China National Center for Bioinformation

从NCBI到CNCB

鲍一明

中国科学院北京基因组研究所
(国家生物信息中心)

PKU/CAAS Course
Oct. 3, 2020 • Online

中国科学院北京基因组研究所
BEIJING INSTITUTE OF GENOMICS CHINESE ACADEMY OF SCIENCES

鲍一明 简介

baoym@big.ac.cn

中国科学院北京基因组研究所 二级研究员

国家生物信息中心 副主任

国家基因组科学数据中心 主任

中国科学院“百人计划”研究员

中国科学院大学健康医疗大数据国家研究院 副院长

1987年获北京大学生物化学学士学位

1994年获英国John Innes Centre (东英大学）遗传学博士学位

1994-2001年，美国Noble Foundation博士后，助理研究员

2001-2017年，美国国家生物信息中心（NCBI）Staff Scientist

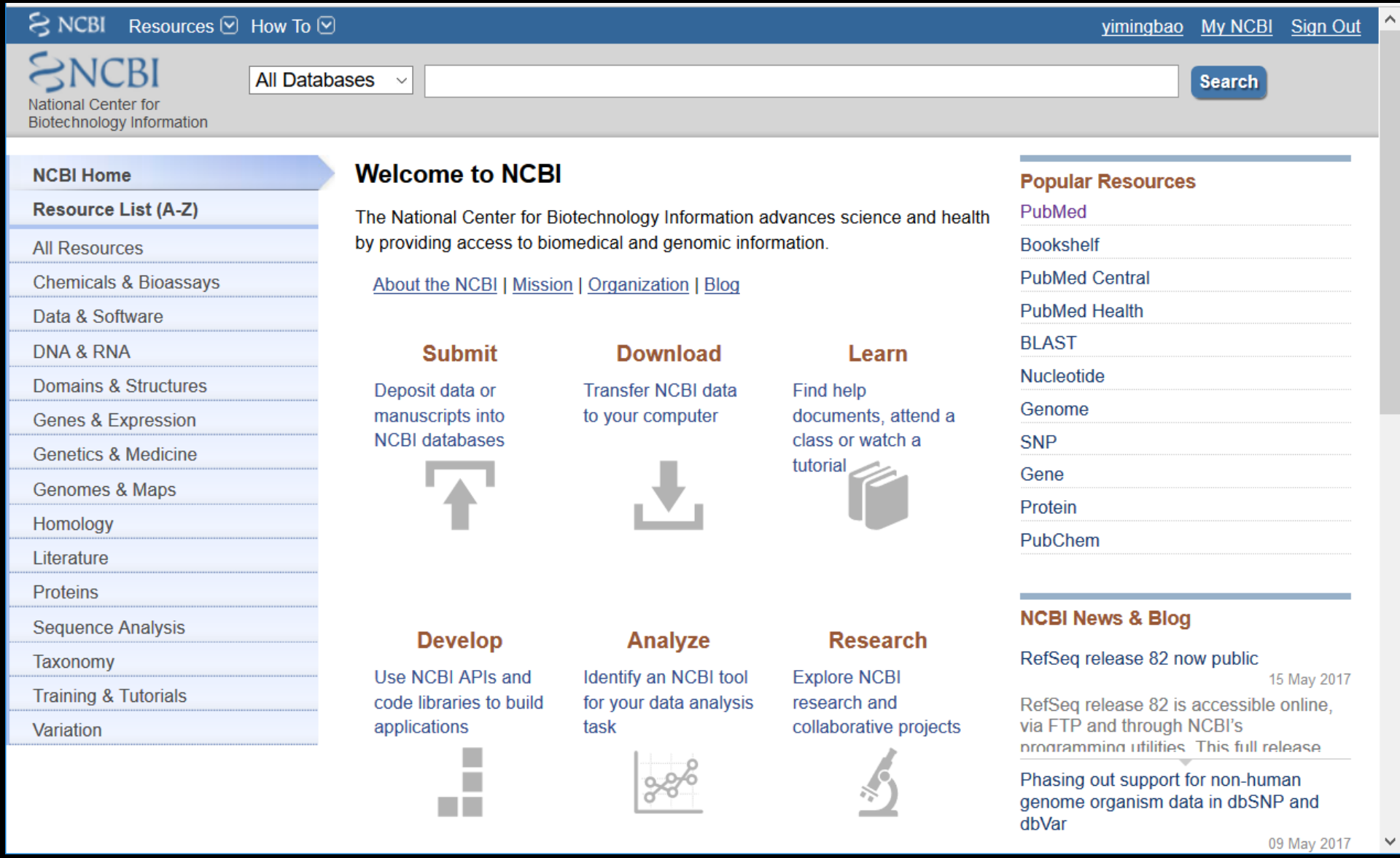
2019-至今，国家基因组科学数据中心 主任

2020-至今，国家生物信息中心 副主任

课题组主要研究方向是生物信息数据库（基因组和表观组），大数据整合与挖掘，病毒基因组注释、病毒进化与分类

NCBI

NCBI, https://www.ncbi.nlm.nih.gov/



National Center for Biotechnology Information (NCBI)

Plant Physiol. (1997) 113: 1015–1022

Update on Bioinformatics

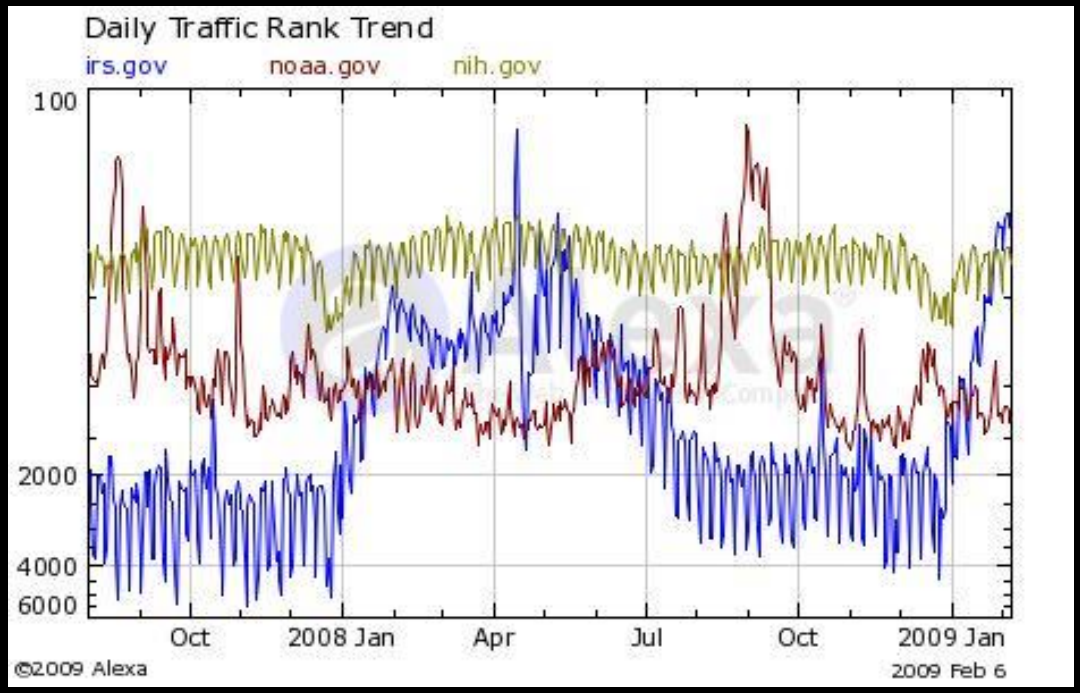
Plant Biology in the Post-Gutenberg Era

Everything You Wanted to Know and More on the World Wide Web

Chris Somerville*, David Flanders, and J. Michael Cherry

is to recover the sequence. This is accomplished by a search of the NCBI at http://www.ncbi.nlm.nih.gov/. This is the mother of all biological databases and, in the eyes of many researchers, is considered to be the single most useful activity of the Federal Government. One of the most useful features is the ENTREZ suite, which allows retrieval of citations and records in the area of molecular biology from

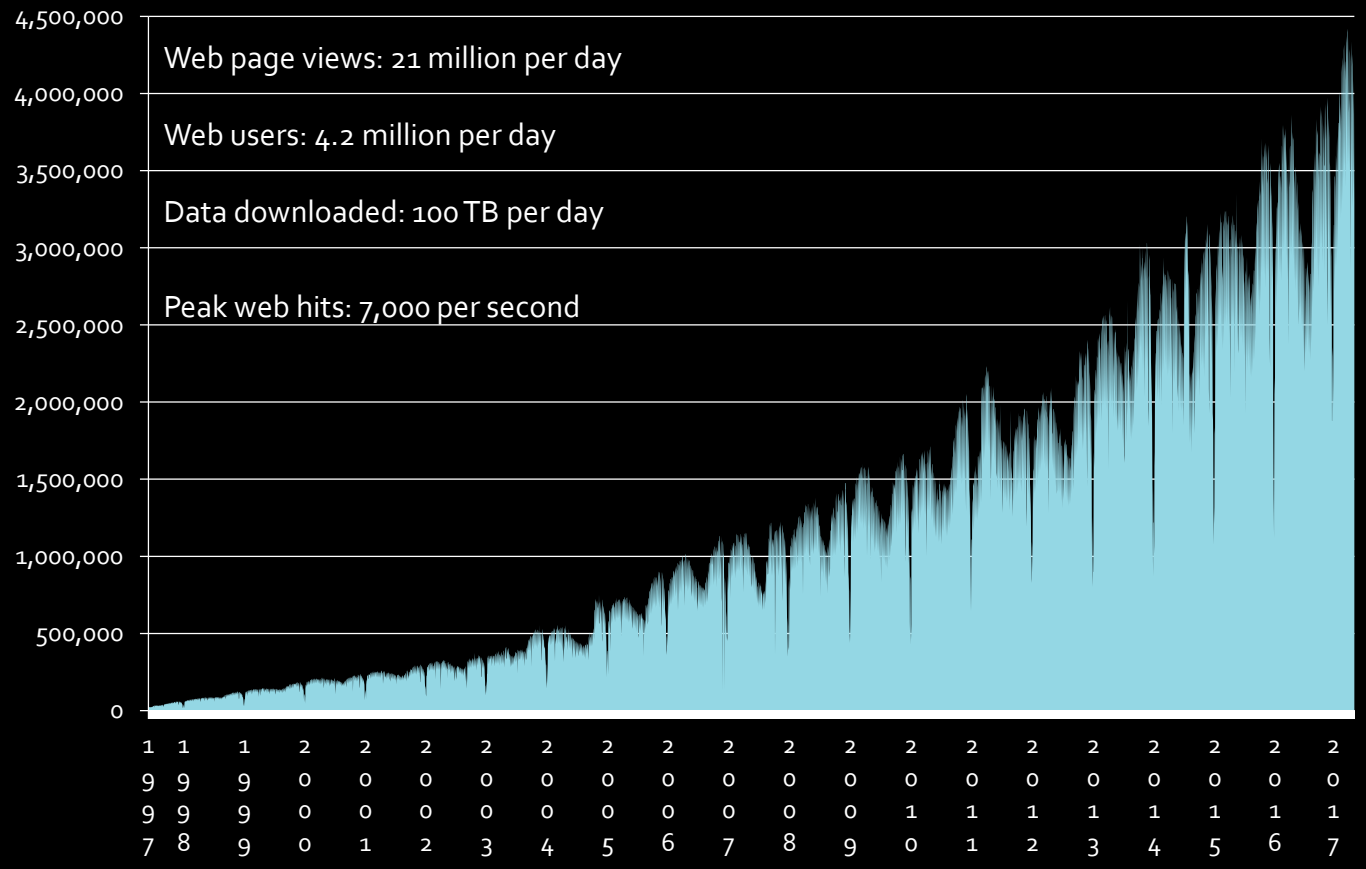
NCBI website ranks



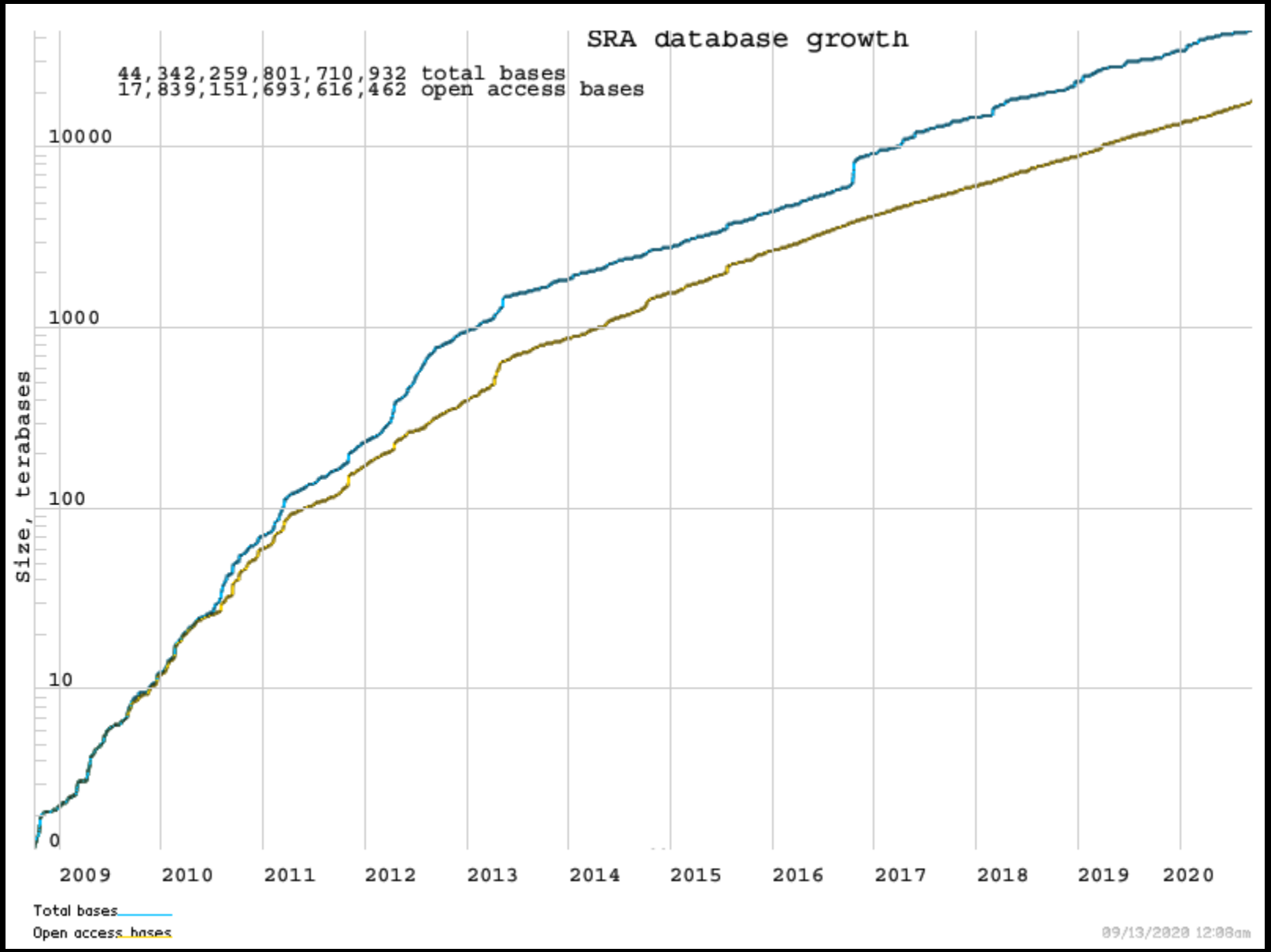
Source:
www.alexa.com

NCBI

NCBI Daily Users



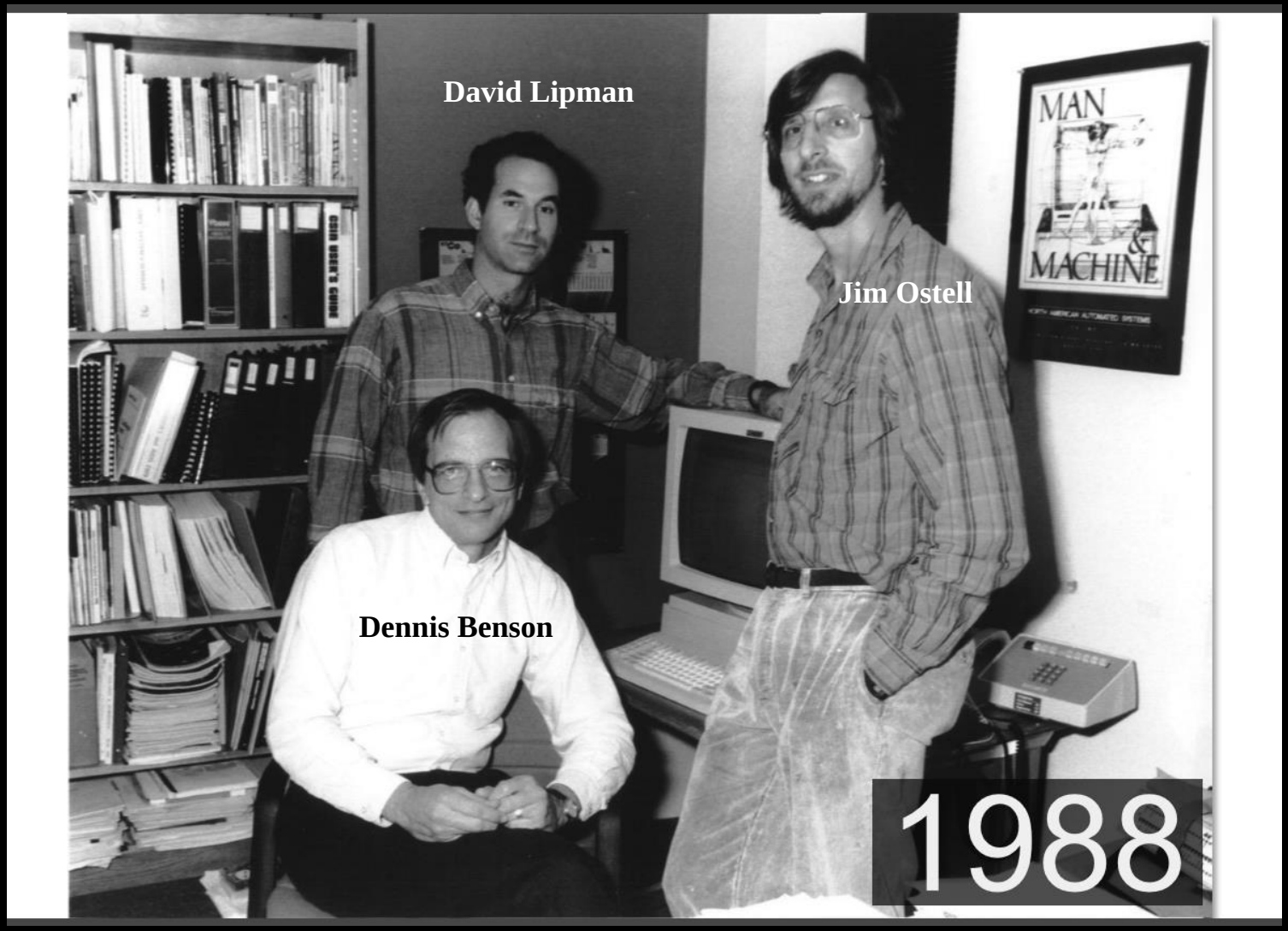
NCBI


SRA database growth
44,342,259,801,710,932 total bases
17,839,151,693,616,462 open access bases
Size, terabases
10000
1000
100
10
0
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
Total bases
Open access bases
09/13/2020 12:08am

NCBI

NCBI成功的秘诀

- 团队



NCBI

NCBI

NCBI概况

- 总人数：～700
- 组织结构：
  - 科学顾问委员会 (BOSC)，发展决策
  - 计算生物学部门 (CBB)，25%，基础研究
  - 信息工程部门 (IEB)，60%，产品实施
  - 信息资源部门 (IRB)，15%，技术支持
- 主要数据库及软件
  - 文献（PubMed, PubMed Central）
  - 一级数据库 （Nucleotide, SRA, PubChem, taxonomy等）
  - 二级数据库（RefSeq, Gene, Influenza Virus Resource等）
  - BLAST
  - E-utilities
  - ORF Finder

NCBI

NCBI成功的秘诀

- 团队
- 政策

Congressional Record

United States of America

PROCEEDINGS AND DEBATES OF THE 113th CONGRESS, FIRST SESSION

Vol. 159 WASHINGTON, MONDAY, OCTOBER 28, 2013 No. 151

Senate

NATIONAL CENTER FOR BIOTECHNOLOGY INFORMATION ANNIVERSARY

Mr. CARDIN. Mr. President, I rise to recognize the 25th anniversary of the National Center for Biotechnology Information—NCBI—part of the National Library of Medicine at the National Institutes of Health, America's world-renowned research institution in Bethesda, MD.

[illegible] build [illegible] separate laws creating individual institutes at NIH. In 1987, while a Member of the House of Representatives and chairman of the Select Committee on Aging's Subcommittee on Health and Long-Term Care, Pepper introduced H.R. 393, the National Biotechnology Information Act, which established the NCBI. At a March 1987 hearing on H.R. 393, Pepper explained that "we are dealing with nothing less than the mystery of human life and the unfolding scroll of knowledge, seeking to penetrate that mystery, which is life itself." He noted that his bill was intended "to facilitate the development of advanced computer and communication systems that will make it possible for the vast expanding knowledge of the gene to be assimilated into a computer system and made available for distribution to researchers and to people generally all over the World."

Soon thereafter, Congress embraced the importance of the biotechnology field, recognized the pressing need to harness the large volume of data emanating from the genetic revolution in science, and endorsed the establishment of NCBI to manage this valuable resource for the benefit of human health. With strong bipartisan support in Congress, Pepper's bill was enacted as part of Senator Ted Kennedy's comprehensive measure, the Health Omnibus Extension Programs of 1988, on November 4, 1988.

Today, biomedical research encompassing genomic and genetic knowledge is a major driver of medical progress. The foresight of Congress in establishing the NCBI, combined with the innovative leadership of Director Dr. David Lipman and the expertise of the agency's dedicated staff, has led to the emergence of an impressive national resource for molecular biology information. In June of this year, Dr. Lipman was [illegible] the White House Office of Science and Technology Policy's Advisory Committee on Creative Government. Dr. Lipman was awarded the Vannevar Bush Award for his work at NCBI. By organizing and integrating genomic data for developing diagnostic and clinical applications, the Center serves as a bridge from research to the medical community. Each day, more than 3 million users access NCBI's 40 interlinked genomic and bibliographic databases and download more than 30 terabytes of data.

I am proud that Congress has continued to support funding for the NCBI over the past 25 years. Recently, by requiring that the results of NIH-funded research be made public through the Center's PubMed Central Database, Congress has opened to everyone the full text of published journal articles that are essential to advancing scientific research and public health.

The biomedical research funded by the NIH provides knowledge essential to combat debilitating diseases, and continuing this research is dependent on the resources and tools that NCBI has developed so successfully for the benefit of the biomedical community. As NIH Director Francis Collins has noted, we are entering an era of precision medicine in which a patient's genetic makeup may determine the exact treatment that is provided. Surely, the NCBI databases and tools will be needed on the front lines of this new effort.

On the occasion of this 25th anniversary, I ask my colleagues to join me in congratulating Dr. Lipman and the outstanding staff of NCBI, who through their skill and vision have built this unique biomedical resource.

1988年，美国国会议员成立NCBI

1997年，美国副总统戈尔启动PubMed

NCBI News

NATIONAL CENTER FOR BIOTECHNOLOGY INFORMATION
National Library of Medicine
National Institutes of Health

August 1997

Vice President Launches PubMed, Lauds Free MEDLINE Access

"MEDLINE...will henceforth be available free to the American people." With those words, Vice President Al Gore inaugurated the PubMed search system at a Capitol Hill press conference on June 26. PubMed, which provides Web access to the National Library of Medicine's (NLM) database of the biomedical journal literature, MEDLINE, was hers[illegible] Tom Hart[illegible] (IA) as [illegible] of a sing[illegible] creative government initiative." The Vice President viewed free access to MEDLINE as consistent with the Clinton administration's other "empowerment" initiatives stating, "This development...may do more to reform and improve the quality of health care in the United States than anything else we've done in a long time."

Searching PubMed

PubMed grew out of NCBI's Entrez project which, since 1992, has offered a subset of MEDLINE records [illegible] enhanced [illegible] search [illegible] sing all of MEDLINE and PreMEDLINE, PubMed retains Entrez's ability to use one article as a "seed" to find other similar articles. By traversing the **See Related Articles'** links, a user can find articles similar in concept with speed and precision. PubMed expands upon Entrez by linking MEDLINE articles to full-text Web sites maintained by publishers. Currently, 95 journals are linked to PubMed, including Cell, Journal of Biological Chemistry, Journal of Cell Biology, New England Journal of Medicine, and Science. Access to [illegible] may require subscriptions or registration.

NCBI Director David Lipman (far left) coaches Vice President Gore (seated) as he searches PubMed. NIH Director Harold Varmus (center) and NLM Director Donald Lindberg (far right) look on.

PubMed Options

PubMed offers the option to search MEDLINE or any of NCBI's molecular biology databases. Users can select from a variety of search fields, including but not limited to: text words, author names, and journal titles. A MEDLINE citation for which there is a corresponding online, full-text article will have a button at the top of the abstract page that links to the publisher's Web site. Additional links point to

Continued on page 2

IN THIS ISSUE

PubMed Launched 1
Using Sequin 2
Structure Neighbors 3
NCBI Data by FTP 3
ORF Finder 4
Electronic PCR 4
Recent Publications 4
CGAP Revolutionizes Research .. 5
Frequently Asked Questions 6

NCBI

国际期刊对数据递交的要求

Distribution of Materials and Data

One of the terms and conditions of publishing with Cell Press is that authors be willing to distribute any materials and protocols used in the published experiments to qualified researchers for their own use. Materials include but are not limited to cells, DNA, antibodies, reagents, organisms, and mouse strains or, if necessary, the relevant ES cells. These must be made available with minimal restrictions and in a timely manner, but it is acceptable to request reasonable payment to cover the cost of maintenance and transport of materials. If there are restrictions to the availability of any materials, data, or information, these must be disclosed in the cover letter and in the STAR Methods section of the manuscript at the time of submission.

Data sets must be made freely available to readers from the date of publication and must be provided to editors and peer reviewers at submission for the purposes of evaluating the manuscript. In addition, we offer the opportunity for authors to make underlying data not included in the paper itself or deposited in a database available to the scientific community by posting them on Mendeley Data and then including a link in the published paper. For more detailed instructions, click here.

For the following types of data, submission of the full data set to a community-endorsed, public repository is mandatory. Accession numbers must be provided in the paper (see "Database Linking" below for specific formatting instructions). Examples of appropriate public repositories are listed below.

DNA and Protein Sequences

Protein Sequences: Uniprot

DNA and RNA Sequences: Genbank/European Nucleotide Archive (ENA)/DDBJ, Protein DataBank, UniProt

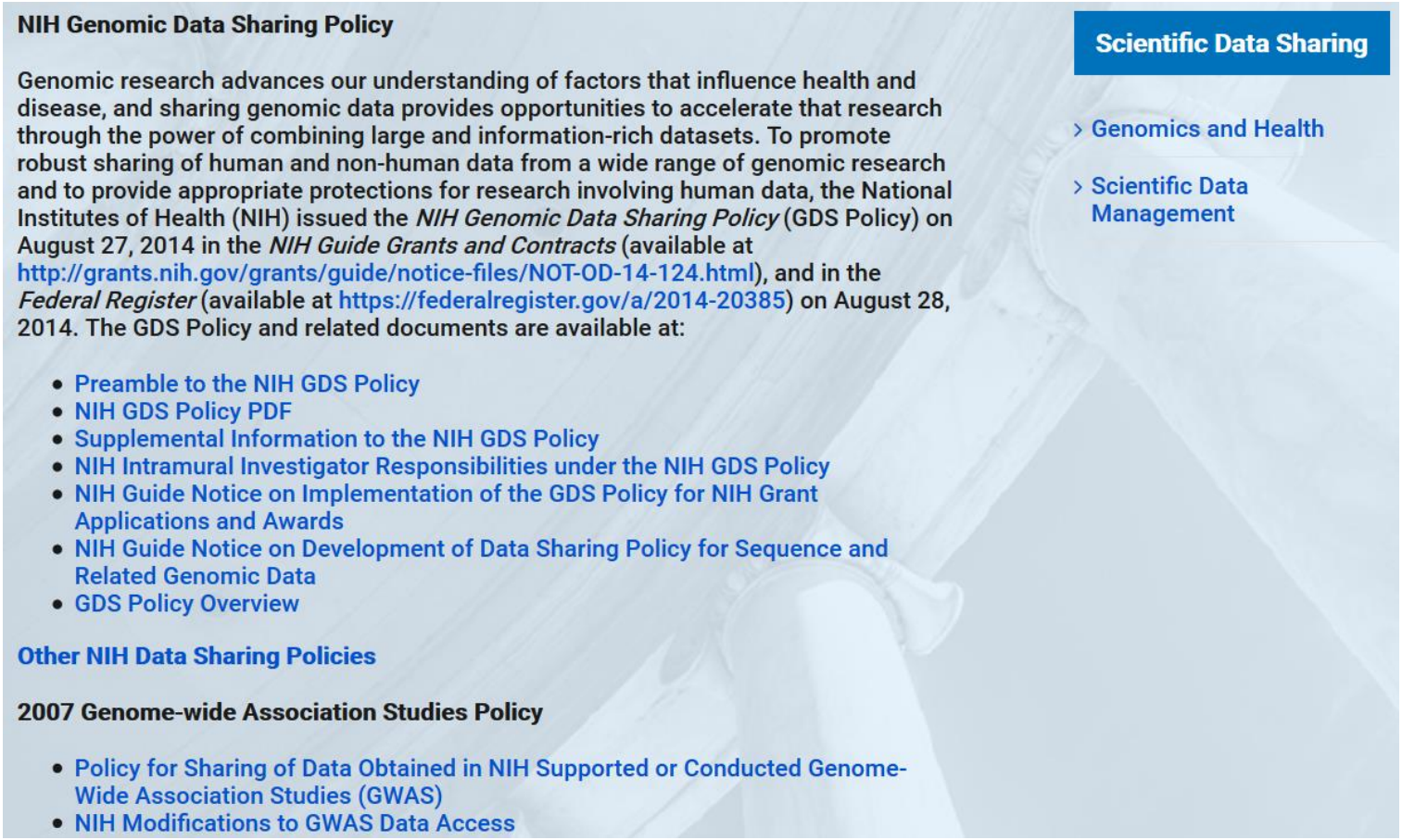
DNA Sequencing Data (traces and short reads): NCBI Trace and Short-Read Archive, ENA's Sequence Read Archive

Deep Sequencing Data: Deposit in GEO or ArrayExpress upon submission to the journal

The sequences of all RNAi, antisense, and morpholino probes must be included in the paper or deposited in a public database with the accession number provided in the paper.

DNA和RNA等序列要递交到NCBI/EBI/DDBJ

NIH要求及时公开被资助项目产出的数据



NIH Genomic Data Sharing Policy

Genomic research advances our understanding of factors that influence health and disease, and sharing genomic data provides opportunities to accelerate that research through the power of combining large and information-rich datasets. To promote robust sharing of human and non-human data from a wide range of genomic research and to provide appropriate protections for research involving human data, the National Institutes of Health (NIH) issued the *NIH Genomic Data Sharing Policy* (GDS Policy) on August 27, 2014 in the *NIH Guide Grants and Contracts* (available at http://grants.nih.gov/grants/guide/notice-files/NOT-OD-14-124.html), and in the *Federal Register* (available at https://federalregister.gov/a/2014-20385) on August 28, 2014. The GDS Policy and related documents are available at:

- Preamble to the NIH GDS Policy
- NIH GDS Policy PDF
- Supplemental Information to the NIH GDS Policy
- NIH Intramural Investigator Responsibilities under the NIH GDS Policy
- NIH Guide Notice on Implementation of the GDS Policy for NIH Grant Applications and Awards
- NIH Guide Notice on Development of Data Sharing Policy for Sequence and Related Genomic Data
- GDS Policy Overview

Other NIH Data Sharing Policies

2007 Genome-wide Association Studies Policy

- Policy for Sharing of Data Obtained in NIH Supported or Conducted Genome-Wide Association Studies (GWAS)
- NIH Modifications to GWAS Data Access

Scientific Data Sharing

› Genomics and Health

› Scientific Data Management

美国NIH对基因组数据共享的政策

NIH要求及时公开被资助项目产出的数据

National Institutes of Health Genomic Data Sharing Policy

B. Non-human Genomic Data

1. Data Submission Expectations and Timeline

Large-scale non-human genomic data, including data from microbes, microbiomes, and model organisms, as well as relevant associated data (e.g., phenotype and exposure data), are to be shared in a timely manner. Genomic data undergo different levels of data processing, which provides the basis for NIH's expectations for data submission. These expectations are provided in the Supplemental Information. In general, investigators should make non-human genomic data publicly available no later than the date of initial publication. However, earlier availability (i.e., before publication) may be expected for certain data or IC-funded projects (e.g., data from projects with broad utility as a resource for the scientific community such as microbial population-based genomic studies).

2. Data Repositories

Non-human data may be made available through any widely used data repository, whether NIH-funded or not, such as the Gene Expression Omnibus (GEO)[13] Sequence Read Archive (SRA),[14] Trace Archive,[15] Array Express,[16] Mouse Genome Informatics (MGI),[17] WormBase,[18] the Zebrafish Model Organism Database (ZFIN),[19] GenBank,[20] European Nucleotide Archive (ENA),[21] or DNA Data Bank of Japan (DDBJ).[22] NIH expects investigators to continue submitting data types to the same repositories that they submitted the data to before the effective date of the GDS Policy (e.g., DNA sequence data to GenBank/ENA/DDBJ, expression data to GEO or Array Express). Data types not previously submitted to any repositories may be submitted to these or other widely used repositories as agreed to by the funding IC.

非人的基因组数据：

必须不晚于文章发表之前公开

可以存放在任何广泛使用的数据库中

NIH要求论文全文共享

NIH Public Access Policy

Search

OER Glossary | Contact us

Home | Training | Policy Details | Managing Papers | FAQs | Special users | My NCBI | NIHMS

When and How to Comply

1 Preparing a manuscript — Address copyright — show me

2 Accepted for publication — Post it to PubMed Central and track it in My NCBI — show me

3 Reporting to NIH — Include PMCID in citations — show me

Overview:

To advance science and improve human health, NIH makes the peer-reviewed articles it funds publicly available on PubMed Central. The NIH public access policy requires scientists to submit final peer-reviewed journal manuscripts that arise from NIH funds to PubMed Central immediately upon acceptance for publication. [more]

Show me specific instructions for my publication

Public Access Policy Video Training

1 NIHMS overview

3 My Bibliography overview

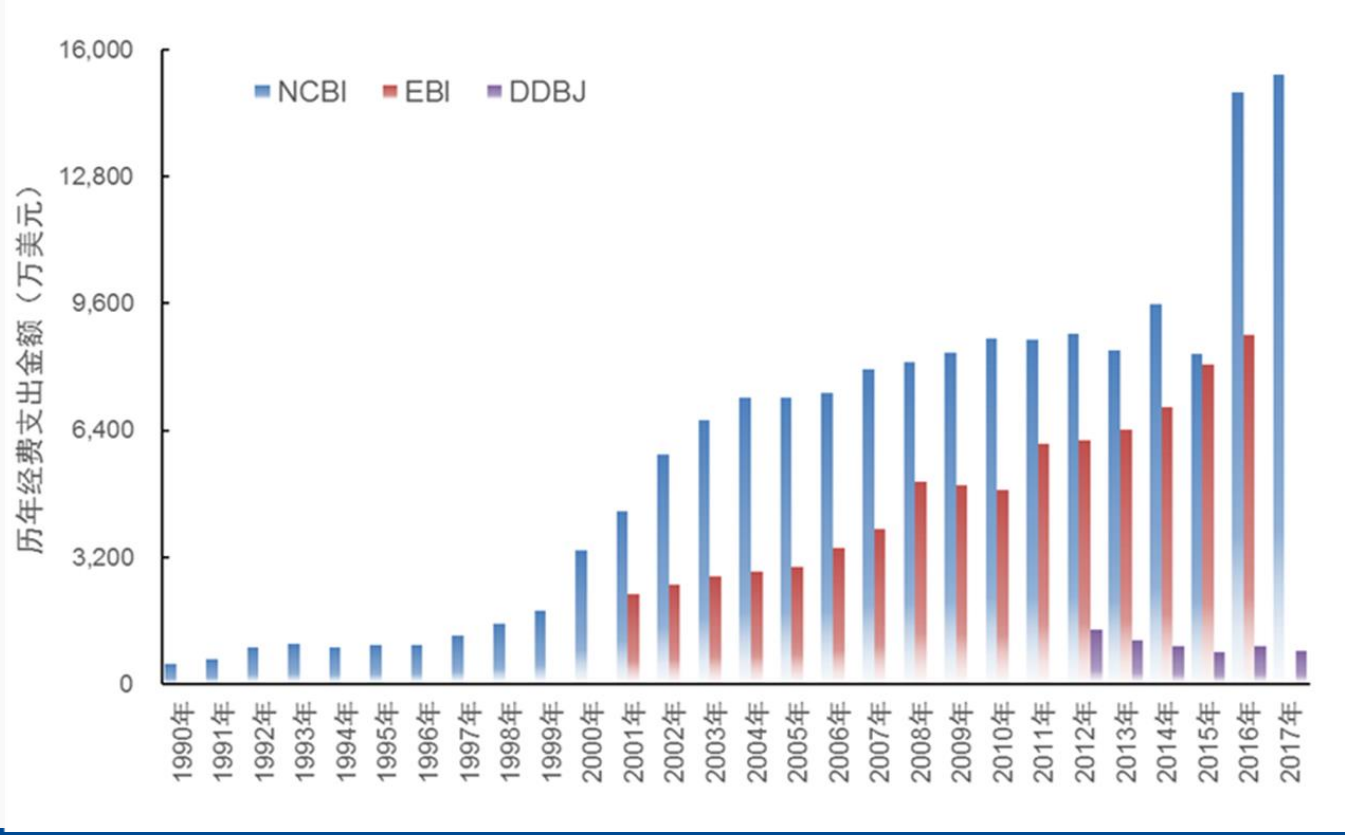
2 My NCBI overview

4 Public Access Compliance

Last Updated: Tuesday, March 18, 2014

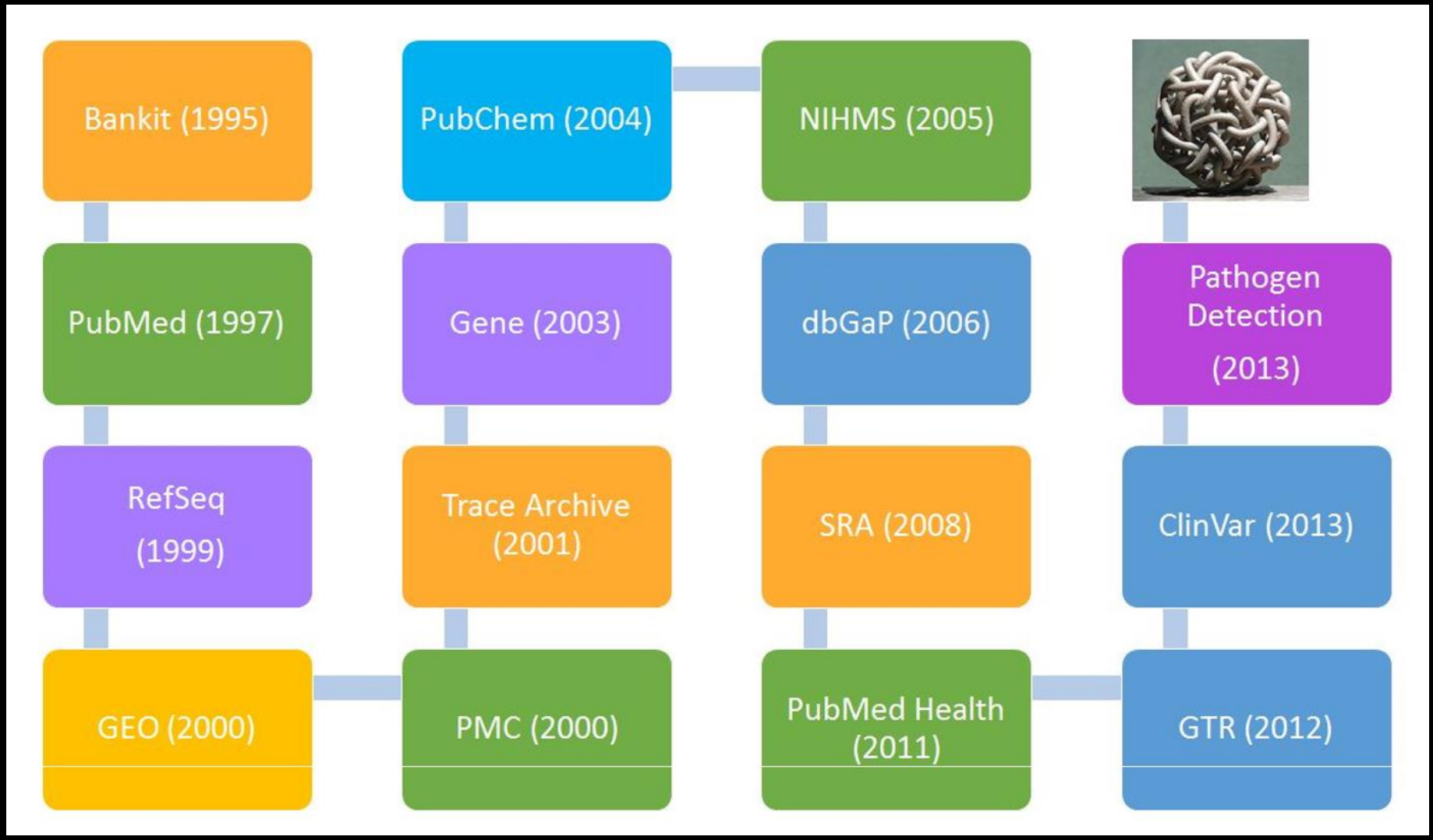
国际信息/数据中心运行经费充足、稳定

- NCBI：约700人，由美国国会固定经费支持，2017年**1.5**亿美元
- EBI：约600人，主要欧盟固定经费+英国政府经费，2016年**8,820**万美元
- DDBJ：约50人，主要由日本政府支持，2016年**891**万美元



NCBI成功的秘诀

- 团队
- 政策
- 创新


Bankit (1995)
PubChem (2004)
NIHMS (2005)
PubMed (1997)
Gene (2003)
dbGaP (2006)
Pathogen Detection (2013)
RefSeq (1999)
Trace Archive (2001)
SRA (2008)
ClinVar (2013)
GEO (2000)
PMC (2000)
PubMed Health (2011)
GTR (2012)

NCBI

Download BLAST
Get BLAST databases and executables

Use BLAST API
Call BLAST from your application

Use BLAST in the cloud
Start an instance at a cloud provider

Specialized searches

SmartBLAST
Find proteins highly similar to your query

Primer-BLAST
Design primers specific to your PCR template

Global Align
Compare two sequences across their entire span (Needleman-Wunsch)

CD-search
Find conserved domains in your sequence

GEO
Find matches to gene expression profiles

IgBLAST
Search immunoglobulins and T cell receptor sequences

VecScreen
Search sequences for vector contamination

CDART
Find sequences with similar conserved domain architecture

Targeted Loci
Search markers for phylogenetic analysis

Multiple Alignment
Align sequences using domain and protein constraints

BioAssay
Search protein or nucleotide targets in PubChem BioAssay

MOLE-BLAST
Establish taxonomy for uncultured or enviromental sequences

NCBI

NCBI成功的秘诀

- 团队
- 政策
- 创新
- 技术

Article types
Clinical Trial
Review
Customize ...

Text availability
Abstract
Free full text
Full text

Publication dates
5 years
10 years
Custom range...

Species
Humans
Other Animals

Clear all

Show additional filters

Format: Summary **Sort by**: Most Recent **Per page**: 20 **Send to**

Filters: Manage Filters

Sort by:
Best match | Most recent

Search results

Items: 1 to 20 of 586

« First ‹ Prev Page 1 of 30 Next › Last »

1. Risk factor analysis of proximal junctional kyphosis after posterior osteotomy in patients with ankylosing spondylitis.
Wang T, Zhao Y, Liang Y, Zhang H, Wang Z, Wang Y.
J Neurosurg Spine. 2018 Apr 20:1-6. doi: 10.3171/2017.11.SPINE17228. [Epub ahead of print]
PMID: 29676671
Similar articles

2. Radiographic and clinical outcomes following MIS-TLIF in patients with adult lumbar degenerative scoliosis.
Zhao Y, Liang Y, Mao K.
J Orthop Surg Res. 2018 Apr 19;13(1):93. doi: 10.1186/s13018-018-0764-7.
PMID: 29673371 **Free Article**
Similar articles

3. miR-223-3p Inhibits Human Osteosarcoma Metastasis and Progression by Directly Targeting CDH6.
Ji Q, Xu X, Song Q, Xu Y, Tai Y, Goodman SB, Bi W, Xu M, Jiao S, Maloney WJ, Wang Y.
Mol Ther. 2018 Mar 14. pii: S1525-0016(18)30117-5. doi: 10.1016/j.ymthe.2018.03.009. [Epub ahead of print]
PMID: 29628305 **Free Article**
Similar articles

4. Cell cycle and complement inhibitors may be specific for treatment of spinal cord injury in aged and young mice: Transcriptomic analyses.
Hao M, Ji XR, Chen H, Zhang W, Zhang LC, Zhang LH, Tang PF, Lu N.

Results by year

Download CSV

Titles with your search terms

[Primary symptoms **of 301** patients with prostate cancer: Nago [Nihon Hinyokika Gakkai Zasshi....]

Male breast carcinoma--a review **of 301** cases from **the** Christie **Hospital** &a [Br J Cancer. 1985]

An interpretation **of** chapter **301**; laws **of** 1959 (prepared jointly by **the** State M [Wis Med J. 1966]

See more...

Find related data

Database: Select

NCBI存在的问题/挑战

- 政府停摆

NCBI Resources How To

yimingbao My NCBI Sign Out

NCBI
National Center for Biotechnology Information

All Databases | Search

Due to the lapse in government funding, the information on this web site may not be up to date, transactions submitted via the web site may not be processed, and the agency may not be able to respond to inquiries until appropriations are enacted. Updates regarding government operating status and resumption of normal operations can be found at opm.gov.

- NCBI Home
- Resource List (A-Z)
- All Resources
- Chemicals & Bioassays
- Data & Software
- DNA & RNA
- Domains & Structures
- Genes & Expression
- Genetics & Medicine
- Genomes & Maps
- Homology
- Literature
- Proteins
- Sequence Analysis
- Taxonomy

Welcome to NCBI

The National Center for Biotechnology Information advances science and health by providing access to biomedical and genomic information.

About the NCBI | Mission | Organization | NCBI News & Blog

Submit

Deposit data or manuscripts into NCBI databases

Download

Transfer NCBI data to your computer

Learn

Find help documents, attend a class or watch a tutorial

Develop

Analyze

Research

Popular Resources

- PubMed
- Bookshelf
- PubMed Central
- PubMed Health
- BLAST
- Nucleotide
- Genome
- SNP
- Gene
- Protein
- PubChem

NCBI News & Blog

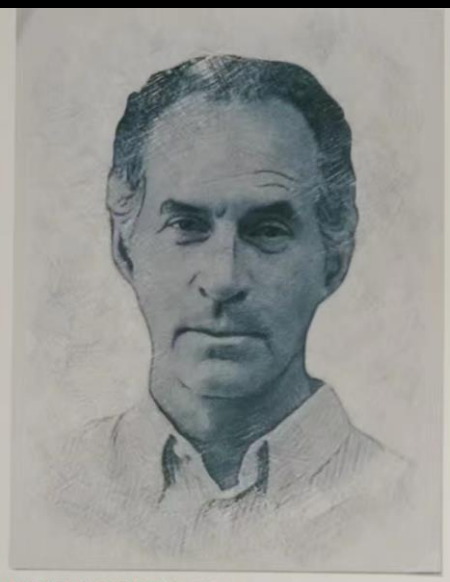
22 Jan 2018

NCBI

NCBI存在的问题/挑战

- 政府停摆
- 数据及用户的急剧增长，尤其是来自中国的
- 软件开发受限
- 数据分析缺乏
- 有些传统束缚

NCBI



DAVID J. LIPMAN, M.D.
Professor

Dr. David Lipman is the Director NCBI, USA, since 1989. He has overseen its growth into one of the most heavily used resources in the world for the search and retrieval of biomedical information. He is one of the developers of BLAST algorithm for fast sequence comparison. Dr. Lipman is the recipient of numerous awards, and is a member of the National Academy of Sciences, the Institute of Medicine, and the American Academy of Arts and Sciences.

NCBI Farewell

Lipman, David (NIH/NLM/NCBI) [E]

To: NLM/NCBI List ncbi

Wednesday, May 03, 2017 10:16 AM

"How lucky I am to have something that makes saying goodbye so hard."

Dear NCBI Colleagues,

After almost 30 years at NCBI, the time has come for me to say goodbye. I'll be leaving later this month to join an old friend, Pat Brown, at his company, Impossible Foods, as Chief Science Officer.

The reorganization efforts we started here last Fall have been successful - Kim, Bart, Fitz, & Al are a great team and David Landsman will continue heading up CBB. Dr. Patti Brennan has asked Jim Ostell to be the Acting Director and she gives him her fullest support as does the NIH Director, Dr. Francis Collins.
So I know I'm leaving NCBI as strong as it has ever been.

I hope you know by now how much I've appreciated working with all of you. We've done amazing things together and I expect greatness from you after I'm gone. Before leaving, I'll make the rounds to chat with as many of you as I can.

Thank you for being such an important part of my life...

- David

NCBI

NCBI

国际数据中心发展趋势

- **云计算模式**：采用商业云计算资源解决大数据存储和加工的效率问题
- **加大数据收集力度**：2014年，EBI以欧洲为主，建立ELIXIR数据联盟
- **重视人的数据收集**：2017年，NCBI dbSNP宣布不接收非人数据

Phasing out support for non-human genome organism data in dbSNP and dbVar

May 9, 2017

29 Votes

This blog post is directed toward people who use dbSNP and dbVar, particularly those who submit non-human data to the two databases.

dbSNP and dbVar archive, process, display and report information related to germline and somatic variations from multiple species. These two databases have grown rapidly as sequencing and other discovery technologies have evolved, and now contain nearly two billion variants from over 360 species.

Based on projected growth and the resources required to archive and distribute the data, continued support for all organisms will become unsustainable for NCBI in the near future. Therefore, NCBI will phase out support for all non-human organisms in dbSNP and dbVar, and will support only human variation.

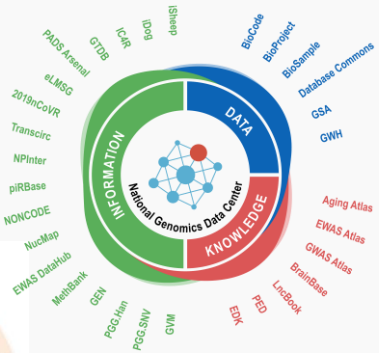
NCBI will phase out support for non-human organisms in dbSNP and dbVar following this timeline:

September 1, 2017 – dbSNP and dbVar stop accepting non-human variant data submissions

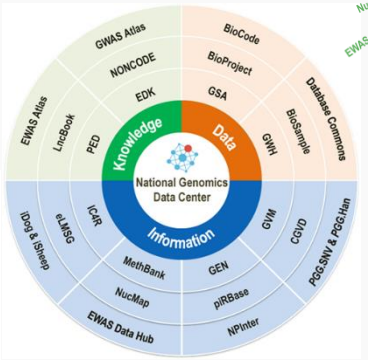
November 1, 2017 – dbSNP and dbVar interactive websites and related NCBI services stop presenting non-human variant data. The data will, however, continue to be available for download on the dbSNP and dbVar FTP sites.

Any non-human data that is already in the databases or that is submitted before September 1, 2017 will continue to be available via the dbSNP and dbVar FTP download sites.

北京基因组研究所大数据中心

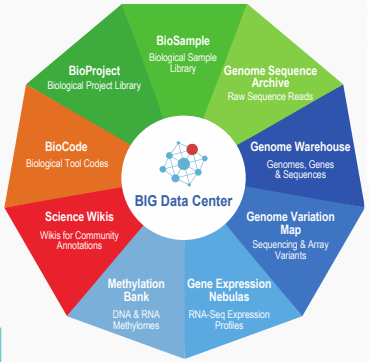


2020



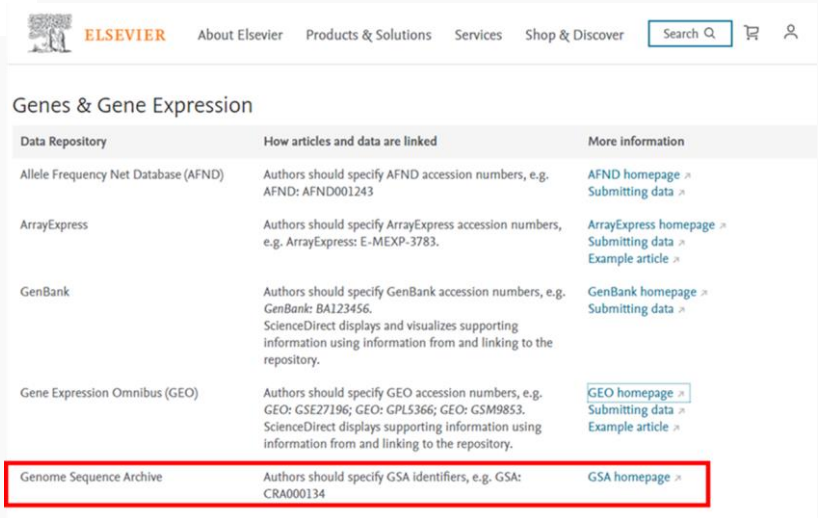
2019

多数据模块
2018



九大数据模块
2017

六大数据模块
2016

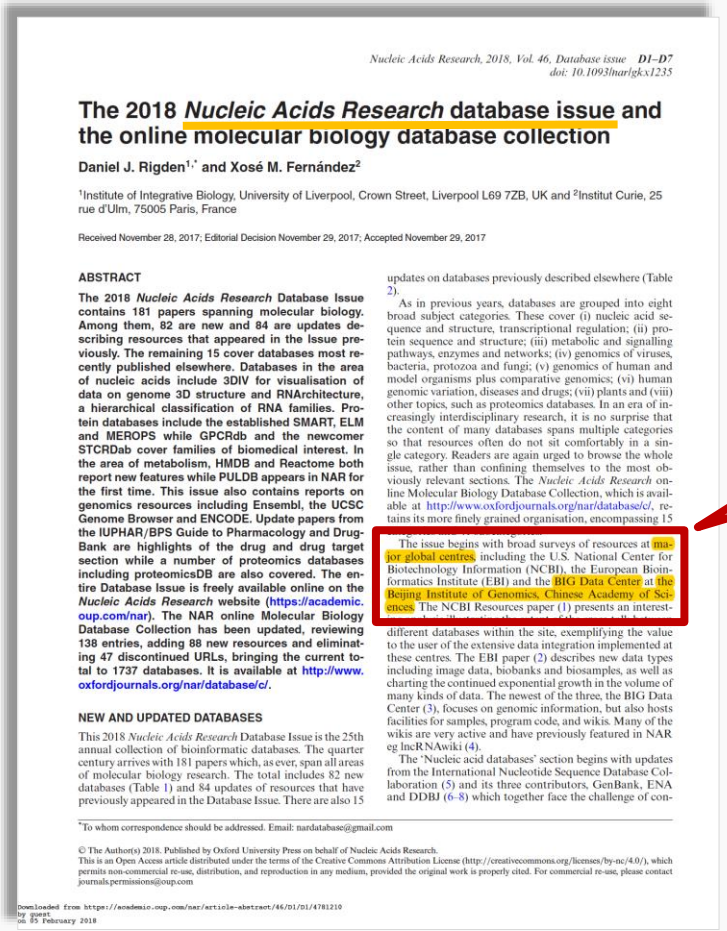


ELSEVIER About Elsevier Products & Solutions Services Shop & Discover Search

Genes & Gene Expression

Data Repository	How articles and data are linked	More information
Allele Frequency Net Database (AFND)	Authors should specify AFND accession numbers, e.g. AFND: AFND001243	AFND homepage Submitting data
ArrayExpress	Authors should specify ArrayExpress accession numbers, e.g. ArrayExpress: E-MEXP-3783.	ArrayExpress homepage Submitting data Example article
GenBank	Authors should specify GenBank accession numbers, e.g. *GenBank: BA123456.* ScienceDirect displays and visualizes supporting information using information from and linking to the repository.	GenBank homepage Submitting data
Gene Expression Omnibus (GEO)	Authors should specify GEO accession numbers, e.g. *GEO: GSE27196; GEO: GPL5366; GEO: GSM9853.* ScienceDirect displays supporting information using information from and linking to the repository.	GEO homepage Submitting data Example article
Genome Sequence Archive	Authors should specify GSA identifiers, e.g. GSA: CRA000134	GSA homepage

生命与健康大数据中心连续三年
被国际同行列为国际主要生物数据中心之一



Nucleic Acids Research, 2018, Vol. 46, Database issue D1–D7
doi: 10.1093/nar/gkx1235

The 2018 *Nucleic Acids Research* database issue and the online molecular biology database collection

Daniel J. Rigden[1,*] and Xosé M. Fernández[2]

[1]Institute of Integrative Biology, University of Liverpool, Crown Street, Liverpool L69 7ZB, UK and [2]Institut Curie, 25 rue d'Ulm, 75005 Paris, France

Received November 28, 2017; Editorial Decision November 29, 2017; Accepted November 29, 2017

ABSTRACT

The 2018 *Nucleic Acids Research* Database Issue contains 181 papers spanning molecular biology. Among them, 82 are new and 84 are updates describing resources that appeared in the Issue previously. The remaining 15 cover databases most recently published elsewhere. Databases in the area of nucleic acids include 3DIV for visualisation of data on genome 3D structure and RNArchitecture, a hierarchical classification of RNA families. Protein databases include the established SMART, ELM and MEROPS while GPCRdb and the newcomer STCRDab cover families of biomedical interest. In the area of metabolism, HMDB and Reactome both report new features while PULDB appears in NAR for the first time. This issue also contains reports on genomics resources including Ensembl, the UCSC Genome Browser and ENCODE. Update papers from the IUPHAR/BPS Guide to Pharmacology and DrugBank are highlights of the drug and drug target section while a number of proteomics databases including proteomicsDB are also covered. The entire Database Issue is freely available online on the *Nucleic Acids Research* website (https://academic.oup.com/nar). The NAR online Molecular Biology Database Collection has been updated, reviewing 138 entries, adding 88 new resources and eliminating 47 discontinued URLs, bringing the current total to 1737 databases. It is available at http://www.oxfordjournals.org/nar/database/c/.

NEW AND UPDATED DATABASES

This 2018 *Nucleic Acids Research* Database Issue is the 25th annual collection of bioinformatic databases. The quarter century arrives with 181 papers which, as ever, span all areas of molecular biology research. The total includes 82 new databases (Table 1) and 84 updates of resources that have previously appeared in the Database Issue. There are also 15 updates on databases previously described elsewhere (Table 2).

As in previous years, databases are grouped into eight broad subject categories. These cover (i) nucleic acid sequence and structure, transcriptional regulation; (ii) protein sequence and structure; (iii) metabolic and signalling pathways, enzymes and networks; (iv) genomics of viruses, bacteria, protozoa and fungi; (v) genomics of human and model organisms plus comparative genomics; (vi) human genomic variation, diseases and drugs; (vii) plants and (viii) other topics, such as proteomics databases. In an era of increasingly interdisciplinary research, it is no surprise that the content of many databases spans multiple categories so that resources often do not sit comfortably in a single category. Readers are again urged to browse the whole issue, rather than confining themselves to the most obviously relevant sections. The *Nucleic Acids Research* online Molecular Biology Database Collection, which is available at http://www.oxfordjournals.org/nar/database/c/, retains its more finely grained organisation, encompassing 15

The issue begins with broad surveys of resources at major global centres, including the U.S. National Center for Biotechnology Information (NCBI), the European Bioinformatics Institute (EBI) and the BIG Data Center at the Beijing Institute of Genomics, Chinese Academy of Sciences. The NCBI Resources paper (1) presents an interest-

different databases within the site, exemplifying the value to the user of the extensive data integration implemented at these centres. The EBI paper (2) describes new data types including image data, biobanks and biosamples, as well as charting the continued exponential growth in the volume of many kinds of data. The newest of the three, the BIG Data Center (3), focuses on genomic information, but also hosts facilities for samples, program code, and wikis. Many of the wikis are very active and have previously featured in NAR eg lncRNAwiki (4).

The 'Nucleic acid databases' section begins with updates from the International Nucleotide Sequence Database Collaboration (5) and its three contributors, GenBank, ENA and DDBJ (6–8) which together face the challenge of con-

*To whom correspondence should be addressed. Email: nardatabase@gmail.com

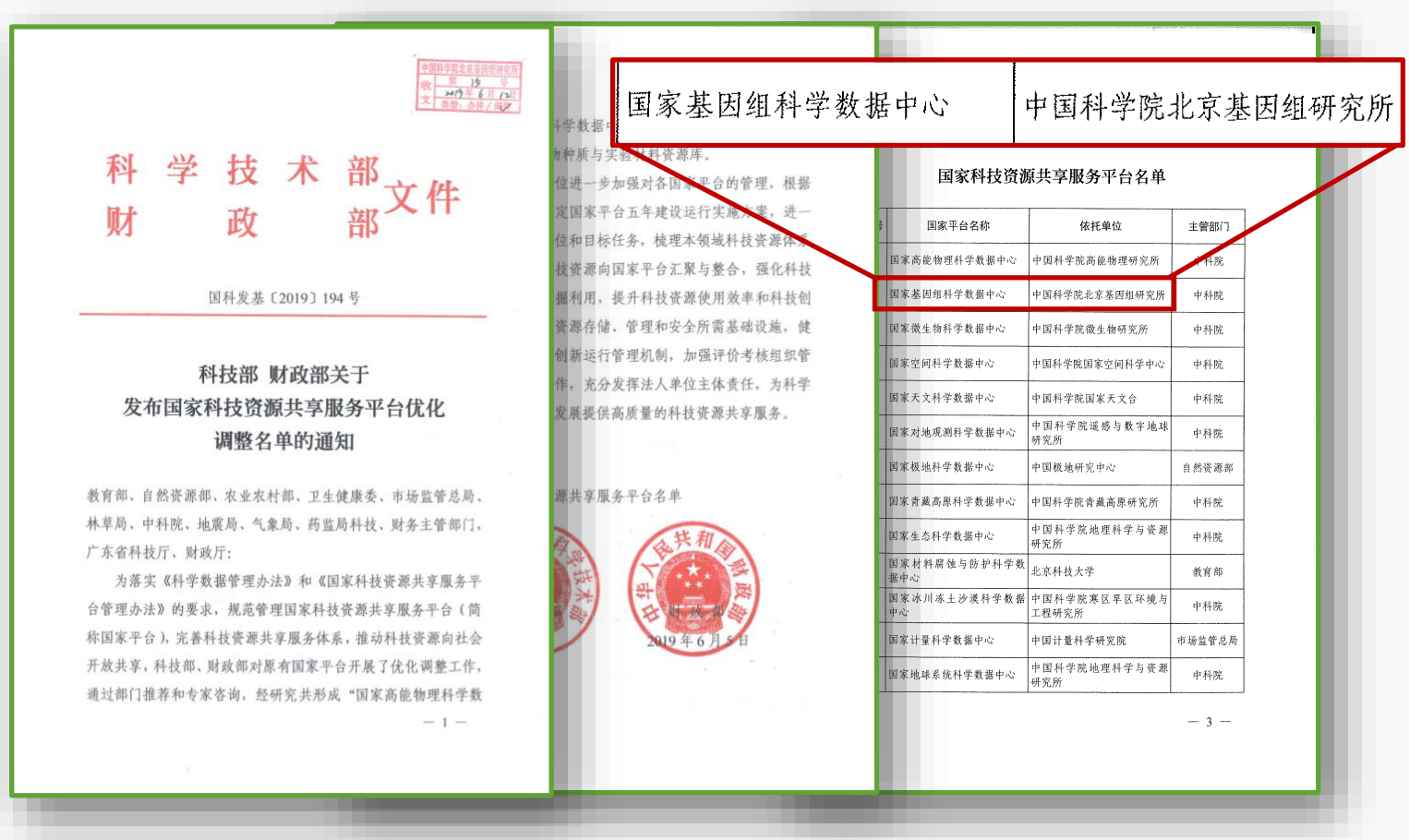
© The Author(s) 2018. Published by Oxford University Press on behalf of Nucleic Acids Research.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

Downloaded from https://academic.oup.com/nar/article-abstract/46/D1/D1/4781210
by guest
on 05 February 2018

Nucleic Acids Research 2018, 46:D1–D7

- 2017 年在*Nucleic Acids Research*发表**1**篇论文
- 2018 年在*Nucleic Acids Research*发表**4**篇研究成果
- 2019 年在*Nucleic Acids Research*发表**7**篇研究成果

共建“国家基因组科学数据中心”- NGDC



科 学 技 术 部
财 政 部 文件

国科发基〔2019〕194 号

科技部 财政部关于
发布国家科技资源共享服务平台优化
调整名单的通知

教育部、自然资源部、农业农村部、卫生健康委、市场监管总局、林草局、中科院、地震局、气象局、药监局科技、财务主管部门，广东省科技厅、财政厅：

为落实《科学数据管理办法》和《国家科技资源共享服务平台管理办法》的要求，规范管理国家科技资源共享服务平台（简称国家平台），完善科技资源共享服务体系，推动科技资源向社会开放共享，科技部、财政部对原有国家平台开展了优化调整工作，通过部门推荐和专家咨询，经研究共形成“国家高能物理科学数

— 1 —

2019 年 6 月 5 日

国家基因组科学数据中心 | 中国科学院北京基因组研究所

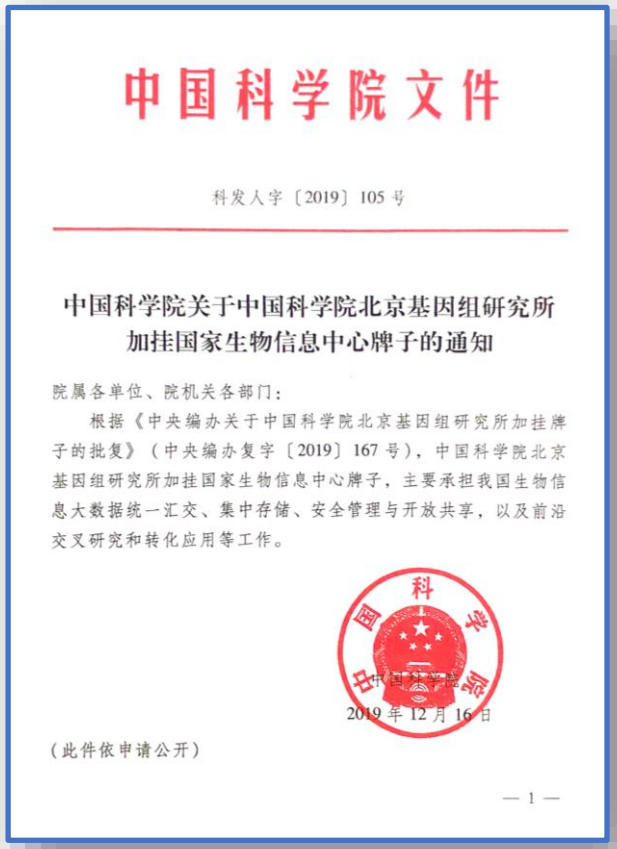
国家科技资源共享服务平台名单

国家平台名称	依托单位	主管部门
国家高能物理科学数据中心	中国科学院高能物理研究所	中科院
国家基因组科学数据中心	中国科学院北京基因组研究所	中科院
国家微生物科学数据中心	中国科学院微生物研究所	中科院
国家空间科学数据中心	中国科学院国家空间科学中心	中科院
国家天文科学数据中心	中国科学院国家天文台	中科院
国家对地观测科学数据中心	中国科学院遥感与数字地球研究所	中科院
国家极地科学数据中心	中国极地研究中心	自然资源部
国家青藏高原科学数据中心	中国科学院青藏高原研究所	中科院
国家生态科学数据中心	中国科学院地理科学与资源研究所	中科院
国家材料腐蚀与防护科学数据中心	北京科技大学	教育部
国家冰川冻土沙漠科学数据中心	中国科学院寒区旱区环境与工程研究所	中科院
国家计量科学数据中心	中国计量科学研究院	市场监管总局
国家地球系统科学数据中心	中国科学院地理科学与资源研究所	中科院

— 3 —

2019.6.5，科技部、财政部联合发布，依托中国科学院北京基因组研究所，成立“**国家基因组科学数据中心**”

建立“国家生物信息中心”



中国科学院文件

科发人字〔2019〕105号

中国科学院关于中国科学院北京基因组研究所加挂国家生物信息中心牌子的通知

院属各单位、院机关各部门：

根据《中央编办关于中国科学院北京基因组研究所加挂牌子的批复》（中央编办复字〔2019〕167号），中国科学院北京基因组研究所加挂国家生物信息中心牌子，主要承担我国生物信息大数据统一汇交、集中存储、安全管理与开放共享，以及前沿交叉研究和转化应用等工作。

中国科学院

2019年12月16日

（此件依申请公开）

— 1 —

2019.11.13，中央编办批复：经研究并报中央编委批准，同意中国科学院北京基因组研究所加挂“**国家生物信息中心**”牌子

2019新型冠状病毒信息库（2019nCoVR）

2020年1月22日上线，提供全球数据整合、冗余去除、质量评估、序列下载、动态监控等

六大模块

- 2019新型冠状病毒序列（138393 条）
- 冠状病毒序列（391258 条）
- 基因组变异（18721 个）
- 智能诊断 & 在线工具
- 临床信息（208 条）
- 文献情报（72140 篇）

https://bigd.big.ac.cn/ncov

服务全球

截至2020年10月2日，已收录全球**138,393**条新冠病毒基因组序列信息，为全球**175个**国家/地区**17万**余访客提供数据服务，累计数据下载**1.48亿**多条，国外访客占比高达**60%**

谢谢！

baoym@big.ac.cn
18611507347