

# SubCELL: the landscape of subcellular compartment-specific molecular interactions

Yintao Zhang<sup>1,2,†</sup>, Wanghao Jiang<sup>1,†</sup>, Teng Li<sup>1</sup>, Hangwei Xu<sup>1</sup>, Yimiao Zhu<sup>1</sup>, Kerui Fang<sup>1</sup>, Xinyu Ren<sup>1</sup>, Shanshan Wang<sup>3</sup>, Yuzong Chen<sup>4,5,\*</sup>, Ying Zhou<sup>1,\*</sup> and Feng Zhu<sup>1,2,\*</sup>

<sup>1</sup>College of Pharmaceutical Sciences, Department of Pharmacy, Second Affiliated Hospital, Zhejiang University School of Medicine, State Key Laboratory of Advanced Drug Delivery and Release Systems, Zhejiang University, Hangzhou 310058, China

<sup>2</sup>Innovation Institute for Artificial Intelligence in Medicine of Zhejiang University, Alibaba-Zhejiang University Joint Research Center of Future Digital Healthcare, Hangzhou 330110, China

<sup>3</sup>Qian Xuesen Collaborative Research Center of Astrochemistry and Space Life Sciences, Institute of Drug Discovery Technology, Ningbo University, Ningbo 315211, China

<sup>4</sup>State Key Laboratory of Chemical Oncogenomics, Key Laboratory of Chemical Biology, The Graduate School at Shenzhen, Tsinghua University, Shenzhen 518055, China

<sup>5</sup>Institute of Biomedical Health Technology and Engineering, Shenzhen Bay Laboratory, Shenzhen 518000, China

\*To whom correspondence should be addressed. Tel: +86 189 8946 6518; Email: zhufeng@zju.edu.cn

Correspondence may also be addressed to Prof. Ying Zhou. Email: zhou\_ying@zju.edu.cn

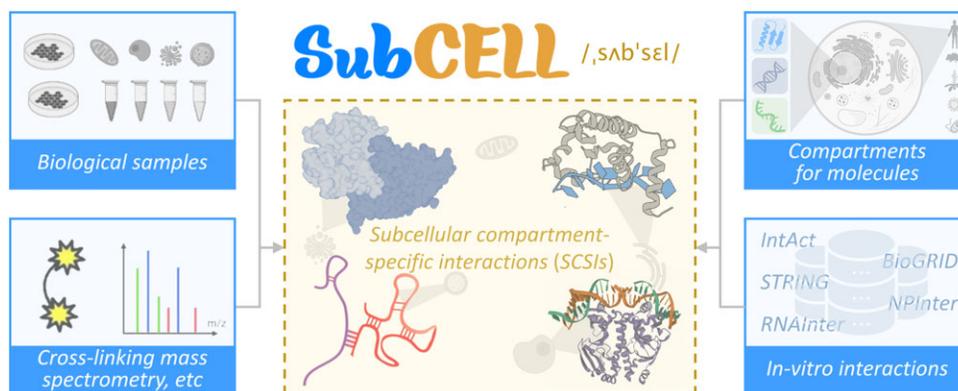
Correspondence may also be addressed to Prof. Yuzong Chen. Email: chenyzong@sz.tsinghua.edu.cn

<sup>†</sup>The first two authors should be regarded as Joint First Authors.

## Abstract

The subcellular compartment-specific molecular interactions (SCSIs) are the building blocks for most molecular functions, biological processes and disease pathogenesis. Extensive experiments have therefore been conducted to accumulate the valuable information of SCSIs, but none of the available databases has been constructed to describe those data. In this study, a novel knowledge base *SubCELL* is thus introduced to depict the landscape of SCSIs among DNAs/RNAs/proteins. This database is UNIQUE in (a) providing, for the first time, the experimentally-identified SCSIs, (b) systematically illustrating a large number of SCSIs inferred based on well-established method and (c) collecting experimentally-determined subcellular locations for the DNAs/RNAs/proteins of diverse species. Given the essential physiological/pathological role of SCSIs, the *SubCELL* is highly expected to have great implications for modern molecular biological study, which can be freely accessed with no login requirement at: <https://idrblab.org/subcell/>.

## Graphical abstract



## Introduction

Cellular functions are tightly regulated by the interactions among DNAs/RNAs/proteins, which are intimately tied with their interacting subcellular compartments (1–3). These in-

teractions serve as building blocks for most molecular functions, biological processes, and disease pathogenesis (4–6). Sophisticated experimental technologies [cross-linking mass spectrometry (7), proximity labeling (8) and so on] are pro-

Received: August 2, 2024. Revised: September 6, 2024. Editorial Decision: September 16, 2024. Accepted: September 20, 2024

© The Author(s) 2024. Published by Oxford University Press on behalf of Nucleic Acids Research.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License

(<https://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact [reprints@oup.com](mailto:reprints@oup.com) for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact [journals.permissions@oup.com](mailto:journals.permissions@oup.com).

posed to discover the subcellular compartment-specific interaction (SCSI). Moreover, an inferring method has also been made available to enable the discovery of SCSIs through integrating experimentally validated *in-vitro* interactions with the corresponding information of interacting molecule's subcellular compartment (9). Both types of SCSI are found critical for clarifying the molecular mechanism underlying disease-induced organelle remodeling (10–12), elucidating the complicated biological process regulated by DNAs/RNAs/proteins (13–16), refining existing interaction networks for deconvoluting the highly complex architecture of cells (17–19), and so on. Additionally, these data provide valuable foundation for the application of AI (20–22) into biomedical studies (23–25). Therefore, it is key to accumulate the data of subcellular compartment-specific molecular interactions to meet the needs of modern biological research.

So far, several databases related to molecular interaction have been developed. Some specialized in protein-protein interaction, such as STRING (26), DIP (27), and HIPPIE (28); some described RNA-centric interactions, like RNAInter (29), NPInter (30) and RISE (31); the remaining gave multiple types of interactions, including IntAct (32), BioGRID (33) and ConsensusPathDB (34). All these databases had accumulated broad interests from worldwide audiences, but none of them included the data of SCSI. Currently, only two databases are available for providing SCSI-related data, ComPPI (35) and IID (36). However, both of them focused only on providing the inferring interactions between proteins, which led to the absence of a huge number of interactions between protein–RNA, protein–DNA, RNA–RNA. Moreover, the data of experimentally-validated SCSIs were completely missing from both databases. Thus, there is an urgent need to have a knowledge base that provides the valuable information of the subcellular compartment-specific interactions (SCSIs, both experimental and inferred) among DNAs/RNAs/proteins.

Herein, a database, named *SubCELL*, was thus developed to provide the landscape of subcellular compartment-specific interactions among DNAs/RNAs/proteins. First, comprehensive review was conducted by searching keyword combination like ‘*subcellular + interactome*’ and ‘*organelle + interactome*’ in PubMed, and 11 559 experimentally-identified SCSIs were accumulated. Second, the subcellular compartment information of 68 081 DNAs/RNAs/proteins were collected from literatures, which covered 52 organelles (such as endoplasmic reticulum, mitochondrion, chloroplast, melanosome, glyoxysome and carboxysome) from 2 291 species. Third, 398 798 experimentally-validated *in-vitro* interactions were compiled from literatures and integrated from existing databases (such as BioGRID (33), and IntAct (32)). Finally, a well-established inferring method (9) was applied to filter the interactions that two molecules located in the same subcellular location, resulting in a refined set of 108 771 inferred SCSIs. As a result, a total of 120 330 SCSIs (including 11 559 experimentally-identified and 108 771 inferred ones) were systematically collected to *SubCELL* database.

All in all, our *SubCELL* (illustrated in Figure 1) is unique in (a) providing, for the first time, the experimentally-identified SCSIs, (b) systematically describing a large number of SCSIs inferred based on a well-established method and (c) comprehensively collecting the experimentally-determined subcellular compartments for 68 081 DNAs/RNAs/proteins of diverse species. All molecules collected into this study were cross-linked to other well-established databases such as

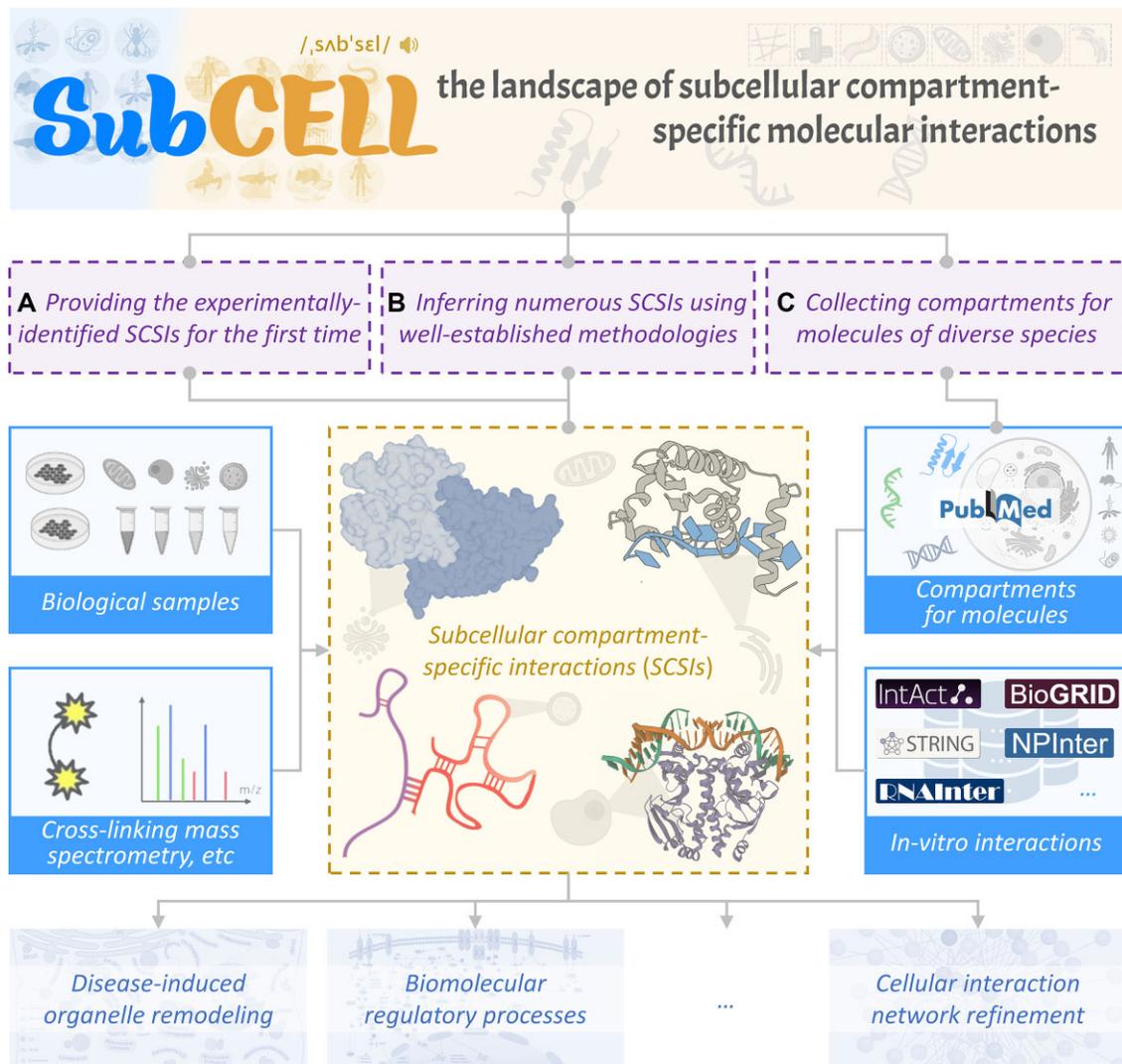
UniProt (37), miRbase (38), TTD (39), DrugMAP (19), PDB (40), AlphaFoldDB (41), Ensembl (42), HGNC (43) and NCBI Taxonomy (44). Since these valuable SCSIs collected in *SubCELL* are building blocks of molecular functions, biological processes and disease pathogenesises, they are anticipated to have great implication for the modern molecular biological research.

## Factual content and data retrieval

### Collection of subcellular compartment-specific interactions

In *SubCELL*, two types of subcellular compartment-specific interactions (SCSIs) were described: the experimentally-identified ones and the inferred ones. Those experimentally-identified SCSIs were discovered by sophisticated experimental techniques (such as proximity labeling and cross-linking mass spectrometry), while the inferred SCSIs were deduced by integrating the subcellular distributions of and the *in-vitro* interactions among the studied DNAs/RNAs/proteins. To collect the experimentally-identified SCSIs, comprehensive literature review was conducted in PubMed (45) using such keyword combinations as: ‘*subcellular compartment + interaction*’, ‘*subcellular location + interactome*’ and ‘*organelle + interactome*’. All retrieved publications were manually checked and a total of 11 559 experimentally-identified SCSIs were carefully recorded alongside the details of the detecting compartment and experimental technique. For the inferred SCSIs, the processes for data collection were as follows. First, the subcellular compartment information of proteins and RNAs were comprehensively collected from literatures and existing databases, such as *UniProt* (37), *HPA* (17), *MitoCarta* (46) and *RNAlocate* (47), which led to a total of 101 192 compartment annotations for 68 081 molecules. Then, 398 798 experimentally-validated *in-vitro* interactions were retrieved from both literature and the existing databases [such as BioGRID (33) and IntAct (32)]. Finally, only those interactions whose interacting molecules located in the same compartment were collected into our *SubCELL*, which resulted in 108 771 inferred SCSIs.

As a result, a total of 120 330 SCSIs and 101 192 compartment annotations for 68 081 molecules were systematically described in our *SubCELL*, which came from 2 291 species (*Homo sapiens*, *Mus musculus*, *Arabidopsis thaliana*, *Fission yeast*, etc.) and 52 organelles (mitochondrion, endoplasmic reticulum, chloroplast, carboxysome, etc.). Detailed distribution of *SubCELL* data was illustrated in Figure 2A. As demonstrated, the experimentally-identified and inferred SCSIs accounted for 10% and 90% of all those collected data, respectively. The top-5 compartments with the highest amount of SCSI data were *Nucleus*, *Cytoplasm*, *Cytosol*, *Mitochondrion* and *Cell Membrane*, and the top-5 species containing the largest number of SCSI information were *Homo sapiens*, *Baker's yeast*, *Mus musculus*, *Arabidopsis thaliana* and *Fission yeast*. Moreover, Figure 2B further described the distribution of compartment annotation data in *SubCELL*, and the annotations for protein, RNA and DNA accounted for 67.1%, 29.1% and 3.8%, respectively. The top-5 compartments containing the most annotations were *Nucleus*, *Cytoplasm*, *Cytosol*, *Secreted* and *Cell Membrane*, and the top-5 organisms having the most annotations were *Homo sapiens*, *Mus musculus*, *Arabidopsis thaliana*, *Baker's yeast* and *Fission yeast*.

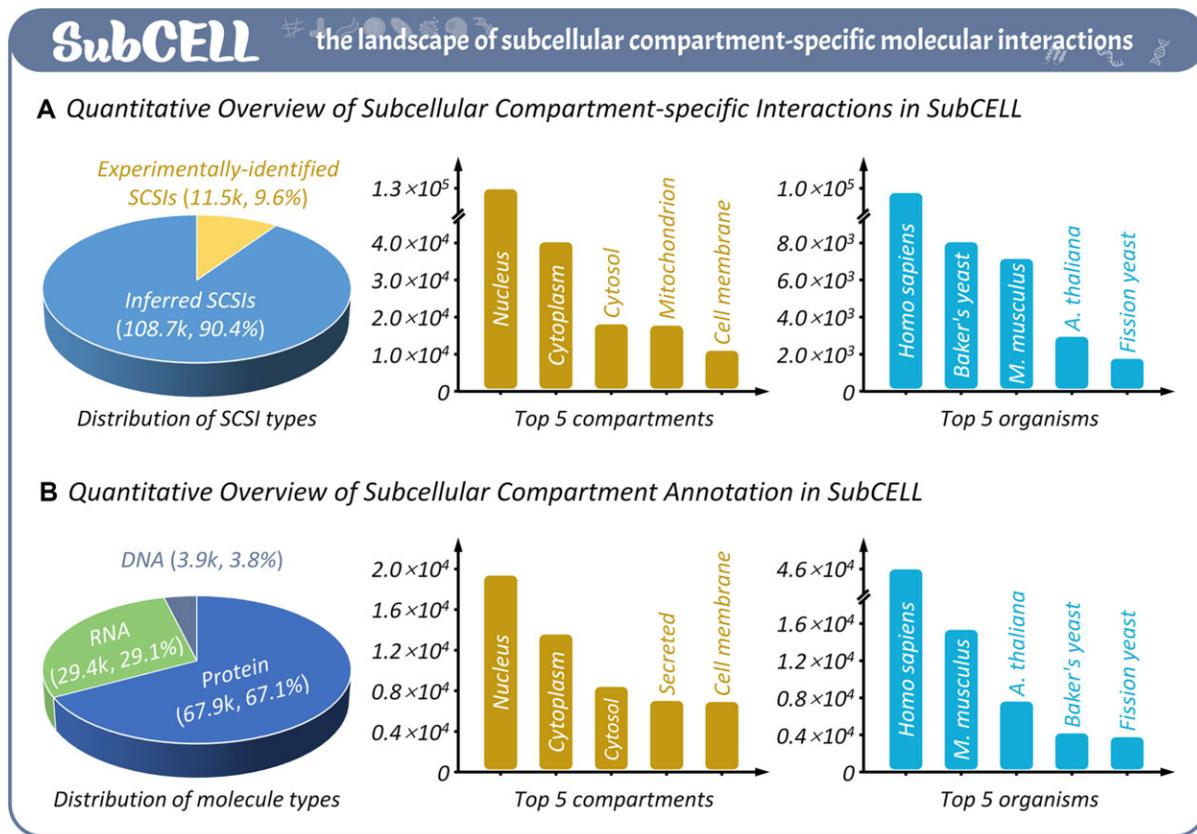


**Figure 1.** The uniqueness of the *SubCELL* and the potential applications of the collected data in *SubCELL*. The *SubCELL* database was constructed to provide the landscape of subcellular compartment-specific interactions (SCSIs) among DNAs/RNAs/proteins, which was unique in (a) providing, for the first time, the experimentally-identified SCSIs, (b) systematically offering a large number of SCSIs inferred based on well-established method and (c) comprehensively collecting experimentally-determined subcellular compartments for DNAs/RNAs/proteins of diverse species. Both experimentally-identified and inferred SCSIs are found critical for finding the molecular mechanisms underlying the disease-induced organelle remodeling, clarifying complicated biological process regulated by DNAs/RNAs/proteins, refining existing interaction network for deconvoluting the highly complex architecture of the cell, and so on.

### Deciphering subcellular compartment-specific interaction profiles of molecules

To execute its cellular functions, DNAs/RNAs/proteins interacted with various molecules in diverse subcellular compartments (48). Unraveling SCSIs landscape of DNAs/RNAs/proteins could provide great insights into how a molecule functioned in various cellular processes (49–51), how it was regulated in specific locations (52,53), and how disruption in these interactions might contribute to a disease/cellular dysfunction (54–57). Taking the tumor suppressor protein p53 as example, which bound to E3 ubiquitin-protein ligase MDM2 in cytosol, thereby maintaining p53 in its inactive state. However, in response to the cellular stress, p53 could dissociate from MDM2, translocated into the nucleus, and acted as a transcription factor to induce apoptosis (58,59).

To facilitate the understanding of the full picture and differences in molecule's interaction profile in different subcellular compartments, a web page for each molecule (as shown in Figure 3) was constructed to provide the compartment data and the SCSIs for each molecule. As described, the molecule page was organized to two sections. The first section (as shown in Figure 3A) presented the general information of the molecule together with its compartments. The general information demonstrated in the left side of Figure 3A included molecule name, synonyms, species, structure, sequence and functions. On the right side of Figure 3A, the corresponding compartments of this molecule were visualized based on *SwissBioPics* (60). The second section (as depicted in Figure 3B) focused on giving the experimentally-identified and the inferred SCSIs. For experimentally-identified ones, the related information was carefully showed, such as the experimental technique and corresponding publication. For the inferred



**Figure 2.** The distribution of the data collected in *SubCELL*. **(A)** A quantitative overview of the subcellular compartment-specific interactions (SCSIs) in *SubCELL*. The experimentally-identified and the inferred SCSIs accounted for 10% and 90% of all collected data, respectively. The five subcellular compartments that have the most interactions in *SubCELL* were nucleus, cytoplasm, cytosol, mitochondrion and cell membrane. The five organisms that have the most interactions in *SubCELL* were *Homo sapiens*, *Baker's yeast*, *Mus musculus*, *Arabidopsis thaliana* and *Fission yeast*. **(B)** A quantitative overview of the subcellular compartment annotation of molecules in *SubCELL*. The subcellular compartment annotations in *SubCELL* for proteins, RNAs and DNAs accounted for 67.1%, 29.1% and 3.8%, respectively. The 5 subcellular compartments that have the most annotations in *SubCELL* were nucleus, cytoplasm, cytosol, secreted and cell membrane. The five organisms that have the most annotations in *SubCELL* were *Homo sapiens*, *Mus musculus*, *Arabidopsis thaliana*, *Baker's yeast* and *Fission yeast*.

ones, the corresponding interactions were grouped by compartments. The user can click the Protein/RNA/DNA Info button to redirect to interacting molecule page for additional information. Moreover, a section entitled ‘*Pathological Context of This Interaction*’ was provided online to demonstrate the pathological contexts of each molecular interaction, and cross links were also provided to existing online databases (28,36).

### Exploring interactions and molecules in each subcellular compartment

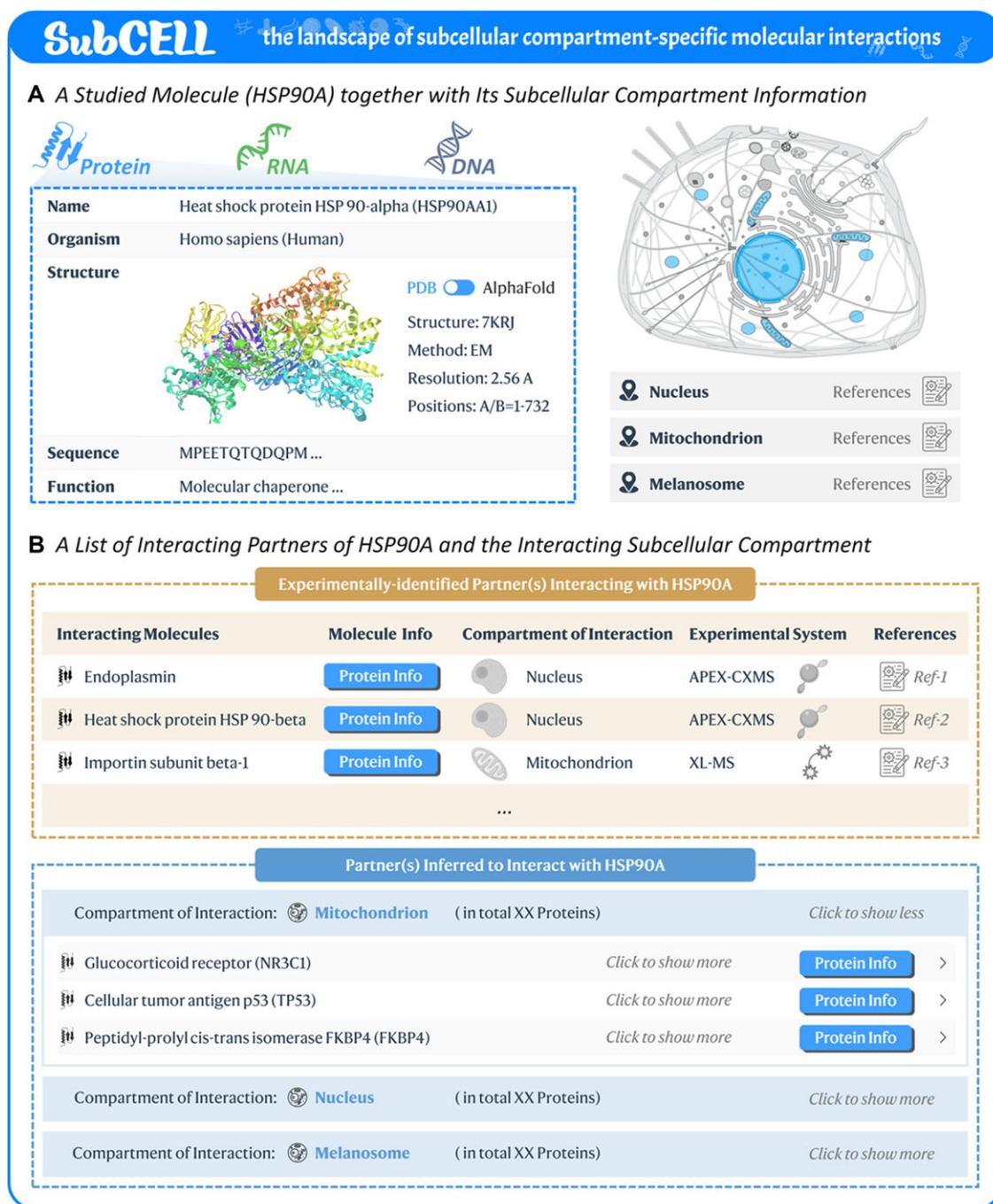
Exploring interactions in a specific subcellular compartment can help to understand the spatial organization of cellular components (61–63), uncover the molecular mechanisms underlying the diseases (10,64), aid in the design of targeted therapies (65–68) and so on. For example, the large-scale identification of interactome within mitochondria provided a valuable tool for advancing the understanding of mitochondrial function and dysfunction (69,70), and served as a powerful complement to new cryo-EM-derived structures of mitochondrial complexes (71).

In *SubCELL*, a dedicated page was designed for each subcellular compartment (as illustrated in Figure 4) to provide the information of proteins and their interactions with the compartment. As depicted in Figure 4A, the first section presented

the general information about the subcellular compartment, such as *synonyms*, *description* and *hierarchy*. The hierarchy of the subcellular compartment was constructed based on all compartment information collected in the *SubCELL*. Taking the *Nucleus* as an example, it can be subdivided into different substructure, such as nucleus membrane, nucleolus, nuclear speckle and nucleoplasm. Some of the substructures can be further divided into tertiary levels, like nucleus inner membrane and nucleus outer membrane. The second section (as described in Figure 4B) systematically provided the subcellular compartment-specific interactions of a given compartment. Users were empowered to filter interactions by selecting certain class(es) of interactions (such as protein–protein interaction and protein–RNA interaction) or keywords search. The third section (as shown in Figure 4C) listed all molecules that located in the subcellular compartment. For each molecule, the name, category, organism, and the detailed compartment information were explicitly provided and users can click the molecule name to jump to the molecule page for more detailed information.

### User-friendly search and download function

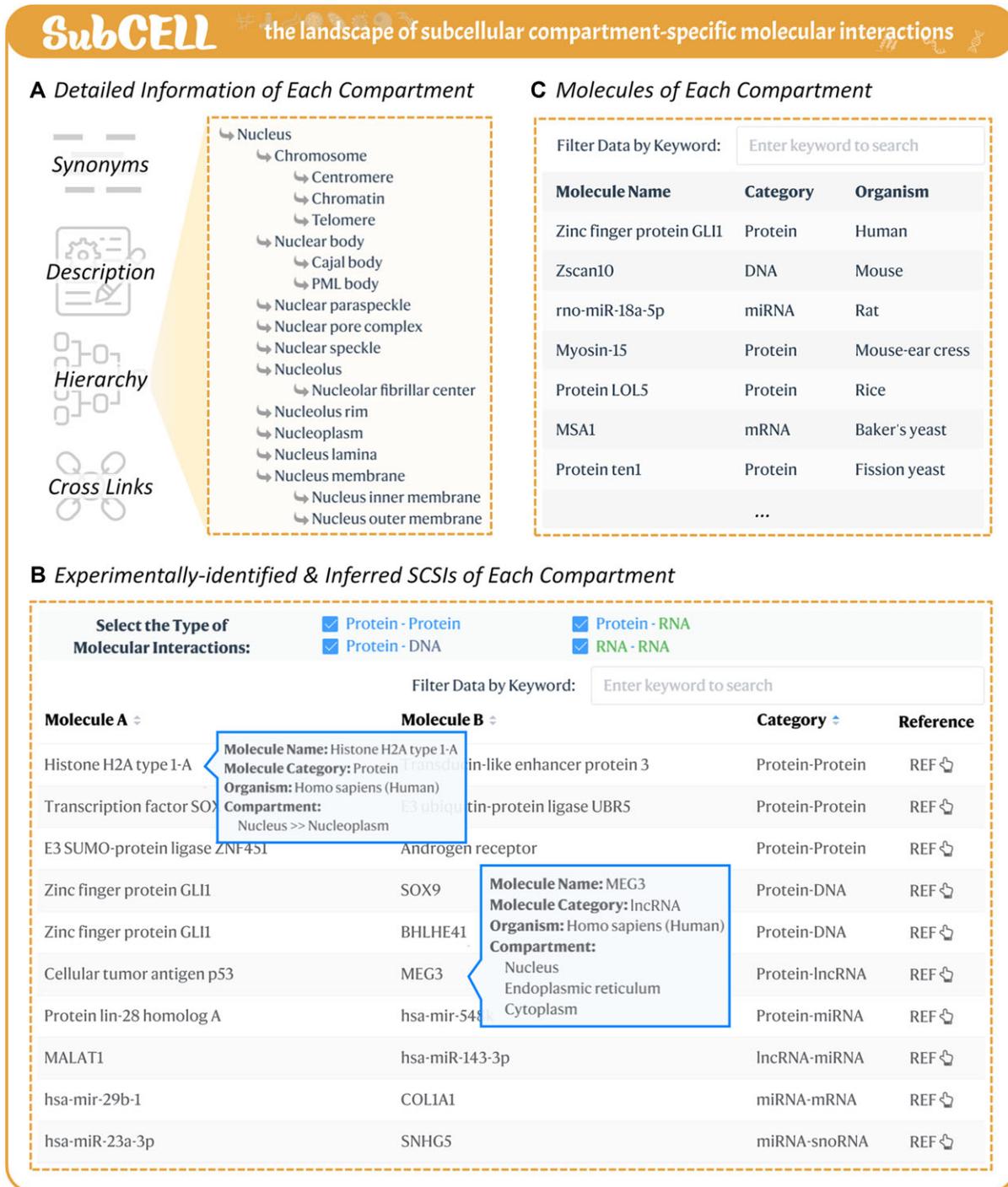
*SubCELL* provided diverse search options and download functions to enhance user experience and access to informa-



**Figure 3.** Diverse data provided in *SubCELL* for each molecule. **(A)** A studied molecule together with its subcellular compartment information in *SubCELL*. On the left side, the general information of each molecule is displayed, included name, synonyms, organism, structure, sequence, functions and so on. On the right side, all subcellular compartments of the molecule reported in publications were displayed using SwissBioPics. **(B)** A list of interacting partners of the molecule and the compartment of interaction. For experimentally-identified partner(s), the compartment of interaction, experimental method and corresponding publication were provided. For inferred Partner(s), the interactions were grouped by subcellular compartments. Users can click the button of *Molecule Info* (*Protein Info*, *RNA Info* and *DNA Info* for protein, RNA and DNA, respectively) to redirect to the interacting molecule page for detailed information.

tion. As shown in Figure 5A, users can explore molecules and subcellular compartment of interests by keywords (such as molecule/compartment name), species, sequence and so on. When the users search for the proteins or subcellular compartments using any of these strategies, the relevant results will be presented in card format (as demonstrated in Figure 5B). Users can seamlessly navigate to specific molecule

or compartment pages for detailed insights by clicking the *Protein/RNA/DNA Info* or *Compartment Info* button. In addition, as shown in Figure 5C, all data collected in *SubCELL* can be downloaded in the Download page, including the subcellular compartment-specific molecular interaction data, the subcellular compartment information, and the general information for molecules. This feature enables users to access



**Figure 4.** Various information offered in *SubCELL* for each subcellular compartment. **(A)** General information of the subcellular compartment. The general information of the subcellular compartment provided in *SubCELL* included synonyms, description, hierarchy and so on. The hierarchy of the subcellular compartment was constructed based on the all compartment information of molecules collected in *SubCELL* database. **(B)** List of molecular interactions in the subcellular compartment. All interactions in the subcellular compartment were listed in the tabular form. For each molecule in a specific interaction, the name, category (DNA/RNA/Protein), organism (such as *Homo sapiens*, *Mus musculus* and *Arabidopsis thaliana*), and the all reported compartment information were explicitly provided. Users can select certain class(es) of interactions (such as protein–protein interaction and protein–RNA interaction) or keywords search to filter the interactions and users can click the molecule name to jump to the molecule page for more detailed information. **(C)** List of molecules located in the subcellular compartment. For each molecule, the name, category (DNA/RNA/Protein), organism (such as *Homo sapiens*, *Mus musculus* and *Arabidopsis thaliana*), and the detailed compartment information of the molecule were explicitly provided. Users can click the molecule name to jump to the molecule page for more detailed information.



**Figure 5.** User-friendly search, browse and download function in *SubCELL*. **(A)** Multiple options for molecule and compartment search in *SubCELL*. Users can search for molecules and subcellular compartments of interests by keywords (such as molecule/compartment name), species, sequence and so on. **(B)** Search result page in *SubCELL*. When users search for protein or subcellular compartment using any of the strategies, the relevant results will be presented in the format of card. Users can click the button of *Protein/RNA/DNA Info* or *Compartment Info* to jump to a specific molecule or compartment page for detailed information. **(C)** Diverse data download in *SubCELL*. All data collected in *SubCELL* can be downloaded in the Download page, including the data of subcellular compartment-specific molecular interactions, the information of subcellular compartment, and the general information for molecules.

and utilize the compiled data for further analysis and research purposes.

#### Database construction and maintenance

*SubCELL* (<https://idrblab.org/subcell/>) was constructed using HTML, JavaScript, CSS and Python on the *Nginx* server. Its front end was constructed using multiple software packages, such as *Vue* and *Element UI*. The current database could be

readily accessed by diverse mainstream browsers, including *Chrome*, *FireFox*, *Safari*, *Edge*, etc. The back end was based on *Django*, a Python web framework. MySQL and Swiss-BioPics were implemented to facilitate the data storage, query and visualization. With the rapid development of experimental technique for detecting the subcellular location-specific interactions (SCSIs), an exponential increase of SCSi data was highly expected, which asked for a persistent update of this database in the coming years. Therefore, it was critical for us

to keep updating *SubCELL* by (a) conducting literature review semiannually to collect the newly discovered SCSIs and (b) incorporating the *large language model* (LLM) as an automated tool to accelerate the data collection process (surely, a group of well-trained pharmacologists and bioinformaticians in our research team will manually validate all data to ensure their quality).

## Conclusion

In this study, a database, named *SubCELL*, was developed to provide the landscape of subcellular compartment-specific interactions (SCSIs) among DNAs/RNAs/proteins. Although a number of molecular interaction-based databases have been constructed, their data scope was limited by the corresponding technology at their time. Particular, several advanced technologies became mature in recent years, which enabled the direct capture of interactions in particular subcellular location. The advancement of those technologies witnessed the exponential growth of the experimentally-validated SCSIs. In other words, at the time of existing databases, their SCSi data were extremely limited, and *SubCELL* covered the largest number of experimentally-validated SCSIs. Moreover, the relationships among interacting partner, interaction and subcellular location were extremely sophisticated, and it was therefore challenging to integrate the subcellular data for those existing databases. In this study, substantial efforts were made to accumulate the most comprehensive set of both subcellular information and interaction data. In other words, our *SubCELL* provided the most diverse molecular types, species origins, and organelles among the available databases. Due to the importance of collected *SubCELL* data, they were expected to have great implications for the modern molecular biological research.

## Data availability

*SubCELL* is publicly accessible without any login requirement at <https://idrblab.org/subcell/>.

## Acknowledgements

We would like to express our gratitude to the members of both the *Innovative Drug Research and Bioinformatics* (IDRB) and the *Bioinformatics and Drug Design* (BIDD) group for their valuable assistance and support throughout this research project. Their contributions have been instrumental in the successful completion of this study. Furthermore, we extend our appreciation to the broader research community working in the fields of molecular interactions and subcellular localization of molecules. Their research has laid a solid groundwork for our project.

## Funding

Funded by National Natural Science Foundation of China [82373790, 22220102001, 81872798, U1909208]; Natural Science Foundation of Zhejiang Province [LR21H300001]; Fundamental Research Funds for Central Universities [2018QNA7023]; National Key R&D Program of China [2022YFC3400501]; Double Top-Class University [181201\*194232101]; Key R&D Program of Zhejiang [2020C03010]; Westlake Laboratory (Westlake Lab of Life

Sciences and Biomedicine); Leading Talent of the ‘Ten Thousand Plan’ National High-Level Talents Special Support Plan of China; Alibaba-Zhejiang University Joint Research Center of Future Digital Healthcare; Alibaba Cloud; The Information Technology Center of Zhejiang University; National Key R&D Program of China, Synthetic Biology Research [2019YFA0905900]; The Startup Fund from Shenzhen Bay Laboratory [21310091]; Ningbo Top Talent Project [215-432094250]. Funding for open access charge: Natural Science Foundation of Zhejiang Province [LR21H300001].

## Conflict of interest statement

None declared.

## References

1. Cho,N.H., Cheveralls,K.C., Brunner,A.D., Kim,K., Michaelis,A.C., Raghavan,P., Kobayashi,H., Savy,L., Li,J.Y., Canaj,H., *et al.* (2022) OpenCell: endogenous tagging for the cartography of human cellular organization. *Science*, **375**, eabi6983.
2. Villanueva,E., Smith,T., Pizzinga,M., Elzek,M., Queiroz,R.M.L., Harvey,R.F., Breckels,L.M., Crook,O.M., Monti,M., Dezi,V., *et al.* (2024) System-wide analysis of RNA and protein subcellular localization dynamics. *Nat. Methods*, **21**, 60–71.
3. Yin,J., Li,F., Zhou,Y., Mou,M., Lu,Y., Chen,K., Xue,J., Luo,Y., Fu,J., He,X., *et al.* (2021) INTEDE: interactome of drug-metabolizing enzymes. *Nucleic Acids Res.*, **49**, D1233–D1243.
4. Qin,W., Cho,K.F., Cavanagh,P.E. and Ting,A.Y. (2021) Deciphering molecular interactions by proximity labeling. *Nat. Methods*, **18**, 133–143.
5. Cheng,F.X., Zhao,J.F., Wang,Y., Lu,W.Q., Liu,Z.H., Zhou,Y.D., Martin,W.R., Wang,R.S., Huang,J., Hao,T., *et al.* (2021) Comprehensive characterization of protein-protein interactions perturbed by disease mutations. *Nat. Genet.*, **53**, 342–353.
6. Xue,W., Fu,T., Deng,S., Yang,F., Yang,J. and Zhu,F. (2022) Molecular mechanism for the allosteric inhibition of the human serotonin transporter by antidepressant escitalopram. *ACS Chem. Neurosci.*, **13**, 340–351.
7. Piersimoni,L., Kastiris,P.L., Arlt,C. and Sinz,A. (2022) Cross-linking mass spectrometry for investigating protein conformations and protein-protein interactions horizontal line a method for all seasons. *Chem. Rev.*, **122**, 7500–7531.
8. Choi,C.R. and Rhee,H.W. (2022) Proximity labeling: an enzymatic tool for spatial biology. *Trends Biotechnol.*, **40**, 145–148.
9. Luck,K., Kim,D.K., Lambourne,L., Spirohn,K., Begg,B.E., Bian,W., Brignall,R., Cafarelli,T., Campos-Laborie,F.J., Charlotteaux,B., *et al.* (2020) A reference map of the human binary protein interactome. *Nature*, **580**, 402–408.
10. Caudal,A., Tang,X., Chavez,J.D., Keller,A., Mohr,J.P., Bakhtina,A.A., Villet,O., Chen,H., Zhou,B., Walker,M.A., *et al.* (2022) Mitochondrial interactome quantitation reveals structural changes in metabolic machinery in the failing murine heart. *Nat Cardiovasc Res*, **1**, 855–866.
11. Zoncu,R. and Perera,R.M. (2022) Built to last: lysosome remodeling and repair in health and disease. *Trends Cell Biol.*, **32**, 597–610.
12. Gan,Z., Fu,T., Kelly,D.P. and Vega,R.B. (2018) Skeletal muscle mitochondrial remodeling in exercise and diseases. *Cell Res.*, **28**, 969–980.
13. Sun,H., Fu,B., Qian,X., Xu,P. and Qin,W. (2024) Nuclear and cytoplasmic specific RNA binding proteome enrichment and its changes upon ferroptosis induction. *Nat. Commun.*, **15**, 852.
14. Guo,A.D., Yan,K.N., Hu,H., Zhai,L., Hu,T.F., Su,H., Chi,Y., Zha,J., Xu,Y., Zhao,D., *et al.* (2023) Spatiotemporal and global profiling of DNA-protein interactions enables discovery of low-affinity transcription factors. *Nat. Chem.*, **15**, 803–814.

15. Xue,W., Yang,F., Wang,P., Zheng,G., Chen,Y., Yao,X. and Zhu,F. (2018) What contributes to serotonin-norepinephrine reuptake inhibitors' dual-targeting mechanism? The key role of transmembrane domain 6 in human serotonin and norepinephrine transporters revealed by molecular dynamics simulation. *ACS Chem. Neurosci.*, **9**, 1128–1140.
16. Yang,Q., Wang,Y., Zhang,Y., Li,F., Xia,W., Zhou,Y., Qiu,Y., Li,H. and Zhu,F. (2020) NOREVA: enhanced normalization and evaluation of time-course and multi-class metabolomic data. *Nucleic Acids Res.*, **48**, W436–W448.
17. Thul,P.J., Akesson,L., Wiking,M., Mahdessian,D., Geladaki,A., Ait Blal,H., Alm,T., Asplund,A., Bjork,L., Breckels,L.M., *et al.* (2017) A subcellular map of the human proteome. *Science*, **356**, eaal3321.
18. Tomkins,J.E. and Manzoni,C. (2021) Advances in protein-protein interaction network analysis for Parkinson's disease. *Neurobiol. Dis.*, **155**, 105395.
19. Li,F., Yin,J., Lu,M., Mou,M., Li,Z., Zeng,Z., Tan,Y., Wang,S., Chu,X., Dai,H., *et al.* (2023) DrugMAP: molecular atlas and pharmacology information of all drugs. *Nucleic Acids Res.*, **51**, D1288–D1299.
20. Wang,Y., Chen,Z., Pan,Z., Huang,S., Liu,J., Xia,W., Zhang,H., Zheng,M., Li,H., Hou,T., *et al.* (2023) RNAIncode: a deep learning-based encoder for RNA and RNA-associated interaction. *Nucleic Acids Res.*, **51**, W509–W519.
21. Mou,M., Pan,Z., Zhou,Z., Zheng,L., Zhang,H., Shi,S., Li,F., Sun,X. and Zhu,F. (2023) A transformer-based ensemble framework for the prediction of protein-protein interaction sites. *Research (Wash D C)*, **6**, 0240.
22. Lu,M.K., Yin,J.Y., Zhu,Q., Lin,G.L., Mou,M.J., Liu,F.Y., Pan,Z.Q., You,N.X., Lian,X.C., Li,F.C., *et al.* (2023) Artificial intelligence in pharmaceutical sciences. *Engineering*, **27**, 37–69.
23. Fu,J., Zhang,Y., Wang,Y., Zhang,H., Liu,J., Tang,J., Yang,Q., Sun,H., Qiu,W., Ma,Y., *et al.* (2022) Optimization of metabolomic data processing using NOREVA. *Nat. Protoc.*, **17**, 129–151.
24. Yang,Q., Li,B., Tang,J., Cui,X., Wang,Y., Li,X., Hu,J., Chen,Y., Xue,W., Lou,Y., *et al.* (2020) Consistent gene signature of schizophrenia identified by a novel feature selection strategy from comprehensive sets of transcriptomic data. *Brief Bioinform.*, **21**, 1058–1068.
25. Tang,J., Fu,J., Wang,Y., Li,B., Li,Y., Yang,Q., Cui,X., Hong,J., Li,X., Chen,Y., *et al.* (2020) ANPELA: analysis and performance assessment of the label-free quantification workflow for metaproteomic studies. *Brief. Bioinform.*, **21**, 621–636.
26. Szklarczyk,D., Kirsch,R., Koutrouli,M., Nastou,K., Mehryar,F., Hachilif,R., Gable,A.L., Fang,T., Doncheva,N.T., Pyysalo,S., *et al.* (2023) The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Res.*, **51**, D638–D646.
27. Xenarios,I., Salwinski,L., Duan,X.J., Higney,P., Kim,S.M. and Eisenberg,D. (2002) DIP, the database of interacting proteins: a research tool for studying cellular networks of protein interactions. *Nucleic Acids Res.*, **30**, D303–D305.
28. Alanis-Lobato,G., Andrade-Navarro,M.A. and Schaefer,M.H. (2017) HIPPIE v2.0: enhancing meaningfulness and reliability of protein-protein interaction networks. *Nucleic Acids Res.*, **45**, D408–D414.
29. Kang,J., Tang,Q., He,J., Li,L., Yang,N., Yu,S., Wang,M., Zhang,Y., Lin,J., Cui,T., *et al.* (2022) RNAInter v4.0: RNA interactome repository with redefined confidence scoring system and improved accessibility. *Nucleic Acids Res.*, **50**, D326–D332.
30. Zheng,Y., Luo,H., Teng,X., Hao,X., Yan,X., Tang,Y., Zhang,W., Wang,Y., Zhang,P., Li,Y., *et al.* (2023) NPInter v5.0: ncRNA interaction database in a new era. *Nucleic Acids Res.*, **51**, D232–D239.
31. Gong,J., Shao,D., Xu,K., Lu,Z., Lu,Z.J., Yang,Y.T. and Zhang,Q.C. (2018) RISE: a database of RNA interactome from sequencing experiments. *Nucleic Acids Res.*, **46**, D194–D201.
32. Del Toro,N., Shrivastava,A., Ragueneau,E., Meldal,B., Combe,C., Barrera,E., Perfetto,L., How,K., Ratan,P., Shirodkar,G., *et al.* (2022) The IntAct database: efficient access to fine-grained molecular interaction data. *Nucleic Acids Res.*, **50**, D648–D653.
33. Oughtred,R., Stark,C., Breitkreutz,B.J., Rust,J., Boucher,L., Chang,C., Kolas,N., O'Donnell,L., Leung,G., McAdam,R., *et al.* (2019) The BioGRID interaction database: 2019 update. *Nucleic Acids Res.*, **47**, D529–D541.
34. Kamburov,A. and Herwig,R. (2022) ConsensusPathDB 2022: molecular interactions update as a resource for network biology. *Nucleic Acids Res.*, **50**, D587–D595.
35. Veres,D.V., Gyurko,D.M., Thaler,B., Szalay,K.Z., Fazekas,D., Korcsmaros,T. and Csermely,P. (2015) CompPI: a cellular compartment-specific database for protein-protein interaction network analysis. *Nucleic Acids Res.*, **43**, D485–D493.
36. Kotlyar,M., Pastrello,C., Ahmed,Z., Chee,J., Varyova,Z. and Jurisica,I. (2022) IID 2021: towards context-specific protein interaction analyses by increased coverage, enhanced annotation and enrichment analysis. *Nucleic Acids Res.*, **50**, D640–D647.
37. UniProt,C. (2023) UniProt: the universal protein knowledgebase in 2023. *Nucleic Acids Res.*, **51**, D523–D531.
38. Kozomara,A., Birgaoanu,M. and Griffiths-Jones,S. (2019) miRBase: from microRNA sequences to function. *Nucleic Acids Res.*, **47**, D155–D162.
39. Zhou,Y., Zhang,Y., Zhao,D., Yu,X., Shen,X., Zhou,Y., Wang,S., Qiu,Y., Chen,Y. and Zhu,F. (2024) TTD: therapeutic target database describing target druggability information. *Nucleic Acids Res.*, **52**, D1465–D1477.
40. Burley,S.K., Bhikadiya,C., Bi,C., Bittrich,S., Chao,H., Chen,L., Craig,P.A., Crichlow,G.V., Dalenberg,K., Duarte,J.M., *et al.* (2023) RCSB Protein Data Bank (RCSB.org): delivery of experimentally-determined PDB structures alongside one million computed structure models of proteins from artificial intelligence/machine learning. *Nucleic Acids Res.*, **51**, D488–D508.
41. Varadi,M., Bertoni,D., Magana,P., Paramval,U., Pidruchna,I., Radhakrishnan,M., Tsenkov,M., Nair,S., Mirdita,M., Yeo,J., *et al.* (2024) AlphaFold protein structure database in 2024: providing structure coverage for over 214 million protein sequences. *Nucleic Acids Res.*, **52**, D368–D375.
42. Harrison,P.W., Amode,M.R., Austine-Orimoloye,O., Azov,A.G., Barba,M., Barnes,I., Becker,A., Bennett,R., Berry,A., Bhai,J., *et al.* (2024) Ensembl 2024. *Nucleic Acids Res.*, **52**, D891–D899.
43. Bruford,E.A., Braschi,B., Denny,P., Jones,T.E.M., Seal,R.L. and Tweedie,S. (2020) Guidelines for human gene nomenclature. *Nat. Genet.*, **52**, 754–758.
44. Schoch,C.L., Ciufu,S., Domrachev,M., Hotton,C.L., Kannan,S., Khovanskaya,R., Leipe,D., McVeigh,R., O'Neill,K., Robertse,B., *et al.* (2020) NCBI Taxonomy: a comprehensive update on curation, resources and tools. *Database*, **2020**, baaa062.
45. Sayers,E.W., Beck,J., Bolton,E.E., Brister,J.R., Chan,J., Comeau,D.C., Connor,R., DiCuccio,M., Farrell,C.M., Feldgarden,M., *et al.* (2024) Database resources of the national center for biotechnology information. *Nucleic Acids Res.*, **52**, D33–D43.
46. Rath,S., Sharma,R., Gupta,R., Ast,T., Chan,C., Durham,T.J., Goodman,R.P., Grabarek,Z., Haas,M.E., Hung,W.H.W., *et al.* (2021) MitoCarta3.0: an updated mitochondrial proteome now with sub-organelle localization and pathway annotations. *Nucleic Acids Res.*, **49**, D1541–D1547.
47. Cui,T., Dou,Y., Tan,P., Ni,Z., Liu,T., Wang,D., Huang,Y., Cai,K., Zhao,X., Xu,D., *et al.* (2022) RNALocate v2.0: an updated resource for RNA subcellular localization with increased coverage and annotation. *Nucleic Acids Res.*, **50**, D333–D339.
48. Bian,W., Jiang,H., Yao,L., Hao,W., Wu,L. and Li,X. (2023) A spatially defined human Notch receptor interaction network reveals Notch intracellular storage and Ataxin-2-mediated fast recycling. *Cell Rep.*, **42**, 112819.
49. Lundberg,E. and Borner,G.H.H. (2019) Spatial proteomics: a powerful discovery tool for cell biology. *Nat. Rev. Mol. Cell Biol.*, **20**, 285–302.

50. Zheng,L., Shi,S., Lu,M., Fang,P., Pan,Z., Zhang,H., Zhou,Z., Zhang,H., Mou,M., Huang,S., *et al.* (2024) AnnoPRO: a strategy for protein function annotation based on multi-scale protein representation and a hybrid deep learning of dual-path encoding. *Genome Biol.*, **25**, 41.
51. Li,B., Tang,J., Yang,Q., Li,S., Cui,X., Li,Y., Chen,Y., Xue,W., Li,X. and Zhu,F. (2017) NOREVA: normalization and evaluation of MS-based metabolomics data. *Nucleic Acids Res.*, **45**, W162–W170.
52. Wang,J., Horlacher,M., Cheng,L. and Winther,O. (2023) RNA trafficking and subcellular localization—a review of mechanisms, experimental and predictive methodologies. *Brief Bioinform.*, **24**, bbad249.
53. Das,S., Vera,M., Gandin,V., Singer,R.H. and Tutucci,E. (2021) Intracellular mRNA transport and localized translation. *Nat. Rev. Mol. Cell Biol.*, **22**, 483–504.
54. Wang,T.H., Wu,C.C., Huang,K.Y., Chuang,W.Y., Hsueh,C., Li,H.J. and Chen,C.Y. (2020) Profiling of subcellular EGFR interactome reveals hnRNP A3 modulates nuclear EGFR localization. *Oncogenesis*, **9**, 40.
55. Yang,Q., Li,B., Chen,S., Tang,J., Li,Y., Li,Y., Zhang,S., Shi,C., Zhang,Y., Mou,M., *et al.* (2021) MMEASE: online meta-analysis of metabolomic data by enhanced metabolite annotation, marker selection and enrichment analysis. *J. Proteomics*, **232**, 104023.
56. Wang,Y., Pan,Z., Mou,M., Xia,W., Zhang,H., Zhang,H., Liu,J., Zheng,L., Luo,Y., Zheng,H., *et al.* (2023) A task-specific encoding algorithm for RNAs and RNA-associated interactions based on convolutional autoencoder. *Nucleic Acids Res.*, **51**, e110.
57. Zhang,S., Amahong,K., Zhang,C., Li,F., Gao,J., Qiu,Y. and Zhu,F. (2021) RNA-RNA interactions between SARS-CoV-2 and host benefit viral development and evolution during COVID-19 infection. *Brief Bioinform.*, **23**, bbab397.
58. Pankow,S., Martinez-Bartolome,S., Bamberger,C. and Yates,J.R. (2019) Understanding molecular mechanisms of disease through spatial proteomics. *Curr. Opin. Chem. Biol.*, **48**, 19–25.
59. Peugot,S., Zhou,X. and Selivanova,G. (2024) Translating p53-based therapies for cancer into the clinic. *Nat. Rev. Cancer*, **24**, 192–215.
60. Le Mercier,P., Bolleman,J., de Castro,E., Gasteiger,E., Bansal,P., Auchincloss,A.H., Boutet,E., Breuza,L., Casals-Casas,C., Estreicher,A., *et al.* (2022) SwissBioPics—an interactive library of cell images for the visualization of subcellular location data. *Database*, **2022**, baac026.
61. Bergmann,L., Lang,A., Bross,C., Altinluk-Hambuchen,S., Fey,I., Overbeck,N., Stefanski,A., Wiek,C., Kefalas,A., Verhulsdonk,P., *et al.* (2020) Subcellular localization and mitotic interactome analyses identify SIRT4 as a centrosomally localized and microtubule associated protein. *Cells*, **9**, 1950.
62. Cho,W.K., Spille,J.H., Hecht,M., Lee,C., Li,C., Grube,V. and Cisse,I.I. (2018) Mediator and RNA polymerase II clusters associate in transcription-dependent condensates. *Science*, **361**, 412–415.
63. Li,F., Zhou,Y., Zhang,Y., Yin,J., Qiu,Y., Gao,J. and Zhu,F. (2022) POSREG: proteomic signature discovered by simultaneously optimizing its reproducibility and generalizability. *Brief Bioinform.*, **23**, bbac040.
64. Suomalainen,A. and Battersby,B.J. (2018) Mitochondrial diseases: the contribution of organelle stress responses to pathology. *Nat. Rev. Mol. Cell Biol.*, **19**, 77–92.
65. Yang,J., Griffin,A., Qiang,Z. and Ren,J. (2022) Organelle-targeted therapies: a comprehensive review on system design for enabling precision oncology. *Signal Transduct. Target Ther.*, **7**, 379.
66. Shao,X., Meng,C., Song,W., Zhang,T. and Chen,Q. (2023) Subcellular visualization: organelle-specific targeted drug delivery and discovery. *Adv. Drug. Deliv. Rev.*, **199**, 114977.
67. Li,Y.H., Li,X.X., Hong,J.J., Wang,Y.X., Fu,J.B., Yang,H., Yu,C.Y., Li,F.C., Hu,J., Xue,W.W., *et al.* (2020) Clinical trials, progression-speed differentiating features and swiftness rule of the innovative targets of first-in-class drugs. *Brief. Bioinform.*, **21**, 649–662.
68. Yin,J., Sun,W., Li,F., Hong,J., Li,X., Zhou,Y., Lu,Y., Liu,M., Zhang,X., Chen,N., *et al.* (2020) VARIDT 1.0: variability of drug transporter database. *Nucleic Acids Res.*, **48**, D1042–D1050.
69. Refolo,G., Vescovo,T., Piacentini,M., Fimia,G.M. and Ciccosanti,F. (2020) Mitochondrial interactome: a focus on antiviral signaling pathways. *Front. Cell Dev. Biol.*, **8**, 8.
70. Bakhtina,A.A., Pharaoh,G.A., Campbell,M.D., Keller,A., Stuppard,R.S., Marcinek,D.J. and Bruce,J.E. (2023) Skeletal muscle mitochondrial interactome remodeling is linked to functional decline in aged female mice. *Nat Aging*, **3**, 313–326.
71. Schweppe,D.K., Chavez,J.D., Lee,C.F., Caudal,A., Kruse,S.E., Stuppard,R., Marcinek,D.J., Shadel,G.S., Tian,R. and Bruce,J.E. (2017) Mitochondrial protein interactome elucidated by chemical cross-linking mass spectrometry. *Proc. Natl. Acad. Sci. U.S.A.*, **114**, 1732–1737.