张之奇,王兴涛,任端阳,等.多维数据链驱动与智能决策协同:智能育种技术体系的系统构建与遗传增益跃迁[J].华中农业大学学报,2026,45(3):1-17.DOI:10.13300/j.cnki.hnlkxb.2026.03.001

多维数据链驱动与智能决策协同:智能育种技术体系的系统构建与遗传增益跃迁

张之奇,王兴涛,任端阳,曹文福,王建军

山西农业大学生态农牧研究所,朔州 036001

摘要 智能育种作为现代种业技术革新的核心驱动力,正推动作物遗传改良从经验依赖型向数据驱动型转变。本文系统梳理了智能育种技术体系的构建路径及其在作物遗传改良中的应用进展,旨在为我国种业智能化转型提供理论依据与实践参考。基于高通量表型与基因型数据采集、人工智能算法决策、多组学平台集成等核心技术模块,结合国内外典型案例,综述了智能育种在复杂性状设计、种质资源挖掘与遗传增益预测等方面的突破性进展。智能育种通过构建“基因型-表型-环境”多维数据驱动模型,显著提升了育种效率与精准性。未来应加强生物-信息-工程交叉融合,开发知识嵌入型算法,构建全球协作平台,推动种质资源数字化共享与商业化育种体系重构,建议通过政策引导建立产学研协同机制,加速技术成果向产业端转化。

关键词 智能育种;人工智能算法;基因编辑;虚拟育种模拟;智能挖掘;遗传增益;深度学习模型

中图分类号 S337 **文献标识码** A **文章编号** 1000-2421(2026)03-0001-17

康奈尔大学爱德华·巴克勒教授在2018年首次提出智能育种概念[1]。智能育种是将生物技术、大数据及人工智能技术融合,实现高效且个性化的作物新品种培育,加速育种过程从依赖统计遗传学向由数据驱动的预测性AI模型增强的模式转变[2]。作为现代农业科技革命的代表性领域,智能育种主要依靠人工智能、大数据分析与生物信息学的深度融合,构建由“基因型-表型-环境”多维数据所驱动的作物遗传改良体系,和传统经验育种、分子标记辅助选择相比,智能育种实现了从以经验驱动到以数据驱动的模式转变,其技术框架囊括基因组解析、表型组采集、算法模型创建和决策优化4个部分。

智能育种在作物育种领域的演进可划分成3个阶段:(1)技术萌芽的阶段。依靠高通量测序技术的突破,全基因组选择模型开始被运用到作物杂种优势预测上,但受计算能力和表型数据规模大小的制约,实际应用大多还处于理论研究状态[3]。(2)技术聚合期。CRISPR基因编辑技术和无人机表型平台发展成熟,推动了作物基因图谱的构建,例如“玉米泛基因组计划”[4]完成了针对B73等骨干亲本的基因组注释。(3)智能决策阶段。以Transformer架构为基础的多模态大模型(如CropSight系统[5])将天气数据、土壤信息和基因组变异图谱进行整合,完成作物(比如玉米)品种的按需化设计。最近,我国也发布了基于生物大数据和AI技术的智能育种平台。例如,中国农业科学院作物科学研究所与阿里巴巴达摩院共同开发了一个全流程智能育种平台,实现了育种数据管理与分析、计算加速以及亲本系和优良品种的AI预测的集成,该平台将基因测序数据中的变异位点计算加速了110倍,群体遗传分析与传统技术相比超过1 000倍,基因组选择算法的准确性平均提高了15%,为未来育种工作的精度和效率奠定了坚实基础[6]。

传统育种方法长期依赖表型和经验选择,针对现代农业的需求,其系统局限性已开始显现,具体表现为:

1)基因型-环境互作效应显著减弱了性状选择的预见性。鉴于作物遗传网络的复杂程度极高,上

收稿日期:2025-06-10

基金项目:国家重点研发计划项目(2024YFD1201305);国家联合攻关项目子课题(NK202307020403);山西农业大学科研项目计划项目(CXGC202449)

张之奇,E-mail:465715272@qq.com

通信作者:王建军,E-mail:mlddym@163.com

位性效应以及微效多基因调控模式使亲本选配存在明显的盲目性。常规杂交组合里，大多数重要农艺性状是由多个基因控制的数量性状，其遗传力较低，表型易受环境影响，导致选择效率低下，而理想基因型的筛选成功率一般低于百万分之一，这导致遗传资源的大量浪费[7]。

2）多代的表型驯化使得育种周期特别长[8]。玉米育种周期要历经8～10 a，其中F_2至F_6代群体稳定阶段所占的比例是60%～75%，后续区试和生试环节额外消耗3～4 a[9]。在当今气候波动加剧的环境中，时间成本进一步放大了育种风险，例如2016、2018、2021年河南南部授粉期碰到高温年份，黄淮海夏玉米区由极端高温引起的授粉障碍，使传统育种材料表型稳定性验证面临更为突出的挑战。

3）表型选择受生物学限制，对性状精准改良形成制约，如玉米株型优化与抗逆性提升往往会产生连锁累赘现象，传统回交育种消除非目标基因的连锁概率不高，CRISPR/Cas9基因编辑技术可使育种效率提升至98%以上[10]。过去30 a，全球种质资源遗传多样性的衰减比例变大，核心种质内突破性大效应等位基因的短缺，已形成了实质性创新壁垒[11]（表1为传统育种与智能育种核心技术对比）。

表1 传统育种与智能育种核心技术对比

Table 1 Comparison of core technologies between traditional breeding and smart breeding

维度 Dimension	传统育种 Conventional breeding	智能育种 Intelligent Breeding
决策模式	经验驱动	数据＋AI算法驱动
周期/a	8～10[9]	1～3[9]
性状选择精度	可稳定遗传力低[9]	＞98%[10]
数据整合能力	单一表型	多组学融合

技术架构方面，智能育种囊括三大核心模块（图1）。

1）数据采集。智能育种中的数据采集融合了多学科先进技术，全面获取作物的表型、基因型和环境信息。在表型数据采集方面，通过无人机搭载多光谱、高光谱、热红外和RGB传感器实现大范围、非破坏性的生长状态监测，结合田间机器人、固定式监测平台及手持设备进行高精度生理参数测量，实现从个体到群体的多尺度表型分析[14]。基因型数据依托高通量测序技术、基因分型芯片和分子标记技术，高效解析作物的遗传变异与基因特征[15]。通过物联网传感器网络和卫星遥感实时采集土壤、气候等环境数据，支撑基因型与环境互作研究[16]。利用地理信息系统、云计算和人工智能算法对表型、基因型与环境多源数据进行融合与智能分析，构建预测模型，推动育种决策由经验驱动向数据驱动转变，显著提升育种效率与精准度[17]。

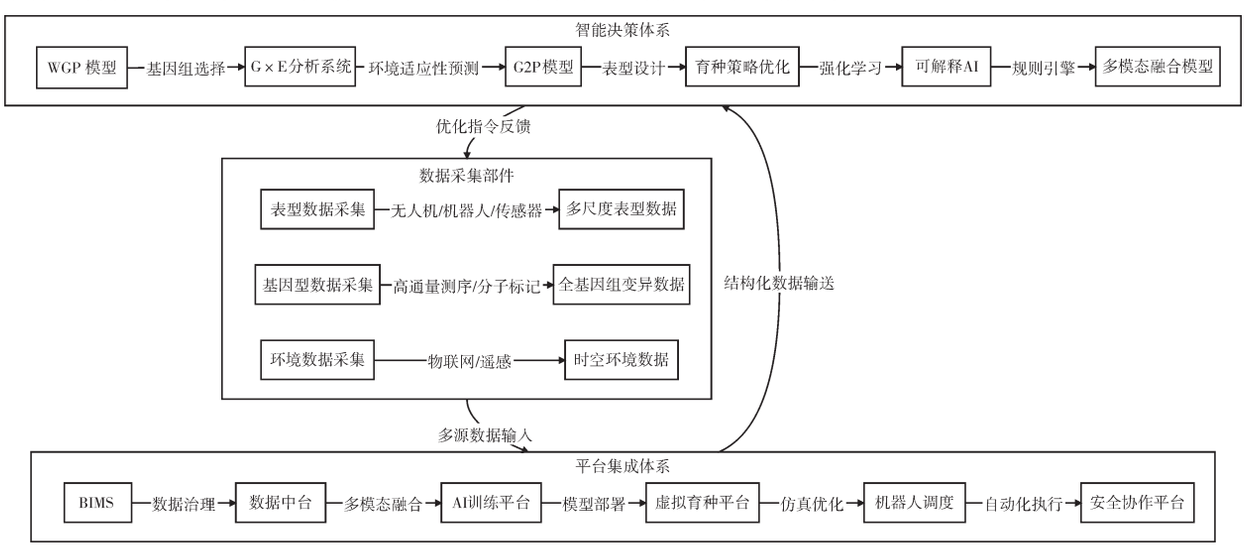


图1 智能育种技术体系框架图

Fig. 1 Framework of smart breeding technology system

2）智能决策体系。它是整个系统的大脑，负责从海量数据中提取知识、做出预测并生成育种策略。它依托人工智能、统计遗传学和系统生物学等方法，将原始数据转化为可执行的育种决策。该体系主要

包括以下6大核心组成部分。

①全基因组预测模型(whole-genome prediction，WGP)。利用基因组-wide SNP标记预测个体的育种值或复杂性状表现。特点是支持早期选择，大幅缩短育种周期，减少田间试验成本。常用模型有GBLUP、Bayesian方法、机器学习(XGBoost)等[18]。

②多环境预测与G×E互作分析系统。系统通过分析基因型与环境(genotype×environment，G×E)互作效应，预测品种在不同气候、土壤条件下的表现，并且其结合气象、土壤、遥感等环境数据，实现“按区域定制育种方案”。支持多环境试验数据的整合分析，提升跨区推广的准确性，Yu等[19]开发了可解释的机器学习框架EXGEP，通过整合基因型、环境数据和SHAP解释技术，实现了玉米产量的高精度预测，并揭示了关键环境与基因型互作特征。

③基因型到表型预测模型(genotype-to-phenotype，G2P)[20]。建立从DNA序列到复杂表型的映射关系，融合基因调控网络、代谢通路模型与深度学习，揭示性状形成的生物学机制，支持“设计育种”—根据目标表型反向推演最优基因组合。

④AI驱动的育种策略优化系统(breeding pipeline optimization)。利用强化学习、运筹学算法优化育种流程，如亲本选配、杂交设计、后代选择强度、测试资源配置等，实现“最大化遗传增益、最小化时间与成本”的智能决策，可模拟不同育种策略的长期遗传进展，辅助制定中长期育种规划[21]。

⑤可解释AI与专家知识融合系统(explainable AI+expert rules)[22]。将AI模型的“黑箱”决策转化为可解释的生物学逻辑，并且融合领域专家经验构建规则引擎，确保AI建议符合农学原理，增强育种家对AI系统的信任，实现“人机协同决策”。

⑥多模态融合分析模型(multimodal fusion models)[23]。整合基因组、转录组、表型组、环境等多源异构数据，提升预测精度。使用多模态门控融合模型、图神经网络等技术，挖掘跨模态关联。实现“从分子到田间”的全链条表型预测。

3)平台集成体系。智能育种中的平台集成体系是连接“数据采集”与“智能决策”的中枢系统，负责实现多源数据的汇聚、管理、分析流程的自动化以及人机交互的协同操作。它本质上是一个高度整合的数字化育种操作系统，支撑整个育种飞轮的高效运转。该体系主要包括以下核心组成部分：

①育种信息管理系统(breeding information management system，BIMS)。系统可以统一管理育种项目中的基因型、表型、系谱、试验设计、田间操作记录等全生命周期数据，特点是支持高通量数据导入、样本追踪、数据版本控制和权限管理，操作系统有BreedBase[24]、TerraPortal等，BIMS是智能育种的“数据库中枢”，确保数据可追溯、可共享、可复用。

②数据中台与多模态数据融合平台[9]。平台可以整合来自基因测序、无人机遥感、地面机器人、传感器网络、实验室检测等多源异构数据，实现时间、空间、尺度对齐，构建标准化的数据库，技术支撑有ETL工具、时空数据库、图数据库、云原生架构，同时可支撑HMD(hyper-dimensional and multimodal data)模块的落地。

③AI模型训练与推理平台(AI engine)。平台可提供机器学习/深度学习模型的开发、训练、验证、部署一体化环境，支持WGP、G2P、图像识别、时序预测等模型的批量运行，集成AutoML、模型版本管理、性能监控等功能，技术支撑有基于TensorFlow/PyTorch的定制化AI pipeline，结合Kubernetes进行分布式计算调度，可以对接“智能决策体系”，实现“数据-模型-预测”的自动化流转。

④虚拟育种与仿真模拟平台(omni-simulated breeding platform)[9]。平台基于生理机制模型、基因调控网络和环境数据，构建“数字植株”或“虚拟田间”，可以模拟不同基因型在多种气候、管理措施下的表现，预演育种结果，支持假设分析，优化杂交组合与选择策略，实行先模拟、后实验，显著降低试错成本。

⑤自动化执行与机器人调度系统(robotics orchestration system)[25]。系统可以控制温室/田间的高通量机器人、无人机、自动灌溉系统等设备；将AI生成的育种指令转化为机器人可执行动作，实现播种、表型扫描、样本采集、授粉等操作的无人化闭环。相应技术有ROS、IoT协议、边缘计算等。支撑PM(peopleless management)模块，打通决策-执行链路。

⑥安全与协同协作平台。平台支持多团队跨地域协作，包括数据加密、知识产权保护、访问审计，API接口开放，便于与外部数据库(ensembl plants[26])或第三方工具对接。

依靠多组学数据融合与虚拟育种模拟，可推动种质资源利用效率指数级上涨。例如崖州湾国家实验室和上海人工智能实验室等单位联合推出首个种业大语言模型“丰登”，它填补了我国AI育种领域的

技术空白，降低了进入育种领域学习的门槛，协助提升从业者研究与实践的效率，同时又为AI助力生物育种寻找了全新的探索路径。中国农业科学院科研团队推出的SeedLLM·Rice（简称SeedLLM）模型，拥有70亿参数，训练数据来自140万篇与水稻相关的出版物，约占全球水稻研究产出的98.24%。初步评估表明，SeedLLM的表现优于通用模型，包括OpenAI的GPT-4和DeepSeek-R1，在水稻特异性任务上的胜率达到57%～88%，SeedLLM已与水稻生物知识图谱相集成[27]。Xie等[28]将生物技术与人工智能深度融合，首次提出作物-机器人协同设计的“双向奔赴”理念，通过基因编辑重新设计作物花型，快速精准创制“机器人友好”的结构型雄性不育系，运用深度学习和人工智能成功研制世界首台可自动巡航杂交授粉的智能育种机器人“吉儿”（GEAIR），打破杂交育种和制种瓶颈，大幅降低育种成本、缩短育种周期、提高育种效率。Li等[29]系统阐述了基因组学、基因编辑、蛋白质设计与高通量表型等多技术协同的作物改良新模式，提出利用蛋白质设计按需打造功能元件的新途径，并描绘了AI辅助的优异种质设计蓝图，有望引领新一轮作物改良革命；中种集团向外界发布“玉米品种晋级与精准定位决策系统”，快速准确地完成品种好坏分析以及适宜种植区域的精准判定，体现了生物技术与信息技术深度融合的优势。

2025年中央一号文件设定的“商业化育种体系重构”目标，正推进智能育种技术由实验室向产业端迅速渗透，促成“数据共享-算法迭代-品种产出”的产学研互助协同机制，这说明我国在冲破种源“卡脖子”技术阻碍的进程中进入了系统化创新阶段。

1 智能育种技术体系构建

1.1 高通量表型组与基因型数据采集技术

现代智能育种体系的根基是搭建精准可控的生物信息采集系统，该技术将先进传感装置与分子检测技术整合在一起，达成从宏观表型到微观基因组的全维度数据采集。该技术的主要突破体现在田间动态监测以及基因组解析这2个关键维度。

1）无人机驱动的田间表型监测。无人机平台通常搭载多光谱、高光谱、热红外、可见光和激光雷达等传感器，能够从空中快速、大范围地获取作物信息。可检测的表型数据主要有以下4个方面。

①生物量与生长动态（biomass & growth dynamics）方面[30]。监测数据主要为植株高度、冠层覆盖度、叶面积指数、生物量等。传感器主要有RGB相机、LiDAR。通量极高，1次飞行可在几分钟到几小时内覆盖数十甚至上百公顷的试验田或农田，实现全生育期的动态监测（如每周或几天1次）。

②营养状况与生理状态（nutritional & physiological status）方面[31]。监测数据主要为叶绿素含量、氮素营养状况、水分胁迫、光合活性。传感器主要是多光谱相机、高光谱相机、热红外相机。通量极高，与生物量监测类似，可快速覆盖大面积，实现田间尺度的营养和胁迫状况空间制图。

③胁迫响应（stress response）方面[32]。监测数据主要为病害、虫害、干旱、盐碱、涝渍等胁迫的早期检测和严重程度评估。传感器主要是高光谱相机（最敏感，可检测病害引起的细微光谱变化）、多光谱/热红外（检测胁迫导致的生理变化，如叶片温度升高、植被指数下降）。通量高，可以快速扫描大片区域，识别出胁迫发生的“热点”区域，指导精准施药或管理。

④产量预测（yield prediction）方面[33]。监测数据为产量的早期预测。方法为结合多个生育期的植被指数、植株高度、冠层结构等数据，通过机器学习模型进行预测。通量高，基于UAV获取的高通量表型数据进行预测，覆盖范围广。

2）物联网驱动的田间表型监测。物联网系统由部署在田间的地面传感器网络（固定或移动）组成，通常包括气象站、土壤传感器、茎流计、微型相机等，通过无线网络实时传输数据。主要有以下3个方面可检测的表型数据。

①微环境监测（micro-environment monitoring）方面[34]。监测数据为实时的空气温湿度、光照强度、风速风向、降雨量、土壤温湿度、土壤电导率、土壤张力。传感器为各类气象和土壤传感器。通量极高，核心优势在于连续、实时的数据采集，时间分辨率可达分钟级甚至秒级。空间分辨率取决于传感器密度。

②植物生理响应（plant physiological response）方面[35]。监测数据为茎秆直径微变化、果实生长速率、叶片角度/运动、气孔导度等。传感器为微位移传感器、小型RGB或近红外相机、茎流计等。通量中等但敏感度极高，通常监测单株或少数植株，但数据是连续、高频的，能捕捉到昼夜节律和对环境变化的瞬时响应。

③精准灌溉与施肥决策支持（precision irrigation & fertilization）方面[36]。监测数据为基于土壤和植物传感器数据，实现按需灌溉和施肥。方法为物联网系统提供实时数据，触发自动化灌溉/施肥系统。通量高，实现了实时、精准的田间管理，是“闭环”智能农业的体现。

依托无人机组网协同的立体监测系统已形成标准化作业模式，具备多光谱成像模块的无人机阵列可同步获取冠层温度、叶面积指数等多项表型参数，它的空间分辨率达到了亚厘米的量级[37]。物联网传感节点网络凭借布设土壤水分、光照强度等环境传感器，建成时间分辨率为 10 min 级的田间微环境动态数据库，最新探索显示，LiDAR 与深度学习模型（尤其是 PointNet＋＋）相融合，能够快速、自动识别高粱植株结构参数，为高通量作物表型分析提供了可靠的技术路线，具有十分可观的应用前景[38]。

3）单细胞测序技术。基于单细胞分辨率的基因组解析技术突破传统批量测序局限，ChIP-Seq[39]联合使用可精确定位调控元件空间构象的动态情况，最新开发的 scATAC-seq＋技术可分析染色质动态调控状况，揭示细胞异质性并提高功能调控研究水平，为深入研究基因表达调控机制提供了更准确且系统的工具[40]。Peng 等[41]首次利用 Long-read ChIA-PET 技术，绘制了玉米活跃表达基因参与的高分辨率三维基因组图谱，鉴定了基因组顺式调控元件三维互作模式，揭示了玉米三维基因组结构调控基因的表达进而影响表型变异的潜在机制。为获得对组织结构和功能的全面理解，Li 等[42]通过 WebAtlas 管线将单细胞和空间转录组数据结合，高效分析了细胞类型映射、空间模式识别及细胞间交互网络构建等方面的数据。Ouyang 等[13]阐明了在植物关键性状形成过程中，DNA 元件、RNA 分子和蛋白质因子之间相互作用的分子机制，并指出了三维基因组结构如何影响基因组稳定性，揭示了空间邻近性对染色体易位的决定作用，证明了三维结构的功能重要性；Ni 等[43]首次报道了大豆的三维基因组图谱，并探讨了未来的发展方向，以期推动大豆 3D 功能基因组学研究和分子育种的进一步提升；Sun 等[44]在拟南芥中利用 Micro-C-XL 技术揭示了染色质互作热点与基因活性、表观修饰的关系，建立了“空间位置-功能”的关联，他们将 Micro-C-XL 技术拓展应用于大豆和水稻，揭示了植物中精细尺度的染色质组织结构以及增强子-启动子间的环状互作，为深入理解植物三维基因组的结构与功能提供了重要见解，也为育种者提供了理论基础；Huang 等[45]阐明了不同组织中亚基因组间的协同与分化过程，为育种者提供了丰富的资源，描绘了基因组的调控架构，有助于深入理解生物学过程，并为棉花育种提供指导；Albert 等[46]通过单细胞测序技术展示了 B 染色体在发育过程中的空间重排，说明三维结构具有动态性，可能参与作物性状与发育调控，为理解性状时空特异性表达提供线索；Schiessl 等[47]从多个尺度对植物染色质组织进行了系统论述，并探讨了三维基因组研究中的新兴观点与概念，重点关注最先进的三维基因组技术、RNA-染色质相互作用、液-液相分离以及染色质的动态变化。

1.2　人工智能算法在育种决策中的应用

从本质上讲，人工智能算法就是告诉计算机如何学习并自主运行的程序，人工智能算法由一套复杂的规则驱动，决定了人工智能程序的执行步骤及其学习能力。没有算法，人工智能就不可能存在[48]。

作为智能育种体系的核心决策中枢，人工智能算法通过解析海量遗传信息与环境互作数据，正在重塑传统育种模式的决策逻辑。在植物育种多环境试验期间，高通量基因分型、表型分析和环境分型技术的整合，为全基因组预测以及处理基因型×环境互作提供了充足的数据支撑，作物生长模型通过整合关键性状及其对环境响应的遗传变异，可模拟不同环境条件下作物的表现，并与 WGP 方法达成耦合，生成基于过程的全基因组预测方法，其中可凭借近似贝叶斯计算等算法实现模型参数的估算。Cooper 等[49]以模拟验证为基础，首次将 CGM（crop growth model）-WGP 运用到 1 个玉米多环境抗旱试验的数据集里，以 1 组双单倍体（DH）群体为对象，在 2 个干旱环境状况下实现了对其杂交种籽粒产量的正向预测，该结果以在 CGM 中引入 5 个事先敲定的重要组分性状的遗传信息为依托，CGM-WGP 不仅具有投入实际应用的可行性，还在应对非加性基因效应的预测任务过程中具备优势；Wei 等[50]构建了基于 8 个 GWAS 队列的水稻数量性状核苷酸综合图谱，该团队开发了 1 个名为 RiceNavi 的革命性基因组导航系统，实现了从基因组知识到育种实践的高效转化。RiceNavi 是一个集成了 QTN 效应信息、群体遗传背景和育种目标的智能决策平台。其核心功能在于 QTN 聚合和育种路径优化，能够为育种家提供精准的指导：选择最优亲本、设计高效的杂交与回交

方案，以最小的世代数和成本，将多个关键的优良等位基因精准地堆叠到同一个优良品种背景中，同时最大限度地规避连锁累赘。

1.3 智能育种平台集成创新

智能育种平台系统集成能力对技术体系落地起着关键支撑作用，其核心是做到数据驱动之下的全流程闭环优化，这里以多组学数据融合和虚拟育种模拟两大维度为切入点展开论述。

1）多组学数据融合分析系统。多组学数据整合已打破单一组学的分析限制，借助构建跨尺度数据关联的知识图谱，实现基因组、转录组、蛋白组与代谢组数据的协同性解析，Subramanian等[51]利用开发自适应数据标准化算法解决了不同实验平台和测序批次间的数据异质性。Pan等[52]开发了1个名为“单细胞分析师”的网络平台，它可针对单细胞RNA测序、单细胞ATAC测序、单细胞免疫分析、单细胞拷贝数变异、飞行时间质谱术、流式细胞术和空间转录组学等6种不同的单细胞组学数据进行处理操作。深度学习技术的发展伴随数据量的增加，为复杂且异构的数据类型融合与分析带来了前所未有的新机遇，不同数据模态下给出的互补信息可被有效利用，对分析研究对象大有裨益。基于深度学习的多组学数据整合采用2种模型，也就是生成式模型和非生成式模型（2个模型间的比较如表2所示）。生成式模型包含变分自编码器、生成对抗网络、生成预训练变换器等，优点是它可借助约束共享来实现某些特定属性或引入先验知识，还可生成或填补缺少的数据模态。Ballard等[53]认为，生成式模型处理不完整数据的表现十分出色，还突破了传统分子组学数据类型的界限，实现了和其他数据模态的有效联合，同时它也可应对数据缺失问题带来的挑战，可借助整合更多类型数据、抓取每个样本的全面特性等形式，快速改善下游任务的性能水平，会对育种工作起到有效支持作用；非生成式模型包括前馈神经网络（feedforward neural networks，FNNs）、图卷积神经网络（graph convolutional networks，GCNs）和自动编码器（autoencoders，AEs）。FNNs可利用信息共享来学习模态间的交互作用，但它不能充分利用样本间的相关性；GCNs可通过相似性网络有效利用组学特征以及样本或数据类型间的相关性来提供生物意义结构，还可以将以前的生物学知识纳入模型；AEs可通过外部生物学网络信息集成，使多个共享同一层的编码器和解码器将模态整合为单一表示。无论是生成式还是非生成式模型，都有助于实现多模态数据的有效整合，并能在一定程度上解决数据不完全引起的问题。

表2 生成式模型与非生成式模型比较

Table 2 Generative models vs. non-generative models

模型类型 Model type	子类别 Subcategory	功能描述 Function description	数据处理能力 Data processing capability	优点 Advantages	缺点 Disadvantages
生成式模型	变分自编码[54]	用于多组学数据整合[55]	能够处理不完整数据，推断缺失信息[54]	鼓励不同模态间的一致性，遵守先验分布[54]	对样本数量要求较高[54]
	生成对抗网络	包括1个生成器和1个判别器，通过对抗过程学习数据分布[56]	处理表格形式的组学数据及成像数据[57]	提升生成类似真实数据的能力，有助于正则化或确保不同数据类型之间的一致性[58]	模型复杂度高，训练难度大，对样本量有较高要求[58]
	生成预训练变换	在大量数据集上预先训练，学习普遍模式后针对具体任务进行微调	—	快速发展的领域，适用于多种任务	—
非生成式模型	前馈神经网络	尝试共享信息以学习模态间的交互作用[59]	不完全利用样本间的相关性[59]	简单易用，适合初步分析[59]	缺乏对样本间关系的充分利用[59]
	图卷积神经网络[60]	利用相似性网络更有效地使用组学特征及样本或数据类型间的相关性[60]	构建样本相似性网络，集成模态[60]	提供生物意义结构，可纳入先前生物学知识[61]	相比FNNs更复杂，但更具灵活性[61]
	自动编码器	使用中间集成方法，在分类之前整合模态	通过多个共享同一层的编码器和解码器将模态整合为单一表示[53]	易于实现，适合处理大规模数据集	在某些情况下可能缺乏生成式模型的灵活性

2）虚拟育种模拟与遗传增益预测。虚拟育种技术依靠建立基因组重组和表型输出的数学映射关系，大幅削减了田间试验的试错成本，基于具有前瞻性的蒙特卡洛算法实施多性状渗入，借助回交与自交手段，把多个供体的优良特性转移到一个受体品种里的过程。其目标为在保留受体原有的全部优良特性之际，引入既定的优良性状，可以对不同杂交组合的性状表现开展量化评价[62]。遗传增益[63]是衡

量育种效果的核心指标，指经过1代或多代选择后，群体中目标性状的平均遗传水平提升的幅度。遗传增益预测[64]是通过数学模型和数据分析，提前估算育种策略在未来世代中可能实现的遗传改进效果。它是现代智能育种的核心技术之一，用于优化育种方案、降低试错成本，并加速品种改良进程。常用的遗传增益预测方法有传统统计模型、全基因组学选择（genomic selection，GS）[65]、计算机模拟，用软件（如AlphaSim）模拟虚拟群体，测试不同选择策略的效果。在遗传增益预测中，动态优化模型采用整合历史育种数据与环境互作参数的办法，完成了多世代选择反应的精准仿真，Sakha[66]提及了一个将机器人自动采集、数字孪生建模与VR可视化相结合的新型植物育种辅助平台。该平台不仅削减了对人工田间调查的依赖，还能辅助育种专家更直观、高效地评判植物表型，加速新品种的选育进程；基于贝叶斯网络与深度神经网络的混合模型，可有效辨认目标的关键调控模块，比如针对番茄叶片的病害进行分类，研究团队先采用定制化的卷积神经网络进行特征提取，然后把这些特征输入到把K最近邻、支持向量机以及随机森林作为基础估计器的体系中，以XGBoost作为最终估计器组建的堆叠模型中，采用这种途径能够有效抓取不同病害特征，增强整体分类的精准度，为科研人员在抗病育种工作中提供助力[67]。

2　智能育种技术应用进展

智能育种技术体系正推动作物遗传改良从依赖经验的“艺术”，转变为可预测、可设计的“科学”。其应用进展的核心体现在构建了一个由多维数据感知为起点、以种质资源挖掘为战略储备、以基因与代谢通路精准设计为执行核心的闭环迭代流程（图2）。该流程并非技术的简单堆砌，而是通过数据-知识-决策的持续转化，将海量生物信息转化为可操作的育种方案，最终实现遗传增益的精准与高效跃迁。

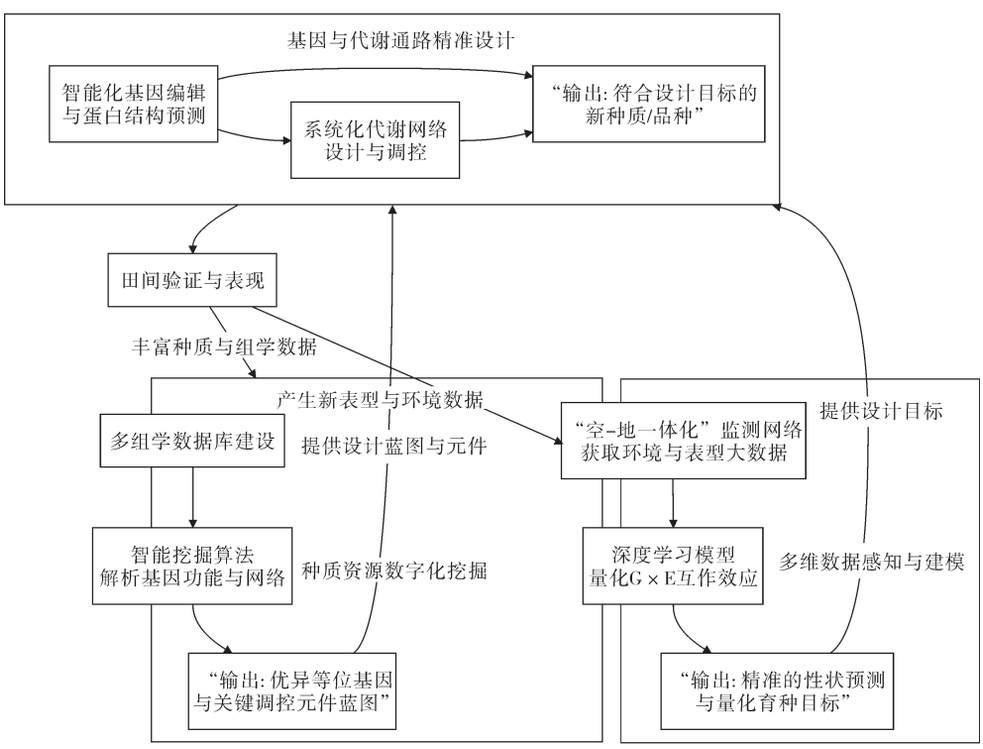


图2　智能育种技术应用的核心闭环流程

Fig. 2　The core closed-loop workflow of smart breeding technology applications

2.1　多维数据感知与基因型-环境互作建模

智能育种体系的运行始于对作物生长全周期、多维度信息的精准感知与解析。传统育种依赖于有限的人工表型记录，而现代智能育种则通过构建空-地一体化的立体监测网络，并借助深度学习模型融合多模态数据，实现对基因型-表型-环境互作关系的量化建模，为后续的精准决策奠定数据基石。

1）空-地一体化的高通量表型与环境感知网络。在空中平台，搭载多光谱、高光谱、热红外及LiDAR等传感器的无人机阵列，实现对大田群体表型的高通量、无损监测，在单次飞行中快速获取植株高度、叶面积指数、冠层温度、叶绿素含量等关键生长和生理参

数,并实现对生物胁迫与非生物胁迫的早期识别与空间定位[32]。其空间分辨率可达亚厘米级,能够精确捕捉株间变异[37]。在地面物联网络,布设于田间的土壤水分、温度、光照强度等传感器构成了1个时间分辨率高达分钟级的微环境动态监测系统[34],可连续记录作物茎秆直径微变化、蒸腾速率等生理响应[35],为解析作物对环境变化的瞬时反应提供了前所未有的时间维度数据。这种无人机大范围普查与物联网定点连续监测相结合的体系形成了优势互补。最新研究表明,将LiDAR获取的三维点云数据与深度学习模型相结合,能够自动化、高精度地解析高粱等作物的植株结构参数[38],标志着作物表型分析进入从二维图像迈向三维结构解析的新阶段。

2)基于深度学习的G×E互作效应建模。海量多模态数据的价值是通过先进的算法模型挖掘出指导育种的规律。基于深度学习的表型-环境互作模型成为预测复杂性状在不同环境下表现的核心工具。例如,Yewle等[68]提出的RicEns-Net深度集成模型,创新性地融合了合成孔径雷达、多光谱遥感影像与气象观测数据,用于预测湄公河三角洲地区的水稻产量,预测精度可达实际平均产量的95%~96%。Mena等[69]开发的多模态门控融合模型MMGF将预测尺度精细至亚田块级别。该模型通过设计的门控融合单元能自适应地为不同地区和作物类型调整遥感图像与气象数据的权重,实现了对异构数据的智能融合。

2.2 基因与代谢通路的精准设计

在智能感知模型明确育种目标与性状-环境响应关系的基础上,智能育种的核心环节在于实现对基因与代谢通路的精准设计。这一过程使育种模式从对自然变异的被动筛选,转向依据生物学机制与多组学蓝图进行主动、理性的遗传改良。当前,该领域的进展主要体现在基因编辑工具的智能化升级与代谢通路的系统化调控2个层面。

在基因编辑工具的智能化升级方面,其核心是提升靶向操作的精准度、效率与可预测性。第三代基因编辑工具CRISPR-Cas12i通过优化其sgRNA设计算法,展现出更宽的PAM兼容性与更高的编辑效率,为多基因协同调控网络的并行操作提供了可能[70]。基因编辑已不再局限于序列本身的修改,而是与蛋白质结构预测等前沿技术深度融合。例如,基于DeepMind开发的AlphaFold3模型能够高精度地预测蛋白质与核酸的相互作用结构[71],这为CRISPR-Cas12i等工具的sgRNA设计及脱靶效应预判提供了至关重要的结构生物学基础,从而将靶点设计的理性化与成功率提升至全新水平。如“一步成系”单倍体诱导编辑等技术创新[72]实现了对优良商业品种的快速、高效定向改良,规避了传统回交育种中连锁累赘的难题,显著加速了设计性状的产业化应用。

在代谢通路的系统化调控方面,精准设计超越了单一基因的编辑,着眼于对复杂性状背后整个生化网络的功能优化。以光合效率为例,对光能捕获复合体蛋白编码基因(如*Lhcb1*等)的靶向编辑,可直接提升光量子利用效率[73]。将叶片三维结构建模与代谢流量分析相整合,利用空间显式的三维叶片分析方法精准识别限制碳同化效率的关键生理节点[74],为构建高光效作物的数字孪生模型、指导多靶点协同设计提供了依据。在品质性状改良中,已建立对籽粒淀粉合成通路调控的动态模型。通过CRISPR/Cas9技术对关键分支酶基因*SBEIIb*与*SBEI*的启动子或编码区进行精确编辑[75],成功实现了水稻和玉米中抗性淀粉与直链淀粉含量的定向提升,展示了通过精准干预代谢通路的关键节点来按需设计谷物营养品质的应用前景。

2.3 种质资源的数字化与智能挖掘

精准设计并非凭空创造,其核心元件在很大程度上源于自然界经过长期进化形成的丰富种质资源。智能育种模式下,种质资源库的角色正从传统的“实物档案馆”向数字化的“战略元件库”与“设计灵感库”转变。通过对种质资源进行系统性的数字化表征与智能挖掘,能够前所未有地揭示其隐藏的遗传宝藏,为定向遗传改良提供精准的靶点与蓝图。

1)多组学数据库建设与系统性基因功能解析。近年来,我国在作物多组学数据库建设方面取得了显著进展。2020年,华中农业大学严建兵团队整合同一玉米群体的基因组、转录组、表型组、代谢组、表观基因组、遗传变异及定位数据,构建了支持云端集成、快速检索与智能分析的ZEAMAP数据库,为玉米分子设计育种奠定了基础,并进一步发布了首个玉米多组学整合网络图谱,利用机器学习技术成功预测调控开花时间等关键性状的功能基因与调控通路[76]。2022年,北京市农林科学院基于全球1 000余份生菜种质的高通量表型与重测序数据,建立了综合性数据库LettuceGDB,为生菜育种提供了重要数字资源[77]。2024年,广州大学等单位整合大豆六大

组学数据，构建了当前最系统的大豆多组学数据库SoyMD[78]；同年，相继发布了黄瓜（Cucumber-DB）[79]和谷子（Setaria-DB）[80]的多组学数据库，显著丰富了我国特色作物遗传研究与智能化，实现了遗传资源的可计算化。

2）野生近缘种与地方品种中有利等位基因的智能发掘。栽培品种经过长期人工选择后，其遗传多样性往往大幅度降低，导致许多具有育种价值的稀有等位基因丢失。智能挖掘技术的强大之处，在于能够从野生近缘种和地方品种这些“遗传宝库”中，系统性地鉴定出能够改良现代品种的关键等位基因。Wang等[81]从玉米野生祖先大刍草中成功克隆出广谱抗病基因*ZmMM*，该研究不仅揭示了玉米广谱抗病性的新机制，还鉴定出1个能够增强ZmMM1蛋白翻译水平的调控序列*qLMchr7C117*。

3　技术挑战与发展方向

3.1　技术瓶颈分析

1）多源数据标准化与互操作难题。智能育种技术的推行十分依赖多维度数据的整合与剖析，然而现阶段各研究机构和技术平台在数据生成、存储以及标注方面差异明显，田间表型监测系统和基因组测序平台输出的数据格式依旧未统一，造成多组学数据融合分析效率欠佳，多模态组织图谱数据集在网络传播与公平获取上面临关键挑战：一是单细胞RNA测序和空间转录组数据一般以多种非统一的测序及图像文件格式保存，这些格式在网页技术方面表现欠佳，目前的软件平台很难支持对多种整合数据模态进行同步浏览[42]。二是全球范围内种质资源数据库异构性的问题凸显，不同国家及企业采用的数据接口标准互不匹配，制约了跨国协作与数据共享机制的开展。

针对多源数据标准化与互操作难题，建议构建三阶融合体系。一是语义层通过领域本体库（如Plant Ontology[82]）建立跨源数据的标准化映射关系，消除核心概念歧义。二是技术层部署轻量化边缘智能网关，实现田间异构数据的实时协议转换与质量清洗，将多模态融合延迟压缩至亚分钟级。三是机制层采用区块链赋权的数据信托模型（比如AIAgriChain），构建“贡献确权-使用清算”的跨国协作网络，解决种质资源主权与共享激励矛盾。

2）算法可解释性与生物学机制脱节。尽管深度学习等人工智能算法在基因组预测和编辑靶点筛选方面体现出高精度的长处，然而其“黑箱”特征造成模型决策逻辑与生物学原理之间缺少可追溯的关联。以全基因组选择优化算法为例，当前模型对基因互作、表观遗传调控等复杂机制的分析能力欠佳，预测结果和田间试验表型屡屡出现偏差。基因编辑这类新型编辑工具的应用案例表明，算法驱动的靶点预测在整合蛋白质三维构象动态变化数据上不充分导致部分理论预测的编辑位点在实际操作中无法实现预期效应，这种算法跟生物学机制的割裂情形，不仅让育种专家对相关技术的信任度下降，更阻碍了智能决策系统在复杂性状改良中的深度应用。例如全基因组选择模型在预测玉米抗旱性状时，虽然在训练集上表现出较高的决定系数（$R^2>0.8$），但在田间验证中出现了显著偏差：模型所筛选出的“高价值”单倍型在实际干旱环境中并未表现出预期表型。进一步的分析表明，该模型未能有效捕捉到玉米抗旱关键基因如*ZmNAC111*和*ZmABP2*之间的上位性互作，以及在水分胁迫下通过ABA信号通路调控气孔开闭的动态表达模式。这种因忽略基因网络动态调控机制而导致的预测偏差，凸显了当前算法在整合多层级生物学知识方面的局限[81]。要解决算法可解释性与生物学机制脱节问题，核心是对智能育种不同层级任务需求的协同满足：复杂模型（如Transformer、Win-Former）因其在高维特征提取与模式识别上的卓越能力而被推崇，主要用于解决感知与预测问题；针对“黑箱”特性问题，在育种策略制定与生物学机制发现这一决策层级需要模型提供透明、可信的生物学洞察。因此，未来的发展方向并非摒弃复杂模型，而是通过引入可解释人工智能（explainable artificial intelligence，XAI）技术，构建“黑箱感知、白箱决策”的混合系统。具体落地方案包括：第一，采用事后解释技术，例如，在基于Transformer的全基因组预测模型中集成SHAP（shapley additive explanations）分析，量化每个SNP标记或单倍型对最终预测结果的贡献度，从而将全局的决策模型分解为局部的、可理解的遗传效应值[19]。第二，设计内在可解释的模型结构，如在图神经网络中显式嵌入已知的基因调控网络或代谢通路作为先验知识，约束模型在符合生物学逻辑的路径内进行推理。第三，建立可解释性输出与实验验证的闭环，将模型识别出的关键遗传位点或调控关系通过基因编辑、转基因或分子实验进行功能验证，将数据驱动的发现转化为坚实的生物学知识，从而在提升预测性能的同

时,弥合算法与生物学机制之间的鸿沟。

3.2 未来突破方向

1)生物-信息-工程跨学科融合创新。在智能育种技术体系中,生物信息学、计算科学和工程技术深度融合成为突破现有技术壁垒的主要驱动力。目前,合成生物学领域取得了最新进展,如动态代谢网络调控技术[83]等,在收集玉米不同发育阶段及所处环境条件下的大量生物数据的同时可按照多维度的数据构建一个综合性的玉米代谢网络图谱,它含有各种分子层次之间的相互作用关系[84]。

2)基于大模型的育种决策系统开发。智能育种语境下的大模型已超越信息领域中原海量文本训练的生成式语言模型,泛指参数规模巨大、能够从高通量多模态数据中学习复杂非线性关系的深度学习模型。其核心特征在于能够整合跨尺度、异构的育种数据,通过端到端的学习方式,直接挖掘从基因型到表型的隐藏规律,从而为亲本选配、后代筛选等关键决策提供高维度、全局性的量化依据。探讨的Win-Former与基于Transformer的作物行检测模型,正是此类大模型在育种前端高通量、自动化表型解析中的具体体现。它们虽不直接输出最终的育种方案,但通过精准解译作物三维结构与田间空间布局,为决策系统提供了传统方法难以获取的精准、定量化表型输入,是构建完整育种决策系统不可或缺的数据感知基石。例如,Sun等[85]提出的Win-Former模型,在玉米植株点云语义分割任务中实现了83.45%的平均交并比,能够高效、准确地自动化提取单株玉米的茎、叶、穗等器官结构参数。

3)全球协作网络与智能育种平台。拜耳公司的FieldView与科迪华等企业所推动的数字农业平台,已超越传统"决策支持工具"的范畴,演变为旨在塑造和主导未来农业市场的综合性生态系统。科迪华与Pairwise的合作展示了基因编辑技术对传统作物育种方法的变革。通过应用CRISPR、碱基编辑和模板编辑等先进的基因编辑工具,科学家能够实现对作物基因的精确修改,以达成增强抗病性、提升营养价值及增加产量等特定目标。相较于传统育种方法需耗费数年至数十年方能培育出具有所需特性的新品种,基因编辑显著缩短了这一进程。这些公司创建的数字农业平台核心战略在于通过系统地集中与整合农业生产全过程中的关键数据,包括产量、水分、投入品使用及作物表型等,构建一个高度整合的农业育种大数据平台。该平台不仅可以为客户提供决策支撑,更使企业能够将自身核心产品的使用数据与田间最终产出结果进行精确关联,从而实现产品价值的量化验证、基于真实反馈的育种优化,并形成显著的客户锁定效应。以FieldView Drive2.0[86]为代表的硬件集成方案,通过即插即用设计和直接传输功能,将不同品牌的农业机械无缝接入其数字生态,进一步巩固了对生产流程的控制力。企业通过战略性合作与投资,将数字平台的数据洞察与基因编辑等前沿生物技术相结合,构建了一个从数据驱动研发、田间验证到市场推广的闭环创新体系。

4 产业转化与政策建议

4.1 商业化育种体系建设中的主要问题

1)育种体系转型动力不足。一是从经验驱动向数据驱动转变缓慢。传统经验主导的育种方式仍占主流,缺乏以大数据为核心驱动力的现代化育种机制。二是单一技术突破难以带动全链条创新。现有技术创新多为点状突破,缺乏从基础研究到市场推广的系统协同。

2)政策支持体系不完善。一是政策覆盖不全。缺乏覆盖基础研究、品种审定、市场推广等各阶段的完整政策扶持链。二是公共科研资源与企业需求脱节。国家重大专项与企业研发基金之间联动机制缺失,导致公共资源无法精准对接产业需求。三是重点技术落地难。如基因编辑产业化、智能育种平台商业化等关键环节缺乏明确政策引导和支持。

3)技术融合存在障碍。一是生物技术、信息技术与装备技术融合度低。三者之间尚未实现有效融合,影响智能育种平台建设效率。二是高通量设备工程化程度不高。田间机器人表型采集系统、高通量基因分型设备等尚处于试验或半成熟阶段,推广应用受限。三是技术门槛过高。标准化、模块化程度低,导致中小育种企业难以进入智能育种领域。

4)数据共享与产业链协同困难。一是"数据孤岛"现象严重。科研机构与企业之间、企业与企业之间的数据壁垒阻碍了信息流动与资源共享。二是核心种质资源数据库未开放。国家作物种质资源库的数据未能有效对外公开,限制了育种单位对遗传信息的利用。三是行业协同机制缺失。缺乏龙头企业牵头的智能育种创新联合体,算法与数据资源未能实现共享。

4.2 智能育种技术推广中的伦理与安全问题

1)生态风险与生物安全问题突出。一是脱靶效

应潜在风险大。基因编辑技术虽能实现精准修饰，但存在脱靶效应，可能通过基因漂移影响野生近缘种群。二是杂草稻去驯化现象揭示适应机制的复杂性。研究表明，杂草稻通过新变异和分子机制适应环境，而非简单恢复野生状态[86]，提示育种释放后可能引发不可控进化路径。三是区域生态评估体系缺失。缺乏基于区域生态特征的动态评估机制，难以预测和管控转基因作物释放后的生态影响。四是监管分级策略滞后。现行法规未能针对不同风险等级制定相应的释放与管控措施。

2)种质资源数字化中的数据主权与利益分配问题。一是全球遗传资源与技术能力不对称。发达国家凭借技术优势掌握发展中国家种质资源数据，形成数据垄断。二是数据贡献国权益受损。如非洲玉米种质资源被跨国公司用于训练商业算法，但原始提供国未能获得知识产权或经济回报。三是惠益分享机制滞后。《生物多样性公约》及数字序列信息相关条款未能有效回应数据跨境使用带来的利益冲突。四是数据溯源与确权困难。缺乏有效的区块链技术支持的数据追踪系统，难以保障数据来源的透明度和权益归属的清晰度。

4.3　商业化育种体系建设

基于现代农业生物技术创新的驱动力量，商业化育种体系正处在从凭借经验驱动向凭借数据驱动、从单一技术突破向全链条协同创新的模式转换进程中，亟须从三个方面实现转变：

一是形成政府引导的市场化技术转化机制。在2025年中央一号文件将“突破性品种选育”与“生物育种产业化”并列提出的导向下，应构建覆盖基础研究、品种审定、市场推广全阶段的政策扶持体系，通过建立生物育种国家重大专项和企业研发基金之间的联动机制，确保公共资源与企业需求准确契合，着重支持基因编辑技术产业化落地、智能育种平台商业化经营等核心体系建设。

二是推进生物技术、信息技术和装备制造深度融合。依靠多组学数据融合分析系统，合并基因型-表型-环境互作的大数据，构建覆盖种质资源鉴定、遗传剖析、品种设计的数字化育种平台，从技术实施角度，要重点推进高通量基因分型装备与田间机器人表型采集系统的工程化集成工作，构建标准化、模块化的智能育种技术组合，切实降低技术应用的准入门槛。

三是造就开放协同的产业链态势。针对现阶段种业企业存在的“数据孤岛”困境，建议借助国家作物种质资源库搭建核心种质重测序数据库共享机制，驱动科研机构积累的基因组数据向育繁推一体化企业定向对外公开，同时号召龙头企业牵头兴办智能育种创新联合体，采用“算法开放＋数据有偿利用”模式，实现全行业对深度学习模型、基因组预测算法等核心工具的共享。例如，中国农业科学院与阿里巴巴达摩院合作推出了1个可覆盖育种全流程的全新智能育种平台，它集成了育种数据的管理与分析、大规模模型及高性能计算优化加速体系，到借助人工智能算法预估最佳亲本组合并筛选优质品种的全方位功能；浙江大学、中国农业科学院和阿里巴巴联合研发的AI模型DeepFDML破译了棉花品种改良“密码”，整合了遗传学、生物大数据与人工智能技术，搭建了1个包含207个品种的棉花全基因组DNA甲基化的图谱，鉴别出约2.87亿个单甲基化多态性位点。目前在所有作物里这个模型的规模最大，从这些数据里，研究人员找出了43个可能参与纤维发育的表达数量性状甲基化位点基因，为棉花育种提供了关键的候选基因清单[87]。崖州湾国家实验室、上海人工智能实验室与中国农业大学联合研发的水稻育种大模型“丰登·水稻”，通过构建覆盖140余万篇文献的科研语料库和全球首个水稻多组学知识图谱，实现了对水稻生物学知识的深度整合。

4.4　智能育种技术推广的伦理考量

智能育种技术的快速发展正在重塑全球农业伦理格局，其引发的争议已超越传统生物技术范畴，呈现出技术复杂性、利益关联性和文化敏感性的三重特征。这里从生态安全、数据主权、社会公平3个维度构建分析框架，并提出基于技术治理的创新解决方案。

1)基因编辑技术的生态风险与生物安全。基因编辑技术的精准性特征掩盖了其潜在的生态级联效应。墨西哥瓦哈卡州传统玉米品种因转基因玉米花粉污染导致92个地方品种等位基因丢失[88]，揭示了低概率事件在生态系统中放大的风险。浙江大学农业与生物技术学院樊龙江教授团队联合中国水稻研究所陆永良课题组的研究表明，中国杂草稻皆是从栽培稻起源，也经历了强大的遗传瓶颈冲击；杂草稻在去驯化的时候并非简单地恢复为野生模样，而是利用新的变异与分子机制去适应外界环境；基于等位基因频率变化的分析显示，已有变异与新突变在不同类型杂草稻进化方面的作用有明显差异，杂草

稻基因组上很多区域存在平衡选择信号，杂草稻群体在适应复杂自然环境时造就了更广泛的遗传多态性；这一过程和水稻驯化过程中经历的正向选择相悖；该研究成果拓展了对作物驯化和去驯化进化遗传机制的认识，也为解读杂草稻通过基因漂移影响野生近缘种群遗传多样性这一现象提供了重要理论依据[86]。

2）种质资源数字化中的数据主权争议。种质资源的数字化正在瓦解传统的生物遗传资源主权概念。根据《名古屋议定书》2024年修正案，数字序列信息的惠益分享机制面临诸多执行困境，例如在数据访问与利益分配不均层面，假设某发展中国家拥有独特的植物遗传资源，并且这种植物基因序列被一家跨国公司获取，如果该公司利用这些序列育成新品种，根据《名古屋议定书》，该发展中国家应获得一定的经济或其他形式的利益分享，然而在实践中，由于缺乏明确的数据跟踪和利益分配机制，该国可能难以证明其对所使用的基因序列的权利，从而导致实际利益未能有效返回。

3）技术垄断与利益分配的结构性失衡。当前全球智能育种技术正加速向数据驱动、算法主导、平台集成的方向演进，以拜耳、科迪华为代表的跨国企业依托长期技术积累与全产业链布局，已建立起高度工程化、标准化的智能育种体系，具体表现在以下3个方面。

①在数据采集与标准化体系方面。跨国企业依托全球化试验网络和先进的传感技术，构建了高精度、多尺度的表型数据采集体系。例如，拜耳的FieldView平台集成卫星遥感、无人机多光谱成像与物联网传感器，实现了高精度、单株分辨率的环境-表型动态监测。科迪华的Encirca系统则利用区块链技术，实现全球育种数据的实时共享与可追溯，保障了数据的完整性与一致性。

②在种质资源利用与底层技术创新能力方面。跨国企业通过全球种质资源并购与专利布局等方式构建技术壁垒。拜耳持有全球68%的玉米耐除草剂相关专利，其智能育种平台已实现“表型-基因型-市场需求”的闭环匹配。科迪华则利用合成生物学技术创制人工染色体，突破野生种杂交障碍，拓展了种质创新边界。

③在技术成熟度与产业转化效率方面。跨国企业已形成“算法-设备-品种”的垂直整合商业模式。拜耳数字育种服务年营收超20亿美元，其智能决策系统可缩短60%的品种审定周期；其育种机器人日均处理样本量高达10万份，展现出强大的工程化能力。

为应对上述挑战，可采取系统性策略，推动智能育种技术的自主可控与普惠发展。一是构建标准化、多模态的数据采集体系。应加快部署高分辨率的表型监测设备，提升田间机器人覆盖率，推动全国范围内的数据格式统一与共享机制建设。可借鉴科迪华Encirca系统的区块链数据溯源经验，建立可信的育种数据链，同时充分发挥我国在种质资源数字化方面的优势，将已建成的玉米三维基因组数据库与表型数据深度融合，提升数据资产价值。二是发展可解释、专用化的AI育种模型。在通用大模型研发尚处追赶阶段的背景下，应重点发展面向中国特色性状的专用模型。崖州湾实验室的物理约束建模路径值得推广，未来应加强AI与生物学机制的融合，提升模型的可解释性。应推动多模态数据融合架构研发，提升在小样本、异质性环境下的预测鲁棒性，避免“算法歧视”。三是强化底层技术创新与种质资源主权保障。应加大对基因编辑底盘细胞、合成基因组组装等底层技术的研发投入，减少对进口试剂的依赖。通过国家科技专项支持，构建自主可控的种质创新工具链。四是推动技术普惠化与制度创新。针对跨国企业通过专利壁垒控制核心算法、导致中小机构被边缘化的问题，建议参考世界知识产权组织相关标准，对涉及粮食安全与公共利益的育种算法实施强制许可。通过公私合作模式构建普惠性技术推广网络，降低技术使用门槛。

4）伦理治理机制的创新路径。构建技术创新与制度创新适应智能育种特性的治理体系。

①在技术治理层面。欧盟“GAIA-X农业云”率先将可解释性作为准入标准，要求全基因组选择模型必须输出SHAP值，特别是环境因子与遗传因子的交互效应权重不得超过预设阈值。这种技术性规定有效防止了MMGF等模型将产量变异过度归因于不可控环境因素；欧盟《AI法案》将育种算法列为高风险系统，强制披露关键决策路径。

②在传统知识保护层面。可以构建自动触发惠益回馈机制，例如2024年在日内瓦举行的外交会议通过了《产权组织知识产权、遗传资源和相关传统知识条约》，对土著社区贡献的等位基因知识进行时间戳存证，确保相关品种商业化时能自动触发惠益回馈机制；构建“技术-法律-社区”三位一体的治理模

式，将技术层开发和伦理评估相结合，实时监控基因编辑作物的表型偏离度。

③全球协作层面。需要重构现行治理架构，在技术标准领域，ISO/TC 307《Blockchain and distributed ledger technologies》拟将编辑工具脱靶率、数据偏差系数等37项参数纳入强制检测；在法律框架方面，建议在《国际植物保护公约》中增设数字序列信息跨境流动的“红灯条款”，禁止未经来源国同意的种质数据跨境训练模型；在文化认同维度，FAO新设立的农业技术伦理委员会通过确保发展中国家占比40%以上的席位分配，防止技术伦理话语权的单边垄断。

4.5　智能育种的经济效益评估与测算

智能育种技术体系的终极价值在于其产业转化能力，而对此进行科学的经济效益评估是推动其商业化应用的关键。尽管目前大规模的商业化实践数据仍在积累，但基于现有案例与技术参数，可以构建其经济回报的量化测算模型。该模型主要围绕单位种子成本、品种上市周期与市场份额3个核心维度展开。

1）在降低单位种子成本方面，智能育种通过“虚拟筛选”大幅缩减了田间试验的规模。在传统育种中，F_2至F_6代需种植数万乃至数十万株群体以进行表型选择，其土地、人力、物料成本高昂。而基于全基因组预测的早期筛选，可将进入田间测试的基因型数量减少70%以上[9]。

2）在缩短上市周期方面，智能育种的影响最为显著。传统玉米育种周期通常为8～10 a，而集成多模态预测和虚拟育种模拟后，可通过优化杂交组合与加速世代周转，将周期压缩至4～6 a[28]。根据新产品上市的“时间价值”模型，品种提前1 a上市所产生的超额利润现值极为可观。

3）在提高市场份额方面，智能育种通过精准设计赋予品种更强的竞争力。例如，借助G×E互作模型开发的品种能更精准地匹配不同生态区的需求，从而提升农户采纳率。可以采用基于算法预测精度的市场份额增长模型：若智能育种将成功推出“爆款”品种（指市占率＞8%）的概率从传统育种的10%提升至25%，那么育种公司的预期市场回报将实现倍增。

5　小　结

智能育种已进入由多维数据链驱动与智能决策协同引领的新阶段。本文阐述了以基因组解析、多组学数据融合、人工智能算法为核心的智能育种技术体系，揭示了其在破解传统育种瓶颈、加速遗传增益跃迁方面的巨大潜力。高精度的三维基因组图谱为理解基因调控机制提供了空间维度新视角；人工智能算法，特别是全基因组预测与作物生长模型的耦合、基因组导航系统及深度学习驱动的多模态数据整合，显著提升了复杂性状预测与育种路径优化的精准性与效率；虚拟育种模拟与遗传增益预测技术的应用则大幅降低了田间试错成本，实现了育种过程的可预测、可设计。然而，该体系的产业化仍面临技术融合壁垒、数据孤岛、生态安全风险及政策制度滞后等多重挑战。为此，应着力推进生物-信息-装备技术的深度融合，构建标准化、模块化的智能育种平台；建立政府引导、市场主导的协同转化机制与全链条政策支持体系；构建科学的生态风险评估与分级管控框架；并创新数据治理模式，保障数据主权与惠益公平。未来，智能育种的发展将依赖于大模型驱动、多主体协同、全链条贯通的生态系统构建。唯有实现科技创新与制度创新的双轮驱动，方能真正释放数据潜能，实现遗传增益的持续跃迁，为保障国家粮食安全与推动农业现代化提供核心科技支撑。

参考文献 References

[1] WALLACE J G, RODGERS-MELNICK E, BUCKLER E S. On the road to breeding 4.0: unraveling the good, the bad, and the boring of crop quantitative genomics [J]. Annual review of genetics, 2018, 52: 421-444.

[2] JING H C, TIAN Z X, CHONG K, et al. Progress and perspective of molecular design breeding [J]. Scientia sinica vitae, 2021, 51(10): 1356-1365.

[3] ZHAO C J, ZHANG Y, DU J J, et al. Crop phenomics: current status and perspectives [J/OL]. Frontiers in plant science, 2019, 10: 714 [2025-06-10]. https://doi. org/10.3389/fpls.2019.00714.

[4] WANG B B, HOU M, SHI J P, et al. *De novo* genome assembly and analyses of 12 founder inbred lines provide insights into maize heterosis [J]. Nature genetics, 2023, 55(2): 312-323.

[5] REYNOLDS D, BALL J, BAUER A, et al. CropSight: a scalable and open-source information management system for distributed plant phenotyping and IoT-based crop management [J/OL]. GigaScience, 2019, 8 (3): giz009 [2025-06-10]. https://doi.org/10.1093/gigascience/giz009.

[6] ZHANG Y, HUANG G M, ZHAO Y X, et al. Revolutionizing crop breeding: next-generation artificial intelligence and big da-

ta-driven intelligent design[J].Engineering,2025,44:245-255.
[7] LANDE R,THOMPSON R.Efficiency of marker-assisted selection in the improvement of quantitative traits[J].Genetics,1990,124(3):743-756.
[8] 种康.作物育种进化史:从野生驯化到智能设计[J].中国科技产业,2025(6):4. ZHONG K.Crop breeding evolution:from domestication to intelligent design[J].Science & technology industry of China,2025(6):4(in Chinese).
[9] 简六梅,肖英杰,严建兵.从头驯化:作物品种设计与培育的新方向[J].遗传,2023,45(9):741-753. JIAN L M,XIAO Y J,YAN J B.De novo domestication:a new way for crop design and breeding[J].Hereditas,2023,45(9):741-753(in Chinese with English abstract).
[10] VAN VU T,DAS S,HENSEL G,et al.Genome editing and beyond:what does it mean for the future of plant breeding?[J/OL].Planta,2022,255(6):130[2025-06-10]. https://doi.org/10.1007/s00425-022-03906-2.
[11] BRESEGHELLO F,COELHO A S G.Traditional and modern plant breeding methods with examples in rice (*Oryza sativa* L.)[J].Journal of agricultural and food chemistry,2013,61(35):8277-8286.
[12] FAUX A M,GORJANC G,GAYNOR R C,et al.AlphaSim:software for breeding program simulation[J/OL].The plant genome,2016,9(3):27902803[2025-06-10]. https://doi.org/10.3835/plantgenome2016.02.0013.
[13] OUYANG W Z,XIONG D,LI G L,et al.Unraveling the 3D genome architecture in plants:present and future[J].Molecular plant,2020,13(12):1676-1693.
[14] RUI Z,ZHANG Z,ZHANG M,et al.High-throughput proximal ground crop phenotyping systems:a comprehensive review[J/OL].Computers and electronics in agriculture,2024,224:109108[2025-06-10].https://doi.org/10.1016/j.compag.2024.109108.
[15] WANG W X,GUO W J,LE L,et al.Integration of high-throughput phenotyping,GWAS,and predictive models reveals the genetic architecture of plant height in maize[J].Molecular plant,2023,16(2):354-373.
[16] CRAIN J,MONDAL S,RUTKOSKI J,et al.Combining high-throughput phenotyping and genomic information to increase prediction and selection accuracy in wheat breeding[J/OL].The plant genome,2018,11:170043[2025-06-10]. https://doi.org/10.3835/plantgenome2017.05.0043.
[17] ZHANG Y,WANG J L,DU J J,et al.Dissecting the phenotypic components and genetic architecture of maize stem vascular bundles using high-throughput phenotypic analysis[J].Plant biotechnology journal,2021,19(1):35-50.
[18] CLARK S A,VAN DER WERF J.Genomic best linear unbiased prediction (gBLUP) for the estimation of genomic breeding values[M]//Genome-Wide Association Studies and Genomic Prediction.Totowa:Humana Press,2013:321-330.
[19] YU T X,ZHANG H,CHEN S K,et al. EXGEP:a framework for predicting genotype-by-environment interactions using ensembles of explainable machine-learng models[J/OL].Briefings in bioinformatics,2025,26(4):bbaf414[2025-06-10]. https://doi.org/10.1093/bib/bbaf414.
[20] 王轲麟.基于深度学习的从基因型到表型预测模型研究[D].北京:中国农业科学院,2022.WANG K L.Research on deep learning-based genotype to phenotype prediction models[D].Beijing:Chinese Academy of Agricultural Sciences,2022(in Chinese with English abstract).
[21] COVARRUBIAS-PAZARAN G,MARTINI J W R,QUINN M,et al.Strengthening public breeding pipelines by emphasizing quantitative genetics principles and open source data management[J/OL].Frontiers in plant science,2021,12:681624[2025-06-10].https://doi.org/10.3389/fpls.2021.681624.
[22] JOSGARCPELVO F,QUEZ-INGELMO A,GARC-HOLGADO A.Explainable rules and heuristics in AI algorithm recommendation approaches:a systematic literature review and mapping study[J].Computer modeling in engineering & sciences,2023,136(2):1023-1051.
[23] HANGLOO S,ARORA B.Multimodal fusion techniques:review,data representation,information fusion,and application areas[J/OL].Neurocomputing,2025,649:130827[2025-06-10].https://doi.org/10.1016/j.neucom.2025.130827.
[24] NICOLAS M,ALEX C O,BRYAN J E,et al.Breedbase:a digital ecosystem for modern plant breeding[J/OL],Genes genomes genetics,2022,12(7):jkac078[2025-06-10],https://doi.org/10.1093/g3journal/jkac078.
[25] TOFFETTI G,MILITANO L,THARAKA R,et al. ROS-based robotic applications orchestration in the compute continuum:challenges and approaches[C]//Proceedings of the IEEE/ACM 16th International Conference on Utility and Cloud Computing. Taormina (Messina),Italy:ACM,2023:1-6.
[26] 丁鹏钧.植物m6A调控网络数据库的构建及应用[D].杨凌:西北农林科技大学,2025.DING P J.Construction and application of plant m6A regulatory network database[D].Yangling:Northwest A&F University(in Chinese with English abstract).
[27] YANG F,KONG H J,YING J,et al.SeedLLM·Rice:a large language model integrated with rice biological knowledge graph[J].Molecular plant,2025,18(7):1118-1129.
[28] XIE Y,ZHANG T H,YANG M H,et al.Engineering crop flower morphology facilitates robotization of cross-pollination and speed breeding[J].Cell,2025,188(21):5809-5830.
[29] LI G T,AN L N,YANG W N,et al.Integrated biotechnological and AI innovations for crop improvement[J].Nature,2025,643(8073):925-937.
[30] MENG L,MING B,LIU Y,et al.Maize biomass estimation by integrating spectral,structural,and textural features from unmanned aerial vehicle data[J/OL].European journal of agronomy,2025,168:127647[2025-06-10].https://doi.org/10.1016/j.eja.2025.127647.
[31] 郭庆华,杨维才,吴芳芳,等.高通量作物表型监测:育种和精准农业发展的加速器[J].中国科学院院刊,2018,33(9):940-946.GUO Q H,YANG W C,WU F F,et al.High-throughput

crop phenotyping: accelerators for development of breeding and precision agriculture[J]. Bulletin of Chinese Academy of Sciences, 2018, 33(9): 940-946(in Chinese with English abstract).

[32] MAHLEIN A K. Plant disease detection by imaging sensors-paradigms and advances [J]. Plant disease, 2016, 100(2): 2228-2242.

[33] ZHOU H K, HUANG F D, LOU W D, et al. Yield prediction through UAV-based multispectral imaging and deep learing in rice breeding trials[J/OL]. Agricultural systems, 2025, 223: 104214 [2025-06-10]. https://doi. org/10.1016/j. agsy.2024.104214.

[34] AYAZ M, AMMAD-UDDIN M, SHARIF Z, et al. Internet-of-things (IoT)-based smart agriculture: toward making the fields talk[J]. IEEE access, 2019, 7: 129551-129583.

[35] AMIR A, BUTT M. Improved sap flow prediction: a comparative deep learning study based on LSTM, BiLSTM, LRCN, and GRU with stem diameter data[J/OL]. Smart agricultural technology, 2025, 12: 101105[2025-06-10]. https://doi.org/10.1016/j.atech.2025.101105.

[36] ADAMO T, CAIVANO D, COLIZZI L, et al. Optimization of irrigation and fertigation in smart agriculture: an IoT-based micro-services framework[J/OL]. Smart agricultural technology, 2025, 11: 100885[2025-06-10]. https://doi.org/10.1016/j.atech.2025.100885.

[37] BONGOMIN O, LAMO J, GUINA J M, et al. UAV image acquisition and processing for high-throughput phenotyping in agricultural research and breeding programs[J/OL]. The plant phenome journal, 2024, 7: e20096[2025-06-10]. https://doi.org/10.1002/ppj2.20096.

[38] PATEL A K, PARK E S, LEE H, et al. Deep learning-based plant organ segmentation and phenotyping of sorghum plants using LiDAR point cloud[J]. IEEE journal of selected topics in applied earth observations and remote sensing, 2023, 16: 8492-8507.

[39] NAKATO R, SAKATA T. Methods for ChIP-seq analysis: a practical workflow and advanced applications [J]. Methods, 2021, 187: 44-53.

[40] GUR E R, HUGHES J R. scATAC-seq generates more accurate and complete regulatory maps than bulk ATAC-seq[J/OL]. Scientific reports, 2025, 15: 3665[2025-06-10]. https://doi.org/10.1038/s41598-025-87351-7.

[41] PENG Y, XIONG D, ZHAO L, et al. Chromatin interaction maps reveal genetic regulation for quantitative traits in maize [J/OL]. Nature communications, 2019, 10: 2632 [2025-06-10]. https://doi.org/10.1038/s41467-019-10602-5.

[42] LI T, HORSFALL D, BASURTO-LOZADA D, et al. WebAtlas pipeline for integrated single-cell and spatial transcriptomic data[J]. Nature methods, 2025, 22(1): 3-5.

[43] NI L B, TIAN Z X. Toward *cis*-regulation in soybean: a 3D genome scope [J/OL]. Molecular breeding, 2023, 43(4): 28 [2025-06-10]. https://doi.org/10.1007/s11032-023-01374-4.

[44] SUN L H, ZHOU J R, XU X, et al. Mapping nucleosome-resolution chromatin organization and enhancer-promoter loops in plants using Micro-C-XL [J/OL]. Nature communications, 2024, 15: 35[2025-06-10]. https://doi.org/10.1038/s41467-023-44347-z.

[45] HUANG X H, WANG Y J, ZHANG S N, et al. Epigenomic and 3D genomic mapping reveals developmental dynamics and subgenomic asymmetry of transcriptional regulatory architecture in allotetraploid cotton[J/OL]. Nature communications, 2024, 15: 10721 [2025-06-10]. https://doi. org/10.1038/s41467-024-55309-4.

[46] ALBERT P S, YANG H, GAO Z, et al. Aneuploidy and ploidy variation conditioned by the B chromosome of maize[J]. Heredity, 2025, 12: 00764-y [2025-06-10]. https://doi. org/10.1038/s41437-025-00764-y.

[47] SCHIESSL S V, KATCHE E, IHIEN E, et al. The role of genomic structural variation in the genetic improvement of polyploid crops[J]. The crop journal, 2019, 7(2): 127-140.

[48] KOCHENDERFER M J. Algorithms for artificial intelligence [M]. California : MIT Press, 2019.

[49] COOPER M, TECHNOW F, MESSINA C, et al. Use of crop growth models with whole-genome prediction: application to a maize multienvironment trial[J]. Crop science, 2016, 56(5): 2141-2156.

[50] WEI X, QIU J, YONG K C, et al. A quantitative genomics map of rice provides genetic insights and guides breeding[J]. Nature genetics, 2021, 53(2): 243-253.

[51] SUBRAMANIAN I, VERMA S, KUMAR S, et al. Multi-omics data integration, interpretation, and its application [J]. Bioinformatics and biology insights, 2020, 14: 1-24.

[52] PAN L, TANG B F, ZHANG X, et al. Comprehensive analysis of multi-omics single-cell data using the single-cell analyst [J/OL]. iMeta, 2025, 4(3): e70038[2025-06-10]. https://doi.org/10.1002/imt2.70038..

[53] BALLARD J L, WANG Z X, LI W R, et al. Deep learning-based approaches for multi-omics data integration and analysis [J/OL]. BioData mining, 2024, 17: 38[2025-06-10]. https://doi.org/10.1186/s13040-024-00391-z.

[54] AHMED K T, SUN J, CHENG S, et al. Multi-omics data integration by generative adversarial network[J]. Bioinformatics, 2021, 38(1): 179-186.

[55] HIRA M T, RAZZAQUE M A, ANGIONE C, et al. Integrated multi-omics analysis of ovarian cancer using variational autoencoders[J/OL]. Scientific reports, 2021, 11: 6265[2025-06-10]. https://doi.org/10.1038/s41598-021-85285-4.

[56] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial networks [J]. Communications of the acm, 2020, 63(11): 139-144.

[57] LIU Y B, YUE L, XIAO S F, et al. Assessing clinical progression from subjective cognitive decline to mild cognitive impairment with incomplete multi-modal neuroimages[J/OL]. Medical image analysis, 2022, 75: 102266 [2025-06-10]. https://

doi.org/10.1016/j.media.2021.102266.

[58] YANG H, CHEN R, LI D D, et al. Subtype-GAN: a deep learning approach for integrative cancer subtyping of multi-omics data[J].Bioinformatics,2021,37(16):2231-2237.

[59] BICA I, VELIČKOVIĆ P, XIAO H. Multi-omics data integration using cross-modal neural network [C]// ESANN 2018 proceedings, European Symposium on Artificial Neural Networks. Bruges: Computational Intelligence and Machine Learning.2018:385-390.

[60] GUO H,LV X,LI Y Z,et al.Attention-based GCN integrates multi-omics data for breast cancer subtype classification and patient-specific gene marker identification[J].Briefings in functional genomics,2023,22(5):463-474.

[61] PENG W, CHEN T L, DAI W. Predicting drug response based on multi-omics fusion and graph convolution[J]. IEEE journal of biomedical and health informatics, 2022, 26(3): 1384-1393.

[62] MOEINIZADE S, HAN Y, PHAM H, et al. A look-ahead Monte Carlo simulation method for improving parental selection in trait introgression[J/OL]. Scientific reports, 2021, 11: 3918 [2025-06-10]. https://doi. org/10.1038/s41598-021-83634-x.

[63] RUTKOSKI J E.A practical guide to genetic gain[J].Advances in agronomy,2019,157:217-249.

[64] JONES D, FORNARELLI R, DERBYSHIRE M, et al. The pursuit of genetic gain in agricultural crops through the application of machine-learning to genomic prediction[J/OL]. Frontiers in genetics,2023,14:1186782[2025-06-10].https://doi.org/10.3389/fgene.2023.1186782.

[65] 刘海岚，夏超，兰海.全基因组选择技术在作物育种中的研究进展[J].华北农学报,2022, 37(S1): 51-58. LIU H L, XIA C , LAN H. The research progress of genomic selection in breeding of crops[J]. Acta agriculturae boreali-sinica.2022, 37(S1): 51-58(in Chinese with English abstract)

[66] SAKHA M M, DAIBER F, ENDERS M, et al. Enhancing plant variety discovery process with visual trait assessment in VR[C]//2025 IEEE Conference Virtual Reality and 3D User Interfaces (VR). March 8-12, 2025. Saint Malo: IEEE, 2025: 125-134.

[67] KHAN B,DAS S,FAHIM N S,et al.Bayesian optimized multimodal deep hybrid learning approach for tomato leaf disease classification[J/OL].Scientific reports,2024,14:21525[2025-06-10].https://doi.org/10.1038/s41598-024-72237-x.

[68] YEWLE A D,MIRZAYEVA L,KARAKUŞ O.Multi-modal data fusion and deep ensemble learning for accurate crop yield prediction[J/OL].Remote sensing applications:society and environment, 2025, 38: 101613[2025-06-10]. https://doi. org/10.1016/j.rsase.2025.101613.

[69] MENA F,PATHAK D,NAJJAR H,et al.Adaptive fusion of multi-modal remote sensing data for optimal sub-field crop yield prediction[J/OL].Remote sensing of environment,2025, 318: 114547 [2025-06-10]. https://doi. org/10.1016/j.rse.2024.114547.

[70] HE Y,LIAO S Y,REN Q R,et al.CRISPR-Cas12i confers efficient genome editing and gene regulation in plants[J/OL]. Plant physiology, 2025, 198(1): kiaf125 [2025-06-10]. https://doi.org/10.1093/plphys/kiaf125.

[71] ABRAMSON J,ADLER J,DUNGER J,et al.Accurate structure prediction of biomolecular interactions with AlphaFold 3 [J].Nature,2024,630(8016):493-500.

[72] TIAN J G,WANG C L,CHEN F Y,et al.Maize smart-canopy architecture enhances yield at high densities[J]. Nature, 2024,632(8025):576-584.

[73] ORDON J, BRESSAN M, KRETSCHMER C, et al. Optimized Cas9 expression systems for highly efficient *Arabidopsis* genome editing facilitate isolation of complex alleles in a single generation[J].Functional & integrative genomics,2020, 20(1):151-162.

[74] EARLES J M,BUCKLEY T N,BRODERSEN C R,et al. Embracing 3D complexity in leaf carbon-water exchange[J]. Trends in plant science,2019,24(1):15-24.

[75] SANG S F,SUN X H,MA T Y,et al.Efficient promoter editing of the SBEIIb gene enables fine-tuning of the resistant starch content in rice[J/OL].International journal of biological macromolecules, 2025, 290: 138904 [2025-06-10]. https://doi.org/10.1016/j.ijbiomac.2024.138904.

[76] GUI S T,YANG L F,LI J B,et al.ZEAMAP,a comprehensive database adapted to the maize multi-omics era[J/OL]. iScience,2020,23(6):101241[2025-06-10].https://doi.org/10.1016/j.isci.2020.101241.

[77] GUO Z L,LI B,DU J J,et al.LettuceGDB:the community database for lettuce genetics and omics[J/OL].Plant communications, 2023, 4(1): 100425 [2025-06-10]. https://doi. org/10.1016/j.xplc.2022.100425.

[78] YANG Z Q,LUO C F,PEI X X,et al.SoyMD:a platform combining multi-omics data with various tools for soybean research and breeding[J].Nucleic acids research,2024,52(D1): D1639-D1650.

[79] GUAN J T,MIAO H,ZHANG Z H,et al.A near-complete cucumber reference genome assembly and Cucumber-DB, a multi-omics database[J].Molecular plant,2024,17(8):1178-1182.

[80] HE Q, WANG C C, HE Q, et al. A complete reference genome assembly for foxtail millet and *Setaria*-db, a comprehensive database for *Setaria*[J]. Molecular plant, 2024, 17(2): 219-222.

[81] WANG H Z,HOU J B,YE P,et al.A teosinte-derived allele of a MYB transcription repressor confers multiple disease resistance in maize[J]. Molecular plant, 2021, 14(11): 1846-1863.

[82] COOPER L,JAISWAL P.The plant ontology:a tool for plant genomics [M]//EDWARDS D. Plant bioinformatics. New York:Springer,2016:89-114.

[83] REAM M, PRATHER K L J. Engineered autonomous dynamic regulation of metabolic flux[J]. Nature reviews bioengineering, 2024, 2(3): 233-243.

[84] GOMEZ-CANO F, RODRIGUEZ J, ZHOU P, et al. Prioritizing maize metabolic gene regulators through multi-omic network integration[PP/OL]. bioRxiv, 2024, 2: 582075(2024-02-26)[2025-06-10]. https://doi.org/10.1101/2024.02.26.582075.

[85] SUN Y, GUO X D, YANG H. Win-former: window-based transformer for maize plant point cloud semantic segmentation [J/OL]. Agronomy, 2023, 13 (11): 2723 [2025-06-10]. https://doi.org/10.3390/agronomy13112723.

[86] QIU J, ZHOU Y J, MAO L F, et al. Genomic variation associated with local adaptation of weedy rice during de-domestication[J/OL]. Nature communications, 2017, 8: 15323[2025-06-10]. https://doi.org/10.1038/ncomms15323.

[87] ZHAO T, GUAN X Y, HU Y, et al. Population-wide DNA methylation polymorphisms at single-nucleotide resolution in 207 cotton accessions reveal epigenomic contributions to complex traits[J]. Cell research, 2024, 34(12): 859-872.

[88] QUIST D, CHAPELA I H. Transgenic DNA introgressed into traditional maize landraces in *Oaxaca*, Mexico[J]. Nature, 2001, 414(6863): 541-543.

Synergy of multi-dimensional data chain-driven and intelligent decision-making: systematic construction of intelligent breeding technology systems and transition of genetic gain

ZHANG Zhiqi, WANG Xingtao, REN Duanyang, CAO Wenfu, WANG Jianjun

Institute of Ecological Animal Husbandry, Shanxi Agricultural University, Shuozhou 036001, China

Abstract Intelligent breeding, as the core driving force of technological innovation in modern seed industry, is facilitating the transformation of crop genetic improvement from experience-based approaches to data-driven paradigms. This article systematically reviewed the path of constructing the intelligent breeding technology framework and the progress on its application in crop genetic improvement to provide theoretical foundations and practical references for the intelligent transformation of seed industry in China. The achieved breakthroughs of intelligent breeding in designing complex trait, mining germplasm resource, and predicting genetic gain were summarized based on combining key technical modules including the acquisition of data from high-throughput phenotyping and genotyping, the decision-making empowered by artificial intelligence (AI) algorithm and the integration of multi-omics platform with typical case studies at home and abroad. Intelligent breeding significantly improved the efficiency and precision of breeding by constructing multi-dimensional data-driven models that integrate genotype, phenotype, and environment (G×P×E). In the future, we should strengthen the interdisciplinary integration of biology, information and engineering, develop knowledge-embeddedalgorithms, establish a global collaboration platform, promote the digital sharing of germplasm resources and the reconstruction of commercial breeding systems. It is recommended to establish policy-guided mechanisms to deepen industry-academia-research collaboration, thereby accelerating the translation of technological innovations into the industrial end.

Keywords intelligent breeding; artificial intelligence (AI) algorithms; gene editing; simulation of virtual breeding; intelligent mining; genetic gain; model of deep learning

(责任编辑：张志钰)