Biochemistry and Biophysics Reports 46 (2026) 102599

Contents lists available at ScienceDirect

Biochemistry and Biophysics Reports

journal homepage: www.elsevier.com/locate/bbrep

Appropriate threshold setting and multiple methods combination may improve reproducibility of gene ontology enrichment analysis

Qi Guo [a], Tian Fu [a], Na Wang [a], Qingyin Yu [b], Yunlu Zhao [b,*]

[a] Department of Physiology, Hebei University of Chinese Medicine, Shijiazhuang, Hebei Province, 050000, China
[b] Research institution of Integrated Traditional Chinese and Western Medicine, Hebei University of Chinese Medicine, Shijiazhuang, Hebei Province, 050000, China

ARTICLE INFO

Keywords:
Transcriptome analysis
Gene ontology
Enrichment analysis
Statistical methods
Single-cell RNA sequencing
Spatial transcriptomic analysis

ABSTRACT

Transcriptome analyses are widely used for biological research. Gene ontology (GO) enrichment analysis is often used for effectively analyzing large data matrices. However, various software and methods can be used for performing GO enrichment analyses, which might lead to different conclusions. To date, there is no agreement whatsoever in the scientific research community about standards and processes for analysis; moreover, the description about such analyses in previous research is often brief, causing difficulties in both research reproducibility and manuscript review. Herein, we introduce the advantages of different tools and methods, while the limitations or problems are stated. We found that the essential key to improving the reproducibility and accuracy of GO enrichment analyses was to set appropriate thresholds in data processing and might combine different GO methods to avoid the limitations of each one. Finally, research associations might need to consider draft-standardized generic transcriptomic analysis standards to promote advances in biological research.

1. Introduction

Transcriptome analyses, including single-cell RNA sequencing (scRNA-seq), spatial transcriptomic analysis, bulk-RNA sequencing (bulkRNA-seq), and RNA micro-array are widely used in current biological research [1,2]. These powerful tools allow researchers for quickly and accurately identifying the differences between two groups, cell clusters, or different spatial locations. In addition, compared with the older biology research methods detecting transcriptome changes, such as real-time quantitative Polymerase Chain Reaction (qPCR), these tools provide broad and overall results of the transcriptome changes, indicating that the most valuable research directions from thousands of genes are changing [3,4], providing potential directions worth noting and elucidating [5,6].

Currently, methods for analyzing differential gene expression (DEG) from transcriptome analyses have made significant progress [7,8]; many studies reported methods for data analysis [8,9], data standardization [10,11], and group comparison, making these tools available to every biological and medical researcher [12,13]. Regarding pairwise comparative analyses, gene ontology (GO) enrichment (or pathway enrichment) analysis is often used for identifying significant changes in biological pathways [14]. However, various software and methods can be used for performing GO enrichment analyses, which might lead to different conclusions. Moreover, there is no agreement whatsoever in the scientific research community about standards and processes for analysis. Furthermore, the description about such analyses in articles is often brief, causing difficulties in both research reproducibility and manuscript review. Therefore, standards for such analysis should be established by the scientific community to improve the accuracy and reproducibility of scientific research.

Herein, we introduce the advantages of different tools and methods and a few limitations or problems found in actual processes [15–19]. The findings in this study might provide some insights and recommendations for the GO analysis process to beginners in GO analysis and tips for senior researchers to review and inspect articles.

2. GO enrichment analyses based on threshold setting

Herein, we summarize the tools of enrichment analyses into two types: one relies on setting a statistical threshold to filter out the gene set, tools such as database for annotation visualization and integrated discovery (DAVID), Metascape, and Genome enrichment (GO) analysis (Fig. 1a). The other type takes gene function and expression differences into the calculation process and summarizes the results with directionality of functional change, indicating directionality of pathways, such as

* Corresponding author.
E-mail address: zhaoyunlu@hebcm.edu.cn (Y. Zhao).

https://doi.org/10.1016/j.bbrep.2026.102599
Received 19 February 2026; Received in revised form 3 April 2026; Accepted 16 April 2026

Abbreviation

scRNA-seq	Single-cell RNA sequencing
bulkRNA-seq	Bulk-RNA sequencing
Realtime-qPCR	Quantitative Polymerase Chain Reaction
GO enrichment analysis	Gene ontology enrichment analysis
DAVID	Database for annotation visualization and integrated discovery
GSEA	Gene set enrichment analysis
IPA	Ingenuity Pathway Analysis
KEGG	(Kyoto Encyclopedia of Genes and Genomes)
FC	Fold change
FDR	False discovery rate

gene set enrichment analysis (GSEA) and Ingenuity Pathway Analysis (IPA) (Fig. 1b).

Methods or tools using gene sets filtered out by a statistical threshold on DEG are commonly used for the analyses of bulk-RNAseq, single-cell RNA sequencing, spatial transcriptomic analysis, and any other transcriptional analyses to summarize the overall changes between the groups [20]. With a large number of pathways or signaling databases, this type of analysis covers almost all known biological activity pathways; analyses using these types of tools and methods are simple and fast to perform.

The DAVID is a popular bioinformatics resource system that includes a web server and service for functional enrichment analysis, published in ten papers since 2003 [21,22]. DAVID's functional annotation tools identify enriched biological themes and GO terms and might create Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway [23–25] maps to visualize gene expression changes. Its main advantage is that it can cluster and classify result terms by level, providing clear relationship information between genes, pathways, and biological processes. However, although the DAVID provides a gene ID conversion function, analyses using gene symbols are difficult to set up and implement.

GO analysis, officially published in 2000 [26], provides a comprehensive and up-to-date computational model for analyzing the significance and relationships of genes, such as proteins and non-coding RNAs [27–29]. Using GO analysis, these aspects can be divided into three categories: Molecular function, biological process, and cellular components. The results are presented in a table that can be obtained immediately from GO analysis. However, visualization of enrichment results must be performed manually after analysis using other software or tools.

Metascape, launched in 2015, provides an integrated and user-friendly web tool designed to facilitate multi-platform OMICs data analysis [30]. The setup of enrichment analysis using Metascape is easy and clear; it can automatically identify various formats of gene ID. In addition, it generates a brief analysis report and a few bar charts sort the pathways or biological processes by *p* values and visualizes enrichment networks. However, reports and charts automatically generated by Metascape combine terms or pathways from different databases, such as GO, Wiki pathway, KEGG, and Reactome Gene Sets, which might result in duplicate pathways from different databases. Manual reorganization of results is necessary before publication.

However, these analytical methods are easy to perform; the database is abundant of almost all known biological processes and pathways. There are two main limitations to these analyses: First, these methods rely on the gene set filtered from the data, which is highly affected by statistical calculation methods and threshold settings (Fig. 2a). Second, the results of these methods only provide information on pathways related to the DEGs but cannot indicate direction of the difference.

3. Appropriate threshold setting may improve the reproducibility and accuracy

To overcome the limitations of enrichment analyses based on threshold settings, an appropriate statistical method and clear description about the method used for data processing are essential for reproducibility. For instance, data processing should consider a quality control procedure and use the false discovery rate (FDR) to correct the *p* value.

FDR, introduced by Benjamini and Hochberg in 1995, describes the rate at which null hypothesis testing produces errors. It has since become an essential tool for interpreting large datasets to reduce type I error [31,32]. Considering transcriptome analysis, the Benjamini and Hochberg FDR correction methods are often used (Fig. 2b). However, these methods are affected by quality control procedures. Different conditions of the quality control procedure will change the total tests as the "m" in the formal of the B–H's method. Ultimately, this caused changes in the filtered gene set for GO enrichment analysis.

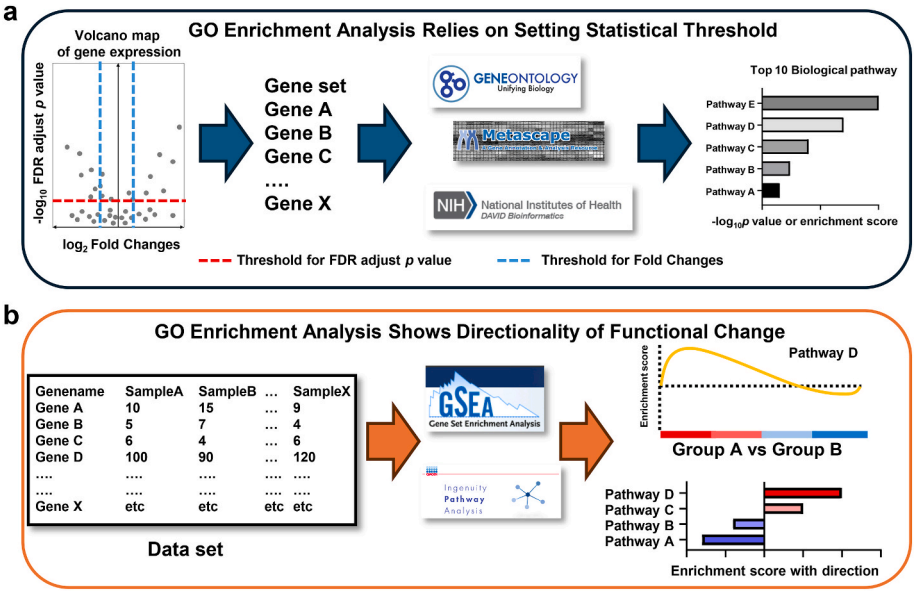

Fig. 1. Different process of Gene Onotolgy (GO) Enrichment analyses (a) The process of GO enrichment analyses based on threshold setting. (b) The process of GO enrichment analyses shows directionality of functional change.

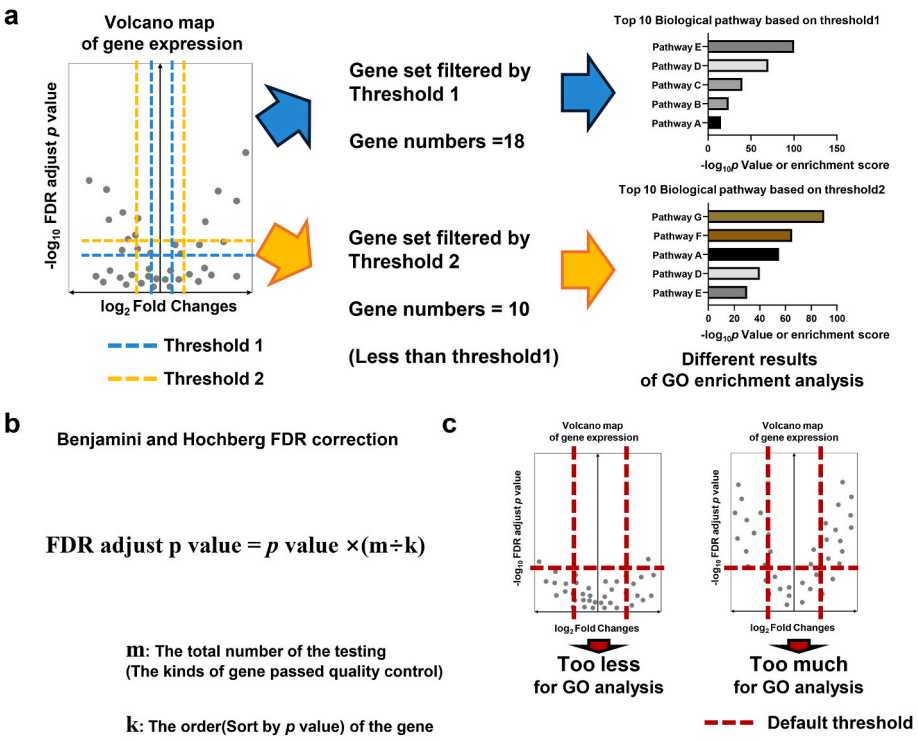

Fig. 2. Threshold Setting will highly affect the results of Gene Onotolgy (GO) Enrichment analyses (a) Different threshold settings will filter different kinds of gene used for GO enrichment analysis, leading to different results. (b) Benjamini and Hochberg FDR correction formula. (c) Unappropriated threshold setting or using default threshold without adjustments may not be suitable for filtering gene set used for GO enrichment analysis.

Usually, Fold change (FC) $\geq$2; p value < 0.05 with 5% FDR are often used and recommended as the default threshold for bulkRNA-seq analysis. However, because the detection method of microarrays is based on fluorescence intensity, its sensitivity is inferior to that of bulk RNA-seq. The 5% FDR correction might be too powerful and cause the gene list to be filtered from the data, thereby less encouraging to proceed with GO enrichment analysis (Fig. 2c). Moreover, such condition (FC $\geq$ 2; p value < 0.05 with 5% FDR) will filter far more genes than needed from the scRNAseq or spatial transcriptomic analysis for GO enrichment analysis (Fig. 2c). Using these analyses, a flexible adjustment of p values is necessary; moreover, the FC can be adjusted as an alternative. However, adjustment of the threshold setting must be clearly described to ensure reproducibility.

4. Enrichment analyses indicating directionality on pathways

In addition to the methods introduced above, GSEA [33] and IPA [34], which could indicate the directionality of pathways, are often used for omics dataset analysis. The results of these methods can provide clear information to readers about how and in which direction the difference exists.

GSEA is a computational method developed in 2003. It provides an a priori defined set of genes S and determines whether the members of S would be randomly distributed throughout L or are primarily found at the top or bottom to determine trends in distribution. Moreover, GSEA might use a dataset from KEGG or GO to display the overall characteristics of the two datasets.

The IPA software was produced by QIAGEN. It provides a platform for the analysis of omics experiments such as RNA-seq, small RNA-seq, metabolomics, proteomics, microarrays including miRNAs, single-nucleotide polymorphisms, and small-scale experiments. Unlike GO analysis based on threshold settings, IPA is mainly based on raw data and can distinguish trends in biological pathway distribution. In addition to biological pathway analysis, it can analyze molecular relationships and automatically visualize results.

However, regardless of whether GSEA or IPA is used, the omics dataset must be reformatted into a specific format for analysis. Additionally, the results of GSEA only showed changes between the two groups and might have overexpanded the difference by a single sample (Fig. 3a). Therefore, a heat map showing related gene expression changes is recommended for proving that the changes are holistic between the two groups.

5. Multiple method combination can be used for cross-validating the findings

Based on the above discussion, each method has its own advantages and disadvantages. To facilitate peer review and knowledge exchange, we used a combination of multiple GO analysis methods for cross-validating our findings. In particular, researchers might use GSEA to analyze specific pathways indicated by the GO enrichment method based on threshold settings to verify each gene change and its contribution to the pathway. Similarly, the results of GSEA might be highly affected by individual samples. To overcome these limitations, genes with significant individual sample differences can be filtered out after correct data processing. In addition, GO enrichment analysis based on threshold settings can be used for verifying the GSEA results (Fig. 3b). Additionally, the results of GO enrichment analyses based on closed-source datasets, such as IPA, cannot be easily reproduced. Therefore, other GO enrichment analysis methods should be used for proving the reproducibility of such analyses.

Lastly, researchers or research associations might need to consider draft-standardized, generic transcriptomic analysis standards to improve reproducibility and accuracy, facilitate academic exchanges, and promote advances in biological research. Currently, researchers might use appropriate threshold settings and multiple method combinations to describe data processing as much as possible for improving the quality and reproducibility of their own studies.

In conclusion, a higher standard of threshold setting and multiple methods combination might largely improve reproducibility of scientific

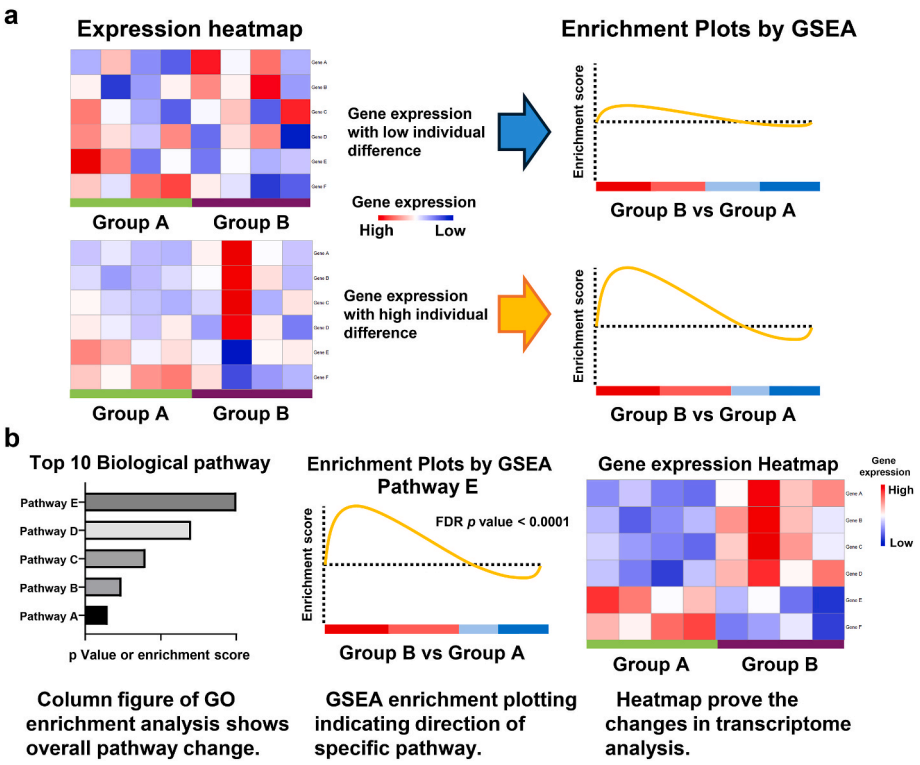

Fig. 3. Multiple method combination can be used to cross-validate the findings (a) The result of gene set enrichment analysis may highly overexpand the difference by a single sample. (b) Results from different gene ontology enrichment analyses can be combined to validate the findings and improve the reproducibility.

research, such standard urgently needs further establishment.

Consent to publish declaration

All listed authors agree to publish the article.

Consent to participate declaration

Not applicable.

Ethics declaration

Not applicable.

CRediT authorship contribution statement

Qi Guo: Investigation, Methodology, Software, Visualization, Writing – original draft. **Tian Fu:** Validation. **Na Wang:** Validation. **Qingyin Yu:** Validation. **Yunlu Zhao:** Conceptualization, Methodology, Software, Supervision, Writing – review & editing.

Declaration of competing interest

In this work, there is no potential conflict of interest to declare.

Acknowledgments

The figures in this article are schematic diagrams and do not include the actual experimental data.

Data availability

No data was used for the research described in the article.

References

[1] D. Jovic, X. Liang, H. Zeng, L. Lin, F. Xu, Y. Luo, Single-cell RNA sequencing technologies and applications: a brief overview, Clin. Transl. Med. 12 (2022) e694, https://doi.org/10.1002/ctm2.694.
[2] E. Zormpas, R. Queen, A. Comber, S.J. Cockell, Mapping the transcriptome: realizing the full potential of spatial data analysis, Cell 186 (2023) 5677–5689, https://doi.org/10.1016/j.cell.2023.11.003.
[3] N. Ahmad, L. Hou, J. Ma, X. Zhou, H. Xia, M. Wang, S. Leal-Bertioli, S. Zhao, R. Tian, J. Pan, C. Li, A. Li, D. Bertioli, X. Wang, C. Zhao, Bulk RNA-Seq analysis reveals differentially expressed genes associated with lateral branch angle in peanut, Genes (Basel) 13 (2022) 841, https://doi.org/10.3390/genes13050841.
[4] A. Janjic, L.E. Wange, J.W. Bagnoli, J. Geuder, P. Nguyen, D. Richter, B. Vieth, B. Vick, I. Jeremias, C. Ziegenhain, I. Hellmann, W. Enard, Prime-seq, efficient and powerful bulk RNA sequencing, Genome Biol. 23 (2022) 88, https://doi.org/10.1186/s13059-022-02660-8.
[5] E. Lecluze, B. Jégou, A.D. Rolland, F. Chalmel, New transcriptomic tools to understand testis development and functions, Mol. Cell. Endocrinol. 468 (2018) 47–59, https://doi.org/10.1016/j.mce.2018.02.019.
[6] N. Wang, Y.X. Huo, Using genome and transcriptome analysis to elucidate biosynthetic pathways, Curr. Opin. Biotechnol. 75 (2022) 102708, https://doi.org/10.1016/j.copbio.2022.102708.
[7] S. Slovin, A. Carissimo, F. Panariello, A. Grimaldi, V. Bouché, G. Gambardella, D. Cacchiarelli, Single-cell RNA sequencing analysis: a step-by-step overview, Methods Mol. Biol. 2284 (2021) 343–365, https://doi.org/10.1007/978-1-0716-1307-8_19.
[8] R. Satija, J.A. Farrell, D. Gennert, A.F. Schier, A. Regev, Spatial reconstruction of single-cell gene expression data, Nat. Biotechnol. 33 (2015) 495–502, https://doi.org/10.1038/nbt.3192.
[9] Y. Hao, T. Stuart, M.H. Kowalski, S. Choudhary, P. Hoffman, A. Hartman, A. Srivastava, G. Molla, S. Madad, C. Fernandez-Granda, R. Satija, Dictionary learning for integrative, multimodal and scalable single-cell analysis, Nat. Biotechnol. 42 (2024) 293–304, https://doi.org/10.1038/s41587-023-01767-y.
[10] M.I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2, Genome Biol. 15 (2014) 550, https://doi.org/10.1186/s13059-014-0550-8.
[11] C. Hafemeister, R. Satija, Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression, Genome Biol. 20 (2019) 296, https://doi.org/10.1186/s13059-019-1874-1.
[12] A. Butler, P. Hoffman, P. Smibert, E. Papalexi, R. Satija, Integrating single-cell transcriptomic data across different conditions, technologies, and species, Nat. Biotechnol. 36 (2018) 411–420, https://doi.org/10.1038/nbt.4096.
[13] S. Mangiola, M.A. Doyle, A.T. Papenfuss, Interfacing seurat with the R tidy universe, Bioinformatics 37 (2021) 4100–4107, https://doi.org/10.1093/bioinformatics/btab404.

[14] J. Reimand, R. Isserlin, V. Voisin, M. Kucera, C. Tannus-Lopes, A. Rostamianfar, L. Wadi, M. Meyer, J. Wong, C. Xu, D. Merico, G.D. Bader, Pathway enrichment analysis and visualization of omics data using g:profiler, GSEA, cytoscape and EnrichmentMap, Nat. Protoc. 14 (2019) 482–517, https://doi.org/10.1038/s41596-018-0103-9.

[15] T. Kasahara, M. Tanaka, Y. Zhao, A. Kamiyoshi, T. Sakurai, Y. Ichikawa-Shindo, H. Kawate, Y. Matsuda, Y. Zhang, Q. Guo, P. Li, K. Hoshiyama, J. Li, N. Onishi, M. Hayashi, H. Sanjo, K. Ishida, S. Tanaka, M. Kawamata, T. Shindo, Receptor activity-modifying proteins of adrenomedullin (RAMP2/3): roles in the pathogenesis of ARDS, Peptides 171 (2024) 171118, https://doi.org/10.1016/j.peptides.2023.171118.

[16] Y. Liu, Y. Zhao, Q. Guo, P. Wang, P. Li, Q. Du, H. Xu, Q. Yu, X. Zhao, W. Zhang, S. An, S. Wu, Sophoricoside reduces inflammation in type II collagen-induced arthritis by downregulating NLRP3 signaling, Biochem. Biophys. Rep. 40 (2024) 101867, https://doi.org/10.1016/j.bbrep.2024.101867.

[17] M. Tanaka, S. Kakihara, K. Hirabayashi, A. Imai, Y. Toriyama, Y. Iesato, T. Sakurai, A. Kamiyoshi, Y. Ichikawa-Shindo, H. Kawate, M. Tanaka, N. Cui, Y. Wei, Y. Zhao, K. Aruga, A. Yamauchi, T. Murata, T. Shindo, Adrenomedullin-receptor activity-modifying protein 2 system ameliorates subretinal fibrosis by suppressing epithelial-mesenchymal transition in age-related macular degeneration, Am. J. Pathol. 191 (2021) 652–668, https://doi.org/10.1016/j.ajpath.2020.12.012.

[18] Q. Guo, E. Kouyama-Suzuki, Y. Shirai, K. Tabuchi, Jun N-terminal kinase inhibitor suppresses CASK deficiency-induced cerebellar granular cell death in MICPCH syndrome model mice, Cells 14 (2025) 750, https://doi.org/10.3390/cells14100750.

[19] Y. Zhao, T. Sakurai, A. Kamiyoshi, M. Tanaka, Y. Ichikawa-Shindo, H. Kawate, Y. Matsuda, Y. Zhang, Q. Guo, P. Li, K. Hoshiyama, M. Hayashi, J. Li, T. Shindo, Adrenomedullin 2/intermedin exerts cardioprotective effects by regulating cardiomyocyte mitochondrial function, Hypertension 82 (2025) e6–e21, https://doi.org/10.1161/HYPERTENSIONAHA.124.23666.

[20] C. Bettembourg, C. Diot, O. Dameron, Optimal threshold determination for interpreting semantic similarity and particularity: application to the comparison of gene sets and metabolic pathways using GO and ChEBI, PLoS One 10 (2015) e0133579, https://doi.org/10.1371/journal.pone.0133579.

[21] B.T. Sherman, M. Hao, J. Qiu, X. Jiao, M.W. Baseler, H.C. Lane, T. Imamichi, W. Chang, David: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update), Nucleic Acids Res. 50 (2022) W216–W221, https://doi.org/10.1093/nar/gkac194.

[22] W. da Huang, B.T. Sherman, R.A. Lempicki, Systematic and integrative analysis of large gene lists using David bioinformatics resources, Nat. Protoc. 4 (2009) 44–57, https://doi.org/10.1038/nprot.2008.211.

[23] M. Kanehisa, M. Furumichi, Y. Sato, M. Kawashima, M. Ishiguro-Watanabe, KEGG for taxonomy-based analysis of pathways and genomes, Nucleic Acids Res. 51 (2023) D587–D592, https://doi.org/10.1093/nar/gkac963.

[24] M. Kanehisa, S. Goto, KEGG: kyoto encyclopedia of genes and genomes, Nucleic Acids Res. 28 (2000) 27–30, https://doi.org/10.1093/nar/28.1.27.

[25] M. Kanehisa, M. Furumichi, M. Tanabe, Y. Sato, K. Morishima, KEGG: new perspectives on genomes, pathways, diseases and drugs, Nucleic Acids Res. 45 (2017) D353–D361, https://doi.org/10.1093/nar/gkw1092.

[26] 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 (2000) 25–29, https://doi.org/10.1038/75556.

[27] P.D. Thomas, D. Ebert, A. Muruganujan, T. Mushayahama, L.P. Albou, H. Mi, PANTHER: making genome-scale phylogenetics accessible to all, Protein Sci. 31 (2022) 8–22, https://doi.org/10.1002/pro.4218.

[28] P.D. Thomas, D.P. Hill, H. Mi, D. Osumi-Sutherland, K. Van Auken, S. Carbon, J. P. Balhoff, L.P. Albou, B. Good, P. Gaudet, S.E. Lewis, C.J. Mungall, Gene ontology causal activity modeling (GO-CAM) moves beyond GO annotations to structured descriptions of biological functions and systems, Nat. Genet. 51 (2019) 1429–1433, https://doi.org/10.1038/s41588-019-0500-1.

[29] S. Carbon, A. Ireland, C.J. Mungall, S. Shu, B. Marshall, S. Lewis, AmiGO hub, web presence working group, AmiGO: online access to ontology and annotation data, Bioinformatics 25 (2009) 288–289, https://doi.org/10.1093/bioinformatics/btn615.

[30] Y. Zhou, B. Zhou, L. Pache, M. Chang, A.H. Khodabakhshi, O. Tanaseichuk, C. Benner, S.K. Chanda, Metascape provides a biologist-oriented resource for the analysis of systems-level datasets, Nat. Commun. 10 (2019) 1523, https://doi.org/10.1038/s41467-019-09234-6.

[31] Y. Benjamini, Y. Hochberg, Controlling the false discovery rate: a practical and powerful approach to multiple testing, J. R. Stat. Soc., B (Methodol.) 57 (1995) 289–300, https://doi.org/10.1111/j.2517-6161.1995.tb02031.x.

[32] S. Zehetmayer, M. Posch, A. Graf, Impact of adaptive filtering on power and false discovery rate in RNA-seq experiments, BMC Bioinf. 23 (2022) 388, https://doi.org/10.1186/s12859-022-04928-z.

[33] A. Subramanian, P. Tamayo, V.K. Mootha, S. Mukherjee, B.L. Ebert, M.A. Gillette, A. Paulovich, S.L. Pomeroy, T.R. Golub, E.S. Lander, J.P. Mesirov, Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles, Proc. Natl. Acad. Sci. U. S. A 102 (2005) 15545–15550, https://doi.org/10.1073/pnas.0506580102.

[34] Z. Shao, K. Wang, S. Zhang, J. Yuan, X. Liao, C. Wu, Y. Zou, Y. Ha, Z. Shen, J. Guo, W. Jie, Ingenuity pathway analysis of differentially expressed genes involved in signaling pathways and molecular networks in RhoE gene-edited cardiomyocytes, Int. J. Mol. Med. 46 (2020) 1225–1238, https://doi.org/10.3892/ijmm.2020.4661.