Received: 27 June 2018 | Revised: 20 August 2018 | Accepted: 25 October 2018

DOI: 10.1002/ame2.12044

ORIGINAL ARTICLE

LasDB: A collective database for laboratory animal strain resources

Qi Kong[1,2,3,4] | Chuan Qin[1,2,3,4]

[1]Institute of Laboratory Animal Science, Chinese Academy of Medical Sciences (CAMS) and Comparative Medicine Center, Peking Union Medical College (PUMC), Beijing, China

[2]NHC Key Laboratory of Human Disease Comparative Medicine, Key Laboratory of Human Diseases Animal Model, State Administration of Traditional Chinese Medicine, Beijing, China

[3]Beijing Key Laboratory for Animal Models of Emerging and Reemerging Infectious Diseases, Beijing, China

[4]Beijing Engineering Research Center for Experimental Animal Models of Human Critical Diseases, Beijing, China

Correspondence
Qi Kong, Institute of Laboratory Animal Science, Chinese Academy of Medical Sciences & Comparative Medical Center, Peking Union Medical College, Beijing, China.
Email: infor@cnilas.org

Funding information
the Central Research Institutes Basic Operating Grants, Grant/Award Number: DWS201512; CAMS Innovation Fund for Medical Sciences (CIFMS), Grant/Award Number: 2016-I2M-2-006-03; National Major Scientific and Technological Special Project for Key Infectious Diseases, Grant/Award Number: 2017ZX10304402-001

Abstract

Background: With the aim of establishing the most comprehensive database of laboratory animal strains, the "laboratory animal strain resources database" (LasDB) was constructed as a searchable online database of all laboratory animal strains, stocks and mutant embryonic stem-cell lines available worldwide, including inbred, outbred, mutant and genetically engineered strains.

Methods: *MySQL* database software was used to construct the LasDB, offering an easy-to-use interface.

Results: To date, LasDB has a collection covering data for 21 596 mouse strains, 2062 rat strains, 13 monkey strains, 2 hamster strains, 5 dog strains, 5 rabbit strains and more than 50 other laboratory animal strains. LasDB will be continually improved with regular updates of new laboratory animal strains from all over the world.

Conclusion: To the best of our knowledge, this is the first database that attempts to systematically integrate all available laboratory animal strain data with the aim of supporting open usage and full resource sharing.

KEYWORDS
comparative medicine, database, laboratory animal, resource, strain

1 | INTRODUCTION

Laboratory animals are commonly used in biomedical research to understand disease mechanisms and develop potential therapies. Because of this, a large and ever-increasing number of animal strains have been developed. The construction of a laboratory animal strain resource database could effectively promote further biomedical research and assist new developments in bio-pharmaceutical industries.[1,2]

The present work reports the establishment of a laboratory animal strain database, named LasDB, which was constructed by collecting all the available laboratory animal strain data from China and overseas. The laboratory animal strains collected from overseas databases or networks are described with a brief introduction and links

or references to the original owners. Our aim was to build the largest and most complete database of laboratory animal species and strains, which would be an invaluable resource for researchers or users in the fields of science, biology, medicine, agriculture, etc.

In this database, we have collected more than 2000 laboratory animal species or strains made or used in China and also some published data that are not included in foreign databases, such as the International Mouse Strain Resource (IMSR), which is limited to mouse strains and does not include other laboratory animal species and contains little data from China.

2 | MATERIAL AND METHODS

The database was constructed following three major steps: (a) data collection (including a data survey of peer researchers in China, a web search, database mining, a literature search, the records of the Institute of Laboratory Animal Science, Chinese Academy of Medical Sciences (ILAS), and author donations), (b) data annotation and analysis, and (c) database construction of a *MySQL* relational database with a web interface for data storage and usage (Figure 1).

2.1 | Data collection

2.1.1 | Data survey

In order to collect data from other researchers and peers in China, we asked 2000 Chinese laboratory animal units to supply their laboratory animal strains. The results were incorporated in the metadata design.

2.1.2 | Data sources

In addition to the comprehensive data survey, we also collected data from websites, databases, and published papers and ILAS's own records, etc. We searched the strain's name and reviewed the interpretation in the metadata.

Web search

Using the web engine Baidu (http://www.baidu.com) to retrieve information, we searched various relevant laboratory animal websites.

Database mining

We retrieved laboratory animal strain information from the United States, Europe, Japan and other countries having laboratory animal strain-associated databases, and extracted the strain names along with related information in the metadata.

Literature search

We searched the China National Knowledge Infrastructure (CNKI) database (http://www.cnki.net), Wanfang database (http://www.wanfangdata.com.cn) and PubMed (http://www.ncbi.nlm.nih.gov/pubmed/) and captured laboratory animal strain information from all the published papers.

ILAS's own records

The Institute of Laboratory Animal Science (ILAS), Chinese Academy of Medical Sciences (CAMS) has more than 1000 laboratory animal strains, including genetically engineered animals, which are not included in other databases.

Author donations

A data entry window for submissions from our domestic and foreign counterparts was provided. Information was filled in using the online form or emailed to the administrator of the database.

2.2 | Data annotation and analysis

We retrieved the names of laboratory animal strains from the above data sources and reviewed the interpretation in the metadata from

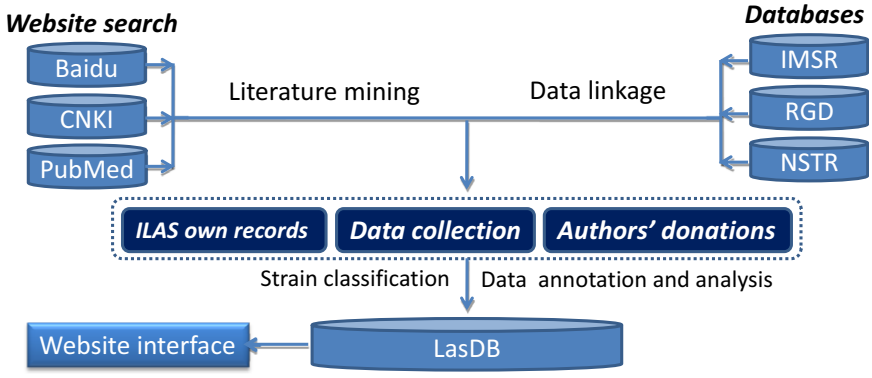

FIGURE 1 Database construction method. For website searches, the web engine Baidu (http://www.baidu.com) was used to retrieve information, and we searched the China National Knowledge Infrastructure (CNKI) database (http://www.cnki.net), and PubMed (http://www.ncbi.nlm.nih.gov/pubmed/) for published Papers. The website interface is linked to ILAS and CALAS as referred in the text. IMSR, International Mouse Strain Resource; RGD, Rat Genome Database; NSTR, National Natural Science & Technology Resources; CNKI, China National Knowledge Infrastructure; ILAS, Institute of Laboratory Animal Science, Chinese Academy of Medical Sciences

these databases, published papers, online websites or author donations by data annotation and analysis.

2.3 | Database construction

The database was constructed using the *MySQL* database software, which offers an easy-to-use interface containing search facilities, and to this we added a tailored input interface.

3 | RESULTS

3.1 | Metadata table

We developed a metadata table, including information about the data format, content, etc., as shown in Table 1.

3.2 | Naming of the laboratory animal strains

The laboratory animal strains in the database were coded as accession numbers, using a six-code format, eg, 31-11-001-0-00-0001. The meaning of the six codes joined by a hyphen (-) is as follows, based on the example code: 31 represents the experimental materials; 11 represents the laboratory animal species; 001 is the code of the laboratory animal species as shown in Table 2; 0, A or Z: is the first number or character of the laboratory animal strain; 00 is the subspecies number; 0001 is the line number.[3-7] Table 2 shows the laboratory animal strains included in the database at present. The database includes all laboratory animal strains such as inbred strains outbred strains and GM mutants.

3.3 | Database name

Laboratory animal strains database (LasDB).

3.4 | Database URL

http://www.cnilas.org/plus/list.php?tid=158.

3.5 | Database usage

3.5.1 | Website interface

There is one quick search method for performing a search of the laboratory animal strains held in the LasDB. Users can search the

TABLE 1 Metadata table of the laboratory animal strain resources database (LasDB)

Field name	Field contents	Header	Optional field
Serial number of laboratory animal strains	The number of lines in the database, according to the data entry order, is automatically formed	SN_AS	No
Accession number of laboratory animal strains	The number of numbered lines in the whole laboratory animal strains, numbered according to our code described in Section 3.2	AC_ AS	No
Name of laboratory animal strains unique identification	The laboratory animal strain's Chinese name, the formal name of the developed strain in accordance with Chinese naming rules	UID	Yes
English name	The laboratory animal strain's name in English, in accordance with international naming rules	EN	Yes
Name_Other	Other common names or aliases	NO	Yes
Description of species	Species described according to the principle of the strains	DS	Yes
Generation	How many generations at present when collected by the database	Generation	Yes
Appearance	Coat colour and other exterior features, can be furnished with a photo or illustration	Appearance	Yes
Genetics status	Genetic status of the strain	GS	Yes
Origin	Lists origin as abroad or domestic cultivation	Origin	Yes
Other related substrains	The next of kin or branch lines	ORS	Yes
Biology status	Unique biological characteristics of the strains, physiology, biochemistry, blood factors, immune status, anatomy, eg, average life expectancy	BS	Yes
Lifespan and disease	Natural life and spontaneous disease, pathology, cancer prevalence rates.	LD	Yes
Research field	Use of a range of applications	RA	Yes
Breeding and reproduction	Breeding characteristics and considerations, breeding conditions, feed, habits	BR	Yes
Institute of colony maintenance	Can provide provenance or sell units	Inst_CM	Yes
Reference	Reference literatures or websites	REF	Yes
Input time	Time of data entry or update in the database can be automatically generated	Input_Time	No
Remarks	Description of other matters	RM	Yes

TABLE 2 Information on the laboratory animal species collected in LasDB

No	Name	Included strains	Total strains	No	Name	Included strains	Total strains
001	Mouse	21 800	26 000	040	Chicken	1	5
002	Rat	2200	2300	041	Duck	1	2
003	Hamster	6	38	042	Goose	1	1
004	Guinea pig	3	12	043	Pigeon	1	1
005	Rabbit	3	8	044	Quail	1	20
006	Dog	1	5	045	Pika	1	1
007	Rhesus monkey	1	1	046	Dasyure	1	1
008	Cynomolgus monkeys	1	1	047	Rainbow trout	1	1
009	Marmoset	1	1	048	Deer	1	1
010	Squirrel monkey	1	1	049	Cow	1	10
011	Green monkey	1	1	050	Sheep	2	10
012	Cotton-top tamarin	1	1	051	Horse	1	1
013	Titi monkey	1	1	052	Ass	1	1
014	Japanese snow monkey	1	1	053	Mule	1	1
015	Vervet monkey	1	1	054	Drosophila	1	5
016	Taiwan monkey	1	1	055	C. elegans	1	11
017	Sooty mangabey	1	5	056	Aplysia	1	1
018	E. patas	1	1	057	Thermophila	1	1
019	Pigtailed macaque	1	1	058	Silkworm	1	1
020	M.arctoides	1	1	059	Kangaroo	1	1
021	M.leonine	1	1	060	Hedgehog	1	1
022	M.thibetana	1	1	061	Armadillo	1	1
023	Cebus capucinus	1	1	062	Bat	1	1
024	M.assamensis	1	1	063	Toad	1	1
025	Leaf monkey	1	1	064	Frog	1	1
026	Baboon	1	5	065	Earthworm	1	1
027	Gibbon	1	5	066	Shrew	1	1
028	Chimpanzee	1	10	067	Porpoise	1	1
029	Tree shrew	1	20	068	Goldfish	1	1
030	Hamster	1	3	069	Zebra fish	1	20
031	Vole	2	2	070	Xiphophorus	1	3
032	Gerbil	2	2	071	Medaka	1	1
033	Marmot	1	1	072	Gobiocypris rarus	1	1
034	Peromyscus	1	1	073	Red carp	1	1
035	Cotton rat	1	2	074	Electric eel	1	1
036	Cat	1	2	075	Crucian carp	1	1
037	Pig	5	12	…	Other	0	130
038	Ferret	1	5				
039	Chinchilla	1	2				

"Included strains" means the strains included in LasDB at present. "Total strains" means the strains published as we planned to be included.

English name of any laboratory animal strains or any other keywords by quick search. At present, the database is only available in Chinese, and the data can be translated by Google (http://translate.google.com/) tools or some web browser. We will try to make an English Version of the database in the future.

3.5.2 | Networking linkages

The database is affiliated with the website of the Institute of Laboratory Animal Science (ILAS), the Chinese Academy of Medical Sciences (CAMS) (http://www.cnilas.org), and the website of the

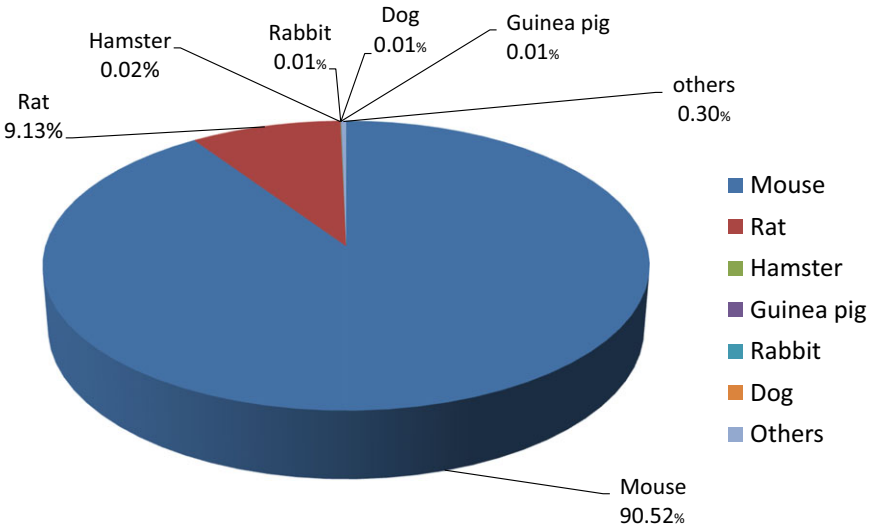

FIGURE 2 Overview of the laboratory animal strains in LasDB. The percentage breakdown of laboratory animals from each species in LasDB at present

Chinese Association for Laboratory Animal Sciences (CALAS) (http://www.calas.org.cn), which offers a free search service.

3.5.3 | Providing data updates

To date, LasDB has a collection covering data for 21 596 mouse strains, 2062 rat strains, 13 monkey strains, 2 hamster strains, 5 dog strains, 5 rabbit strains and more than 50 other laboratory animal strains (Figure 2). The database is open, and professionals can contributed the data of laboratory animal species or strains that are not already included in the database, although they will not be uploaded as an official data release on to the database until they have first been reviewed and verified. As such, the size of the database can be expected to continue to increase over time.

4 | DISCUSSION

Laboratory animal resources are increasingly important for life science research. In this work, we describe the construction of a laboratory animal strain database (LasDB). The data resources in the LasDB are free for biomedical-related scientific and technical personnel to use, and it is envisaged will particularly be vital in promoting and implementing bio-industry research.

A large number of laboratory animal resources have been collected from China and abroad, in order to facilitate information sharing among the scientific community. Examples of resources include a series of databases for mice strains constructed by The Jackson Laboratory in the United States, the *Rat Genome Database* (RGD) constructed by The National Heart Lung and Blood Institute (NHLBI) in the USA, which gathers information together on rat strains,[8] and the *Rodent Genome Database*, *Mouse Cells Genetic Map* and *Freak Mouse Homology Database* built in the UK. Some other databases storing the genetic information of laboratory animals like mice, rats, and related disease models have also been constructed, and are important resources for biomedical systems and related research.

The sharing of laboratory animal information is a global trend. Several information-sharing-based non-profit coalition laboratory animal resources have been formed, mainly related to laboratory rodents, including the International Mouse Strain Resource (IMSR), the Federation of International Mouse Resources (FIMRe) and the International Knockout Mouse Consortium (IKMC),[9] but they are limited to mouse strains, while other species are also useful and commonly used. Researchers are not finding it easy to get information about other species. The LasDB aims the meet the requirement of including all laboratory animal species and strains.

In 1997, the Chinese government started moves to establish a laboratory animal information network and databases. To date, we have established a national laboratory animal statistical reporting system, including the construction of a database on laboratory animal production, supply, information exchange networks and related databases. In 2001-2003 we built an animal model of human diseases database, holding about 100 data samples. In 2008, the National Natural Science & Technology Resources (NSTR) set up a laboratory animal resource database, with a data collection of 188 strains from eight species. The Model Animal Research Institute and Nanjing University together set up the China Resource Sharing Alliance, which mainly provides a genetically engineered mouse strains "one-stop" inquiry and communication sharing platform for alliance members. With the development of its economy and science base, China has become one of the biggest communities using laboratory animals, with 25 million laboratory animals currently produced and with more than 200 000 professionals in the field of laboratory animal science.[2] The need for laboratory animal strain resources is growing, and this is the greatest stimulus for the construction of LasDB.

In China, laboratory animal databases generally contain information about the biological characteristics of the domestic laboratory animal strains, but provide little information on practical applications. A global database though could include all laboratory animal strain resources. Consequently, construction of a worldwide database of laboratory animal strains is essential for domestic researchers and will help them make more efficient use of the existing laboratory animal resources. The LasDB is now open to researchers and also provides consulting services for the public.

ACKNOWLEDGEMENTS

We thank Oxford Language Editing Group for providing language support. This work was supported by the Central Research Institutes Basic Operating Grants (DWS201512), the National Major Scientific and Technological Special Project for Key Infectious Diseases (2017ZX10304402-001) and a CAMS Innovation Fund for Medical Sciences (CIFMS) grant (2016-I2M-2-006-03).

CONFLICT OF INTEREST

None.

AUTHOR CONTRIBUTIONS

QK contributed to the database design, collection of the data-set, and the analysis of the data and in the editing of the manuscript. CQ also contributed to the database design. Both the authors read and approved the final manuscript.

ORCID

Qi Kong http://orcid.org/0000-0003-2867-7382

REFERENCES

1. Kong Q, Qin C. Analysis of current laboratory animal science policies and administration in China. *ILAR J*. 2009;51:e1-e11.
2. Kong Q, Qin C. Laboratory animal science in China: current status and potential for the adoption of alternatives. *Altern Lab Anim*. 2010;38:53-69.
3. Shimoyama M, De Pons J, Hayman GT, et al. The Rat Genome Database 2015: genomic, phenotypic and environmental variations and disease. *Nucleic Acids Res*. 2015;43:D743-D750.
4. FELASA Working Group, Rülicke T, Montagutelli X, Pintado B, Thon R, Hedrich HJ. FELASA guidelines for the production and nomenclature of transgenic rodents. *Lab Anim*. 2007;41:301-311.
5. Davisson MT. Rules and guidelines for genetic nomenclature in mice: excerpted version. Committee on Standardized Genetic Nomenclature for Mice. *Transgenic Res*. 1997; 6:309-319.
6. Maltais LJ, Blake JA, Eppig JT, Davisson MT. Rules and guidelines for mouse gene nomenclature: a condensed version. International Committee on Standardized Genetic Nomenclature for Mice. *Genomics*. 1997;45:471-476.
7. Montoliu L, Whitelaw CB. Using standard nomenclature to adequately name transgenes, knockout gene alleles and any mutation associated to a genetically modified mouse strain. *Transgenic Res*. 2011;20:435-440.
8. Wang SJ, Laulederkind SJ, Hayman GT, et al. Analysis of disease-associated objects at the Rat Genome Database. *Database (Oxford)*. 2013; 2013:bat046.
9. Marks C. Mouse models of human cancers consortium (MMHCC) from the NCI. *Dis Model Mech*. 2009;2:111.

How to cite this article: Kong Q, Qin C. LasDB: A collective database for laboratory animal strain resources. *Animal Model Exp Med*. 2018;1:266–271. https://doi.org/10.1002/ame2.12044