



工作人员在国家基因库内整理全血样本。新华社发

# 中国 诺亚方舟起航

## ——揭秘搭载生物遗传密码的国家基因库

### 基因库里都有啥？

在西方神话中，诺亚建造了一艘方舟，带着各种牲畜、鸟类等，躲过了大洪水，安然渡过“世界末日”。

一粒种子、一个细胞、一管血液、一口唾沫、一段脱氧核糖核酸、一条数据……这些不起眼的“现在”可能是构建未来生物科技和产业的砖石。在现实世界中，美国、欧盟和日本都拥有世界级基因库，由此掌握生命经济时代的战略性资源。

前不久在深圳开始运营的国家基因库正是带着“留存现在、缔造未来”的使命诞生。中国将拥有这样一艘“诺亚方舟”，承载人类及其他生物的遗传样本和密码。

### 我们的“生命银行”

“高大上”的国家基因库看似很遥远，但其实跟每个人都息息相关。华大基因董事长汪建说，每个人一生中所有关键阶段的标本都应该永久保存起来：从出生时的干细胞，到20岁时的免疫细胞，到30岁时的生殖细胞……而国家基因库，就是储存这

它是人类博物馆、植物园、动物园、海洋馆、微生物馆的集合，已存样本1000万份

复同意深圳依托华大基因研究院组建国家基因库。

国家基因库位于深圳市大鹏新区观音山脚下，一期占地面积4.75万平方米。这里几乎与外界隔离，恍若世外桃源。走进建在层层梯田之上、室内绿意盎然的建筑，从一层上三层，依次会看到庞大的高通量基因测序房、超级计算房以及冷冻资源房。

“三库两平台”，是国家基因库主任梅永红描述基因库内容时使用频次最高的一个概念。

据梅永红介绍，与美国、欧盟、日本其他三大世界级的基因库不同，中国国家基因库不仅仅是数据库，而是国际上现有的各类生物样本库、数据

库、生物多样性库、疾病库等的综合升级版。除了“干库”（即基因、蛋白、分子、影像等多组学生物信息数据库）、“湿库”（多样性生物样本和物种遗传资源库）之外，国家基因库还引入了“活库”，即生物活体库，包括动物资源、植物资源、微生物资源和海洋资源等。

不妨把它想象成这样一个载体：它是人类博物馆、植物园、动物园、海洋馆、微生物馆的集合，这里存储的将不仅是一个个标本原型和实物，还有一粒种子、一个细胞、一管血液、一口唾沫、一段脱氧核糖核酸、一条数据……

“两平台”指的是数字化平台和

合成与编辑平台。“三库两平台”的建设，最终为的是解决生命数据的“存、读、懂、写、用”5个作用。

“通俗地说，就是我们要把全球的生物资源都收集起来，用测序仪读取万物的遗传数据，用超级计算机算出结果，用合成与编辑平台写出生命代码，最后用来为人类服务。”梅永红说。

目前，国家基因库已存储多种生物资源样本1000万份；初步建立基因信息数据和生物样本采集、存储、管理相关标准和技术规范，发布深圳市地方标准5项，申请国际、国内标准10项，申请国内外专利46项，出版基因资源专著8本。目前已开展国际/国家重点科研项目20余项，合作项目共发表论文140余篇，其中在《自然》《科学》《细胞》等知名科研杂志上发表论文30余篇。

### 超越国界的平台

人类越来越认识到基因资源以及保护地球生物多样性的重要。国际上，挪威斯瓦尔巴全球种子库、美国自然历史博物馆、英国生物样本库等应运而生，尤其是美国、欧洲、日本先后建立了大型基因数据库，这三大库里的生物信息数据几乎涵盖所有已知的脱氧核糖核酸、核糖核酸和蛋白质数据。

国家基因库执行主任徐讯说：“中国亟须这样一个平

台，从国家层面对具有中国特色的生物样本和基因数据进行有效保存、管理、合理利用。”

“基因库是生命科学的‘国库’，比银行的金库还要宝贵。”梅永红说，农耕时代的核心资源是耕地，工业时代是能源，而生命科学时代则是基因。基因资源目前已成为重要的国家战略资源，如精准医学很大程度上依赖对基因解读的结果。“未来精准医学方面的发展和竞争，一定程度上是基因资源的获得与解析”。

国家基因库将保持开放姿态。目

前，国家基因库已与联合国粮食及农业组织等全球100多个组织和科研机构建立战略合作关系

些样本和数据的地方，像我们的“生命银行”。

“这类似于在人生的不同阶段给自己拍照片，如果我们不拍，那就永远拍不到了。”汪建说，“老照片只能拿来作留念，但是我们在不同阶段存进‘生命银行’的年轻样本，却是在我

们越老的时候越有用，甚至可能在关键时刻救命。”

存储这些数据到底有什么作用？

从数年前开始，汪建就开始有意识地存储自己的健康数据。

这样做的结果是，他对自己身体变化状况了如指掌，并且根据这些数据设计自己

的饮食、运动和生活节奏，“对抗”衰老。他说，随着国家基因库存储容量的增加，未来将会有更多人可以储存和掌握自己的健康数据，过上更健康的生活。

除了对个人，国家基因库对保护生物多样性也有重要意义。目前，保护人类生存环境，保护物种的多样性已刻不容缓。汪建说，对濒危的物种，我们需要尽快地将这些资源存储起来，向子孙后代做一个交代。

多样性，“我们所有人都有职责来保护这些养活我们的种子”。

“国家基因库是开放的平台，一个基因库拥有的资源总是有限的，我们希望通过更好的机制，让大自然馈赠给人类的基因资源能被我们充分地认知和利用。”梅永红说，未来国家基因库一定会超越国界，在全球范围内集聚更多资源。

“国家基因库也面临着很大的挑战，尤其是如何跟国际公约、标准接轨。”联合国粮食及农业组织助理总干事王初说，“但是我们相信，国家基因库将在未来对基因科技的发展产生深远影响。”

记者 马晓澄 孙飞 肖思思

（据新华社深圳电）

### 国家基因库

#### 资源样本库

生物资源：保存种质资源、保存珍贵特有遗传资源

生物样本：人类、动植物、微生物、组织、细胞、血液、DNA、RNA、蛋白质、小分子

#### 生物信息库

数据库：生物信息检索

生物信息数据中心：云计算

#### 产业应用

农业：人工智能、基因编辑、人工智能预测

医学健康：健康数字化、精准医疗

制图/红胄

### 相关链接

#### 世界三大基因数据库

美国、欧盟、日本等发达国家和地区先后建立了NCBI（美国国立生物技术信息中心）、EMBL-EBI（欧洲生物信息研究所）、DDBJ（日本DNA数据库），这三大基因数据库的生物信息数据几乎涵盖了所有已知的DNA、RNA和蛋白质数据，将在未来生命科学大数据时代对基因资源的开发利用形成优势。

#### 美国国立生物技术信息中心

美国国立生物技术信息中心（NCBI）成立于1988年，其目标是发展新的信息学技术来帮助对那些控制健康和疾病的基本分子和遗传过程的理解。

NCBI所提供的可用数据涵盖了部分基因的代表性短序列、完整的基因组、蛋白质结构以及一些遗传疾病的临床描述。NCBI提供了一系列的计算工具以帮助分析各种类型的数据。

#### 欧洲生物信息研究所

欧洲生物信息研究所（EMBL-EBI），是一个非盈利性的学术机构，致力于以信息学手段解答生命科学问题。该所建立于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。

EMBL-EBI拥有超过20年生物信息学研究和服务经验，是全球收集和传播生物数据、提供免费生物信息服务的欧洲节点。该所管理维护着世界最全面的分子生物数据库，其中很多是生物学家熟悉的数据库。

#### 日本DNA数据库

日本DNA数据库（DDBJ）建于1984年。DDBJ主要向研究者收集DNA序列信息并赋予其数据存取号，信息来源主要是日本的研究机构，亦接受其他国家呈递的序列，数据库通过WWW环球网、匿名FTP、e-mail或Gopher方式为广大研究人员服务。

DDBJ开发了SOMate工具，用来搜索基因或蛋白质中短的碱基或氨基酸序列区域，并建立了简便且易操作的SOAP（simple object access protocol）服务器。

（史辑）

国家基因库外观  
新华社发