

· 述 评 ·

人类生物战略资源库现状

崔春舜, 余 辉, 高东平, 杨 渊

[摘要] 随着精准医学、转化医学等以人类生物战略资源为主要资源需求的学科的兴起和发展, 人类生物战略资源库的建设越来越受到关注, 且近年来国内外资源库的建设规模和数量也在逐渐扩大。资源库联盟作为协作平台的出现更加规范了资源库的建设和管理标准, 库中储存的资源如遗传资源、干细胞资源等的质量也在不断提升。作者对国内外主要的人类生物战略资源库的建设现状进行情报调研, 旨在为国内的资源库建设提供参考。

[关键词] 人类; 生物战略资源库; 现状

[中图分类号] R-05

[文献标志码] A

[文章编号] 2095-3097(2018)04-0193-05

doi: 10.3969/j.issn.2095-3097.2018.04.001

The present status of human biological strategic resource repository

CUI Chunshun, YU Hui, GAO Dongping, YANG Yuan

(Institute of Medical Information, Chinese Academy of Medical Science & Peking Union Medical College, Beijing 100020, China)

[Abstract] With the emergence and development of precision medicine, translational medicine and other disciplines that take human biological strategy resources as the main resource needs. The construction of human biological strategic repository is attracting more and more attention. In recent years, the scale and quantity of repository is gradually expanding, and resource alliances as collaboration platform are aimed to standardize the standard construction and management of the repository. The quality of resources such as genetic resources, stem cell resources is constantly improving. The present status of the main human biological strategic repositories is reviewed in this paper, aimed at providing reference for the construction of domestic repository.

[Key words] Human; Biological strategic resource repository; Present status

人类生物资源是人类社会生存与可持续发展不可或缺、生命科学原始创新、获得知识产权以及健康产业的物质基础, 是保障国家生态安全、人类健康安全的战略性资源^[1]。人类生物资源已经成为经济和社会可持续发展的重要战略资源, 包括人类遗传资源、生物样本资源以及干细胞资源等。由于其重要的战略作用, 美、英等发达国家正在全世界范围内以寻求科研合作的方式获取各地人类生物资源建设资源库, 印度、巴西等发展中国家也在不断投入资金支持本国人类生物资源的开发、保护和利用。我国56个民族源远流长, 保持有纯家谱系的人类遗传资源, 同时, 我国疾病谱复杂多变, 因此建设人类遗传资源库存储保护我国人类遗传资源、建设人体生物

样本库, 对了解复杂疾病背后的危险因素和作用机制等具有重要战略意义。另外, 围绕开发、保护和利用我国人类生物资源的目标, 掌握国内外重要人类生物战略资源库的建设进展十分重要。作者对人类遗传资源库、生物样本库以及干细胞库3种类型的资源库的建设现状进行情报调研, 通过文献调研、各资源库的官方统计与年度报告分析国内外人类生物战略资源库建设中的问题, 并给出国内的资源库建设相应的对策建议。

1 国内外人类遗传资源库建设现状

1.1 国外人类遗传资源库建设现状 欧洲生物信息研究所(European Bioinformatics Institute, EBI) 建立于1994年, 位于英国剑桥南部的维康信托基因园, 是欧洲分子生物学实验室的一部分。该组织致力于为科学研究提供免费的公共生物数据、一系列服务和工具以及进行生物信息学基础研究的培训。

[基金项目] 中国医学科学院医学与健康科技创新工程经费资助(2016-I2M-2-004)

[作者单位] 100020 北京, 中国医学科学院北京协和医学院医学信息研究所(崔春舜, 余 辉, 高东平, 杨 渊)

[通讯作者] 杨 渊, E-mail: yang.yuan@imicams.ac.cn

2016年,来自63个国家的624个机构进行的186个研究项目加入了EBI,有来自8个国家的22个公司是EBI的成员。EBI 2016年新加入2个国家,总成员国17个。2016年期间,数据库每天处理2 700万个数据请求,每月运行1 270万个计算任务。EBI构建的元基因组学数据库在2016年数据量增长了11倍。EBI构建了一个名字为“BioStudies”的数据库用来存储非传统数据如图像、单细胞表达集以及基因表达规律等。截至2016年底,EBI存储生物信息数据量:核酸序列数据库存储数量5.91 PB;物种和菌株的基因组42 529个;元基因组样本超过9万个;基因表达试验220万;蛋白质序列7 100万;大分子结构1.2万多个。EBI免费提供的最新的分子数据资源,可通过ENSEMBL数据库访问核酸信息资源,UniProt数据库访问蛋白质信息资源,并提供包括Clustal Omega序列对比工具、InterProScan蛋白质特征序列搜索工具等分析工具^[2]。

日本的DNA数据库(DNA Data Bank of Japan, DDBJ)始建于1986年,由国立遗传学研究所负责数据库的建设、维护及数据的传播。截至2015年底,DDBJ的核酸数据库中DDBJ序列阅读档案存储了2.51 PB序列数据,其中包括1.7 PB的sra格式文件和0.81 PB的fastq格式文件。2015年提交至DDBJ有1 156 001条数据记录,数据库中增加的记录有1 282 165条^[3]。DDBJ提供免费可用的核苷酸序列数据,通过在getentry中输入登录号、在ARSA输入关键字以及TXSearch分类数据搜索方式获取序列信息,研究人员可通过D-Way通道提交自己的序列数据^[4]。

GenBank是美国国立生物技术信息中心(National Center for Biotechnology Information, NCBI)建立的DNA序列数据库,旨在向研究人员提供最新的和全面的基因序列信息。截至2018年2月底,GenBank数据库共收录2 536亿多个碱基和2亿多个核酸序列。每2个月数据库会发布一个说明文档,此文档内容是关于当前数据库版本的详细信息以及接下来数据库更新的变化情况。NCBI对GenBank数据的使用或分发没有任何限制,用户可通过核苷酸序列检索、BLAST序列对比、使用NCBI e-utilities程序以编程方式检索序列等方式获取数据^[5]。

EBI、GenBank以及DDBJ合作,采用统一的数据标准,3个数据库每天交换数据,保持数据一致性,3大基因库在存储功能外均提供计算、模拟实验以及序列对比分析功能。

1.2 我国人类遗传资源库建设现状 2016年9月22日,总部位于深圳的中国国家基因库正式投入运

行。中国国家基因库计划建成“三库两平台”的业务结构和功能,“三库”由生物样本资源库、生物信息数据库和生物活体库组成,“两平台”为数字化平台、合成与编辑平台。国家基因库已存储了众多生物资源样本,一些样本正在陆续从原来分散储存的地点运往基因库,截至2016年底,存储总量已达1 000万份。在这些数据当中,除了基本的生物资源样本和信息之外,还储存了人类疾病和物种多样性的相关数据。目前,罕见病数据库已检测了3 000多个家庭,新发现400多种罕见病,癌症数据库中包含了上万份不同种类的数据。2017年9月,国家基因库又启动了“国家基因库青海高原活体库”植物资源采集和科学考察项目以及国家基因库眼基因库项目^[6]。国内其他基因库建设项目包括从2009年开始,华大基因与地方政府展开合作,到2016年总计完成27万多份样本的检测,包括产前无创、遗传性耳聋、地中海贫血、新生儿遗传代谢病、宫颈癌筛查等^[7]。

由国家卫健委主管,国家卫健委科学技术研究所牵头建设的国家人类遗传资源共享服务平台由中华民族遗传资源中心、国家重大疾病遗传资源中心、国家生殖遗传资源中心、特殊人群遗传资源中心、自然人群遗传资源中心、极端环境遗传资源中心、干细胞遗传资源中心以及分子影像资源中心8个资源中心组成,每个中心下由不同的资源专题组成,比如中华民族遗传资源中心下有中华民族永生细胞专题、中华民族分子画像遗传资源专题等4个专题,每个专题下列出所属的数据库信息。平台现已建成北京、上海以及华南3个创新中心,推动我国人类遗传资源的标准化整合、共享服务体系的进一步完善^[8]。

2 国内外生物样本库建设现状

2.1 国外生物样本库建设现状 英国生物样本库始建于1999年,为大型的前瞻性人类遗传队列生物样本库。在2006—2010年间募集了50万名40~69岁的英国人志愿者,存储志愿者的血液、尿液和唾液样本以供进一步分析^[9]。2016年4月宣布建立内脏器官扫描图像数据库项目,计划使用核磁共振成像以及其他技术了解心脏、大脑、脂肪、骨骼和血管疾病成因与发病机理,涉及癌症、老年痴呆症、关节炎和骨质疏松症、冠状动脉心脏病等疾病,该项目计划扫描10万人群的图像数据,截至2018年1月,已招募21 332名项目参与者。英国生物样本库对每项申请采取250英镑的固定收费模式,学科研究员与准入小组委员会对所有申请进行评估通过后提供样本^[10]。

丹麦国家生物样本库建立于2012年,截至2018年2月,存储有963万余份生物样本,存储量前3位的生物样本分别为血清420万份、干血斑样本220万份以及血浆130万份。获取生物样本需先获得研究伦理委员会的支持,然后向丹麦健康数据管理的科学服务部门提交申请,申请由丹麦国家生物样本库审查后提供生物样本资源^[11]。

美国国家癌症研究所牵头建立的人类组织协作网络(Cooperation Human Tissue Network, CHTN),始建于1987年,目前由东部分部宾夕法尼亚大学医学院、大西洋中部分部弗吉尼亚大学医学院、中西部分部俄亥俄大学维克斯纳医学中心、儿科全国儿童医院、南部分部伯明翰阿拉巴马大学医学院以及西部分部范德比尔特大学医学院6个部分组成。自成立以来,CHTN向3400多名科研人员提供了100多万份标本,在2013—2017年间,CHTN向美国和加拿大批准的研究人员以及国际上的少数研究人员提供了235710份生物样本。2016年,CHTN向433名通过了伦理审查委员会评审研究人员提供了49832份样本,其中,72%的样本由CHTN部门提供,而其余28%的样本由CHTN医疗中心联盟提供。研究人员提交CHTN申请获取生物样本,其申请由外科病理学家审查可行性通过后CHTN向研究人员提供生物样本^[12]。

泛欧洲生物样本库与生物分子资源研究中心(Biobanking and Bio-Molecular Resources Research Infrastructure, BBMRI)拥有16个成员国、3个观察国以及1个组织,涵盖了超过30个国家的225个机构成员的生物样本库。该联盟最新开放的4.0版目录公布其样本库网络中涵盖了1431个生物样本库和生物样本集合^[13]。

韩国国家研究资源中心由科学部的信息、通信和技术与未来规划以及韩国国家研究基金会支持建立。目前分为5个核心中心(人类资源、动物、植物、微生物和融合物),包括有36个资源中心,其中人类资源6个、动物8个、植物3个、微生物12个以及融合物7个^[14]。

加拿大肿瘤信息库网络于2004年在加拿大健康研究所癌症研究所的资助下成立,旨在促进癌症决定因素的研究,更好地了解癌症预防和早期发现,并预测和改进药物反应以及寻找新的靶向药物^[15]。

国际上的生物样本库各有特色,如英国生物样本库以固定目的的项目的方式收集队列人群的生物样本资源;BBMRI则致力于多国样本库的标准化工作;韩国国家研究资源中心则发展多物种的生物样本资源库。同时各国在建设生物样本库的时候均十

分重视伦理审查。

2.2 国内生物样本库建设现状 1994年,在国家自然科学基金重大项目“中华民族基因组中若干位点基因结构的研究”第一子课题“中华多民族基因组的保存”的资助下,建立了中华民族永生细胞库,收集了具有代表性的47个民族70个群体的样本,建立了存有3982株永生细胞株的细胞库和近7210例个体的DNA库。2002年,中山大学肿瘤防治中心肿瘤资源库建成了世界上最大规模的鼻咽癌样本库,截至2017年年初,肿瘤资源库平台已采集并保存血标本及相关信息超过16.5万例、组织标本及相关信息3.6万份以上^[16]。2009年5月,北京市科委正式启动疾病资源库项目建设工作,并委托首都医科大学牵头成立项目管理委员会、项目专家委员会和项目工作委员会,统一协调首都医科大学附属北京天坛医院、佑安医院、地坛医院等11家研究机构分别承担脑血管病、肝炎、艾滋病等12种疾病的资源库建设工作^[17]。2016年6月,由生物芯片上海国家工程研究中心牵头,一期投资近1亿元人民币的上海张江生物银行开始建设,生物样本资源库的一期工程已经完成,已达100万样本存储能力。二期工程于2017年10月完工,达到1000万样本存储能力。该项目三期工程启动在建,预计总投资3~5亿元人民币,样本储存量达到5000万^[18]。2007年,复旦大学与江苏泰州开展全面科技合作,在中国医药城建立了复旦大学泰州健康科学研究院,以泰州500万常驻人口为中国人群的代表人群,以其中35~65周岁的城乡社区居民作为研究对象,关注中国人群高发的多种慢性疾病(如心脑血管疾病,多种代谢性疾病,消化道肿瘤等),到目前为止,已建成约20万人的社区健康人群队列,并建设了与之相配套的大型队列样本库,库存约150万份,种类包括血液、唾液、齿缝菌、尿液、大便及固体组织样本等^[19]。

3 国内外人类干细胞库建设现状

3.1 国外人类干细胞库建设现状 世界骨髓库(The World Marrow Donor Association, WMDA)建立于1994年,总部位于荷兰莱顿市。WMDA是一个志愿组织,各国骨髓库都可以自愿参加,旨在消除跨国查询、捐献和移植的障碍,让各国骨髓库交流、讨论和共同发展。截至2016年8月底,WMDA由50个国家的74个骨髓库和36个国家的57个脐血库组成。据最新的2016年年报,其脐血库中共保存脐血70万余份,2016年增加脐血存储3.6万份,人类淋巴细胞抗原(human lymphocyte antigen, HLA)-A、HLA-B以及HLA-DR型干细胞捐献者达到9400多

万,其中捐献数量超过100万的国家为美国、中国、德国、波兰和英国5个国家^[20]。

美国骨髓库是WMDA里最大的骨髓库成员,1986年成立,至今已有700多万名志愿者,捐献方式有骨髓捐献和外周血造血干细胞捐献,每年的捐献量为4000多例,截至2016年底已完成8万多例骨髓移植供给^[21]。

德国骨髓库于1991年建立,是WMDA里规模第二大的骨髓库成员,截至2018年2月,共有796万名志愿者,向6万多名患者提供了骨髓^[22]。

英国干细胞库由英国医学研究理事会于2003年耗资900万英镑建成,是全球第一个干细胞库,它的建设目标是收集全球所有的细胞系,并提供给科研人员开展研究,目前已存储了24个可供研究的细胞系^[23]。

澳大利亚生物平台下的干细胞数据库项目由澳大利亚研究理事会于2012年投资建立,研究人员将通过该项目与澳大利亚的其他研究人员、组织和行业实体合作,并开展国际合作,其中,已开展合作研究的机构包括:美国的哈佛干细胞共同研究中心和加州干细胞研究所,加拿大的Lunenfeld-Tanenbaum研究所,英国Sheffield大学,日本京都大学和理化学研究所以及荷兰Leiden大学。2016年,英联邦政府国家合作研究基础设施战略也为澳大利亚生物平台下协作性多能干细胞数据计划捐款了100万美元用于干细胞数据库建设^[24]。

越来越多的干细胞库成为WMDA的一员,成为其联盟成员,可以更好地满足干细胞需求者的移植需求,更有利于资源库的标准化建设,英国干细胞库则注重细胞系储存的全面性。

3.2 国内人类干细胞库建设现状 中国造血干细胞捐献者资料库(China Marrow Donor Program, CMDP),简称“中华骨髓库”,它的前身是1992年经卫生部批准建立的“中国非血缘关系骨髓移植供者资料检索库”,截至2017年底,CMDP捐献造血干细胞采集量突破7000例,登记有240多万造血干细胞志愿捐献者。据最新的CMDP 2016年报,CMDP总库容已超过234万人份,数据更新并上传至WMDA数据库,在库参与检索库容量达99万人份,累计为临床提供造血干细胞6198例,其中向国(境)外捐献247例。2016年全年,为临床提供初次检索查询服务8590人次,接受复查申请2888人次,通过WMDA网站为临床提供错配检索1356人次;年内完成造血干细胞捐献828例,同比提升19%,其中向国(境)外捐献24例^[25]。

中国干细胞信息平台下包括北方、南方、华东和

中科院4个干细胞资源库,截至2017年9月,北方干细胞库已存储有不同物种的多能干细胞、成体干细胞及体细胞673株,并建立了临床级的人胚胎干细胞、成体干细胞及体细胞26株;南方干细胞库利用已建立的高效稳定的诱导性多能干细胞(induced pluripotent stem, iPS)技术平台,建立了羊水、脐带、骨膜、尿液、皮肤、牙周膜、牙髓、牙龈等不同组织来源的iPS细胞系,以及地中海贫血、脊髓性肌萎缩、特纳综合征、肝豆状核变性、系统性红斑狼疮、家族性出血性肾炎、血友病、帕金森、唐氏综合征等疾病的iPS细胞系,总计超过50种;华东干细胞库建立并储存了人胚胎干(embryonic stem, ES)细胞株、疾病患者iPS细胞株、大鼠ES样细胞株、小鼠ES细胞株、小鼠iPS细胞株、小鼠孤雌干细胞株等,总计约40株。同时,对所建的干细胞系进行了系统的鉴定,其中ShhES2和ShhES3 2株Hesc细胞系完全符合国际标准,已通过检验并被收入英国干细胞库的总库。中科院干细胞库已收集、保存了100余种各类胚胎干细胞、iPS细胞、成体干细胞、间充质干细胞和其他细胞系,总库容超过1万株^[26]。

4 我国人类生物战略资源库存在的问题及解决建议

4.1 资源库应建设统一的信息管理系统 近年来我国人类生物战略资源库的信息化建设不断取得进展,但尚未形成统一的信息管理系统,不同存储资源的库之间信息共享机制并未建立。资源库的信息化建设是其资源利用避免成为“死库”的关键进程,成熟的信息管理系统能够对库中存储的核酸、细胞以及生物样本等实现包括搜索、保存、分析和管理等功能。采用分布式架构的信息管理系统,针对不同的专业资源库进行不同功能权限的设置,主系统则提供公共计算服务。

4.2 资源库的建设应进行合理规划 国内各地人类生物战略资源库建设如火如荼进行,如2001年天津协和干细胞库正式运营、2009年北京市科学技术委员会启动“北京重大疾病临床数据和样本资源库项目”等,而在什么样的自然以及社会环境的地点建设资源库,建设什么类型的资源库,多大规模的资源库才能满足当地人民的健康需求等诸如此类的问题都需要当地政府甚至是国家层面根据当地人口密度、存储地自然环境、医疗科研需求以及人民健康需要进行合理规划。

4.3 建立完善的安全机制 人类生物战略资源是当前国际生物技术及其产业发展竞争的焦点,国外研究机构以科研合作的名义获取国内基因资源的事件屡见不鲜,因此在资源库建设过程中资源以及信

息的安全保护机制必不可少。在管理系统建设方面通过设置身份认证、对不同身份用户设置不同权限、使数据发送后不可抵赖的不可否认性、保证数据的完整性以及对数据进行加密等方式保护信息数据的安全,同时对管理人员进行规范培训。在资源储存方面通过入库之前进行微生物检测等保证资源质量、建立起一套资源库管理标准来规范人类生物资源从采集到入库以及储存过程中的管理与质控。

4.4 提高资源库资源供给与利用能力 建设资源库的目的就是利用资源,国内目前很多商业用途甚至公益用途的生物样本库仅仅提供存储服务,致力于延长存储时间,但从科研人员的角度出发,存储的目的是用于科研,过长时间的存储并不是建资源库的主要目的,而应该是资源的供给和分析利用。与国内的超算中心如天津中心、广州中心以及深圳中心等建立合作,利用其计算能力结合自身存储资源的特点开发科研平台提高资源库的资源数据分析利用能力,这不仅可以辅助科研,更能更好的满足人民的健康需求和支持相关产业发展。

4.5 建立资源库联盟整合现有资源 在已有的人类生物战略资源库的基础上搭建资源库联盟,以资源库之间的信息整合为主线,为资源库信息的共享、科研协作、人才交流以及标准支撑提供平台。比如WMDA、CHTN、BBMRI等联盟组织就致力于打造合作平台,提供信息共享、分析等功能,越来越多的资源库加入联盟,为人类健康事业的发展贡献力量。

我国拥有丰富的人类生物战略资源,尽管国内人类生物战略资源库的数量以及规模也在不断扩张,存储的资源质量也在不断提升,但通过了解对比国内外资源库的建设现状,发现国内尚缺乏统一的管理和规划,质量提升的同时也应配套规范的标准。因此,在不断加快资源库的建设规模的同时提供高质量的资源,科学规范的管理资源、快速安全的共享资源以及保证资源库的可持续性,是我们亟待解决的问题。

【参考文献】

- [1] 刘旭.我国生物资源开发利用发展战略[J].生物技术,2008,(6):1.
- [2] European Bioinformatics Institute.The European Bioinformatics Institute Impact Report 2016[R/OL].Cambridge:EBI,2016.
- [3] National Institute of Genetics.DDBJ Annual Report 2015[R/OL].Yata:NIG,2016.
- [4] DDBJ Center[EB/OL].[2018-5-8].https://www.ddbj.nig.ac.jp/index-e.html.

- [5] GenBank[EB/OL].[2018-1-24].https://www.ncbi.nlm.nih.gov/genbank/statistics/.
- [6] 国家基因库[EB/OL].[2018-1-26].https://www.cngb.org/news/683d1ae97dbb43d0aa48e2e776d1ea28.html.
- [7] 崔正波.中国国家基因库:功能及意义[J].生态经济,2016,32(12):10-13.
- [8] 国家人类遗传资源共享服务平台[EB/OL].[2018-5-8].http://www.egene.org.cn/cms/g-index.jhtml.
- [9] UK Biobank[EB/OL].[2018-2-5].http://www.ukbiobank.ac.uk/about-biobank-uk/.
- [10] UK Biobank Imaging study[EB/OL].[2018-2-5].https://imaging.ukbiobank.ac.uk/.
- [11] Danmarks Nationale Biobank[EB/OL].[2018-3-2].http://danishnationalbiobank.com/downloads.html.
- [12] Cooperation Human Tissue Network[EB/OL].[2018-3-2].https://www.chtn.org/.
- [13] BBMRI-ERIC[EB/OL].[2018-3-2].https://directory.bbmri-eric.eu/menu/main/home.
- [14] Korea National Research Resources Center[EB/OL].[2018-3-2].http://www.knrrc.or.kr/english/rrc/operation.jsp.
- [15] The Canadian Tumour Repository Network[EB/OL].[2018-3-26].http://www.ctrnet.ca/about/us.
- [16] 中山大学肿瘤防治中心肿瘤资源库[EB/OL].[2018-3-2].http://www.sysucc.org.cn/Department/ArticleList.aspx?CID=168.
- [17] 北京重大疾病临床数据和样本资源库[EB/OL].[2018-5-8].http://139.129.128.68:8080/zdjbclcsjbyzyk/index.shtml.
- [18] 叶扬,张可浩,郜恒骏.生物银行:集约化生物样本库新模式[J].实用器官移植电子杂志,2017,5(6):415-417.
- [19] 复旦大学泰州健康科学研究院[EB/OL].[2018-5-8].http://fdtzihs.org.cn/sy.
- [20] The World Marrow Donor Association[EB/OL].[2018-3-2].https://statistics.wmda.info/.
- [21] National Marrow Donor Program[EB/OL].[2018-3-2].https://bethematch.org/workarea/downloadasset.aspx?id=8755.
- [22] DKMS[EB/OL].[2018-3-2].https://www.dkms.de/en.
- [23] The UK Stem Cell Bank[EB/OL].[2018-3-5].http://www.nibsc.org/ukstemcellbank.
- [24] Bioplatforms Australia Infrastructure[EB/OL].[2018-3-26].http://www.bioplatforms.com/stem-cells-pr/.
- [25] 中华骨髓库[EB/OL].[2018-3-5].http://www.cmdp.org.cn.
- [26] 中国干细胞资源库[EB/OL].[2018-3-5].http://www.chinastemcell.org.