

# 农业育种数据的保护与利用策略

蒋海英<sup>1</sup> 陈玉冲<sup>2</sup>

(<sup>1</sup>广西民族大学相思湖学院, 南宁 530225; <sup>2</sup>广西壮族自治区农业科学院, 南宁 530007)

**摘要:** 育种数据既是农业育种研究的基础,也是推动农业可持续发展的重要资源。传统的田间试验记录和基因组、转录组、翻译组、表观组、蛋白质组、代谢组等组学数据以及表型数据是育种数据的主要来源,具有多维性、动态性和共享性等特点。当前农业育种数据保护与利用存在数据分散与重复、数据产权归属不明确、数据共享开放利用与保护不平衡、海量数据带来数据存储与管理挑战等困境。针对这些问题,提出了完善数据产权制度、建立统一的农业育种数据库、推动数据共享机制、提高数据分析技术、推动国际合作与交流等农业育种数据保护与利用策略。随着人工智能技术的飞速发展,我国农业育种也将进入智能化时代,农业育种数据在智能化育种和农业现代化进程中将发挥更重要的科学价值。

**关键词:** 农业育种; 数据保护; 数据产权; 开放共享; 数据利用

## Protection and Utilization Strategies of Agricultural Breeding Data

JIANG Haiying<sup>1</sup>, CHEN Yuchong<sup>2</sup>

(<sup>1</sup>Xiangsihu College of Guangxi Minzu University, Nanning 530225; <sup>2</sup>Guangxi Academy of Agricultural Sciences, Nanning 530007)

随着全球人口的不断增长和气候变化的加剧,农业作为支撑人类生存和发展的基础产业,面临着前所未有的挑战。种子是农业的“芯片”,也是农业现代化的核心要素<sup>[1]</sup>。农业育种是农业生产发展、粮食安全和环境保护的重要手段,在现代农业生产中发挥着关键作用。近年来,我国在生物育种和智慧育种创新领域取得了显著进展,已成为全球第二

大育种创新强国。然而,在新一代智慧育种技术领域,我国仍与美国存在明显差距<sup>[1]</sup>,而农业育种数据在这一过程中扮演着重要角色。农业育种数据是对农业育种过程进行记录并可以鉴别的符号,包括通过基础研究、应用研究和试验开发等产生的数据,以及通过考察调查、观测监测、检验检测等方式取得并用于农业育种活动的原始数据及其衍生数据。它既是农业育种研究的重要基础,也是推动农业可持续发展的重要资源。随着农业育种工作的不断深入,数据的积累和应用日益广泛,农业育种数据的有效

基金项目:广西民族大学相思湖学院2023年度校级科研项目(2023XJKY19)

通信作者:陈玉冲



1066

[25] 孙清明,李永忠,向旭,陈道明,杨晓燕,方静,吴绪波,周东辉,马帅鹏,马文朝. 利用 SNP 和 EST-SSR 分子标记鉴定荔枝新种质御金球. 分子植物育种,2013,11(3):403-414

[26] Ding F, Li H R, Wang J Y, Peng H X, Chen H B, Hu F C, Lai B, Wei Y Z, Ma W Q, Li H L, He X H, Zhang S W. Development of molecular markers based on the promoter difference of *LcFT1* to discriminate easy- and difficult-flowering litchi germplasm resources and its application in crossbreeding. BMC Plant Biology,

2021,21(1):539-549

[27] Chen J D, He W Z, Chen S I, Chen Q Y, Ma J Q, Jin J Q, Ma C L, Moon D G, Ercisli S, Yao M Z, Chen L. TeaGVD: A comprehensive database of genomic variations for uncovering the genetic architecture of metabolic traits in tea plant. Frontiers in Plant Science, 2022, 28(11):1-6

[28] 陈小央,李燕,吴早贵. 浙江省农作物种质资源保护实践和路径优化研究. 浙江农业科学, 2024, 65(12):2812-2818

(收稿日期:2025-02-05)

保护和合理开发利用成为了目前学界和产业界共同关注的问题。本文将从农业育种数据的来源与特点、保护与利用中的困境以及相应的策略等方面展开探讨。

## 1 农业育种数据的来源与特点

在农业新品种选育过程中,对育种各项数据信息的掌握和了解直接关系到育种的成败。从18世纪首次获得人工杂交种以来,长期的育种过程积攒了大量的农业育种数据。20世纪以来,随着遗传学的发展,特别是分子生物学技术的广泛应用以及高通量测序技术的发展,产生了海量作物育种相关基因及其表达数据,农业育种进入了数据驱动的新时代。总体而言,农业育种数据既有传统的田间试验记录,也有大量的组学数据、室内检测等实验数据。

**1.1 田间数据** 农业育种是一项复杂的系统工程,涵盖种质资源鉴定与创新、新基因发掘、育种技术研发、品种培育、种子生产及产业化等多个环节,围绕新品种选育过程产生了大量的田间数据,包括各农艺性状的田间调查结果,也包括土壤、气候、水分等动态农业环境数据以及农业生产管理数据<sup>[2]</sup>。其中,品种的农艺性状和生长环境等农田数据,为农业生产提供了最直接的科学依据。

**1.2 实验数据** 除了田间试验和调查,科学实验也是农业育种的重要手段。分子标记、转基因、分子设计、基因编辑等现代育种技术的发展不仅极大地推动了农业育种进程,也产生了海量的基因组、转录组、翻译组、表观组、蛋白质组、代谢组等组学数据<sup>[3]</sup>以及由现代智能仪器自动化记录的表型数据。这些数据成为现代农业育种数据的重要组成部分。

**1.3 农业育种数据的特点** 农业育种数据涵盖遗传特性、环境适应性等多方面信息,数据属性复杂,横跨多个学科领域。同时农业育种数据既有田间试验观测、调查的数据,也有开展科学实验的实验数据,类型多样而且数量庞大,具有类型多、来源广、数据海量、时空分布广、结构复杂的特点。农业育种服务范围广,在服务农业科学研究的同时也服务于政府管理决策和农业生产经营。因此,农业育种数据也需要在农业科技人员、政府部门管理者、企业经营者和农民间广泛共享,具有多维性、动态性和共享性的特点。这些特点使得农业育种数据在优化农业生产结构和提升粮食安全方面发挥着基础性的作用,

同时也是目前农业育种数据保护与利用困境的重要因素。

## 2 农业育种数据保护与利用的困境

**2.1 数据分散与重复** 我国农业育种数据分布广泛,不同的研究机构、企业、个人以及田间试验点都有各自的数据记录。由于缺乏统一的标准和平台,农业育种数据往往存在分散、重复且难以整合的情况,存在严重的“数据孤岛”和“数据海洋”问题<sup>[2]</sup>。我国曾在育种领域建立过政府主导的全国大协作的科研体制,这种体制强调全国一盘棋,育种单位集中调配人力、物力资源,无私共享种质资源和技术成果,集中力量取得了杂交水稻等重大突破<sup>[1]</sup>。目前我国科研机构分散,信息交流和育种资源共享不足,科研人员与企业长期合作开发意愿不高,种业企业将育种材料、试验数据、育种方法等育种数据视为商业机密严格保护<sup>[4]</sup>,原有的举国体制优势逐渐弱化。农业育种数据的分散和重复严重影响数据的综合应用和共享,进而影响育种工作的效率和新品种的培育速度。

**2.2 数据产权归属不明确** 在农业育种中,育种数据涉及多个方面的贡献,包括科研机构、企业和个人等,不同主体可能对同一数据有不同的贡献。由于我国数据产权立法缺位,使得数据产权的归属与分配无法可依。农业育种数据归谁所有、由谁使用、使谁获益等产权归属问题尚且缺乏明确的科学界定,保护路径不够畅通<sup>[5]</sup>。数据产权归属不明确不但影响了农业育种数据的开放共享和应用效率,制约农业育种工作的进一步发展,而且容易导致产权流失,使得农业育种数据缺少妥善保存,甚至被不当披露或收编入他人权利的控制范围。

**2.3 数据开放共享利用与保护不平衡** 农业育种数据作为一种重要的农业信息资源,实现其共享不仅是现代种业科技发展的基础,也是农业科技发展的牵动力。除了生产过程复杂、数据主体多元、数据存放分散等瓶颈问题外,数据的产权保护也是严重制约农业育种数据共享利用深度和广度的重要因素。特别是在商业化育种中,数据的共享范围需要合理界定。忽视知识产权保护,随意使用数据将打击数据所有者的积极性,而过度强调保护则会限制数据流动,造成资源浪费<sup>[6]</sup>。只有通过育种数据的科学合理共享,才能够整合各方面的研究和成果,提

高育种效率,并推动农业育种领域的合作与交流,激发更多的创新和突破。2018年3月国务院办公厅印发《科学数据管理办法》,明确了各部门、单位科学数据开放共享的职责以及共享与利用的相关要求。但总体而言我国数据开放共享的知识产权保护体系尚未健全、数据共享管理制度还不够完善,农业育种数据生产、管理、使用等利益攸关方“不愿、不敢、不会”的问题突出,数据在拥有群体“内部封闭式循环”现象仍较为普遍,对外开放共享较难。

**2.4 数据存储与管理的技术挑战** 育种技术的发展大致经历了4个阶段:“1.0”时代为驯化选育;“2.0”时代为杂交育种;“3.0”时代为分子育种;“4.0”时代则是智慧育种,即融合生命科学、信息科学与育种科学的智能设计育种<sup>[1]</sup>。随着现代生物技术的发展,农业育种数据量呈指数级增长,尤其是组学数据和表型数据的规模越来越大。如何高效存储、管理和分析这些数据成为新的课题。科学数据是科技创新系统中最活跃的要素之一,通过数据驱动创新进而推动经济社会发展,利用科学数据打造科技创新优势已成为国际社会的普遍共识<sup>[7]</sup>。然而,育种数据量的快速增长确实对数据保护和应用产生了多方面的影响。积极方面,更多的数据意味着更全面的信息,有助于提高育种工作的准确性;丰富的大数据为人工智能等技术提供了更多支持,促进了育种研究的进步;多样化的数据可以更好地反映不同环境下的农业情况,提升育种的适应性。但消极方面,随着数据量的增加,数据保护的难度也相应增大。例如,在数据分类、存储和管理上需要更多的资源和更高效的机制。此外,随着数据量的增加,确保数据的质量和准确性变得更加重要,这对数据处理提出了更高的要求。同时,数据之间的关联性可能变得更加复杂,需要更深入的分析 and 整理。

### 3 农业育种数据保护与利用的策略

**3.1 完善数据产权制度** 数字经济已成为继农业经济、工业经济之后的主要经济形态,是重组全球资源、重塑经济结构和改变竞争格局的关键力量<sup>[8]</sup>。中国数字经济规模连续多年位居世界第二,数字经济已经成为我国经济发展的新引擎、新动能,对培育发展新质生产力,提升产业链、供应链韧性具有强大支撑作用。数据作为数字经济时代重要的生产要素,已成为国家基础性战略资源。数据所有权、数据生

产权、数据使用权、数据出版权、数据交易权等数据权属问题是网络 and 信息技术发展中的新挑战,其根源在于个人、企业和国家在数字化和虚拟化过程中,数据的流动性和资源属性不断增强<sup>[5]</sup>。在保护与利用之间找到平衡点是关键。建议建立明确的数据产权归属机制,例如通过法律或协议明确各方权利和义务,特别是在商业化育种中,确保数据在共享的同时,其创造者也能得到合理回报。

**3.2 建立统一的农业育种数据库** 在数字经济时代,科学数据的融合汇聚是形成新质生产力的关键<sup>[9]</sup>。为了更好地促进数据的共享与利用,可以考虑建立覆盖全国的统一农业育种数据库。该数据库需要具备灵活性和扩展性,能够涵盖不同维度的数据。如通过新型举国体制突破种业战略前沿和关键核心技术,对于应用技术、数据模拟、测试试验等不适合以论文、专利呈现成果的支撑性服务,以任务完成进度和质量为考核标准,促进全国大协作<sup>[1]</sup>。利用大规模农业育种数据库、富集算法和生物数字模型,结合创新方法应用,通过基因挖掘和育种技术,打造以用户需求为导向的大数据育种技术平台<sup>[2]</sup>。可将高价值的描述项置于显著位置,并优化呈现方式,例如提供数据快照、关键图表缩略图等直观信息,并折叠用户关注度较低的描述项,以突出关键信息,从而建立用户友好型的数据共享平台<sup>[10]</sup>。

**3.3 推动数据共享机制** 构建互惠共赢的数据开放共享机制,通过数据交换、服务交换、资助性状功能鉴定评价、数据标记以及购买数据等方式,引导种业企业共享数据,引入国际育种数据资源,丰富数据来源,并持续优化育种数据的质量和结构<sup>[1]</sup>,建立便利的数据获取机制<sup>[11]</sup>。同时也要注意,未共享本应共享的数据会造成资源浪费,而违规共享不应共享的数据则会威胁数据安全<sup>[12]</sup>。因此明确产权、维护主权对于保障农业育种数据的安全共享至关重要。在明确育种数据的归属和使用权的基础上,鼓励数据的开放共享,平衡数据所有权和使用权的关系,促进数据的流动和应用,参考国际经验,探索多样化的数据使用模式,如数据的开放获取、数据银行等。

**3.4 提高数据分析技术** 面对海量数据,需要进一步提升数据管理与分析的技术水平。引入人工智能、大数据处理等技术能帮助更好地挖掘数据价值,并为农业育种提供更精准的支持。加速布局多维组学

大数据和智慧育种基础平台建设,系统整合各育种基地和研究机构的育种数据<sup>[1]</sup>。通过优化数据分析方法全面挖掘数据中的信息,解决育种数据量的增加和数据类型的多样化问题;通过引入先进的分析工具和技术,如机器学习和人工智能,提高分析效率和准确性;通过标准化的数据格式和统一的标准,确保不同来源的数据能够有效整合和比较;通过加强跨学科的合作,结合生物学、统计学等领域的知识,共同推动育种数据分析技术的发展。

**3.5 推动国际合作与交流** 农业育种数据的保护与利用不仅关系到国内农业生产,也对全球粮食安全具有重要意义。通过加强国际间的合作与交流,可以共同应对数据保护与利用的挑战。通过建立国际协作网络、举办国际会议和研讨会等方式,各国科研机构和企业可以分享育种数据,促进育种学术交流和育种经验分享,共同应对气候变化和粮食安全挑战。

#### 4 农业育种数据的应用前景

在人工智能发展的历程中,每一次范式变革不仅体现了技术手段的迭代升级,更深刻地揭示了人类在自我认知、社会组织模式及技术文明路径构建上的根本重塑<sup>[13]</sup>。如杭州深度求索人工智能基础技术研究有限公司于2025年1月发布的DeepSeek-R1推理模型,在有限的GPU资源下,在较短的时间内成功训练出世界一流的闭源推理大模型,并以开源形式提供给全球研究人员和开发者使用,展示了真正的开放精神,开启了一扇通往无限可能的大门<sup>[14]</sup>。数字科技的快速发展为我国缩小与种业强国的差距、构建精准高效的智能设计育种技术体系提供了机遇<sup>[4]</sup>,中国种业借助数字科技的力量,有望实现跨越式的发展<sup>[2]</sup>。随着人工智能技术的发展,农业育种也将进入智能化时代。如生猪全域数智育种系统融合了人工智能与基因技术,不仅显著提升了育种效率,还大幅降低了技术门槛,使育种过程更加简单高效<sup>[15]</sup>。可以预见,中国按照市场化、产业化模式,聚焦多组学、基因编辑、全基因组选择、智能设计育种等关键技术,开展原创性技术攻关,创建基于基因型—表型—环境多维大数据的智能化育种方案,将有力提升中国育种科技创新能力,推动中国种业迈向“4.0”时代<sup>[16]</sup>。在数字技术的支撑下,通过建立统一的数据标准和共享机制,建立科

学的数据管理体系,加强国际合作,更好地发挥育种数据的作用,为保障粮食安全和促进农业可持续发展提供有力支持。在这个数据驱动的时代,只有有效保护和合理利用育种数据,才能确保我国在国际农业科技竞争中占据优势地位。实现数据的高效整合与应用,为品种选育提供科学依据,这不仅加快了育种进程,提高了育种效率,而且有助于培育出更多适应不同环境条件的优良品种,推动农业现代化向着更高层次发展。

#### 参考文献

- [1]程郁,叶兴庆,宁夏,殷浩栋,伍振军,陈凯华.中国实现种业科技自立自强面临的主要“卡点”与政策思路.中国农村经济,2022(8):35-51
- [2]樊龙江,王卫娣,王斌,叶楚玉,舒庆尧,张辉.作物育种相关数据及大数据技术育种利用.浙江大学学报:农业与生命科学版,2016,42(1):30-39
- [3]王彦青,陈婷婷,张思思,朱军伟,陈焕新,肖景发,宋述慧,章张,赵文明,鲍一明.生命组学大数据安全管理实践.农业大数据学报,2024,6(3):325-332
- [4]范贝贝,李瑾,马晨.我国作物种业数字化发展:成效、困境与前瞻.中国农业科技导报,2022,24(12):25-32
- [5]庄严,杨帅,刘照坤,樊景超,周舒雅.农业科学观测数据权属与保护路径研究.农业大数据学报,2020,2(4):107-112
- [6]冯媛.科学数据开放共享的知识产权治理研究.图书馆,2023(2):56-62
- [7]施栩婕.学科服务视角下高校科研人员科学数据管理问题及解决策略研究.大学图书馆学报,2025,43(1):17-26
- [8]孙姗.我国农业数据发展现状及确权必要性.农业工程,2024,14(4):26-35
- [9]欧阳日辉,李晓壮.科学数据形成新质生产力的理论逻辑、典型案例与发展策略.科学管理研究,2024,42(6):61-71
- [10]范智萱,王健,撒旭,张贵兰.用户视角下农业科学数据描述信息的“结构-效用”研究.农业图书馆学报,2022,34(10):57-69
- [11]林秀芹.数据治理的域外经验与启示.人民论坛·学术前沿,2023(6):48-65
- [12]罗娇,刘细文.知识产权视角下科学数据安全管理的策略选择.图书馆情报工作,2021,65(12):38-46
- [13]令小雄.DeepSeek开启后ChatGPT时代——基于数字范式革新及其运演哲思.西北工业大学学报:社会科学版,2025(2):51-59
- [14]张慧敏.DeepSeek-R1是怎样炼成的?.深圳大学学报:理工版.<https://link.cnki.net/urlid/44.1401.N.20250210.1628.002>
- [15]高远飞.生猪全域数智育种系统补齐行业短板.猪业科学,2024,41(9):33-34
- [16]范贝贝,李瑾,冯献.农业强国目标下作物育种科技与装备创新:态势、挑战与路径.科技导报,2023,41(16):23-31

(收稿日期:2025-02-24)