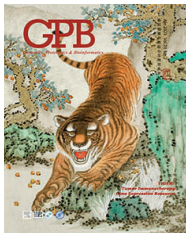
Genomics Proteomics Bioinformatics 21 (2023) 267–277

Genomics Proteomics Bioinformatics

www.elsevier.com/locate/gpb
www.sciencedirect.com



DATABASE

iHypoxia: An Integrative Database of Protein Expression Dynamics in Response to Hypoxia in Animals

Ze-Xian Liu [1,#], Panqin Wang [2,#], Qingfeng Zhang [1,#], Shihua Li [1,2], Yuxin Zhang [2], Yutong Guo [2], Chongchong Jia [2], Tian Shao [2], Lin Li [1], Han Cheng [2,*], Zhenlong Wang [2,*]

[1] *State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Sun Yat-sen University Cancer Center, Sun Yat-sen University, Guangzhou 510060, China*
[2] *School of Life Sciences, Zhengzhou University, Zhengzhou 450001, China*

Received 3 June 2021; revised 7 November 2022; accepted 2 December 2022
Available online 9 December 2022

Handled by Tiannan Guo

KEYWORDS

Hypoxia;
Expression dynamics;
Low-throughput experiment;
High-throughput experiment;
Functional annotation

Abstract Mammals have evolved mechanisms to sense **hypoxia** and induce hypoxic responses. Recently, high-throughput techniques have greatly promoted global studies of protein expression changes during hypoxia and the identification of candidate genes associated with hypoxia-adaptive evolution, which have contributed to the understanding of the complex regulatory networks of hypoxia. In this study, we developed an integrated resource for the **expression dynamics** of proteins in response to hypoxia (iHypoxia), and this database contains 2589 expression events of 1944 proteins identified by **low-throughput experiments** (LTEs) and 422,553 quantitative expression events of 33,559 proteins identified by **high-throughput experiments** from five mammals that exhibit a response to hypoxia. Various experimental details, such as the hypoxic experimental conditions, expression patterns, and sample types, were carefully collected and integrated. Furthermore, 8788 candidate genes from diverse species inhabiting low-oxygen environments were also integrated. In addition, we conducted an orthologous search and computationally identified 394,141 proteins that may respond to hypoxia among 48 animals. An enrichment analysis of human proteins identified from LTEs shows that these proteins are enriched in certain drug targets and cancer genes. Annotation of known posttranslational modification (PTM) sites in the proteins identified by LTEs reveals that these proteins undergo extensive PTMs, particularly phosphorylation, ubiquitination, and acetylation. iHypoxia provides a convenient and user-friendly method for users

* Corresponding authors.
E-mail: wzl@zzu.edu.cn (Wang Z), chenghan@zzu.edu.cn (Cheng H).
# Equal contribution.

Peer review under responsibility of Beijing Institute of Genomics, Chinese Academy of Sciences / China National Center for Bioinformation and Genetics Society of China.

https://doi.org/10.1016/j.gpb.2022.12.001

to obtain hypoxia-related information of interest. We anticipate that iHypoxia, which is freely accessible at https://ihypoxia.omicsbio.info, will advance the understanding of hypoxia and serve as a valuable data resource.

Introduction

Oxygen is an essential element for the survival of aerobic organisms, and a decreased oxygen concentration (hypoxia) may threaten the survival of cells, tissues, and individuals. To cope with hypoxia, mammals have evolved sophisticated oxygen-sensing systems and adaptive mechanisms [1], including erythropoiesis [2], mitochondrial respiration [3], angiogenesis [4], and metabolic adaptations [5]. In 1995, Wang et al. identified hypoxia-inducible factor 1 (HIF-1), which is a heterodimer consisting of a constitutively expressed HIF-1β subunit and a HIF-1α subunit that is regulated and expressed in an oxygen-dependent manner [6]. Under normoxic conditions, HIF-1α is hydroxylated by prolyl hydroxylases and then targeted for proteasomal degradation through von Hippel–Lindau (VHL)-mediated ubiquitination [7–9]. During hypoxia, accumulating HIF-1α dimerizes with HIF-1β and then binds to hypoxia response elements to trigger the hypoxic response [10]. In addition to HIF-1, many proteins and pathways are involved in hypoxia signaling [11] and thus play important roles in various physiological and pathological processes, such as the regulation of immunity [12], embryonic development [13], cancer [14], cardiovascular diseases [15], and anemia [16]. The use of hypoxia-related strategies for the treatment of disease is being investigated. For example, Chen et al. found that roxadustat, an oral inhibitor of HIF prolyl hydroxylase, is useful for treating patients with anemia and long-term kidney disease [17]. Therefore, a deeper understanding of hypoxia signaling is critical, and an exploration of the expression dynamics of proteins in response to hypoxia is the first step.

Several databases related to hypoxia are currently available. In 2012, Pankaj et al. compiled HypoxiaDB, a database of hypoxia-regulated proteins containing 3500 human proteins with detailed annotations that were identified from high-throughput experiments (HTEs) [18]. However, HypoxiaDB has not been updated since then. In 2017, Rashid et al. developed the HRGFish database, which contains 818 hypoxia-responsive genes collected from 38 fishes [19]. During the last decade, many studies have contributed to dissecting the protein expression dynamics in cellular/animal models in response to hypoxia under various experimental/physiological conditions [20–22]. Therefore, a comprehensive database that integrates the quantitative events of protein expression in response to hypoxia will boost the reutilization of previous investigations and provide a helpful data resource for studying the molecular mechanisms of hypoxia-related biology.

Herein, we presented iHypoxia, an integrative database for the expression dynamics of proteins in response to hypoxia in animals. At present, the iHypoxia database has curated and hosts 425,142 quantitative expression events of 33,893 proteins from five mammals in response to hypoxia, 8788 candidate genes involved in adaptive evolution, and 394,141 computationally detected proteins in 48 animals. Detailed annotations, including posttranslational modifications (PTMs), subcellular locations, protein–protein interactions, and drug–target relations, are also provided. Search and browse services have also been implemented for user convenience. An enrichment analysis of human proteins identified in low-throughput experiments (LTEs) showed that proteins that respond to hypoxia were significantly enriched in certain drug targets and oncogenes. By mapping known PTM sites to proteins identified in LTEs, we confirmed that these proteins can undergo extensive PTMs, particularly phosphorylation, ubiquitination, and acetylation. We anticipate that the iHypoxia database will be beneficial to the research community, and this online service is freely available at https://ihypoxia.omicsbio.info.

Database implementation

To develop a comprehensive resource for the expression dynamics of proteins in response to hypoxia in animals, we collected and integrated published (through January 2020) datasets from various sources, such as PubMed, Gene Expression Omnibus (GEO), HypoxiaDB [18], and Gene Ontology (GO) [23]. A workflow of the construction of iHypoxia is depicted in **Figure 1**.

Curation of hypoxia-related proteins identified by LTEs

To obtain validated proteins and their expression patterns in response to hypoxia, we searched for hypoxia-related scientific reviews in PubMed using the keyword "hypoxia". Then, we first read the abstracts to confirm whether there were hypoxia-related proteins. If so or if we were unsure, we checked the full texts and traced them back to the original literature. By carefully reading the literature, we manually extracted the hypoxic expression of proteins identified from LTEs, such as Western blot (WB), Northern blot (NB), polymerase chain reaction (PCR), and other experiments. Considering that proteins show different expression patterns in various tissues/cells, in the presence of different oxygen concentrations, and in response to different durations of hypoxia, we retrieved as many details as possible, including the hypoxic experimental conditions, expression patterns (up-regulation or down-regulation), sample types (tissues or cells), evidence levels (proteins or transcripts), experimental methods (*e.g.*, WB and NB), and a brief description of the proteins from the original articles. All the protein expression experiments were checked by two data curators. In addition, we integrated data from HypoxiaDB [18], GO resources [23], Reactome [24], MsigDB [25], Kyoto Encyclopedia of Genes and Genomes (KEGG) [26], GeneCards [27], and GenAtlas [28]. Regarding the GO resources, we collected only the proteins whose GO terms were related to hypoxia, as marked by experimental evidence codes, which indicated that the annotation of the gene was directly supported by experimental evidence. From pathway databases such as Reactome and MsigDB, hypoxia-related datasets were also downloaded. We then identified the original literature to obtain the protein expression pattern, the hypoxic conditions, and other information. For the data from HypoxiaDB, detailed information was employed without further changes.

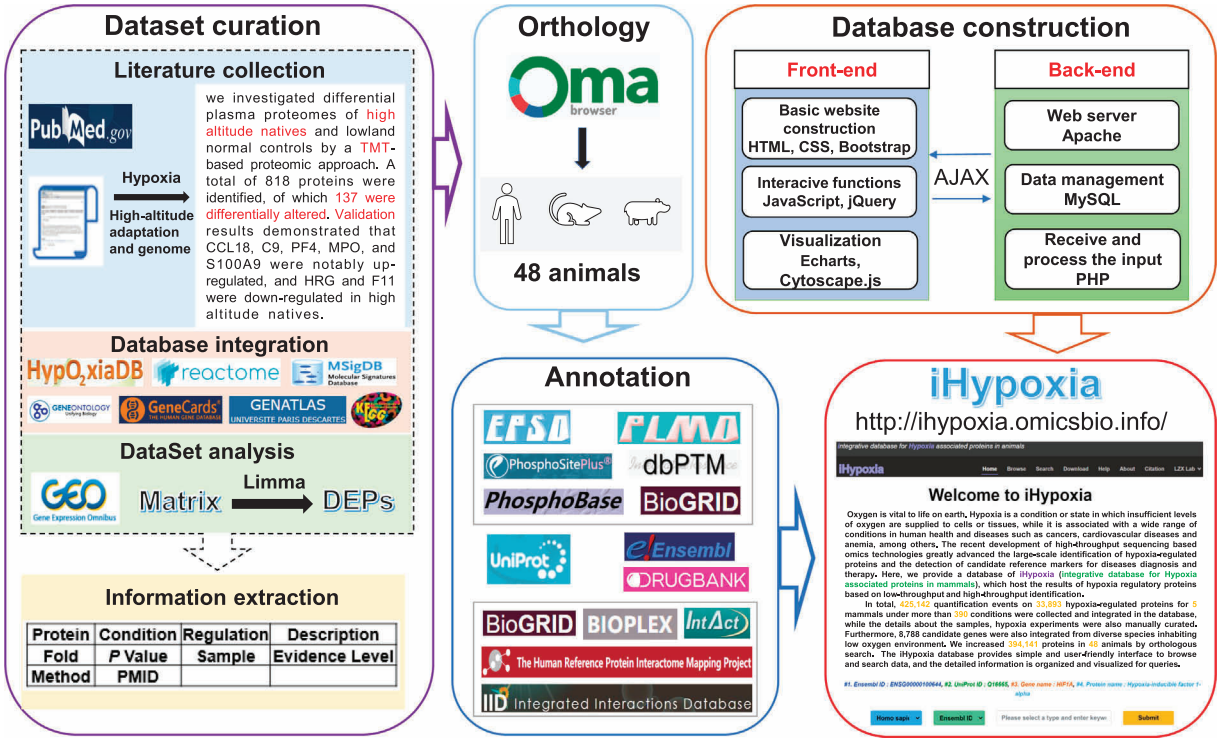


Figure 1 Workflow of the construction of the iHypoxia database
First, we collected and integrated expression data of proteins in response to hypoxia from the literature, HypoxiaDB, GO resource, Reactome, MSigDB, GeneCards, GenAtlas, and GEO. We then performed ortholog detection using the OMA dataset. In addition to collecting basic information, we annotated proteins in four aspects: PTMs, subcellular locations, protein–protein interactions, and drug–target relations. Ultimately, the iHypoxia website was constructed. TMT, tandem mass tag; GO, Gene Ontology; GEO, Gene Expression Omnibus; DEP, differentially expressed protein; PMID, PubMed ID; OMA, Orthologous Matrix; PTM, posttranslational modification.

Integrating high-throughput protein expression profiles in response to hypoxia

Proteomics is a highly effective approach for identifying alterations in protein levels. To identify the expression at the protein level of genes that respond to hypoxia, we searched PubMed with various keywords, such as "(hypoxia OR anoxia) AND (proteome OR proteomic)". To control the quality of the data, we manually examined all the literature retrieved and the corresponding supplemental materials. Information pertaining to the hypoxic conditions, regulation patterns, *P* adjust values, sample types, and other details was then extracted.

Transcriptomic approaches have been widely used to research transcriptional expression during hypoxic responses, and GEO hosts many samples with expression data related to hypoxic responses. Therefore, we searched the GEO database with the keywords "hypoxia" and "anoxia" with the screening conditions "Expression profiling by array" and "Expression profiling by high-throughput sequencing". Hypoxia-related expression datasets were then downloaded and analyzed using the limma package [29]. If the fold change of mRNA expression in hypoxic cells/tissues compared with control (normoxia) cells/tissues was greater than 2 and the *P* adjust value was < 0.05, the protein was included in the iHypoxia database. Detailed information, including the hypoxic experimental conditions, sample types, and other details, was also retrieved from the original studies.

After completing the aforementioned data collection, the "Retrieve/ID mapping" tool in the Universal Protein Resource (UniProt) database was used to convert the different identifiers of the collected proteins into unified Ensembl IDs and UniProt IDs [30].

Curation of hypoxia adaptation-associated genes identified by genomic analysis

Many animals living in low-oxygen environments, especially high-altitude habitats, have evolved genetic mechanisms of adaptation to hypoxia, and genomic analysis has identified numerous candidate genes involved in hypoxia adaptation at the genetic level. To obtain these candidate genes, we searched PubMed with keywords such as "hypoxia AND (genome OR genomic) AND (high-altitude adaptation)". By manually mining abstracts and full texts, we collected candidate genes and relevant information from 127 studies.

Ortholog detection

To identify computationally detected proteins in other animals that may respond to hypoxia, the "Pairwise orthologs" data were downloaded from the Orthologous Matrix (OMA) database (https://omabrowser.org/oma/current/) [31]. The collected proteins were then mapped to the OMA dataset to obtain their orthologs. Through screening, we obtained 48 ani-

mal ortholog datasets, and the Ensembl IDs and UniProt IDs were used as identifiers.

Annotation data obtained from public databases

To provide abundant annotations for the collected proteins, information obtained from several public databases, including basic information, PTMs, subcellular locations, protein–protein interactions, and drug–target information, was integrated.

First, essential information on the proteins, such as the gene name, protein name, sequence, and function, was downloaded from UniProt [32]. PTMs play vital roles in regulating the structure and function of proteins; thus, PTM data from Biological General Repository for Interaction Datasets (BioGRID) [33], dbPTM [34], eukaryotic phosphorylation site database (EPSD) [35], PhosphoBase [36], PhosphoSitePlus [37], and protein lysine modifications database (PLMD) [38] were integrated and matched to the collected proteins. Proteins function in specific subcellular locations, and these locations provide a particular environment and a set of interacting chaperone proteins that are necessary for protein function. Knowledge of the subcellular localization of proteins is important for understanding cellular processes. Therefore, we acquired the subcellular locations from UniProt [32] and protein–protein interactions from BioGRID [33], BioPlex [39], Integrated Interactions Database (IID) [40], IntAct [41], and Human Reference Interactome and Literature Benchmark (HuRI) [42]. Based on expression studies, some proteins have also been proposed as potential diagnostic or therapeutic biomarkers for human hypoxic diseases, especially cancer. To determine whether the proteins that we collected could be used as drug targets or are encoded by cancer-related genes, we downloaded data from DrugBank [43] and a list of oncogenes from the Cancer Gene Census in the Catalog Of Somatic Mutations in Cancer (COSMIC) database [44] and then matched them.

Database construction

All the data in iHypoxia were stored in a MySQL database, and the website was deployed with an Apache server. At the front end, Bootstrap, an open-source user interface framework based on HTML, CSS, and JavaScript, was used to build the basic layout of the website. JavaScript and its tool library Jquery were used to implement the interactive functions. The visualization of the data statistics and PTMs was achieved with Echarts, whereas the network of protein–protein interactions was displayed using Cytoscape.js. Data processing at the back end was performed through PHP, and the interaction between the front and back ends was achieved through Asynchronous JavaScript and XML (AJAX). Moreover, to provide a stable and adaptive service, we tested the iHypoxia website using various web browsers, including Mozilla Firefox, Google Chrome, and Internet Explorer.

Web interface

User interface and functions

To rapidly access the detailed information on a protein of interest, a simple search function is available on the home page for searching in a user-specified area or all areas under "Any Field" (**Figure 2**A). We provided examples for each keyword option to improve the ease of the search process for users. Examples include "ENSG00000100644", "Q16665", "*HIF1A*", and "hypoxia inducible factor 1 subunit alpha", which are the Ensembl ID, UniProt ID, gene name, and protein name, respectively (Figure 2A). On the search page, we provide two search approaches, including an advanced search and a batch search. For the "Advanced search" option, various keywords can be combined through the operators "AND", "OR", and "NOT" to perform a query with different search keywords (Figure 2B). For the "Batch search" option, users can paste numerous keywords, such as Ensembl IDs and gene names, in a line-by-line format or upload a file for batch searching (Figure 2C). The database uses a fuzzy query function, and all matched results are listed and displayed in a table at the bottom of the search page (Figure 2D). The table showing the results contains basic information, *i.e.*, Ensembl ID, UniProt ID, gene name, protein name, and organism, and detailed information for the searched protein can be viewed on the result page by clicking the "More" option (Figure 2D).

On the browse page, all animal species in iHypoxia are presented in alphabetical order (**Figure 3**A). To view the proteins in species of interest, users can click the corresponding species, and a table with basic information (Ensembl ID, UniProt ID, gene name, and experiment types) will be shown in the list below (Figure 3B). Users can click the "Detail" link to view detailed information. Additionally, an evolutionary tree showing the phylogenetic relationships of these 48 animals is shown, and users can directly click the animal of interest to view corresponding proteins (Figure 3C).

The result page for a protein consists of eleven parts. The first part shows the basic information of the protein, including the organism, Ensembl ID, UniProt ID, gene name, protein name, protein function, and sequence (**Figure 4**A). The other six parts detail the expression dynamics of the proteins in response to hypoxia identified in LTEs and HTEs, hypoxia adaptation evidence from genomic analyses, and information obtained from protein ortholog searches (Figure 4B–G). The table of evidence from LTEs and HTEs includes the experimental conditions, regulation under hypoxic conditions, $\log_2$ ratios, *P* adjust values, false-discovery rates (FDRs), and evidence levels (Figure 4B and C). The table of evidence from genomic analyses includes the adaptation types, description, and evidence levels (Figure 4D). The tables for proteins obtained from ortholog searches are similar to the tables of evidence from LTEs, HTEs, and genomic analyses (Figure 4E–G). By clicking on the "plus" button, users can see more information, which includes brief descriptions of the experiments, sample types, detection methods, publication, or dynamic processes (Figure 4H). Moreover, these data can be downloaded by clicking the "Export" button in the upper right corner of the table. The remaining four parts are "Post-translational modification(s)", "Protein–protein interaction(s)", "Subcellular location(s)", and "Drug–target relation(s)", and these parts correspond to all PTMs we collected for the searched protein, the network in which the protein interacts with other LTE-identified proteins, subcellular location information, and potential inhibitors, respectively (Figure 4I–L). On the download page, users can freely acquire all the data for their own academic research.

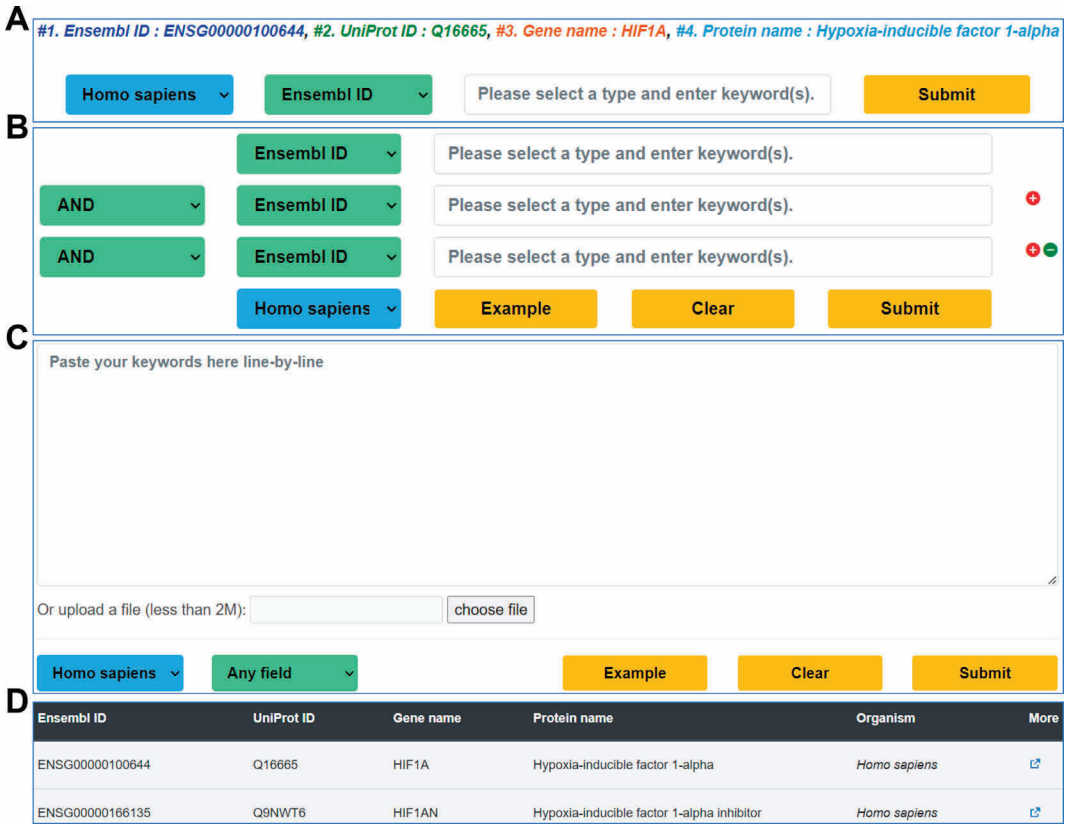


Figure 2 Search options in iHypoxia

A. Simple search options on the home page. **B.** Advanced search options on the search page. **C.** Batch search options on the search page. **D.** Results returned from a search for proteins.

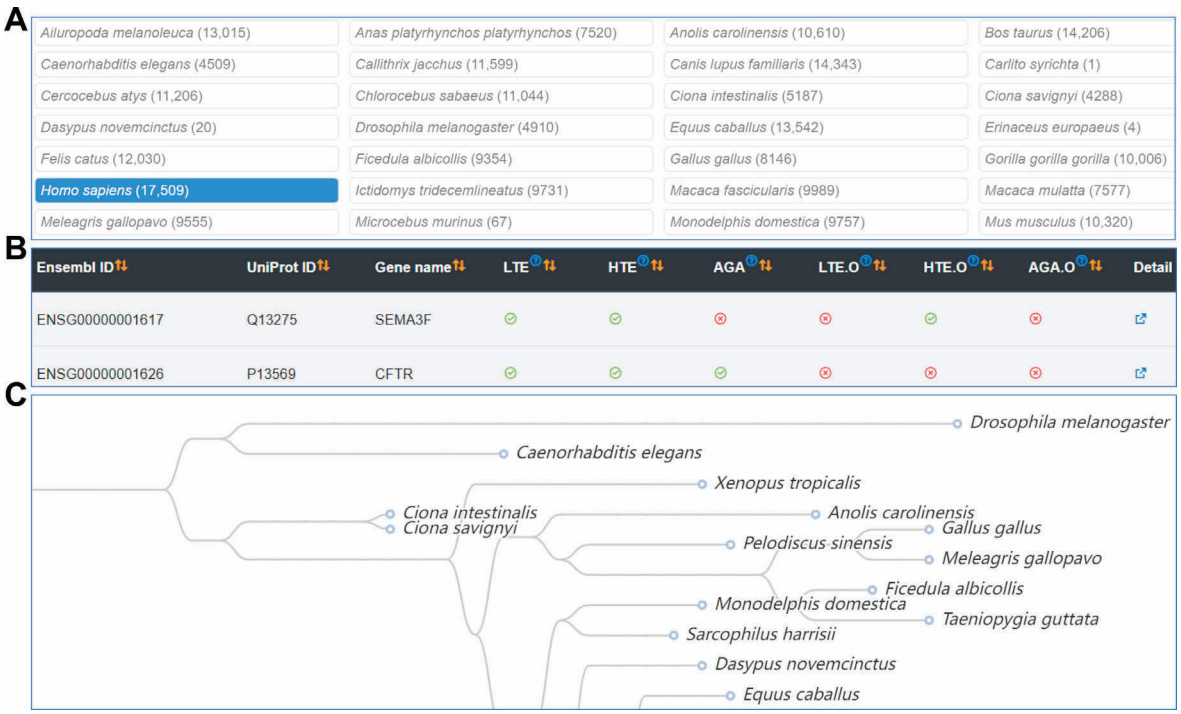


Figure 3 Browsing options in iHypoxia

A. Browse by species in a table. **B.** Results in a table for all proteins of the chosen organism. **C.** Browse by species in an evolutionary tree.

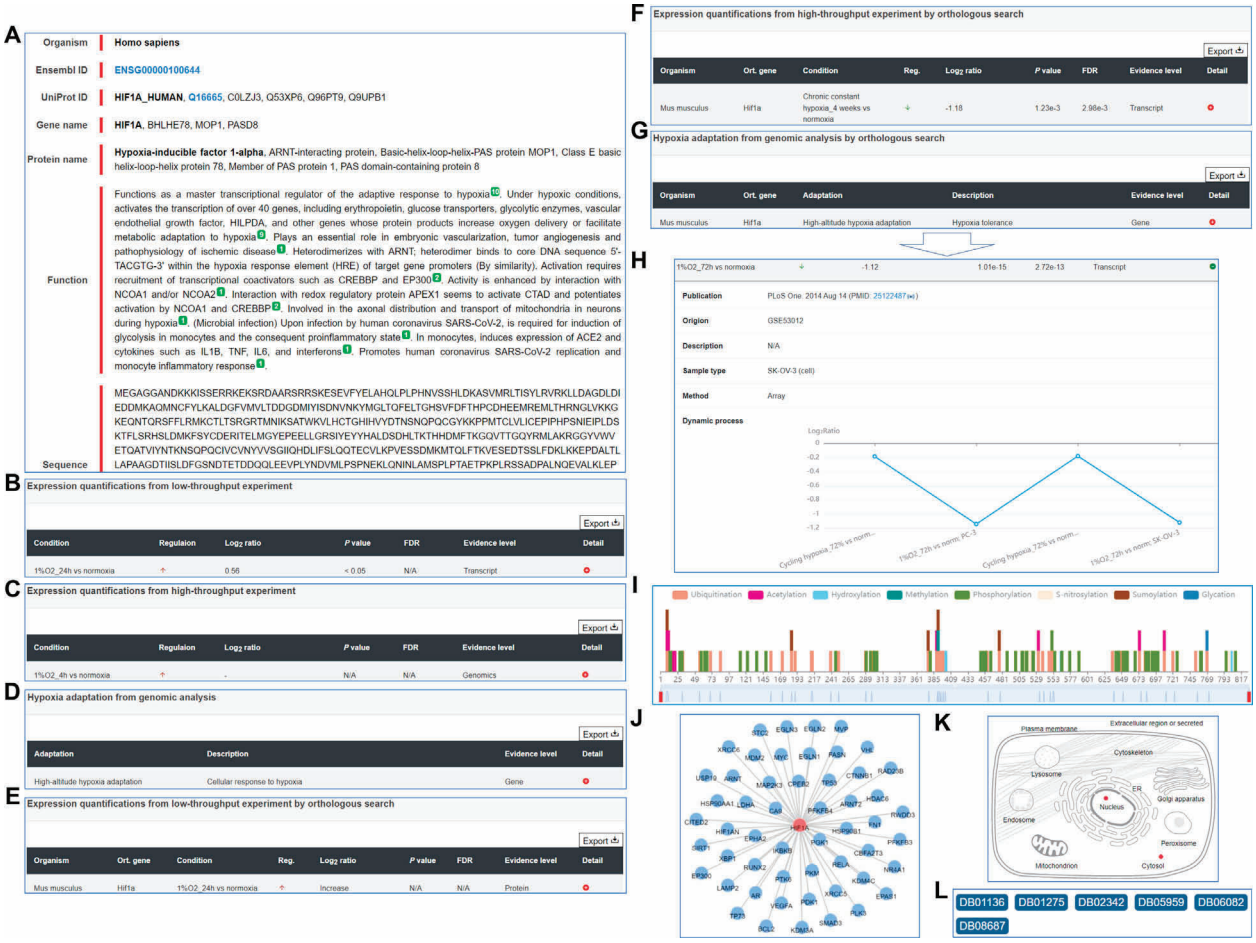


Figure 4 Results page obtained for a searched protein

A. Fundamental annotations of a searched protein. **B.** LTE evidence for a searched protein. **C.** HTE evidence for a searched protein. **D.** Genomic evidence for a searched protein. **E.** LTE evidence for an orthologous search for a searched protein. **F.** HTE evidence for an orthologous search for a searched protein. **G.** Genomic evidence for an orthologous search for a searched protein. **H.** More detail for a searched protein. **I.** PTMs for a searched protein. **J.** Protein–protein interactions of a searched protein. **K.** Subcellular location of a searched protein. **L.** Drug–target relationship of a searched protein. LTE, low-throughput experiment; HTE, high-throughput experiment.

Statistical analysis of iHypoxia

Through literature curation and public database integration, we obtained 425,142 quantitative expression events of 33,893 proteins that respond to hypoxia in five mammals: *Homo sapiens* (human), *Mus musculus* (mouse), *Rattus norvegicus* (rat), *Sus scrofa* (pig), and *Bos taurus* (cow). Based on the alterations in the expression of the proteins in hypoxia compared with normoxia, the proteins were divided into up-regulation and down-regulation categories (**Figure 5**A and B). Because proteins that respond to hypoxia present different expression patterns in various tissues/cells or under various hypoxic states, some proteins may be up-regulated and down-regulated as conditions change, and these proteins were annotated "up/down" (Figure 5A and B). In general, the protein expression quantifications obtained in LTEs were reliable, and those obtained in HTEs needed further validation; therefore, we curated the expression quantifications from LTEs and HTEs separately. In this study, for five mammals, we obtained 1944 and 33,559 unique proteins whose expression levels showed changes in response to hypoxia in LTEs (Figure 5A) and HTEs (Figure 5B), respectively. Regardless of whether the proteins were obtained from HTEs or LTEs, proteins detected in humans, mice, and rats accounted for most proteins, and few proteins were found in cows and pigs. We observed that 82.72% (1608/1944) of the proteins identified in LTEs were also detected in HTEs (Figure 5C), and 13,007 proteins showed altered expression at both the protein and transcription levels (Figure 5D). The distributions of experimental types used to identify proteins in LTEs and HTEs are shown in Figure 5E and F, respectively. PCR and arrays were the most commonly used techniques for the detection of hypoxia-related protein expression. As shown in Figure 5G and H, we observed that cancer cell lines and tissues such as hippocampus tissue were frequently used to study hypoxia responses. Furthermore, 8788 candidate genes associated with hypoxia adaptation by the genomic analysis were derived from diverse species inhabiting low-oxygen environments, such as Tibetan antelope and pika (Table S1). More than 1900 candidate genes showed altered expression at both the protein and transcription levels. In addition, because the data for one species should be helpful for studies in another species because

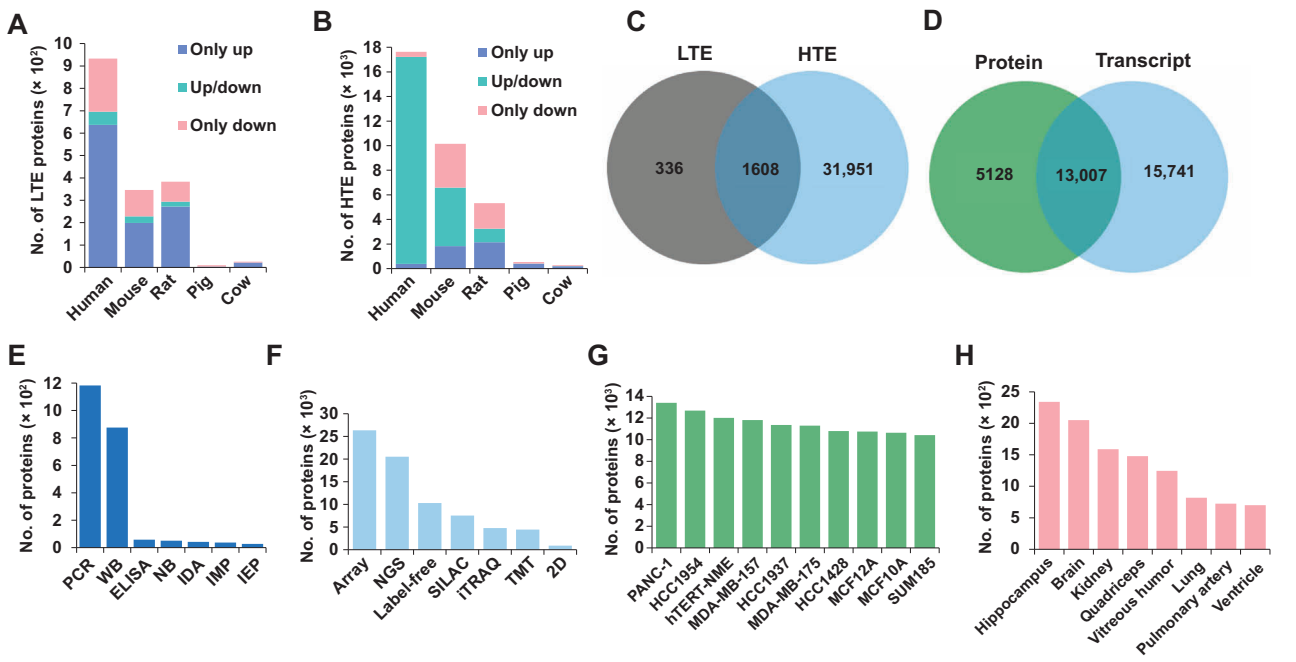


Figure 5 Data statistics of iHypoxia

A. Distribution of proteins identified from LTEs. **B.** Distribution of proteins identified from HTEs. **C.** Overlap between proteins identified from LTEs and HTEs. **D.** Overlap between collected proteins identified at transcript and protein levels. **E.** Number of proteins identified by different LTE methods. IDA, IMP, and IEP are the experimental evidence codes in the GO database. **F.** Number of proteins identified by different HTE methods. **G.** Number of proteins identified in different cell lines. **H.** Number of proteins identified in different tissues. Smaller amounts of data are not shown in the figure. PCR, polymerase chain reaction; WB, Western blot; ELISA, enzyme linked immunosorbent assay; NB, Northern blot; IDA, inferred from direct assay; IMP, inferred from mutant phenotype; IEP, inferred from expression pattern; NGS, next generation sequencing; SILAC, stable isotope labeling by amino acids in cell culture; iTRAQ, isobaric tag for relative and absolute quantitation; 2D, two-dimensional gel electrophoresis.

cells might utilize the same or similar mechanisms across species, we further increased 394,141 proteins in 48 animals based on an ortholog search (Table S2), which could provide helpful information for non-model animals that scientists might be interested in.

To further understand the biological and functional properties of LTE-identified proteins that are differentially expressed in hypoxia, we performed GO [23] and KEGG [26] pathway enrichment analyses. For simplicity, only the top 10 significant GO biological process (P adjust < 9.33E–39), cellular component (P adjust < 7.09E–22), and molecular function (P adjust < 6.33E–11) terms and KEGG terms (P adjust < 1.03E–15) are shown in **Figure 6**. The enriched biological process terms were regulation of vasculature development (GO:1901342), regulation of angiogenesis (GO:0045765), response to oxidative stress (GO:0006979), and GO terms related to oxygen levels such as response to decreased oxygen levels (GO:0036293), response to hypoxia (GO:0001666), and cellular response to hypoxia (GO:0071456) (Figure 6). Many studies have found that the levels of reactive oxygen species (ROS) increase during hypoxia [45]. Systemically, organisms have evolved many vital mechanisms to adapt to hypoxia, and these mechanisms include blood vessel growth, which delivers as much oxygen and nutrients as possible to meet the metabolic demand [4,46]. With respect to cellular component terms, most of the proteins were enriched in the terms proteasome complex (GO:0000502), endopeptidase complex (GO:1905369), and vesicle lumen (GO:0031983). The enriched molecular function terms were growth factor binding (GO:0019838), growth factor activity (GO:0008083), cytokine receptor binding (GO:0005126), and others. The prevalence of "binding" terms suggests that proteins that respond to hypoxia play extremely important roles in complex protein interactions. In addition, the most enriched GO molecular function term was "growth factor binding", which plays a momentous role in modulating angiogenic processes in response to hypoxia [47].

Hypoxia is a complex condition and a feature of physiological and pathological states. Organisms sense and adapt to hypoxic stress via various intricate pathways [1]. In addition to the classical and well–known HIF-1 signaling pathway (hsa04066), the top enriched KEGG pathway terms included proteasome (hsa03050) [48], AGE–RAGE signaling pathway in diabetic complications (hsa04933) [49], pancreatic cancer (hsa05212) [50], prostate cancer (hsa05215) [51], proteoglycans in cancer (hsa05205) [52], and the PI3K–AKT signaling pathway (hsa04151) [53]. In particular, many cancer-related pathways were found to be enriched in human proteins identified from LTEs.

We then performed an enrichment analysis with a hypergeometric test by mapping the human proteins identified in LTEs to the druggable proteins acquired from the DrugBank database [43] and to cancer genes found in COSMIC [44]. As a result, 449 and 122 proteins identified in LTEs were found to match with drug targets and cancer genes, respectively (**Table 1**). Therefore, the dysregulated proteins identified in

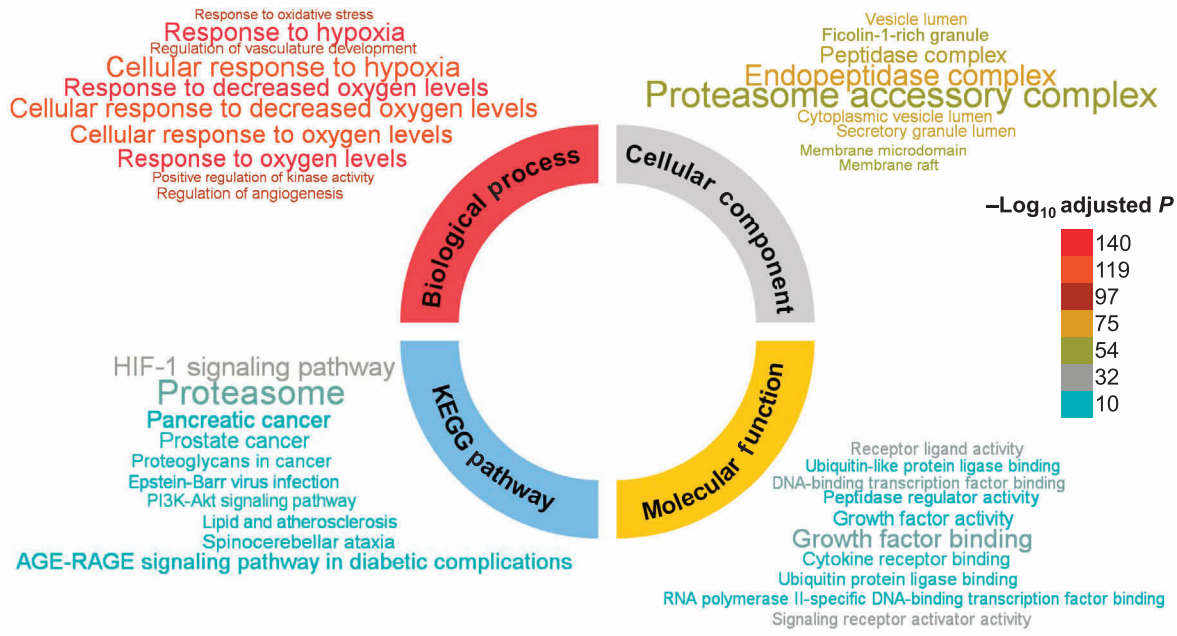


Figure 6 Enrichment of GO and KEGG terms for human LTE proteins
The results were visualized with WocEA (version 1.0). The font size denotes the enrichment ratio of each term, and the color denotes the approximate *P* adjust value. KEGG, Kyoto Encyclopedia of Genes and Genomes; WocEA, word cloud for the enrichment analysis.

Table 1 Enrichment analyses of drug targets and cancer genes for human LTE proteins

	Human proteins in LTEs (total count n = 1096)	Proteins in human proteome (total count n = 22,784)	E-ratio	*P* value
No. of matched drug proteins	449	2895	3.22	1.28E–129
No. of matched cancer genes	122	711	3.57	1.21E–35

Note: LTE, low-throughput experiment.

LTEs may be essential in tumorigenesis and tumor development and may be therapeutic targets.

PTMs play vital roles in regulating a wide range of biological processes. One of the remarkably enriched KEGG pathways from our analysis, the PI3K-AKT signaling pathway, is mainly regulated by phosphorylation. In the iHypoxia database, we provide abundant annotations derived from several PTM resources (Figure 1). We performed PTM analysis by mapping known PTM sites from six databases to proteins identified in LTEs. In total, we obtained 58,790 PTM sites on 1643 substrates, affecting 84.52% (1643/1944) of the proteins identified in LTEs (**Figure 7**A). More than 30 PTMs were found to be involved, and these included 39,224 phosphorylation sites of 1590 proteins, 12,777 ubiquitination sites of 1110 proteins, and 6662 acetylation sites of 942 proteins (Figure 7B; Table S3). Different PTMs can engage in crosstalk with each other to synergistically orchestrate specific biological processes, and we found that a fairly large number of the proteins identified in LTEs could be regulated by multiple PTMs (Figure 7C). Numerous studies have reported that PTMs can regulate relevant pathways to respond to hypoxia and also can be regulated by hypoxia [54–56].

Discussion

Oxygen is vital to life on Earth. However, hypoxia is a common stress frequently encountered by aerobic species, for example, in high altitudes, aquatic habitats, and underground burrows. In addition, cells/tissues within the body are also exposed to a hypoxic microenvironment. Therefore, mammals have evolved adaptive mechanisms to cope with hypoxia [2–5]. Hypoxia is also associated with human diseases such as cancers, cardiovascular diseases, and anemia [15,16,57]. The development of high-throughput technologies has greatly enhanced the large-scale identification of proteins whose expression at the transcript or protein levels is modified during hypoxia, and an increasing number of proteins involved in regulating hypoxia have been verified. Considering that the data from these studies are scattered throughout the literature and are difficult to obtain and utilize, a resource housing the expression dynamics of proteins in response to hypoxia is needed.

In this study, we developed iHypoxia, an integrated database for storing the expression dynamics of proteins that respond to hypoxia in animals. At present, iHypoxia contains

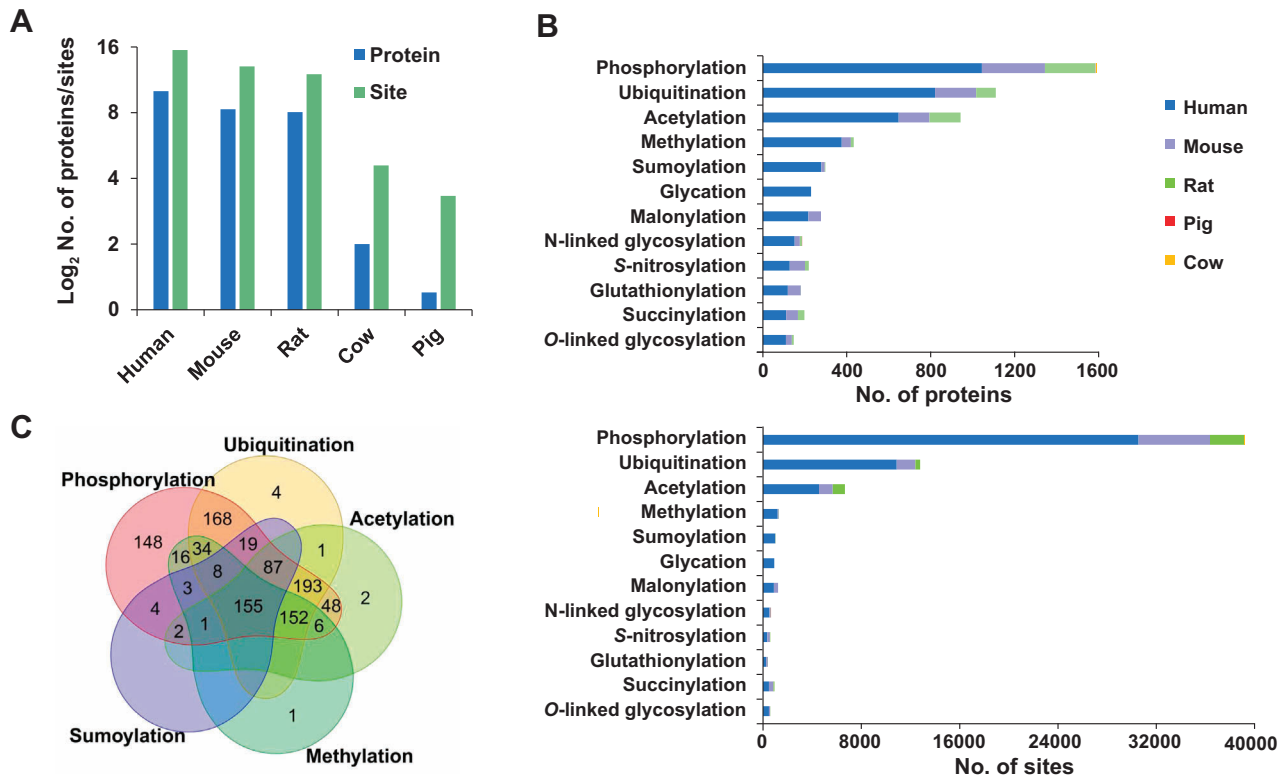


Figure 7 PTM analyses of proteins identified from LTEs

A. Distribution of modified proteins and sites. **B.** Distribution of the mapped proteins and sites in terms of phosphorylation, ubiquitination, acetylation, and others (sites with fewer than 100 sites are not shown). **C.** Overlap of five major types of PTMs of proteins identified in LTEs.

2589 expression events of 1944 proteins identified in LTEs, 422,553 quantitative expression events of 33,559 proteins identified in HTEs, 8788 candidate genes, and 394,141 computationally detected proteins. Compared with previous databases, iHypoxia not only contains more data but also has several advanced features. First, iHypoxia is the first integrated animal database focused on the expression dynamics of proteins in response to hypoxia and provides evidence from HTEs and LTEs at the transcript and protein levels for five mammals. Moreover, this resource is well annotated and can provide abundant information on collected proteins, including the hypoxic experimental conditions, expression patterns, sample types, and other information from the original studies, as well as PTMs, protein–protein interactions, drug–target relations, and other information from public databases. Furthermore, the iHypoxia website is convenient and user-friendly, and users can rapidly and accurately find the information of interest or download datasets.

In the future, iHypoxia will be regularly maintained and updated when new hypoxia data are published. Additionally, we plan to adopt computational methods such as text mining to establish an automated system that extracts LTE-related data from the literature. Furthermore, hypoxia not only transcriptionally and translationally activates genes to modulate cellular adaptations but also influences many other processes, such as protein PTMs and non-coding RNA expression [53,58,59]. Therefore, additional expression data and data derived from more species will be integrated to provide a more comprehensive and useful resource to facilitate related research.

Data availability

The data in iHypoxia database are available at http://ihypoxia.omicsbio.info and https://doi.org/10.6084/m9.figshare.18881372.

Competing interests

The authors have declared no competing interests.

CRediT authorship contribution statement

Ze-Xian Liu: Conceptualization, Methodology, Investigation, Writing – review & editing, Project administration, Funding acquisition. **Panqin Wang:** Data curation, Investigation, Writing – original draft. **Qingfeng Zhang:** Data curation, Software, Writing – original draft. **Shihua Li:** Data curation, Validation. **Yuxin Zhang:** Data curation, Validation. **Yutong Guo:** Data curation, Validation. **Chongchong Jia:** Data curation, Validation. **Tian Shao:** Data curation, Validation. **Lin Li:** Validation. **Han Cheng:** Conceptualization, Writing – review & editing,

Funding acquisition. **Zhenlong Wang:** Conceptualization, Writing – review & editing, Funding acquisition. All authors have read and approved the final manuscript.

Acknowledgments

This work was supported by grants from the National Key R&D Program of China (Grant No. 2021YFA1302100 to Ze-Xian Liu), the National Natural Science Foundation of China (Grant No. U2004152 to Zhenlong Wang, Grant Nos. 81972239 and 91953123 to Ze-Xian Liu), the Fostering Fund of Fundamental Research for Young Teachers of Zhengzhou University, China (Grant No. JC21343016 to Han Cheng), the Program for Guangdong Introducing Innovative and Entrepreneurial Teams, China (Grant No. 2017ZT07S096 to Ze-Xian Liu), and the Tip-Top Scientific and Technical Innovative Youth Talents of Guangdong Special Support Program, China (Grant No. 2019TQ05Y351 to Ze-Xian Liu).

Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.gpb.2022.12.001.

ORCID

ORCID 0000-0001-9698-0610 (Ze-Xian Liu)
ORCID 0000-0001-8003-7337 (Panqin Wang)
ORCID 0000-0001-5580-9999 (Qingfeng Zhang)
ORCID 0000-0002-5785-6171 (Shihua Li)
ORCID 0000-0003-0281-5409 (Yuxin Zhang)
ORCID 0000-0003-4423-9008 (Yutong Guo)
ORCID 0000-0001-9982-8155 (Chongchong Jia)
ORCID 0000-0002-1333-2878 (Tian Shao)
ORCID 0000-0002-7542-4049 (Lin Li)
ORCID 0000-0002-6684-317X (Han Cheng)
ORCID 0000-0001-5946-2806 (Zhenlong Wang)

References

[1] Lee P, Chandel NS, Simon MC. Cellular adaptation to hypoxia through hypoxia inducible factors and beyond. Nat Rev Mol Cell Biol 2020;21:268–83.
[2] Cernaro V, Coppolino G, Visconti L, Rivoli L, Lacquaniti A, Santoro D, et al. Erythropoiesis and chronic kidney disease-related anemia: from physiology to new therapeutic advancements. Med Res Rev 2019;39:427–60.
[3] Fuhrmann DC, Brüne B. Mitochondrial composition and function under the control of hypoxia. Redox Biol 2017;12:208–15.
[4] Pugh CW, Ratcliffe PJ. Regulation of angiogenesis by hypoxia: role of the HIF system. Nat Med 2003;9:677–84.
[5] Nakazawa MS, Keith B, Simon MC. Oxygen availability and metabolic adaptations. Nat Rev Cancer 2016;16:663–73.
[6] Wang GL, Jiang BH, Rue EA, Semenza GL. Hypoxia-inducible factor 1 is a basic-helix-loop-helix-PAS heterodimer regulated by cellular O_2 tension. Proc Natl Acad Sci U S A 1995;92:5510–4.
[7] Maxwell PH, Wiesener MS, Chang GW, Clifford SC, Vaux EC, Cockman ME, et al. The tumour suppressor protein VHL targets hypoxia-inducible factors for oxygen-dependent proteolysis. Nature 1999;399:271–5.
[8] Ivan M, Kondo K, Yang H, Kim W, Valiando J, Ohh M, et al. HIFalpha targeted for VHL-mediated destruction by proline hydroxylation: implications for O_2 sensing. Science 2001;292:464–8.
[9] Jaakkola P, Mole DR, Tian YM, Wilson MI, Gielbert J, Gaskell SJ, et al. Targeting of HIF-alpha to the von Hippel-Lindau ubiquitylation complex by O_2-regulated prolyl hydroxylation. Science 2001;292:468–72.
[10] Semenza GL, Nejfelt MK, Chi SM, Antonarakis SE. Hypoxia-inducible nuclear factors bind to an enhancer element located 3′ to the human erythropoietin gene. Proc Natl Acad Sci U S A 1991;88:5680–4.
[11] Ratcliffe PJ. Oxygen sensing and hypoxia signalling pathways in animals: the implications of physiology for cancer. J Physiol 2013;591:2027–42.
[12] Taylor CT, Colgan SP. Regulation of immunity and inflammation by hypoxia in immunological niches. Nat Rev Immunol 2017;17:774–85.
[13] Ducsay CA, Goyal R, Pearce WJ, Wilson S, Hu XQ, Zhang L. Gestational hypoxia and developmental plasticity. Physiol Rev 2018;98:1241–334.
[14] Jing X, Yang F, Shao C, Wei K, Xie M, Shen H, et al. Role of hypoxia in cancer therapy by regulating the tumor microenvironment. Mol Cancer 2019;18:157.
[15] Rowan SC, Keane MP, Gaine S, McLoughlin P. Hypoxic pulmonary hypertension in chronic lung diseases: novel vasoconstrictor pathways. Lancet Respir Med 2016;4:225–36.
[16] Machogu EM, Machado RF. How I treat hypoxia in adults with hemoglobinopathies and hemolytic disorders. Blood 2018;132:1770–80.
[17] Chen N, Hao C, Peng X, Lin H, Yin A, Hao L, et al. Roxadustat for anemia in patients with kidney disease not receiving dialysis. N Engl J Med 2019;381:1001–10.
[18] Khurana P, Sugadev R, Jain J, Singh SB. HypoxiaDB: a database of hypoxia-regulated proteins. Database 2013;2013:bat074.
[19] Rashid I, Nagpure NS, Srivastava P, Kumar R, Pathak AK, Singh M, et al. HRGFish: a database of hypoxia responsive genes in fishes. Sci Rep 2017;7:42346.
[20] Feng E, Jiang L. Peptidomic analysis of hippocampal tissue for explore leptin neuroprotective effect on the preterm ischemia-hypoxia brain damage model rats. Neuropharmacology 2020;162:107803.
[21] Pașca AM, Park JY, Shin HW, Qi Q, Revah O, Krasnoff R, et al. Human 3D cellular model of hypoxic brain injury of prematurity. Nat Med 2019;25:784–91.
[22] Du X, Zhang R, Ye S, Liu F, Jiang P, Yu X, et al. Alterations of human plasma proteome profile on adaptation to high-altitude hypobaric hypoxia. J Proteome Res 2019;18:2021–31.
[23] The Gene Ontology Consortium. The Gene Ontology Resource: 20 years and still GOing strong. Nucleic Acids Res 2019;47:D330–8.
[24] Jassal B, Matthews L, Viteri G, Gong C, Lorente P, Fabregat A, et al. The Reactome pathway knowledgebase. Nucleic Acids Res 2020;48:D498–503.
[25] Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular Signatures Database (MSigDB) hallmark gene set collection. Cell Syst 2015;1:417–25.
[26] Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K. KEGG: new perspectives on genomes, pathways, diseases and drugs. Nucleic Acids Res 2017;45:D353–61.
[27] Safran M, Rosen N, Twik M, BarShir R, Stein TI, Dahary D, et al. The GeneCards Suite. In: Abugessaisa I, Kasukawa T, editors. Practical guide to life science databases. Singapore: Springer Nature Singapore; 2021, p.27–56.
[28] Frézal J. Genatlas database, genes and development defects. C R Acad Sci III 1998;321:805–17.

[29] Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 2015;43:e47.
[30] Pundir S, Martin MJ, O'Donovan C, UniProt Consortium. UniProt tools. Curr Protoc Bioinformatics 2016;53:1.29.1–15.
[31] Altenhoff AM, Train CM, Gilbert KJ, Mediratta I, de Farias TM, Moi D, et al. OMA orthology in 2021: website overhaul, conserved isoforms, ancestral gene order and more. Nucleic Acids Res 2021;49:D373–9.
[32] UniProt Consortium. UniProt: a worldwide hub of protein knowledge. Nucleic Acids Res 2019;47:D506–15.
[33] Oughtred R, Stark C, Breitkreutz BJ, Rust J, Boucher L, Chang C, et al. The BioGRID interaction database: 2019 update. Nucleic Acids Res 2019;47:D529–41.
[34] Huang KY, Lee TY, Kao HJ, Ma CT, Lee CC, Lin TH, et al. dbPTM in 2019: exploring disease association and cross-talk of post-translational modifications. Nucleic Acids Res 2019;47: D298–308.
[35] Lin S, Wang C, Zhou J, Shi Y, Ruan C, Tu Y, et al. EPSD: a well-annotated data resource of protein phosphorylation sites in eukaryotes. Brief Bioinform 2021;22:298–307.
[36] Kreegipuu A, Blom N, Brunak S. PhosphoBase, a database of phosphorylation sites: release 2.0. Nucleic Acids Res 1999;27:237–9.
[37] Hornbeck PV, Kornhauser JM, Latham V, Murray B, Nandhikonda V, Nord A, et al. 15 years of PhosphoSitePlus®: integrating post-translationally modified sites, disease variants and isoforms. Nucleic Acids Res 2019;47:D433–41.
[38] Xu H, Zhou J, Lin S, Deng W, Zhang Y, Xue Y. PLMD: an updated data resource of protein lysine modifications. J Genet Genomics 2017;44:243–50.
[39] Huttlin EL, Bruckner RJ, Paulo JA, Cannon JR, Ting L, Baltier K, et al. Architecture of the human interactome defines protein communities and disease networks. Nature 2017;545:505–9.
[40] Kotlyar M, Pastrello C, Malik Z, Jurisica I. IID 2018 update: context-specific physical protein-protein interactions in human, model organisms and domesticated species. Nucleic Acids Res 2019;47:D581–9.
[41] Orchard S, Ammari M, Aranda B, Breuza L, Briganti L, Broackes-Carter F, et al. The MIntAct project – IntAct as a common curation platform for 11 molecular interaction databases. Nucleic Acids Res 2014;42:D358–63.
[42] Luck K, Kim DK, Lambourne L, Spirohn K, Begg BE, Bian W, et al. A reference map of the human binary protein interactome. Nature 2020;580:402–8.
[43] Wishart DS, Feunang YD, Guo AC, Lo EJ, Marcu A, Grant JR, et al. DrugBank 5.0: a major update to the DrugBank database for 2018. Nucleic Acids Res 2018;46:D1074–82.
[44] Tate JG, Bamford S, Jubb HC, Sondka Z, Beare DM, Bindal N, et al. COSMIC: the Catalogue Of Somatic Mutations In Cancer. Nucleic Acids Res 2019;47:D941–7.
[45] Smith KA, Waypa GB, Schumacker PT. Redox signaling during hypoxia in mammalian cells. Redox Biol 2017;13:228–34.
[46] Krock BL, Skuli N, Simon MC. Hypoxia-induced angiogenesis: good and evil. Genes Cancer 2011;2:1117–33.
[47] Chen L, Endler A, Shibasaki F. Hypoxia and angiogenesis: regulation of hypoxia-inducible factors via novel binding factors. Exp Mol Med 2009;41:849–57.
[48] Rahimi N. The ubiquitin-proteasome system meets angiogenesis. Mol Cancer Ther 2012;11:538–48.
[49] Gunton JE. Hypoxia-inducible factors and diabetes. J Clin Invest 2020;130:5063–73.
[50] Liu M, Zhong J, Zeng Z, Huang K, Ye Z, Deng S, et al. Hypoxia-induced feedback of HIF-1α and lncRNA-CF129 contributes to pancreatic cancer progression through stabilization of p53 protein. Theranostics 2019;9:4795–810.
[51] Marignol L, Rivera-Figueroa K, Lynch T, Hollywood D. Hypoxia, notch signalling, and prostate cancer. Nat Rev Urol 2013;10:405–13.
[52] Ampofo E, Schmitt BM, Menger MD, Laschke MW. The regulatory mechanisms of NG2/CSPG4 expression. Cell Mol Biol Lett 2017;22:4.
[53] Bossler F, Hoppe-Seyler K, Hoppe-Seyler F. PI3K/AKT/mTOR signaling regulates the virus/host cell crosstalk in HPV-positive cervical cancer cells. Int J Mol Sci 2019;20:2188.
[54] Song H, Feng X, Zhang M, Jin X, Xu X, Wang L, et al. Crosstalk between lysine methylation and phosphorylation of *ATG16L1* dictates the apoptosis of hypoxia/reoxygenation-induced cardiomyocytes. Autophagy 2018;14:825–44.
[55] Mennerich D, Kubaichuk K, Kietzmann T. Dubs, hypoxia, and cancer. Trends Cancer 2019;5:632–53.
[56] Adav SS, Sze SK. Hypoxia-induced degenerative protein modifications associated with aging and age-associated disorders. Aging Dis 2020;11:341–64.
[57] Rankin EB, Giaccia AJ. Hypoxic control of metastasis. Science 2016;352:175–80.
[58] Petrushanko IY, Mitkevich VA, Lakunina VA, Anashkina AA, Spirin PV, Rubtsov PM, et al. Cysteine residues 244 and 458–459 within the catalytic subunit of Na, K-ATPase control the enzyme's hydrolytic and signaling function under hypoxic conditions. Redox Biol 2017;13:310–9.
[59] Xue M, Chen W, Xiang A, Wang R, Chen H, Pan J, et al. Hypoxic exosomes facilitate bladder tumor growth and development through transferring long non-coding RNA-UCA1. Mol Cancer 2017;16:143.