Published online 7 June 2022 *Nucleic Acids Research, 2022, Vol. 50, Web Server issue* ***W405–W411***
https://doi.org/10.1093/nar/gkac480

Heml 2.0: an online service for heatmap illustration

Wanshan Ning[1,2,†], Yuxiang Wei[2,†], Letian Gao[2], Cheng Han[2], Yujie Gou[2], Shanshan Fu[2], Dan Liu[2], Chi Zhang[2], Xinhe Huang[2], Sicheng Wu[2], Di Peng[2], Chenwei Wang[2] and Yu Xue[2,3,*]

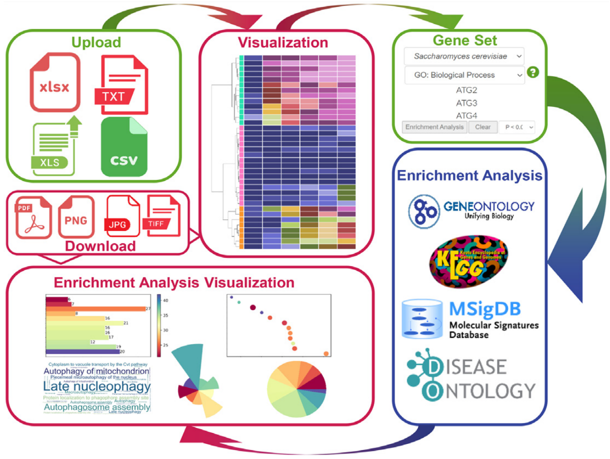
[1]Department of Clinical Laboratory, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei 430022, China, [2]Key Laboratory of Molecular Biophysics of Ministry of Education, Hubei Bioinformatics and Molecular Imaging Key Laboratory, Center for Artificial Intelligence Biology, College of Life Science and Technology, Huazhong University of Science and Technology, Wuhan, Hubei 430074, China and [3]Nanjing University Institute of Artificial Intelligence Biomedicine, Nanjing 210031, China

Received March 02, 2022; Revised May 15, 2022; Editorial Decision May 23, 2022; Accepted May 24, 2022

ABSTRACT

Recent high-throughput omics techniques have produced a large amount of biological data. Visualization of big omics data is essential to answer a wide range of biological problems. As a concise but comprehensive strategy, a heatmap can analyze and visualize high-dimensional and heterogeneous biomolecular expression data in an attractive artwork. In 2014, we developed a stand-alone software package, Heat map Illustrator (Heml 1.0), which implemented three clustering methods and seven distance metrics for heatmap illustration. Here, we significantly improved 1.0 and released the online service of Heml 2.0, in which 7 clustering methods and 22 types of distance metrics were implemented. In Heml 2.0, the clustering results and publication-quality heatmaps can be exported directly. For an in-depth analysis of the data, we further added an option of enrichment analysis for 12 model organisms, with 15 types of functional annotations. The enrichment results can be visualized in five idioms, including bubble chart, bar graph, coxcomb chart, pie chart and word cloud. We anticipate that Heml 2.0 can be a helpful web server for visualization of biomolecular expression data, as well as the additional enrichment analysis. Heml 2.0 is freely available for all users at: https://hemi.biocuckoo.org/.

GRAPHICAL ABSTRACT

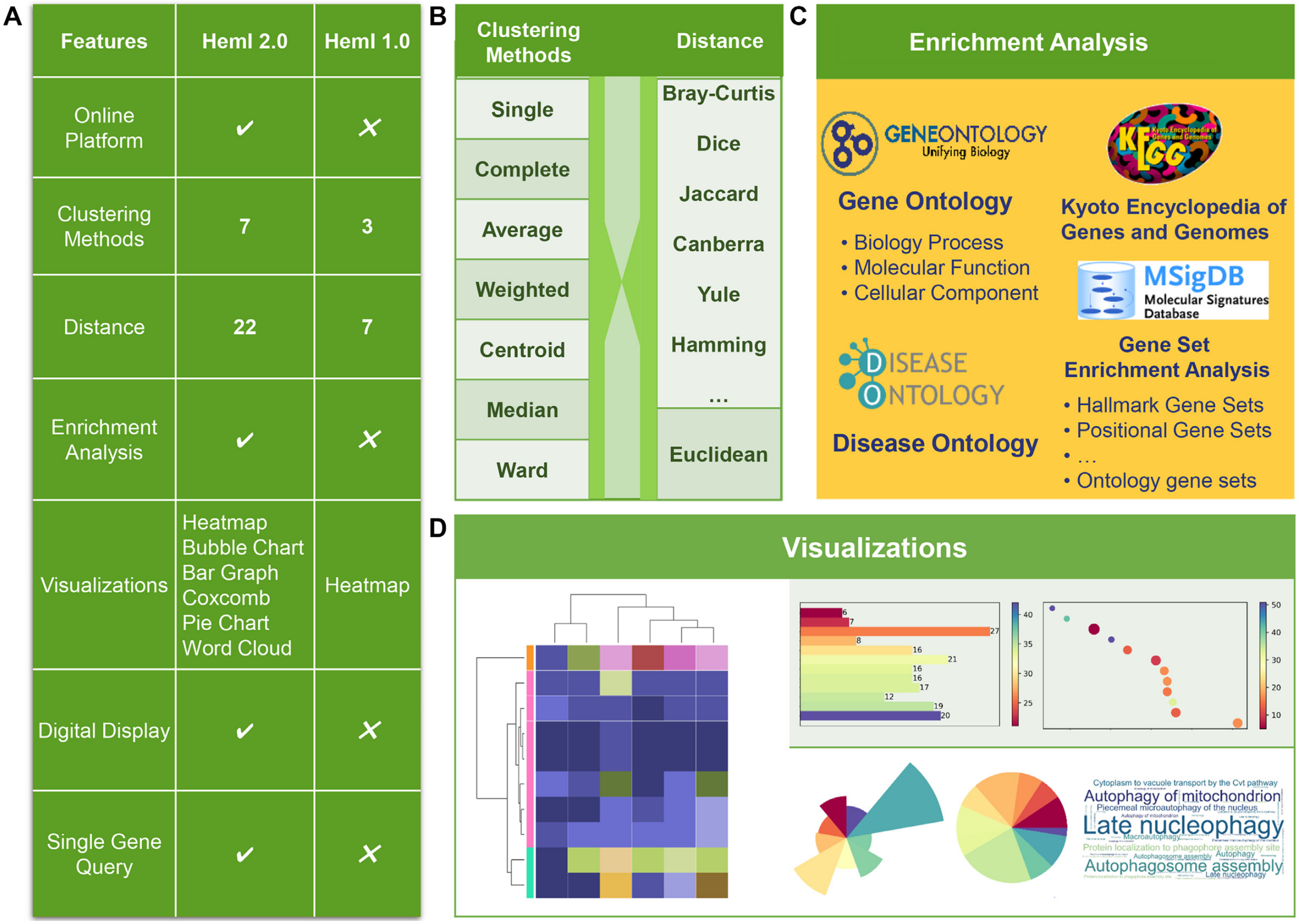


INTRODUCTION

With the rapid increase of big biological data generated by high-throughput omics technology, the demand for visualization of multi-dimensional and numeric data is urgently increasing (1,2). How to quickly and intuitively retrieve the information contained in the big data becomes more and more important. A concise, delicate and precise picture can afford a great advantage over description in words alone.

Visualization of a two-dimensional data matrix in a heatmap is a simple but efficient approach for data analysis and interpretation. The values in the matrix may represent any measurable properties such as biomolecular expression values. To estimate how many papers have heatmaps, we carefully curated all original research articles published from January 2017 to December 2021 in five professional journals, including *Nature Biotechnology*, *Cancer Cell*, *Genome Research*, *Genome Biology* and *Molecular & Cellular Proteomics*. The statistical results showed that

[*]To whom correspondence should be addressed. Tel: +86 27 87793903; Email: xueyu@hust.edu.cn
[†]The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.



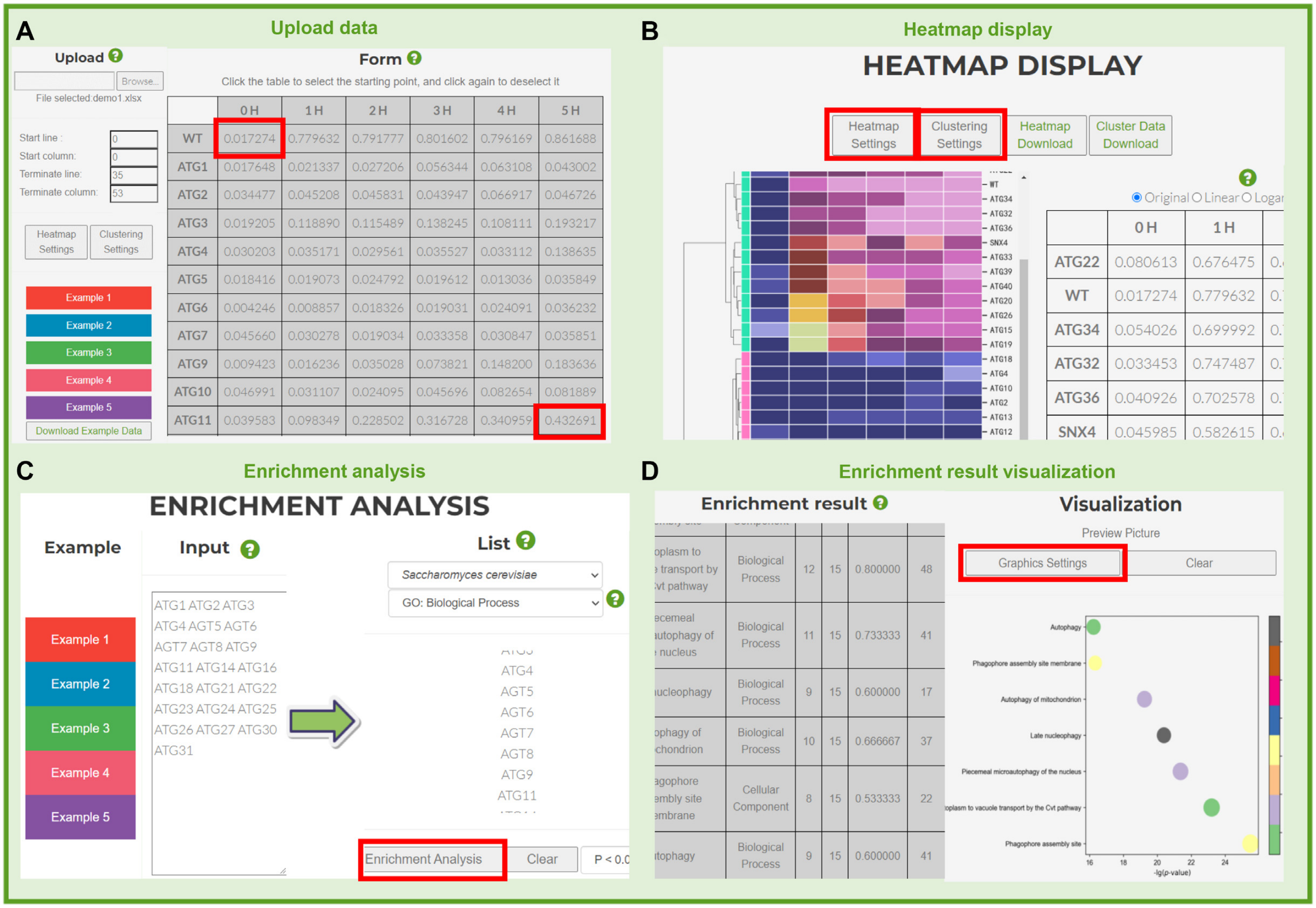
Features	HemI 2.0	HemI 1.0
Online Platform	✔	✕
Clustering Methods	7	3
Distance	22	7
Enrichment Analysis	✔	✕
Visualizations	Heatmap Bubble Chart Bar Graph Coxcomb Pie Chart Word Cloud	Heatmap
Digital Display	✔	✕
Single Gene Query	✔	✕

Figure 1. The implementation of HemI 2.0. (**A**) New features in the online service of HemI 2.0. The clustering methods, distance metrics and visualization styles were increased from 3, 7 and 1 in 1.0 to 7, 22 and 6 in 2.0, respectively. Three additional options, including enrichment analysis, digital display, and single gene query, were implemented for analyzing the data, showing the original expression data, and viewing GO annotations of individual genes. (**B**) There were 7 clustering methods and 22 types of distance metrics implemented in HemI 2.0. (**C**) The enrichment analysis option of HemI 2.0, with 15 types of functional annotations. (**D**) Five idioms for visualization of enrichment results.

about 44% (2454) of 5565 papers contained at least one heatmap (Supplementary Table S1), supporting the importance of heatmaps in visualization of big biological data.

Heatmaps can be illustrated by a considerable number of programs or tools, such as R package gpplot2, Treeview (3), D3, SPSS, GraphPad Prism, ClustVis (4), Perseus (5), TBtools (6) and MetaboAnalyst 5.0 (7) (Supplementary Table S2). However, these tools were developed for more general purposes, while SPSS and GraphPad Prism are commercial software packages. To date, there have been 41 R/Bioconductor/JAVA packages, stand-alone software packages and web servers, specifically designed for heatmap illustration (Supplementary Table S2). Programming skills are needed for using R/Bioconductor/JAVA packages such as pheatmap, ComplexHeatmaps (8) and JHeatChart. In 2014, we developed a JAVA-based stand-alone software package, Heat map Illustrator (HemI) 1.0, for biologists who are not familiar with programming. In HemI 1.0, 3 clustering methods and seven distance metrics were implemented, and heatmaps can be visualized, recolored, rescaled, rotated or exported in a customized manner (9). Later, Babicki *et al.* also implemented a highly useful and interactive heatmap viewer Heatmapper, which could rapidly generate expression, correlation or pairwise distance heatmaps, with four clustering methods and five distance metrics (10). Beyond heatmap illustration, no additional options were provided for further data processing and analysis in currently available heatmap tools.

During the past years, HemI 1.0 has become a useful tool for biologists, and many of them communicated with us and pointed out the pros and cons of this tool. Based on the users' demands and suggestions, here we significantly improved HemI, and developed the HTML5-based online service of 2.0, in which seven clustering methods and 22 types of distance metrics were included. For further data analysis, we compiled 15 types of functional annotations, and added an option of enrichment analysis based on the hypergeometric test for 12 model organisms, as well as the visualization of enrichment results. In particular, publication-quality heatmaps and the clustering and enrichment results could be exported for academic usage. We believe that HemI 2.0 can be a much more



File selected:demo1.xlsx

Start line : 0
Start column: 0
Terminate line: 35
Terminate column: 53

Click the table to select the starting point, and click again to deselect it

	0 H	1 H	2 H	3 H	4 H	5 H
WT	0.017274	0.779632	0.791777	0.801602	0.796169	0.861688
ATG1	0.017648	0.021337	0.027206	0.056344	0.063108	0.043002
ATG2	0.034477	0.045208	0.045831	0.043947	0.066917	0.046726
ATG3	0.019205	0.118890	0.115489	0.138245	0.108111	0.193217
ATG4	0.030203	0.035171	0.029561	0.035527	0.033112	0.138635
ATG5	0.018416	0.019073	0.024792	0.019612	0.013036	0.035849
ATG6	0.004246	0.008857	0.018326	0.019031	0.024091	0.036232
ATG7	0.045660	0.030278	0.019034	0.033358	0.030847	0.035851
ATG9	0.009423	0.016236	0.035028	0.073821	0.148200	0.183636
ATG10	0.046991	0.031107	0.024095	0.045696	0.082654	0.081889
ATG11	0.039583	0.098349	0.228502	0.316728	0.340959	0.432691

● Original ○ Linear ○ Logar

	0 H	1 H	
ATG22	0.080613	0.676475	0.
WT	0.017274	0.779632	0.
ATG34	0.054026	0.699992	0.
ATG32	0.033453	0.747487	0.
ATG36	0.040926	0.702578	0.
SNX4	0.045985	0.582615	0.

ATG1 ATG2 ATG3
ATG4 AGT5 AGT6
AGT7 AGT8 ATG9
ATG11 ATG14 ATG16
ATG18 ATG21 ATG22
ATG23 ATG24 ATG25
ATG26 ATG27 ATG30
ATG31

Saccharomyces cerevisiae
GO: Biological Process

ATG4
AGT5
AGT6
AGT7
AGT8
ATG9
ATG11

P < 0.0

oplasm to e transport by vt pathway	Biological Process	12	15	0.800000	48
ecemeal autophagy of e nucleus	Biological Process	11	15	0.733333	41
ucleophagy	Biological Process	9	15	0.600000	17
ophagy of chondrion	Biological Process	10	15	0.666667	37
agophore embly site embrane	Cellular Component	8	15	0.533333	22
tophagy	Biological Process	9	15	0.600000	41

Figure 2. Usage of HemI 2.0. (**A**) The numeric expression data can be directly loaded, whereas the data area can be selected. The row/column labels are editable by left-clicking on them. A heatmap can be automatically generated after clicking on the "Submit" button. (**B**) Multiple options are available to manipulate the heatmap by clicking on the "Heatmap Settings" button. Up to three layers of row/column annotations can be added by right-clicking on them. (**C**) The genes or proteins for visualizing a heatmap can be selected for further enrichment analysis. (**D**) The enrichment results can be visualized in any of the five idioms and the generated pictures can be easily re-sized and re-colored in a customized manner by clicking on the "Graphic Settings" button.

helpful tool for visualization and analysis of big biological data.

IMPLEMENTATION

The new features in HemI 2.0 were shown in Figure 1A. In HemI 2.0, the data normalization method and color mode were not changed (9). Briefly, a 256 color mode with red, green, and blue tricolor was adopted, and the scale of each color should be pre-defined from n to m (n, m range from 0 to 255, and $n < m$). Prior to visualization, the inputted biomolecular expression data can be linearly normalized as below:

$$NV = \frac{OV - MinV}{MaxV - MinV} \times (m - n + 1)$$

Where NV denotes the normalized value, and OV denotes the original value of the data ($OV > 0$). $MinV$ and $MaxV$ represent the minimal and maximal values of all OVs, respectively ($MaxV > MinV$ and $MinV > 0$).

More frequently, biologists like to analyze the logarithmic relations between different conditions and expression data. Thus, the logarithmic normalization could be implemented as below:

$$NV = \frac{log_a(OV) - log_a(MinV)}{log_a(MaxV) - log_a(MinV)} \times (m - n + 1)$$

Where a could be customized as 2 (default), 10 or e.

For each color, the NV was calculated, whereas the data point was visualized based on the tricolor NVs.

To better meet the enormous demands for analysis of the data in heatmaps, we implemented seven commonly-used methods for data clustering, including average, single, complete, weighted, centroid, median and ward linkage clustering methods (Figure 1B, Supplementary Table S3). Also, we included 22 types of distance metrics, including Euclidean, Bray-Curtis, Canberra, Chebyshev, Manhattan, Correlation, Cosine, Dice, Hamming, Jaccard, Jensen-Shannon, Kulsinski, Mahalanobis, Matching, Minkowski, Rogers-Tanimoto, Russell-Rao, Standardized Euclidean, Sokal-Michener, Sokal-Sneath, Squared Euclidean and Yule distance metrics (Figure 1B, Supplementary Table S4). In

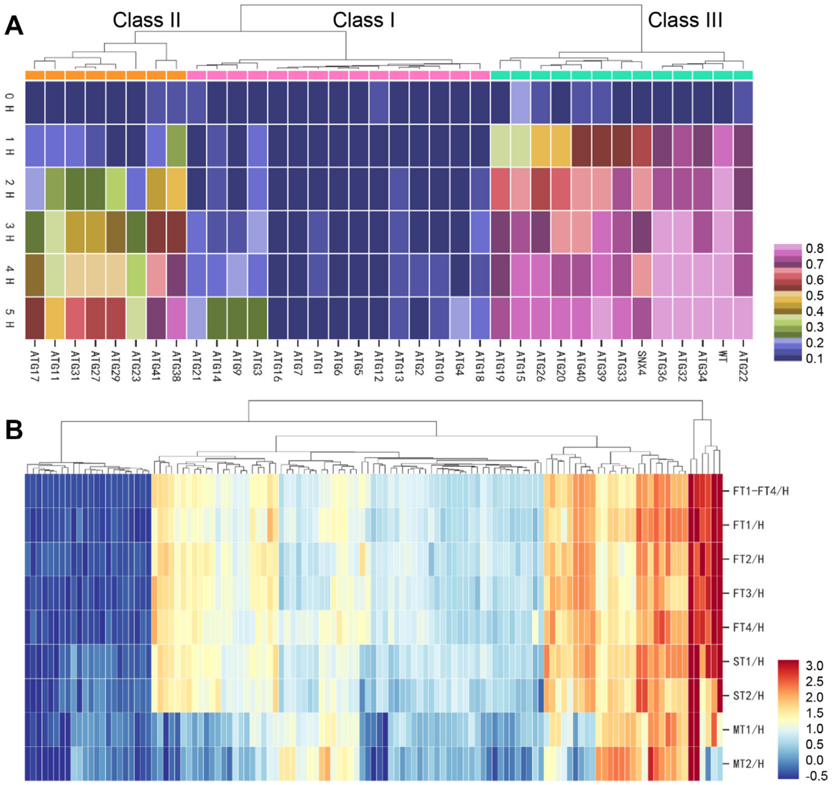


Figure 3. Illustrating heatmaps by HemI 2.0. (**A**) The autophagic phenotypes of 36 yeast strains were illustrated and clustered by HemI 2.0. A higher proportion of autophagic cells represents a stronger autophagy activity. (**B**) The 195 DEPs of COVID-19 were illustrated and clustered by HemI 2.0. A higher score denotes a higher fold change in the fatal group against the healthy cases.

HemI 2.0, the average linkage clustering method and Euclidean distance metric were selected as the default settings.

In omics-related studies, biologists frequently performed enrichment analyses of selected genes after data clustering and heatmap illustration (11,12). To fulfill this demand, we also implemented an option of enrichment analysis for 12 model species (Figure 1C), including *Homo sapiens*, *Mus musculus*, *Rattus norvegicus*, *Saccharomyces cerevisiae*, *Drosophila melanogaster*, *Arabidopsis thaliana*, *Sus scrofa*, *Canis lupus familiaris*, *Bos taurus*, *Gallus gallus*, *Caenorhabditis elegans* and *Danio rerio*. We compiled 15 types of functional annotations, including four sets of Gene Ontology (GO) annotations (All, biology processes, molecular functions, and cellular components) (13,14), Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways (15), Disease Ontology (DO) terms (16) and nine sets of Hallmark Gene Sets taken from Molecular signatures database (MSigDB) (Positional Gene Sets, Curated Gene Sets, Regulatory Target Gene Sets, Computational Gene Sets, Ontology Gene Sets, Oncogenic Signature Gene Sets, Immunologic Signature Gene Sets, Cell Type Signature Gene Sets, and All Gene Sets mixed) (17–19). The gene sets of GO biological processes were selected as the default settings. For the enrichment analysis, the hypergeometric test was used to calculate an enrichment ratio (E-ratio) and a *P*-value for each category of functional annotations. To intuitively visualize the enrichment results, we provided five idioms, including bubble chart, bar graph, coxcomb chart, pie chart and word cloud (Figure 1D). In addition, we implemented an option of digital display to show the numeric expression data in the heatmap, while individual genes were clickable to view their GO annotations (Figure 1A). Compared to other existing enrichment analysis tools such as DAVID (20) and EnrichR (21) (Supplementary Table S5), HemI 2.0 has fewer gene sets for analysis but instead offers more customizable visualizations (Supplementary Table S5).

The heatmap illustration, data clustering and enrichment analysis were implemented in Python 3.9. We developed the web interface using Django 3.2.9 and JavaScript, and deployed the online service of HemI 2.0 on cloud servers through Nginx 1.20.2. For convenience, we tested the online service on a variety of internet browsers, including Internet Explorer, Mozilla Firefox and Google Chrome.

USAGE

The online service of HemI 2.0 was designed in an easy-to-use mode. To explicitly demonstrate the usage of HemI 2.0, we performed a case study based on the data from our previously published study (22). AuTophaGy-related (*ATG*) genes have been reported to be involved mainly or exclusively in yeast autophagy. To analyze the autophagic phenotypes of the 35 *atg* knockout strains, the average autophagy activities were measured for wild-type (WT) S. cerevisiae and 35 *atg* knock-out (KO) strains at different time points from 0 to 5 h, and the corresponding data set was used as an example for HemI 2.0 (22).

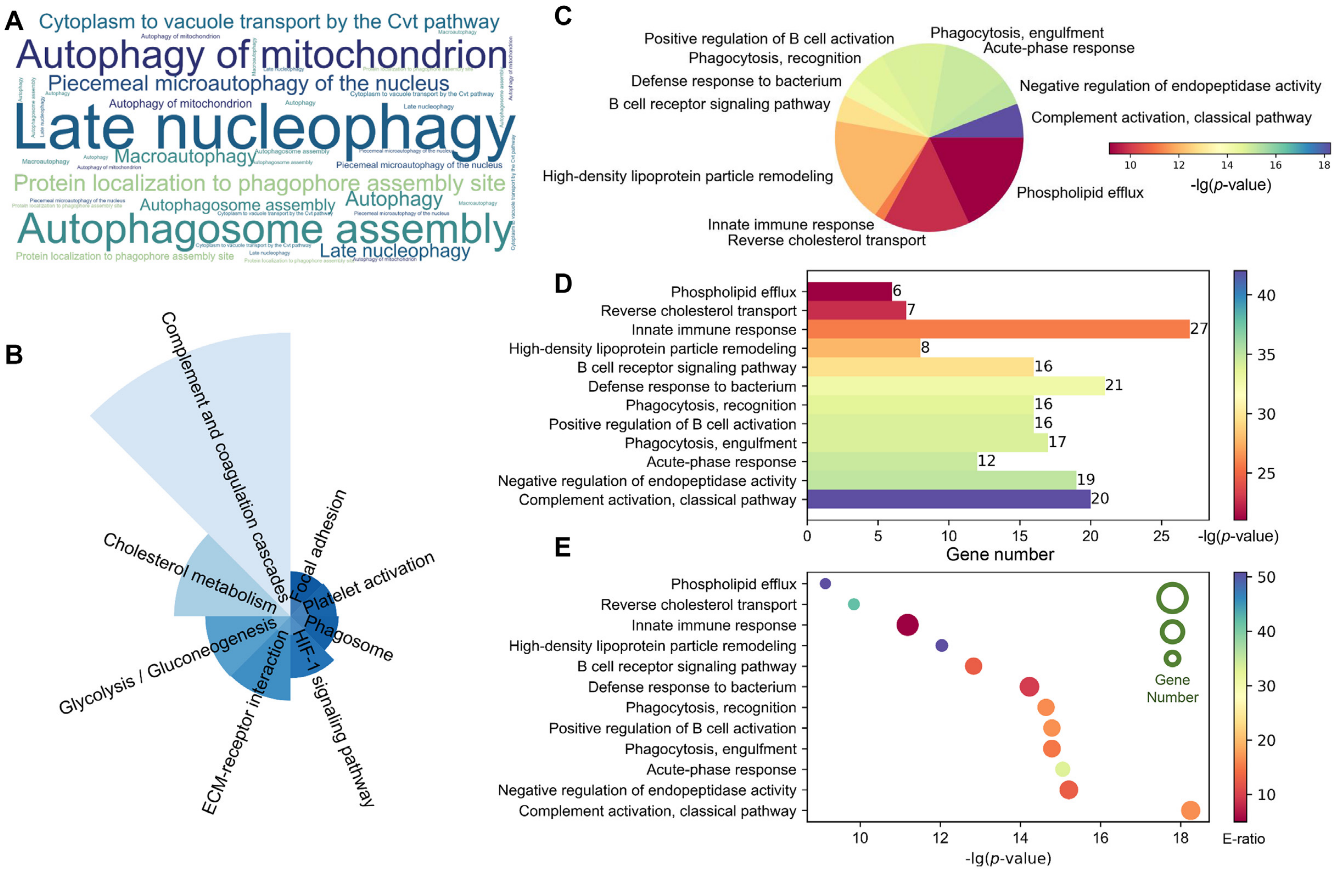


Figure 4. The concise illustrations of enrichment results by HemI 2.0. (**A**) GO enrichment analysis was performed for *ATG* genes. A considerable number of over-represented autophagy-related processes were intuitively illustrated by a word cloud. (B–E) Four additional idioms, including coxcomb chart (**B**), pie chart (**C**), bar graph (**D**) and bubble chart (**E**), were implemented for visualizing the enrichment results of 195 DEPs of COVID-19.

First, the numerical input data could be prepared either in Microsoft Excel spreadsheet (.xls or .xlsx), Tab Separated Value (TSV) or Comma Separated Value (CSV) format (<1 MB). For convenience, users' data will be stored on our server for 24 h. Then, users can select the numerical data area for visualizing a heatmap with the mouse-clicking manipulation (Figure 2A). By clicking on the row/column titles, the corresponding labels can be edited. A heatmap will be automatically generated after clicking on the "Submit" button. The generated heatmap can be easily manipulated in a customized manner by clicking on the "Heatmap Settings" button (Figure 2B). The corresponding data can be clustered for either or both of row and column, while clustering methods and distance metrics can be selected by clicking on the "Clustering Settings" button (Figure 2B). By right-clicking on the row or column labels, up to three layers of annotations can be added into the heatmap. Moreover, the genes or proteins for visualizing a heatmap can be selected by left-clicking on one or multiple row labels for further enrichment analysis (Figure 2C). After selecting organisms and functional annotations, enrichment analysis can be performed by clicking on the "Enrichment Analysis" button (Figure 2C). The enrichment results can be visualized in any of the five idioms (Figure 2D). The number of genes, E-ratio, and *P*-value can be shown, and the generated pictures can be easily re-sized and re-colored in a customized manner by clicking on the "Graphic Settings" button (Figure 2D).

To obtain publication-quality figures, the all generated pictures can be exported and different resolutions can be selected for outputting figures. We provided 4 picture formats, including JPG, PNG, PDF and TIFF, to satisfy the different requirements. The clustering results and enrichment results can also be exported. The whole procedure was carefully implemented into a video with ∼2 minutes on our website (https://hemi.biocuckoo.org/documentation/).

DATA ANALYSIS AND INTERPRETATION

To further demonstrate the usefulness of HemI 2.0, the heatmap of WT and 35 *atg* KO strains (22) was re-illustrated by HemI 2.0 (Figure 3A). The autophagic phenotypes of 36 yeast strains were unambiguously classified into three categories, namely, Class I, Class II and Class III. Consistent with previous results, we observed that autophagy activity was almost fully blocked, significantly prolonged and slightly delayed and in class I, class II and class III mutants, respectively (22).

Recently, we identified 195 differentially expressed proteins (DEPs) of COVID-19 by analyzing plasma proteins that underwent significant fold changes in fatal cases compared with those of healthy subjects (23). We re-illustrated the heatmap of 195 DEPs, and the degree of differential expression of DEPs was obviously reduced in the mild group compared with severe or fatal group (Figure 3B), indicating that the alterations of plasma proteins became more extensive in more severe or deteriorated conditions (23).

Furthermore, we performed GO enrichment analyses for these *ATG* genes and DEPs of COVID-19, respectively. For *ATG* genes, we observed that a considerable number of over-represented autophagy-related processes were intuitively illustrated by a word cloud (22) (Figure 4A). DEPs of COVID-19 were highly enriched in processes or pathway involved in inflammation, complement system and platelet activation (Figure 4B–E), which was visualized by four concise idioms and demonstrated that acute inflammation and excessive immune cell infiltration are associated with the severity of COVID-19 patients (23).

CONCLUSIONS

Visualization of big biological data is an urgent demand in life and biomedical sciences (1,2). As a concise and intuitive approach, illustration of biomolecular expression data in a heatmap has emerged to be highly attractive for data analysis and interpretation, in a simple but comprehensive manner (3–10). Previously, we developed HemI 1.0, and its targeting users were non-expert biologists in computational programming or with difficulty in using complicated and professional packages (10). As a stand-alone software package specifically designed for heatmap illustration, HemI 1.0 has served to the community for 8 years. Besides the local packages supporting Windows, Linux/Unix and Mac, here we further developed the online service of HemI 2.0, which was significantly improved with many features to meet increasing demands from users.

There are several limitations in HemI 2.0. First, although HemI 2.0 is an interactive heatmap viewer, the heatmap cannot be manipulated in a real-time manner, which has been implemented in Heatmapper (10). Besides three types of data-matrix heatmaps, Heatmapper also provided image-based heatmaps, such as choropleth and geospatial heatmaps (10). Thus, we will include more visualization types of heatmaps, and enable a real-time interaction in the next version of HemI. Third, only 15 types of gene sets together with the hypergeometric test were provided for enrichment analysis. More gene sets and more enrichment analysis methods will be included in the near future. Fourth, some tools such as Metascape provided network analysis of the interactome, beyond heatmap visualization and enrichment analysis (24). Thus, the analysis of networks and modules will be also added in our tool. In addition, more normalization methods, clustering strategies and heatmap-related analyses, such as differential expression analysis and molecular subtyping, will be included.

Taken together, we believe that HemI 2.0 can be a much more useful tool for visualization and analysis of biomolecular expression data. The online service of HemI 2.0 will be continuously maintained and refined upon users' comments and feedbacks.

DATA AVAILABILITY

HemI 2.0 is freely available for all users at: https://hemi.biocuckoo.org/ and the source at https://github.com/Ning-310/HemI.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We would like to thank numerous users of HemI for their valuable comments and communications with us during the past years.

FUNDING

Funding for open access charge: National Key R & D Program of China [2021YFF0702000, 2021ZD0201300]; Natural Science Foundation of China [31930021, 31970633, 81701567]; Hubei Innovation Group Project [2021CFA005]; China Postdoctoral Science Foundation [2018M642816, 2019T120648, 2020M682395, 2021M701337]; Fundamental Research Funds for the Central Universities [2017KFXKJC001, 2019kfyRCPY043].

Conflict of interest statement. None declared.

REFERENCES

1. Meyer,R.D. and Cook,D. (2000) Visualization of data. *Curr. Opin. Biotechnol.*, **11**, 89–96.
2. Su,Y., Shi,Q. and Wei,W. (2017) Single cell proteomics in biomedicine: high-dimensional data acquisition, visualization, and analysis. *Proteomics*, **17**, 1600267.
3. Saldanha,A.J. (2004) Java Treeview–extensible visualization of microarray data. *Bioinformatics*, **20**, 3246–3248.
4. Metsalu,T. and Vilo,J. (2015) ClustVis: a web tool for visualizing clustering of multivariate data using Principal Component Analysis and heatmap. *Nucleic Acids Res.*, **43**, W566–W570.
5. Tyanova,S., Temu,T., Sinitcyn,P., Carlson,A., Hein,M.Y., Geiger,T., Mann,M. and Cox,J. (2016) The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat. Methods*, **13**, 731–740.
6. Chen,C., Chen,H., Zhang,Y., Thomas,H.R., Frank,M.H., He,Y. and Xia,R. (2020) TBtools: an integrative toolkit developed for interactive analyses of big biological data. *Mol. Plant*, **13**, 1194–1202.
7. Pang,Z., Chong,J., Zhou,G., de Lima Morais,D.A., Chang,L., Barrette,M., Gauthier,C., Jacques,P., Li,S. and Xia,J. (2021) MetaboAnalyst 5.0: narrowing the gap between raw spectra and functional insights. *Nucleic Acids Res.*, **49**, W388–W396.
8. Gu,Z., Eils,R. and Schlesner,M. (2016) Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics*, **32**, 2847–2849.
9. Deng,W., Wang,Y., Liu,Z., Cheng,H. and Xue,Y. (2014) HemI: a toolkit for illustrating heatmaps. *PLoS One*, **9**, e111988.
10. Babicki,S., Arndt,D., Marcu,A., Liang,Y., Grant,J.R., Maciejewski,A. and Wishart,D.S. (2016) Heatmapper: Web-enabled heat mapping for all. *Nucleic Acids Res.*, **44**, W147–W153.
11. Liu,X., Hu,P., Huang,M., Tang,Y., Li,Y., Li,L. and Hou,X. (2016) The NF-YC-RGL2 module integrates GA and ABA signalling to regulate seed germination in Arabidopsis. *Nat. Commun.*, **7**, 12768.
12. Chung,K.P., Hsu,C.L., Fan,L.C., Huang,Z., Bhatia,D., Chen,Y.J., Hisata,S., Cho,S.J., Nakahira,K., Imamura,M. *et al.* (2019) Mitofusins regulate lipid metabolism to mediate the development of lung fibrosis. *Nat. Commun.*, **10**, 3390.
13. Gene Ontology Consortium. (2021) The Gene Ontology resource: enriching a GOld mine. *Nucleic Acids Res.*, **49**, D325–D334.

14. Ashburner,M., Ball,C.A., Blake,J.A., Botstein,D., Butler,H., Cherry,J.M., Davis,A.P., Dolinski,K., Dwight,S.S., Eppig,J.T. *et al.* (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.*, **25**, 25–29.
15. Kanehisa,M. and Goto,S. (2000) KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.*, **28**, 27–30.
16. Schriml,L.M., Mitraka,E., Munro,J., Tauber,B., Schor,M., Nickle,L., Felix,V., Jeng,L., Bearer,C., Lichenstein,R. *et al.* (2019) Human Disease Ontology 2018 update: classification, content and workflow expansion. *Nucleic Acids Res.*, **47**, D955–D962.
17. 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.* (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. U.S.A.*, **102**, 15545–15550.
18. Liberzon,A., Subramanian,A., Pinchback,R., Thorvaldsdóttir,H., Tamayo,P. and Mesirov,J.P. (2011) Molecular signatures database (MSigDB) 3.0. *Bioinformatics*, **27**, 1739–1740.
19. Liberzon,A., Birger,C., Thorvaldsdottir,H., Ghandi,M., Mesirov,J.P. and Tamayo,P. (2015) The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst.*, **1**, 417–425.
20. Sherman,B.T., Hao,M., Qiu,J., Jiao,X., Baseler,M.W., Lane,H.C., Imamichi,T. and Chang,W. (2022) DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). *Nucleic Acids Res.*, gkac194.
21. Kuleshov,M.V., Jones,M.R., Rouillard,A.D., Fernandez,N.F., Duan,Q., Wang,Z., Koplev,S., Jenkins,S.L., Jagodnik,K.M., Lachmann,A. *et al.* (2016) Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.*, **44**, W90–W97.
22. Zhang,Y., Xie,Y., Liu,W., Deng,W., Peng,D., Wang,C., Xu,H., Ruan,C., Deng,Y., Guo,Y. *et al.* (2020) DeepPhagy: a deep learning framework for quantitatively measuring autophagy activity in Saccharomyces cerevisiae. *Autophagy*, **16**, 626–640.
23. Shu,T., Ning,W., Wu,D., Xu,J., Han,Q., Huang,M., Zou,X., Yang,Q., Yuan,Y., Bie,Y. *et al.* (2020) Plasma Proteomics Identify Biomarkers and Pathogenesis of COVID-19. *Immunity*, **53**, 1108–1122.
24. Zhou,Y., Zhou,B., Pache,L., Chang,M., Khodabakhshi,A.H., Tanaseichuk,O., Benner,C. and Chanda,S.K. (2019) Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun*, **10**, 1523.