO in the Cytoscape API Swagger

Examples: <http://www.ici.upmc.fr/cluego/ClueGOcyRESTExample.R> and <http://www.ici.upmc.fr/cluego/ClueGOcyRESTExample.py>.

Order	Feature description	Type	Link. Base link: /v1/apps/cluego/cluego-manager
1.0	Select the ClueGO Organism to analyze	PUT	/organisms/set-organism/organismName
1.1	Get ClueGO organisms*	GET	/organisms/get-all-installed-organisms
1.2	Get all info for installed organisms*	GET	/organisms/get-all-organism-info
2.0	Upload IDs for a specific Cluster	PUT	/cluster/upload-ids-list/inputPanelIndex
2.1	Select the ClueGO ID type*	PUT	/ids/set-id-type/idTypeName
2.2	Refresh ClueGO source files*	POST	/ids/refresh-cluego-id-files
2.3	Get ClueGO ID types*	GET	/ids/get-all-installed-id-types
2.4	Set the number of Clusters (ID lists)*	PUT	/cluster/max-input-panel/maxInputPanelNumber
2.5	Set analysis properties for a Cluster*	PUT	/cluster/set-analysis-properties/inputPanelIndex/nodeShape/clusterColor... .../minNumberOfGenesPerTerm/minPercentageOfGenesMapped/noRestrictions
2.6	Select visual style*	PUT	/cluster/select-visual-style/visualStyle
3.0	Select Ontologies	PUT	/ontologies/set-ontologies
3.1	Get all available Ontologies*	GET	/ontologies/get-ontology-info
3.2	Select Evidence Codes*	PUT	/ontologies/set-evidence-codes
3.3	Get all available Evidence Codes*	GET	/ontologies/get-evidence-code-info
3.4	Set min, max GO tree level*	PUT	/ontologies/set-min-max-levels/minLevel/maxLevel/allLevels
3.5	Use GO term significance cutoff*	PUT	/ontologies/useSignificanceCutoff/pValueCutoff
3.6	Use GO term fusion*	PUT	/ontologies/useGOTermFusion
3.7	Set statistical parameters*	PUT	/stats/enrichmentType/multipleTestingCorrection/useMidPvalues/useDoubling .../useCustomReferenceSet/fullPathToReferenceSet
3.8	Set the Kappa Score level*	PUT	/ontologies/set-kappa-score-level/kappaScore
3.9	Set grouping parameters*	PUT	/grouping/doGrouping/coloringType/groupLeadingTerm/groupingType... .../initGroupSize/percGroupsForMerge/percTermsForMerge
4.0	Run ClueGO Analysis	GET	/analysisName
4.1	Get ClueGONetworkIDs*	GET	/get-all-cluego-networks
4.2	Get ClueGO result table*	GET	/analysis-results/get-cluego-table/clueGONetworkSUID
4.3	Get ClueGO genes and main functions*	GET	/analysis-results/clueGONetworkSUID/numberOfFunctionsToAdd
4.4	Get gene result table*	GET	/analysis-results/get-gene-table/clueGONetworkSUID
4.5	Get Kappascore Matrix*	GET	/analysis-results/get-kappascore-matrix/clueGONetworkSUID
4.6	Get binary Gene-Term Matrix*	GET	/analysis-results/get-binary-gene-term-matrix/clueGONetworkSUID
4.7	ClueGO Result Chart*	GET	/analysis-results/get-cluego-result-chart/clueGONetworkSUID... .../inputPanelIndex/chartType/exportImageType
4.8	Remove ClueGO analysis result*	DELETE	/remove-cluego-analysis-result/clueGONetworkSUID
4.9	Remove all ClueGO analysis results*	DELETE	/remove-all-cluego-analysis-results

References

- [1] P Shannon, A Markiel, O Ozier, N S Baliga, J T Wang, D Ramage, N Amin, B Schwikowski, T Ideker. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*, 13:2498–2504, 2003.
- [2] S. Lotia, J. Montojo, Y. Dong, G. D. Bader, A. R. Pico. Cytoscape app store. *Bioinformatics*, 29(10):1350–1351, May 2013.
- [3] M Ashburner, C A Ball, J A Blake, D Botstein, H Butler, J M Cherry, A P Davis, K Dolinski, S S Dwight, J T Eppig, M A Harris, D P Hill, L Issel-Tarver, A Kasarskis, S Lewis, J C Matese, J E Richardson, M Ringwald, G M Rubin, G Sherlock. Gene ontology: tool for the unification of biology The Gene Ontology Consortium. *Nat Genet*, 25:25–29, 2000.
- [4] M Kanehisa, S Goto, S Kawashima, A Nakaya. The KEGG databases at GenomeNet. *Nucleic Acids Res*, 30:42–46, 2002.
- [5] A. Fabregat, S. Jupe, L. Matthews, K. Sidiropoulos, M. Gillespie, P. Garapati, R. Haw, B. Jassal, F. Korninger, B. May, M. Milacic, C. D. Roca, K. Rothfels, C. Sevilla, V. Shamovsky, S. Shorsler, T. Varusai, G. Viteri, J. Weiser, G. Wu, L. Stein, H. Hermjakob, P. D’Eustachio. The Reactome Pathway Knowledgebase. *Nucleic Acids Res.*, 46(D1):D649–D655, Jan 2018.
- [6] T. Kelder, M. P. van Iersel, K. Hanspers, M. Kutmon, B. R. Conklin, C. T. Evelo, A. R. Pico. WikiPathways: building research communities on biological pathways. *Nucleic Acids Res.*, 40(Database issue):D1301–1307, Jan 2012.
- [7] L. Cooper, P. Jaiswal. The Plant Ontology: A Tool for Plant Genomics. *Methods Mol. Biol.*, 1374:89–114, 2016.
- [8] B. L. Aken, P. Achuthan, W. Akanni, M. R. Amode, F. Bernsdorff, J. Bhai, K. Billis, D. Carvalho-Silva, C. Cummins, P. Clapham, L. Gil, C. G. Giron, L. Gordon, T. Hourlier, S. E. Hunt, S. H. Janacek, T. Juettemann, S. Keenan, M. R. Laird, I. Lavidas, T. Maurel, W. McLaren, B. Moore, D. N. Murphy, R. Nag, V. Newman, M. Nuhn, C. K. Ong, A. Parker, M. Patricio, H. S. Riat, D. Sheppard, H. Sparrow, K. Taylor, A. Thormann, A. Vullo, B. Walts, S. P. Wilder, A. Zadissa, M. Kostadima, F. J. Martin, M. Muffato, E. Perry, M. Ruffier, D. M. Staines, S. J. Trevanion, F. Cunningham, A. Yates, D. R. Zerbino, P. Flicek. Ensembl 2017. *Nucleic Acids Res.*, 45(D1):D635–D642, 01 2017.
- [9] R. Agarwala, T. Barrett, J. Beck, D. A. Benson, C. Bollin, E. Bolton, D. Bourexis, J. R. Brister, S. H. Bryant, K. Canese, M. Cavanaugh, C. Charowhas, K. Clark, I. Dondoshansky, M. Feolo, L. Fitzpatrick, K. Funk, L. Y. Geer, V. Gorelenkov, A. Graeff, W. Hlavina, B. Holmes, M. Johnson, B. Kattman, V. Khotomlianski, A. Kimchi, M. Kimelman, M. Kimura, P. Kitts, W. Klimke, A. Kotliarov, S. Krasnov, A. Kuznetsov, M. J. Landrum, D. Landsman, S. Lathrop, J. M. Lee, C. Leubsdorf, Z. Lu, T. L. Madden, A. Marchler-Bauer, A. Malheiro, P. Meric, I. Karsch-Mizrachi, A. Mnev, T. Murphy, R. Orris, J. Ostell, C. O’Sullivan, V. Palanigobu, A. R. Panchenko, L. Phan, B. Pierov, K. D. Pruitt, K. Rodarmer, E. W. Sayers, V. Schneider, C. L. Schoch, G. D. Schuler, S. T. Sherry, K. Siyan, A. Soboleva, V. Sousoff, G. Starchenko, T. A. Tatusova, F. Thibaud-Nissen, K. Todorov, B. W. Trawick, D. Vakarov, M. Ward, E. Yaschenko, A. Zasytkin, K. Zbicz. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.*, 46(D1):D8–D13, Jan 2018.

- [10] P. Jones, D. Binns, H. Y. Chang, M. Fraser, W. Li, C. McAnulla, H. McWilliam, J. Maslen, A. Mitchell, G. Nuka, S. Pesseat, A. F. Quinn, A. Sangrador-Vegas, M. Scheremetjew, S. Y. Yong, R. Lopez, S. Hunter. InterProScan 5: genome-scale protein function classification. *Bioinformatics*, 30(9):1236–1240, May 2014.
- [11] D. Szklarczyk, A. Franceschini, S. Wyder, K. Forslund, D. Heller, J. Huerta-Cepas, M. Simonovic, A. Roth, A. Santos, K. P. Tsafou, M. Kuhn, P. Bork, L. J. Jensen, C. von Mering. STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.*, 43(Database issue):D447–452, Jan 2015.
- [12] M. Giurgiu, J. Reinhard, B. Brauner, I. Dunger-Kaltenbach, G. Fobo, G. Frishman, C. Montrone, A. Ruepp. CORUM: the comprehensive resource of mammalian protein complexes-2019. *Nucleic Acids Res.*, 47(D1):D559–D563, Jan 2019.
- [13] S. Orchard, M. Ammari, B. Aranda, L. Breuza, L. Briganti, F. Broackes-Carter, N. H. Campbell, G. Chavali, C. Chen, N. del Toro, M. Duesbury, M. Dumousseau, E. Galeota, U. Hinz, M. Iannucci, S. Jagannathan, R. Jimenez, J. Khadake, A. Lagreid, L. Licata, R. C. Lovering, B. Meldal, A. N. Melidoni, M. Milagros, D. Peluso, L. Perfetto, P. Porras, A. Raghunath, S. Ricard-Blum, B. Roechert, A. Stutz, M. Tognolli, K. van Roey, G. Cesareni, H. Hermjakob. The MIntAct project—IntAct as a common curation platform for 11 molecular interaction databases. *Nucleic Acids Res.*, 42(Database issue):D358–363, Jan 2014.
- [14] P. D. Karp, R. Billington, R. Caspi, C. A. Fulcher, M. Latendresse, A. Kothari, I. M. Keseler, M. Krummenacker, P. E. Midford, Q. Ong, W. K. Ong, S. M. Paley, P. Subhraveti. The BioCyc collection of microbial genomes and metabolic pathways. *Brief. Bioinformatics*, Aug 2017.
- [15] J. Huerta-Cepas, K. Forslund, L. P. Coelho, D. Szklarczyk, L. J. Jensen, C. von Mering, P. Bork. Fast Genome-Wide Functional Annotation through Orthology Assignment by eggNOG-Mapper. *Mol. Biol. Evol.*, 34(8):2115–2122, 08 2017.
- [16] K. Ono, T. Muetze, G. Kolishovski, P. Shannon, B. Demchak. CyREST: Turbocharging Cytoscape Access for External Tools via a RESTful API. *F1000Res*, 4:478, 2015.
- [17] B. Mlecnik, J. Galon, G. Bindea. Automated exploration of Gene Ontology term and pathway networks with ClueGO-REST. *Bioinformatics*, Mar 2019.