

作物表型组大数据技术及装备发展研究

温维亮^{1,2,3}, 郭新宇^{1,2,3}, 张颖^{1,2,3}, 顾生浩^{1,2,3}, 赵春江^{1,2,3*}

(1. 北京市农林科学院信息技术研究中心, 北京 100097; 2. 国家农业信息化工程技术研究中心, 北京 100097;
3. 数字植物北京市重点实验室, 北京 100097)

摘要: 集成自动化平台装备和信息化技术手段, 获取多尺度、多生境、多源异构的作物表型组大数据, 将极大地促进作物功能基因组学、数字育种、智慧栽培的研究进程。本文分析了作物表型组大数据技术及装备的应用需求、产业发展形势, 从物理、传输、数据、知识、应用 5 个层面详细梳理了相应研发现状; 从作物表型组大数据高通量获取、作物表型组大数据智能解析技术两方面着手, 剖析了我国相关技术、装备、产业应用方面的问题和态势。研究建议, 从底层芯片层面突破作物表型传感器关键技术, 在可控开源的基础上形成自主化的表型解析技术体系, 加强作物表型组大数据技术及装备标准体系建设, 提出基因-表型-环境多维大数据驱动的数字育种和智慧栽培创新模式, 建设作物表型组大数据技术及装备人才队伍和协作网络。

关键词: 作物表型组学; 表型大数据; 表型技术及装备; 多组学

中图分类号: S-1 文献标识码: A

Technology and Equipment of Big Data on Crop Phenomics

Wen Weiliang^{1,2,3}, Guo Xinyu^{1,2,3}, Zhang Ying^{1,2,3}, Gu Shenghao^{1,2,3}, Zhao Chunjiang^{1,2,3*}

(1. Information Technology Research Center, Beijing Academy of Agriculture and Forestry Sciences, Beijing 100097, China;
2. National Engineering Research Center for Information Technology in Agriculture, Beijing 100097, China;
3. Beijing Key Laboratory of Digital Plant, Beijing 100097, China)

Abstract: Automatic equipment and information technologies make it possible to acquire multi-scale and multi-source heterogeneous data of crops under different growth conditions, forming big data on crop phenomics. This will greatly promote the research progress of crop functional genomics, digital breeding, and smart cultivation. In this paper, the demand for and industrial development of technology and equipment of big data on crop phenomics are analyzed. Then, the current situation of research and development in this area is summarized from five aspects: data acquisition hardware, data transmission, data analysis, knowledge formation, and applications. The problems and developmental trends of relevant technologies, equipment, and industrial application in China are analyzed from the perspectives of high-throughput acquisition and intelligent analysis of big data on crop phenomics. At last, the following suggestions are proposed: achieving breakthroughs regarding key crop phenotyping sensor technologies from the underlying chip level, forming an autonomous phenotyping extraction technology system on the basis of controllable open source,

收稿日期: 2023-04-21; 修回日期: 2023-06-12

通讯作者: *赵春江, 北京市农林科学院信息技术研究中心研究员, 中国工程院院士, 研究方向为农业信息技术与智能装备;

E-mail: zhaocj@nercita.org.cn

资助项目: 国家重点研发计划项目(2022YFD2002300); 中国工程院咨询项目“生物育种数字化发展战略研究”(2021-JJZD-04), “安徽省智慧农业发展战略研究”(2021-DFZ-17)

本刊网址: www.engineering.org.cn/ch/journal/sscae

strengthening the standards system construction for big data on crop phenomics, creating a new model of genotype-phenotype-environment big data-driven digital breeding and smart cultivation, and building a talent pool and collaborative network for crop phenomics.

Keywords: crop phenomics; phenotyping big data; technology and equipment for phenotyping; multi-omics

一、前言

作物表型指能够反映作物细胞、组织、器官、植株、群体结构及功能特征的物理、生理、生化性状，实质上是作物基因图谱的时序三维表达、地域分异特征、代际演进规律^[1]。随着作物科学、计算机科学与工程等领域协同研究程度的不断提高，通过多源传感、理化分析等方式采集的作物表型数据积累形成了多维度、多尺度的作物表型组大数据。作物表型性状的复杂程度决定了表型组大数据的获取、解析、管理、应用是极大的科学系统工程。

当前，作物表型组大数据已成为国际农业科学、生命科学领域的战略前沿方向^[2]，也被视为种业科技发展的核心竞争力^[3]。在我国，与信息技术、智能装备深度融合的作物表型组学成为抢占未来农业产业发展制高点的关键内容，相关研究发展趋势鲜明：由新兴学科建设转向学科规模化建设，由单一尺度及关键时间点的表型解析转向综合化、智能化及全生命周期的连续检测，表型数据组内关联转向多组学协同创新。

近年来，研究者从作物表型获取技术、表型平台装备、表型解析算法、多组学数据挖掘分析等方面着手，分析了作物表型组学的发展历史、面临挑战、未来趋势^[4-10]。然而，立足国情并把握行业发展实际，围绕我国作物表型组大数据技术及装备开展的综合性发展研究依然缺乏。为此，本文立足国内外作物表型组学已有研究成果，结合我国作物表型组大数据技术及装备的研发现状和产业发展实际，明晰定位、梳理现状、剖析问题并形成技术发展建议，以期作物表型组学及农业科技发展研究提供基础参考。

二、作物表型组大数据技术及装备的发展需求与产业牵引

作物表型组大数据技术及装备正处于快速发展阶段，既是我国农业科研和生产的需要，也得益于作物表型产业的牵引。

(一) 作物表型组大数据技术及装备发展需求分析

1. 实现种业科技自立自强的需要

开展种业“卡脖子”技术攻关、打好种业翻身仗、建设种业强国，离不开作物表型组大数据技术及装备的关键支撑。表型数据的检测贯穿于种业生产的品种繁育、品种测试、示范推广各个环节，但传统的表型采集和鉴定方法费时耗力、标准化程度不足，迫切需要表型组大数据技术装备有力支撑育种决策。现代作物表型组大数据技术及装备的发展最早是由巴斯夫股份公司、拜耳集团、孟山都公司等全球种业巨头推动的，旨在通过信息感知、自动控制、智能解析技术提升作物表型检测的通量、效率和标准化程度。在人工气候室内设置理想种植条件，通过调节光周期促进植物生长达到加代目的，在一年内得到4~6代，大幅缩短育种周期^[11]。借助作物表型组大数据技术装备和分子育种辅助技术，有望实现加速育种技术的革命性突破^[12]。

2. 推进作物种质资源表型性状精准化、规模化鉴定的需要

农作物种质资源是保障国家粮食安全与重要农产品供给的战略性资源，农业科技原始创新与现代种业发展的物质基础。作物表型鉴定通过表征作物种质资源的农艺性状、产量性状、品质性状、抗性性状等特性，旨在发掘具有生产利用价值但未被充分利用的特异性种质及其基因^[13]。我国农作物种质资源评价多为单一性状、单一环境下的鉴定结果，缺乏基因信息和综合评价，限制了种质资源在育种中的有效利用（《全国农作物种质资源保护与利用中长期发展规划（2015—2030年）》）。目前，我国农作物种质资源80%以上来自国内，遗传多样性不足；在长期保存的52万份种质资源中，开展基因型和表型精准鉴定的数量不到10%；在现存的2.5万份玉米、4万份小麦和8万份水稻种质资源中，开展深度鉴定的只有5%^[14]。此外，作物的抗旱、抗倒、高光效等综合表型是多基因控制的复杂数量性状，对表型鉴定提出了较高要求。亟需借助作物表型组大数据技术及装备，开展表型性状规模化、精准化鉴定，以有效推进作物种质资源的保护与利用。

3. 加快突破作物重要性状形成机制基础研究的需要

农作物高产、优质、高效及其协同改良是作物科学重要的研究内容。综合利用遗传学、基因组学、分子生物学等技术手段,挖掘株型、穗型、种子大小等产量性状,蛋白质、油分、硬度等品质性状,养分吸收、转运、代谢等养分高效利用性状的关键调控基因,阐明高产、优质、高效协同改良的分子调控网络,对创制产量、品质、资源高效利用均显著提升的优异新基因资源具有重要意义。然而,传统的表型测定方法存在通量低、标准不一致、深度不足等问题,限制了重要性状形成机制的突破进程^[2]。借助作物表型组大数据技术装备,可以从作物近缘种、野生种、地方品种中找到克服各类逆境的性状,深入开展控制优异性状的关键基因和遗传网络解析,进而克隆相关基因^[5],为分子设计育种^[6]的定向改良提供基因资源。

4. 提升作物智慧生产信息感知技术装备的需要

制定作物生产中耕、种、管、收等关键环节的作业方案,离不开对农田长势状况的判定。当前,我国农作物在生产过程中主要依靠人工到田间地头利用肉眼进行观测、凭借经验进行决策,造成农业生产效率不高,还会存在误判和漏判的风险,直接影响作物产量和农民效益的提升。理论和实践均表明,作物表型组大数据技术及装备可应用于粮食、蔬菜、花卉等大多数农作物生产管理环节的长势诊断,通过移动式或固定式平台搭载可见光、光谱、激光雷达、红外热成像等传感器,快速分析不同尺度农田作物的长势、水分、氮营养、病虫害、产量等信息;结合作物生长模型^[7]和知识管理模型,定量估算作物对水、肥、药的需求量,进而指导农业机械进行精准作业,实现作物产量、品质、资源利用效率的协同提升^[8]。

(二) 以作物表型组大数据技术及装备为核心的表型产业方兴未艾

1. 市场规模迅速增长

植物表型市场规模正在经历快速增长。Coherent Market Insights 公司的研究报告显示,2021 年全球植物表型市场规模约为 1.84 亿美元,预计 2028 年为 3.87 亿美元,复合年增长率为 11.2%^[9]。该市场的迅速增长主要得益于以下方面:一是全球气候变化

和人口增长,使国际粮食安全面临严峻挑战,迫切需要通过科技创新提升粮食产能^[20];二是信息技术的快速发展,如机器学习、计算机视觉、传感器网络、空天遥感的不断进步,使植物表型检测和鉴定变得更加自动化、数字化、高通量,提高了评估效率和准确性^[9];三是育种基础性研究需求的持续增加,对作物种质资源表型性状鉴定提出了多生境、时序性、优质多抗的要求^[13];四是政策和资金的支持,为市场主体在技术研发和推广应用方面创造了良好环境。

2. 产业链条逐步延伸

当前,作物表型组大数据技术及装备主要应用于基因鉴定^[21]、种质资源表型鉴定^[13]、生物和非生物胁迫的无损检测等作物科学研究^[22]。面对保障国际粮食安全和农产品质量安全、促进农业增产与农民增收等世界各国的共同诉求,作物表型组大数据技术装备还可应用于农业生产智能化管理、农产品质量无损检测、农产品智能收获等多类场景,为不同农业产业市场主体(如家庭农场、植物育种公司、农业科研机构、智慧农业公司、农产品贸易商)提供信息化技术和智能化装备支撑。

3. 商业模式不断涌现

全球作物表型产业的业务和产品主要以表型获取基础设施平台的搭建为主,如大田的龙门吊式平台、温室的悬挂式轨道平台、地面移动式平台、低空无人机平台等^[7]。按照产品品类口径,2021 年全球植物表型市场中硬件设备占主导,超过软件和传感器之和。也要注意,基于传感器数据的解析指标主要是形态指标,缺乏对深度表型和综合性状的解析,因而提供智能化的表型解析软件和在线化的表型分析平台将成为新的商业模式。作物表型平台是涉及到自动控制、数据传输、设备调试、系统集成等环节的复杂系统^[23],其运行保障、硬件维护、软件服务等具有一定的技术门槛,将成为作物表型产业的另一类商业模式。

4. 育种龙头企业发展的必由途径

构建以商业化育种流程管理软件、表型高通量精准鉴定平台、表型组大数据管理系统等信息技术和智能装备为核心的商业化育种技术体系,是育种企业数字化转型升级、实现数字技术赋能企业生产经营的重要内容,具有分段式、流程化、标准化的特点,有助于缩短大规模育种周期、促进企业提质

增效。国际种业巨头如拜耳集团、先锋公司、巴斯夫股份公司等，高度重视作物表型组学的发展及其产业应用，纷纷部署作物表型高通量获取平台及设施并纳入商业化育种和发展数字种业的业务流程。作为我国种业龙头企业的北大荒垦丰种业股份有限公司已于2016年建设运行了高通量表型精准鉴定平台，华为技术有限公司正在积极探索表型机器人和智慧育种多组学大模型，袁隆平农业高科技股份有限公司、甘肃省敦煌种业集团有限公司等企业也在积极建设相关能力。

三、作物表型组大数据技术及装备研发现状

作物表型组大数据是一项科学系统工程，以作物育种和栽培的实际需求为导向，依赖传感器、表

型平台、无线通信、数据库、大数据分析等现代信息技术和机械装备，需要农学、植物学、自动化、机械工程、图形图像、计算机科学等多学科紧密协作，才能将作物表型组大数据最终转化为生物学和农学新知识。借鉴大数据知识工程发现模式^[24]，作物表型组大数据从数据的产生到最后的的应用，可分为物理层、传输层、数据层、知识层、应用服务层（见图1）。

（一）物理层

21世纪以来，在作物表型数据高通量获取方面，以各类新型物理、化学和生物（生理）传感器、图形图像技术、人工智能技术、物联网技术为代表的新一代表型获取技术体系，正在为作物研究提供海量表型和环境数据源^[25]。物理层指利用传感

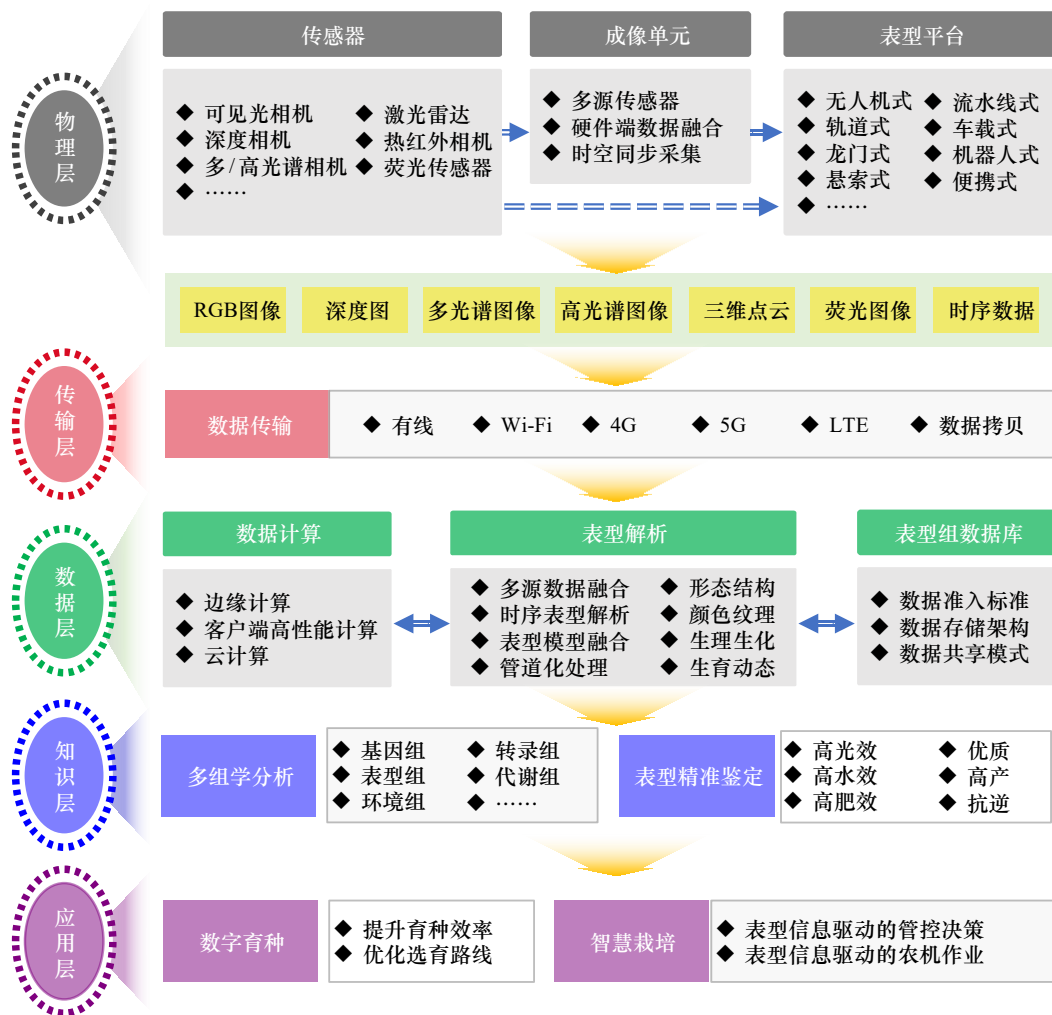


图1 作物表型组大数据技术及装备从研发到应用路线图

注：RGB表示红、绿、蓝3种颜色；Wi-Fi表示无线保真；4G表示第四代移动通信技术；5G表示第五代移动通信技术；LTE表示通用移动通信技术的长期演进。

器、表型平台等数据获取手段，获取作物的图像、点云、光谱等初始表型数据，重点关注数据的高通量获取；具体又可分为传感器、成像单元、表型平台三方面。

1. 作物表型传感器

作物表型常用的传感器有可见光相机、深度相机、多光谱相机、高光谱相机、激光雷达^[26]、热红外相机、叶绿素荧光传感器等。目前，可见光相机是应用最为广泛的传感器，其次是多光谱相机、高光谱相机、激光雷达等^[10]。可见光相机满足多数条件下的作物表型成像需求，具有分辨率高、采集快的特点，但在复杂田间条件下连续采集的稳定性仍待提升。近年来，激光雷达、深度相机、多光谱相机、热红外相机等传感器的分辨率、精度和稳定性显著提升，成本逐步降低，为作物表型提供了多种选择方案，但整体而言分辨率和精度仍有较大提升空间。高光谱相机、叶绿素荧光传感器成本较高且成像范围有限，在表型研究中的使用率有待提升。

2. 作物表型成像单元

多源表型数据的时空同步获取是表型信息采集的重要趋势，有利于提升效率和精度。作物表型成像单元意在整合多传感器以实现表型数据的时空同步采集，可分为多源成像单元、同源多个传感器组成的成像单元。典型的多源成像单元如 Altum 三合一传感器可以实现多光谱、热成像、RGB 信息的同步采集；PlantEye 可以实现点云、RGB、多光谱数据的同步获取与实时解析；CropLidar 集成了激光雷达、多光谱、可见光传感器，可以挂载于无人机或轨道式平台实现多源数据同步采集。典型的同源成像单元是多目立体视觉系统，成像单元可以挂载于无人机、轨道式、车载式等平台。高集成度、小型化、同步控制、数据实时融合解析是作物成像单元的难点问题。

3. 作物高通量表型平台

作物高通量表型平台系统整合了传感器或成像单元，通过传动装置、系统控制、数据传输与计算等实现高通量、自动化的表型数据采集。近年来，室内植物表型平台、大田植物表型平台、便携式表型采集设备、地面及航空机器人采集平台等多层次表型获取技术发展迅猛^[7, 27]。国内作物表型平台建设及应用发展迅速，但主要依靠进口。中国农业科学院生物技术研究所、中国科学院遗传与发育生物

学研究所、北大荒垦丰种业股份有限公司、上海市农业科学院等分别购置并部署了 LemnaTec 高通量表型平台，实现最多 680 盆作物植株的高通量数据采集。南京农业大学作物表型交叉研究中心构建了挂载多个 PlantEye 传感器的田间高通量表型平台 FieldScan、温室传送型高通量表型平台 PhenoConveyor。在自主研发方面，华中农业大学研制了国内第一套水稻表型自动检测分析系统^[28]，陆续发展了室内流水线式作物单株表型平台、水稻考种机等；南京农业大学自主设计并建造了田间作物表型舱、人工智能气候舱；北京市农林科学院信息技术研究中心先后研发了温室轨道式^[29]、大田轨道式^[23]、室内流水线式、便携式^[30]、无人机和果穗考种^[31]等表型平台。

（二）传输层

作物表型传感器、表型平台获取到表型数据后，需要将之传输到具有高性能计算能力的服务器进行数据解析。作物表型数据量庞大，数据传输受到距离、速度等诸多因素的限制。虽然在采集端基于边缘计算^[32]的作物表型实时解析是降低数据传输任务量的解决方案和未来趋势，但尚未广泛应用，也在一定程度上降低了高分辨率数据的应用质量。数据上云并在云端解析是表型数据处理的发展方向，故表型数据的传输仍是作物表型组大数据链条中的重要问题。目前，固定区域的表型平台可通过网络传输，但多无法达到实时水平。例如，在田间或温室轨道式表型平台附近构建控制室，通过有线或无线的方式进行数据传输^[23]；对于小型化的表型平台，可构建局域网减少控制和数据传输所用线缆^[33]，提升平台整体性能。对于非固定区域的表型平台如无人机、便携式平台等，多采用人工拷贝的方式下载数据。5G 技术是未来表型平台数据传输的理想解决方案，但需部署基站而增加表型平台的建设成本。

（三）数据层

数据层（又称信息层）指由获取的原始数据到含有语义信息、高度结构化表型性状的过程，具体包括表型解析、数据计算、数据库构建三部分。

1. 表型解析算法

在作物表型智能解析方面，将初始数据转化为

具有生物学意义的表型性状至关重要。近年来，各类计算机视觉算法、机器学习和深度学习方法在表型数据解析中得到大规模应用，极大地推动了表型大数据的分类、解析与可视化。通过融合专家先验知识，从各种结构化和非结构化的数据中实现了作物形态结构^[34, 35]、颜色纹理^[36]、生理生化^[37]、生育动态^[38]等重要表型性状的解析。由于图像数据便于获取、数据标注方便、面向图像的深度学习框架较为丰富，基于图像的作物表型解析算法应用进展良好^[39]，在实时性和稳定性方面达到较好的效果；而基于点云和光谱等数据的作物表型解析算法研究及应用相对滞后。

2. 数据计算

表型解析的计算效率决定了后续应用的实时性。目前，作物表型数据主要是在采集后通过实验室中的高性能工作站进行计算解析。在采集端进行实时计算解析以图像表型和测距为主，多在数据的精度上有所折衷；但这种实时计算在以栽培为应用场景的实时决策中非常重要，需逐步引入边缘计算以提升数据计算的实时性。此外，由于获取的初始表型数据体量较大，需在自动化、管道化的表型解析算法基础上，由部署在云端的算力进行数据的云计算。

3. 表型数据库构建

作物表型数据库主要用于系统地整理结构化和半结构化的作物表型大数据^[40]。研究者致力于开发表型数据管理系统^[41]并建立包含多物种、多品种、多生境、多尺度的作物表型数据库。受表型数据获取手段、数据解析算法、作物种质资源等因素的限制，构建的作物表型数据库多是围绕单一物种的单一尺度表型组数据库，如显微尺度的玉米茎秆维管束表型数据库^[42]、用于分布式植物表型分析的开源信息管理系统——基于物联网的表型数据分析平台（CropSight）^[43]等。

随着高通量检测技术的快速发展，作物生物学数据朝着多组学、多维度的层面快速积累。多组学信息的系统整合将进一步加速作物遗传研究、改良作物的优异农艺性状。2020年，华中农业大学成功整合了来自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、遗传定位结果等多组学数据，构建了玉米定制化多组学数据库（ZEAMAP）^[44]；通过多维度生物组学数据的联合分

析，挖掘株型、产量等性状相关的重要基因和遗传变异，实现了玉米多组学数据“云端”集成、快速检索、智能分析，为分子设计育种提供了理论基础^[45]。2022年，北京市农林科学院通过人工和前沿表型组学技术对全球超过1000份生菜种质资源进行表型数据获取，结合超过1000种生菜品种的重测序数据，构建了综合生菜数据库（LettuceGDB）^[46]。

（四）知识层

知识层旨在从多源海量的作物表型组大数据中挖掘出能用于解决特定领域问题的知识，通过数据知识化实现数据增值^[24]，主要分为表型精准鉴定、多组学分析两类。

1. 表型精准鉴定

表型信息的获取是作物种质资源精准鉴定的重要环节。传统人工检测的问题日益突出，必须综合考量准确性、通量、成本之间的平衡。随着作物表型技术及装备的发展完善，表型获取通量、指标解析精度、效率等得到显著提升，为种质资源的规模化、批量化鉴定评价提供了硬件基础和技术支撑。国内外科研单位、种业优势企业研发了以“作物表型高通量获取—智能解析—模型计算—大数据分析—表型精准鉴定”为主线的技术体系，在高光效^[47]、抗倒伏^[48]、高水效^[49]等综合性状表型精准鉴定中得到成熟应用。我国农作物资源丰富，但在种质资源精准鉴定方面与国外相比还有较大差距。随着高通量表型技术的发展与完善，在完成主要作物基因组精细图谱绘制、作物种质资源基因组重测序的基础上，系统且精准地鉴定作物遗传资源重要农艺性状（如产量、品质、抗病、耐逆、养分高效等），深入挖掘重要农艺性状优异等位变异并阐明其形成的分子机制，不仅可为我国作物育种取得新突破、保障农业可持续发展提供优异资源和重要基因，而且可为解析作物驯化与改良的分子机制提供理论依据。

2. 多组学分析

随着高通量测序技术的发展与完善，单组学研究日趋成熟，而整合多组学数据研究植物生长发育的工作方兴未艾。多组学研究在作物重要基因挖掘、全基因组关联分析、基因表达调控网络构建、作物全基因组选择、系统生物学研究等方面发挥着日益重要的作用。基于丰富的多组学数据，差异表达基因识别、复杂表型的转录因子识别或代谢物富

集等遗传调控研究进展迅速,代表性工作有:利用显微电子计算机断层扫描(CT)表型技术结合全基因组关联分析(GWAS),鉴定到调控维管束数目、维管束分布密度等性状的特异候选基因^[42];利用高通量表型平台对玉米进行连续无损检测,揭示玉米抗旱的遗传基础以及潜在抗旱位点^[50];基于三维全自动高通量表型分析平台,通过提取图像性状(i-traits)结合全基因组关联研究(GWAS),解析玉米株高形成的动态遗传基础和调控网络^[51];结合表型数据和预测模型,揭示调控玉米节间发育的基因表达模式,发掘一批参与节间数和节间长度形成的候选基因,验证关键候选基因ZmD1调控玉米节间发育的分子机制^[52];基于二维图像根系高通量表型平台,发掘81个玉米根系构型候选基因,利用转基因玉米明确其中两个候选基因(ZmRSA3.1、ZmRSA3.2)的功能^[53]。

2020年, *Nature Review Genetics* 期刊以“遗传学和基因组学的未来之路”为题刊登述评,将“解码多因素表型”列为未来重点方向之一。基因调控网络(GRN)对多个基因表达的调控以及GRN的变化如何引起特定细胞、组织、器官、植株的反应,将是解码多因素表型的有效途径。涵盖微观表型与宏观表型多尺度的作物表型组大数据,将大规模发现基因、表型、环境之间,微观表型与宏观表型之间的跨尺度关联及相互作用,精准解析表型与分子机制之间未发现的联系,从而构建精准的“基因-表型-环境”调控网络^[45]。

(五) 应用层

1. 数字育种

多组学的智能设计育种是新一代育种核心技术。作物智能设计育种在作物基因组学、表型组学等大数据的基础上,通过机器学习等智能算法构建目标作物品种的性状预测模型,预测杂交种的各种农艺性状表现,能够优化品种选育技术路线、提高精准育种效率、快速实现育种目标^[54,55]。孟山都公司在其创建的现代农业育种流程中,利用高通量表型精准鉴定并与基因组数据偶联对植物进行改良,选育出许多优异种质材料,显著提高了育种效率。目前,利用高通量表型实现作物育种的应用案例还较少。这是因为,一方面一线育种学家更倾向于围绕具有直观生物学意义的表型性状(breeder-oriented)

开展育种研究,另一方面利用图像、点云、光谱数据解析得到的综合表型(technology-oriented)所表征的性状较为复杂,难以直接与育种目标性状建立关联,尚未在育种中广泛应用^[56]。

2. 智慧栽培

作物智慧栽培旨在建立覆盖耕、种、管、收作物生产全程,人、机、物全生产要素互联互通并可互操作的智能管控平台,实现良田-良种-良法、信息-农艺-农机的有机融合,最终达到作物高产、优质、高效、生态、安全的生产目标^[57]。作物智慧栽培对作物的感知即利用作物表型组大数据技术及装备,对作物生产过程中的作物参量进行实时采集,进而为定量作物生产系统各要素的关系、优化生产技术规程及管理方案提供数据支撑。当前,利用无人机获取田间作物长相长势表型信息可指导田间作物病害的防控和管理,利用图像深度学习方法可实现田间机器人的除草作业,但受表型获取技术装备时效性的限制,尚难以将农机与表型技术装备高效整合。因此,整合农机与表型技术装备、在田间打通“联接-感知-认知-管控”通道、实现田间栽培管理的实时感知和决策,是未来作物表型组大数据技术及装备开展智慧栽培应用的难点和方向。

3. 实用性分析

从应用角度出发,用户主要关心作物表型大数据技术及装备的数据获取自动化水平、数据处理在线化水平、最终得到表型性状的数据精度水平。然而,这些实用性指标受到诸多因素的影响(见图2):①目标作物本身的形态结构复杂性以及栽培/部署的一致性,对于表型获取和解析具有根本性的影响;②表型数据获取环节的传感器、成像单元和表型平台决定了数据获取的无人化水平;③数据传输、数据预处理和表型解析算法的智能化水平决定了表型数据处理的在线化水平。在表型数据“获取-解析-应用”整个链条中,每个环节都对下一步有着重要影响,会形成误差累积并逐级放大。因此,提升各环节的数据质量和算法精度,对于整体性提升作物表型组大数据技术及装备的实用性至关重要。

四、我国作物表型组大数据技术及装备的发展问题与态势分析

作物表型组大数据技术及装备仍处于快速发展

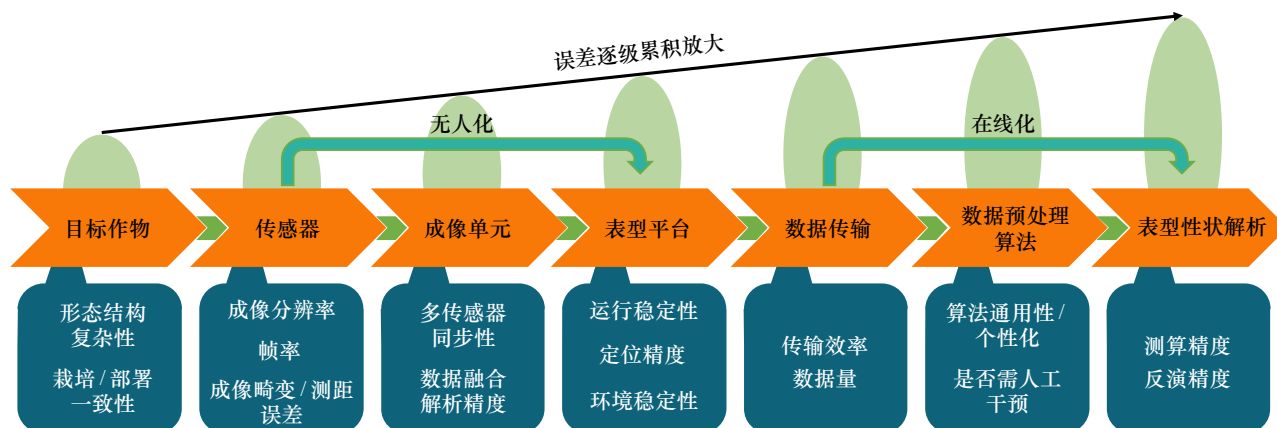


图2 作物表型组大数据技术及装备实用性分析示意图

过程之中，我国在作物表型高通量信息获取与解析方面已有一定积累，但整体上处于跟踪模仿阶段，未能打破表型高通量获取装备主要被欧美国家所垄断的局面；特别是多样化表型配套设施和低成本表型传感器及设备的自主研发、表型大数据实时与高效传输技术、表型多尺度数据融合与组学大数据挖掘理论方法等，亟待解决或突破。

（一）作物表型组大数据高通量获取发展问题与态势分析

我国在作物表型组大数据获取方面整体处于“跟跑”“并跑”阶段，主要表现在以下三方面。

1. 传感器

高分辨率、高精度的传感器和核心部件主要靠进口，存在着价格昂贵、软硬件升级改造受限、关键技术无法定制、表型数据安全隐患等问题。自主研发的传感器光学成像方式单一，在稳定性、数据分辨率方面与国外先进产品仍有差距，尤其是在复杂田间条件下长时间连续获取数据的传感器存在严重的稳定性问题。

2. 成像单元

由于单项传感器的研发水平相对滞后，受制于传感器的集成尺寸、接口协议等因素，国内研发的成像单元与国外先进产品差距显著，主要表现在质量及尺寸大、系统集成性不高、整体运行稳定性低等方面，尚未见到可以同步获取点云、图像、光谱数据，在传感器层面高度集成的作物表型成像单元。为实现高质量多源数据的同步连续采集，只能进口如PlantEye等成像单元，但成本高、后续表型解析个性化程度低，难以满足国内大量多源表型数

据获取的实际需求。

3. 高通量表型平台

国外作物高通量表型平台和基础设施发展较早，具有系统集成度高、稳定性好等特点，我国作物高通量表型平台和基础设施整体仍以进口为主。进口产品购置、运营和维护成本高，平台在使用过程中出现问题后反馈周期长，导致部分进口平台尤其是中大型表型平台“用不起来”，平台考虑普适性而难以对特定作物给出高精度、定制化的表型获取解决方案，硬件控制核心算法和设计“黑盒子化”也难以满足后续表型数据处理与解析软件对不同作物的表型获取需求。近年来，国内高校及科研院所自主研发了多生境的作物高通量表型平台，虽然在美观程度、稳定性、自动化水平方面略逊于进口平台，但因研发者与用户沟通便捷、反馈周期短、可开展定制化表型解析和软件开发等诸多优势，形成了一批真正实用的表型平台产品，如华中农业大学的轨道式表型平台^[28,50]、北京市农林科学院的轨道式表型平台^[23,58]和玉米果穗流水线考种系统^[31]等。

（二）作物表型组大数据智能解析技术发展问题与态势分析

近十年是作物表型组大数据技术装备快速发展阶段，国际表型组学研究与应用仍面临诸多的问题和挑战，主要表现在数据传输、表型解析算法软件、表型信息的挖掘与应用等方面。可以认为，国内外在这些方面形成了“并跑”之势。

1. 表型智能解析

当前，作物表型解析算法多针对主要性状以少

量数据进行测试,虽可以解析到目标性状,但多需要人工交互,存在普适性差的问题,在处理作物表型组大数据过程中表现得尤为明显。迫切需要开发无需人工交互、自动化、管道化的表型解析算法,适应作物表型组大数据批量化处理的需求。作物表型解析的研究重点是增加可观测、可量化、具有明确生物学含义的作物表型性状数量,提高可量化作物表型解析的精度和效率;针对不同作物的形态结构和生理生态功能,研发相关算法开展定制化的表型解析。当前作物表型解析的时效性不足,无法满足诸如面向农机作业实时决策等智慧栽培的需求。通过多源数据融合提升表型解析精度^[58]、通过表型模型融合增强解析过程的可解释性、通过优化解析算法的鲁棒性实现管道化处理,是表型解析未来发展的趋势。

2. 表型大数据融合分析

当前,国内外作物表型组大数据技术及装备仍处于研发阶段,表型大数据分析、应用与服务能力不足。随着多生境表型平台的快速发展,多维度、多尺度、多源表型大数据正在大量累积;如果多尺度、多模态、多生境的表型大数据不进行有效整合,将很难充分发挥大数据的潜在价值。因此,面向表型大数据的融合分析、知识挖掘和应用服务是国内外需要解决的共性难题。

五、我国作物表型组大数据技术及装备发展建议

以包括人工智能在内的新一代信息技术为依托,通过表型数据的无人化获取、在线化解析来实现作物表型组大数据的高通量及精准积累,是发展作物表型组大数据技术及装备的主导技术路径。建设作物表型组大数据技术及装备方面的大科学工程,形成作物表型组大数据“获取-解析-利用”技术体系,实现“基因-表型-环境”多维组大数据整合与分析利用,从而在数字育种、智慧栽培方向提出中国方案。相关目标可分为3个实施阶段:①积极在传感器、成像单元、表型平台、基础设施建设方面形成标志性产品,努力在大数据融合、表型解析关键技术方向取得重要突破,使部分技术与应用达到世界领先水平;②作物表型组大数据技术及装备的国产化率高于95%,形成数字育种、智慧

栽培的代表性应用案例,使核心技术及装备处于与国际“并跑”态势;③建成作物表型组大数据技术及装备的自主创新体系,整体性实现自主产品替代进口,形成面向数字育种、智慧栽培的商业化服务模式并构建作物表型服务新业态,整体上处于“领跑”态势。针对上述目标,提出我国作物表型组大数据技术及装备发展的具体建议。

(1)从底层芯片层面突破作物表型传感器关键技术,解决基础传感器成本高、整合难、采购受限等“卡脖子”问题;研发能用、好用、用得起的表型技术及装备产品,形成稳定性好、高度自主化、规模化的成像单元、表型平台设备及相关基础设施,实现可快速定制、面向不同作物、多生境、个性化的作物表型高通量获取解决方案,带动我国在新一代农业传感器和表型平台技术装备创新从“跟跑”“并跑”转向“领跑”。

(2)在可控开源的基础上,融入小样本学习、预训练大模型、知识图谱等人工智能技术,突破多源数据融合、系列作物多维表型性状智能解析、时序表型解析等关键技术,形成自主化的表型解析技术体系;构建作物表型组大数据解析“大脑”,使高通量获取的作物表型组大数据成为农学家切实可行的数据,促进数据和知识增值赋能。

(3)加强作物表型组大数据技术及装备的标准体系建设,提升作物表型组大数据结构化水平,降低数据噪声及获取和使用成本,提高表型数据可用性;建设国家级和区域尺度的表型组大数据搜索引擎、新型基础资源服务平台,提升作物表型组大数据的整合程度和应用安全性。

(4)提出“基因-表型-环境”多维大数据驱动的数字育种和智慧栽培创新模式,实施组学大数据与表型精准鉴定设施的大科学工程;逐步形成智能化、无人化、在线化的表型工厂服务模式,构建基于作物表型组大数据的数字育种和智慧栽培协同创新平台。

(5)建设作物表型组大数据技术及装备的人才队伍和协作网络,形成多学科交叉合作与协同创新机制,培育涵盖技术装备研发、表型软硬件产品应用、设备设施运维管理、数据与知识挖掘、作物育种与栽培应用环节的人才链条,力争在作物表型组大数据技术及装备领域开创“人无我有、人有我强”的引领式发展格局。

利益冲突声明

本文作者在此声明彼此之间不存在任何利益冲突或财务冲突。

Received date: April 21, 2023; **Revised date:** June 12, 2023

Corresponding author: Zhao Chunjiang is a research fellow from the Information Technology Research Center, Beijing Academy of Agriculture and Forestry Sciences, and a member of the Chinese Academy of Engineering. His major research field is agricultural information technology and intelligent equipment. E-mail: zhaocj@nercita.org.cn

Funding project: National Key R&D Program of China (2022YFD2002300); Chinese Academy of Engineering projects “Strategic Research on the Digital Development of Biological Breeding” (2021-JJZD-04), “Strategic Research on Smart Agriculture Development in Anhui Province” (2021-DFZ-17)

参考文献

- [1] 赵春江. 植物表型组学大数据及其研究进展 [J]. 农业大数据学报, 2019, 1(2): 5–18.
Zhao C J. Big data of plant phenomics and its research progress [J]. Journal of Agricultural Big Data, 2019, 1(2): 5–18.
- [2] Ninomiya S, Baret F, Cheng Z M. Plant phenomics: Emerging transdisciplinary science [J]. Plant Phenomics, 2019, 2019: 2765120.
- [3] Araus J L, Cairns J E. Field high-throughput phenotyping: The new crop breeding frontier [J]. Trends in Plant Science, 2014, 19(1): 52–61.
- [4] 周济, Tardieu F, Pridmore T, 等. 植物表型组学: 发展、现状与挑战 [J]. 南京农业大学学报, 2018, 41(4): 580–588.
Zhou J, Tardieu F, Pridmore T, et al. Plant phenomics: History, present status and challenges [J]. Journal of Nanjing Agricultural University, 2018, 41(4): 580–588.
- [5] Zhao C, Zhang Y, Du J, et al. Crop phenomics: Current status and perspectives [J]. Frontiers in Plant Science, 2019, 10: 714.
- [6] Yang W N, Feng H, Zhang X H, et al. Crop phenomics and high-throughput phenotyping: Past decades, current challenges, and future perspectives [J]. Molecular Plant, 2020, 13(2): 187–214.
- [7] Jin X, Zarco-Tejada P, Schmidhalter U, et al. High-throughput estimation of crop traits: A review of ground and aerial phenotyping platforms [J]. IEEE Geoscience and Remote Sensing Magazine, 2020, 9(1): 200–231.
- [8] Pieruschka R, Schurr U. Plant phenotyping: Past, present, and future [J]. Plant Phenomics, 2019, 2019: 7507131.
- [9] 胡伟娟, 傅向东, 陈凡, 等. 新一代植物表型组学的发展之路 [J]. 植物学报, 2019, 54(5): 558–568.
Hu W J, Fu X D, Chen F, et al. A path to next generation of plant phenomics [J]. Chinese Bulletin of Botany, 2019, 54(5): 558–568.
- [10] Tao H, Xu S, Tian Y, et al. Proximal and remote sensing in plant phenomics: 20 years of progress, challenges, and perspectives [J]. Plant Communications, 2022, 3(6): 100344.
- [11] Watson A, Ghosh S, Williams M J, et al. Speed breeding is a powerful tool to accelerate crop research and breeding [J]. Nature Plants, 2018, 4(1): 23–29.
- [12] Watt M, Fiorani F, Usadel B, et al. Phenotyping: New windows into the plant for breeders [J]. Annual Review of Plant Biology, 2020, 71: 689–712.
- [13] 王晓鸣, 邱丽娟, 景蕊莲, 等. 作物种质资源表型性状鉴定评价: 现状与趋势 [J]. 植物遗传资源学报, 2022, 23(1): 12–20.
Wang X M, Qiu L J, Jing R L, et al. Evaluation on phenotypic traits of crop germplasm: Status and development [J]. Journal of Plant Genetic Resources, 2022, 23(1): 12–20.
- [14] 全国人民代表大会常务委员会专题调研组. 关于加强种质资源保护和育种创新情况的调研报告 [R/OL]. (2021-10-21)[2023-04-15]. <http://www.npc.gov.cn/npc/c30834/202110/b05ec4244ddb4fab3eb560c5452258d.shtml>.
Standing Committee of the National People’s Congress Special Research Group. Research report on strengthening the conservation of germplasm resources and breeding innovation [R/OL]. (2021-10-21)[2023-04-15]. <http://www.npc.gov.cn/npc/c30834/202110/b05ec4244ddb4fab3eb560c5452258d.shtml>.
- [15] Song P, Wang J L, Guo X Y, et al. High-throughput phenotyping: Breaking through the bottleneck in future crop breeding [J]. Crop Journal, 2021, 9(3): 633–645.
- [16] 万建民. 作物分子设计育种 [J]. 作物学报, 2006 (3): 455–462.
Wan J M. Perspectives of molecular design breeding in crops [J]. Acta Agronomica Sinica, 2006 (3): 455–462.
- [17] 朱艳, 汤亮, 刘蕾蕾, 等. 作物生长模型(CropGrow)研究进展 [J]. 中国农业科学, 2020, 53(16): 3235–3256.
Zhu Y, Tang L, Liu L L, et al. Research progress on the crop growth model CropGrow [J]. Scientia Agricultura Sinica, 2020, 53(16): 3235–3256.
- [18] 曹卫星, 朱艳, 田永超, 等. 作物精确栽培技术的构建与实现 [J]. 中国农业科学, 2011, 44(19): 3955–3969.
Cao W X, Zhu Y, Tian Y C, et al. Development and implementation of crop precision cultivation technology [J]. Scientia Agricultura Sinica, 2011, 44(19): 3955–3969.
- [19] Coherent Market Insights. Plant phenotyping market [EB/OL]. (2021-08-01)[2023-04-15]. <https://www.coherentmarketinsights.com/market-insight/plant-phenotyping-market-4584>.
- [20] Gojón A, Nussaume L, Luu D T, et al. Approaches and determinants to sustainably improve crop production [J]. Food and Energy Security, 2023, 12(1): e369.
- [21] Araus J L, Kefauver S C, Zaman-Allah M, et al. Translating high-throughput phenotyping into genetic gain [J]. Trends in Plant Science, 2018, 23(5): 451–466.
- [22] Singh A, Ganapathysubramanian B, Singh A K, et al. Machine learning for high-throughput stress phenotyping in plants [J]. Trends in Plant Science, 2016, 21(2): 110–124.
- [23] Fan J C, Li Y L, Yu S, et al. Application of Internet of Things to agriculture—The LQ-FieldPheno platform: A high-throughput platform for obtaining crop phenotypes in field [J]. Research, 2023, 6(2): 0059.
- [24] 郑庆华, 刘欢, 龚铁梁, 等. 大数据知识工程发展现状及展望 [J]. 中国工程科学, 2023, 25(2): 208–220.
Zheng Q H, Liu H, Gong T L, et al. Development and prospect of big data knowledge engineering [J]. Strategic Study of CAE, 2023, 25(2): 208–220.
- [25] Sun D, Robbins K, Morales N, et al. Advances in optical phenotyping of cereal crops [J]. Trends in Plant Science, 2021, 27(2):

- 191–208.
- [26] Jin S C, Sun X L, Wu F F, et al. Lidar sheds new light on plant phenomics for plant breeding and management: Recent advances and future prospects [J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2021, 171: 202–223.
- [27] 张慧春, 李杨先, 周宏平, 等. 植物表型平台与图像分析技术研究进展与展望 [J]. *农业机械学报*, 2019, 51(3): 1–22.
Zhang H C, Li Y X, Zhou H P, et al. Research progress and prospect in plant phenotyping platform and image analysis technology [J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2019, 51(3): 1–22.
- [28] Yang W N, Guo Z L, Huang C L, et al. Combining high-throughput phenotyping and genome-wide association studies to reveal natural genetic variation in rice [J]. *Nature Communications*, 2014, 5: 5087.
- [29] Du J J, Fan J C A, Wang C A Y, et al. Greenhouse-based vegetable high-throughput phenotyping platform and trait evaluation for large-scale lettuces [J]. *Computers and Electronics in Agriculture*, 2021, 186: 106193.
- [30] Wu S, Wen W, Wang Y, et al. MVS-Pheno: A portable and low-cost phenotyping platform for maize shoots using multiview stereo 3D reconstruction [J]. *Plant Phenomics*, 2020, 2020: 1848437.
- [31] 杜建军, 郭新宇, 王传宇, 等. 基于全景图像的玉米果穗流水线考种方法及系统 [J]. *农业工程学报*, 2018, 34(13): 195–202.
Du J J, Guo X Y, Wang C Y, et al. Assembly line variety test method and system for corn ears based on panoramic surface image [J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2018, 34(13): 195–202.
- [32] 施巍松, 孙辉, 曹杰, 等. 边缘计算: 万物互联时代新型计算模型 [J]. *计算机研究与发展*, 2017, 54(5): 907–924.
Shi W S, Sun H, Cao J, et al. Edge computing—An emerging computing model for the Internet of everything era [J]. *Journal of Computer Research and Development*, 2017, 54(5): 907–924.
- [33] Wu S, Wen W L, Gou W B, et al. A miniaturized phenotyping platform for individual plants using multi-view stereo 3D reconstruction [J]. *Frontiers in Plant Science*, 2022, 13: 897746.
- [34] Ma Z, Du R, Xie J, et al. Phenotyping of silique morphology in oilseed rape using skeletonization with hierarchical segmentation [J]. *Plant Phenomics*, 2023, 5: 0027.
- [35] Li Y, Wen W, Miao T, et al. Automatic organ-level point cloud segmentation of maize shoots by integrating high-throughput data acquisition and deep learning [J]. *Computers and Electronics in Agriculture*, 2022, 193: 106702.
- [36] Du J J, Li B, Lu X J, et al. Quantitative phenotyping and evaluation for lettuce leaves of multiple semantic components [J]. *Plant Methods*, 2022, 18(1): 54.
- [37] Zavafer A, Bates H, Mancilla C, et al. Phenomics: Conceptualization and importance for plant physiology [J]. *Trends in Plant Science*, 2023: 2439.
- [38] Jin S, Su Y, Zhang Y, et al. Exploring seasonal and circadian Rhythms in structural traits of field maize from LiDAR time series [J]. *Plant Phenomics*, 2021, 2021: 9895241.
- [39] Jiang Y, Li C Y. Convolutional neural networks for image-based high-throughput plant phenotyping: A review [J]. *Plant Phenomics*, 2020, 2020: 4152816.
- [40] 王璟璐, 张颖, 潘晓迪, 等. 作物表型组数据库研究进展及展望 [J]. *中国农业信息*, 2018, 30(5): 13–23.
Wang J L, Zhang Y, Pan X D, et al. Research progress and prospect on crop phenomics database [J]. *China Agricultural Informatics*, 2018, 30(5): 13–23.
- [41] Neveu P, Tireau A, Hilgert N, et al. Dealing with multi-source and multi-scale information in plant phenomics: The ontology-driven Phenotyping Hybrid Information System [J]. *New Phytologist*, 2019, 221(1): 588–601.
- [42] Zhang Y, Wang J, Du J, et al. Dissecting the phenotypic components and genetic architecture of maize stem vascular bundles using high-throughput phenotypic analysis [J]. *Plant Biotechnology Journal*, 2020, 19(1): 35–50.
- [43] 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]. *GigaScience*, 2019, 8(3): 11.
- [44] Gui S T, Yang L F, Li J B, et al. ZEAMAP, a comprehensive database adapted to the maize multi-omics era [J]. *Iscience*, 2020, 23(6): 31.
- [45] Mcguire A L, Gabriel S, Tishkoff S A, et al. The road ahead in genetics and genomics [J]. *Nature Reviews Genetics*, 2020, 21(10): 581–596.
- [46] Guo Z L, Li B, Du J J, et al. LettuceGDB: The community database for lettuce genetics and omics [J]. *Plant Communications*, 2023, 4(1): 100425.
- [47] Cabrera-Bosquet L, Fournier C, Bricchet N, et al. High-throughput estimation of incident light, light interception and radiation-use efficiency of thousands of plants in a phenotyping platform [J]. *New Phytologist*, 2016, 212(1): 269–281.
- [48] Wen W, Gu S, Xiao B, et al. In situ evaluation of stalk lodging resistance for different maize (*Zea mays* L.) cultivars using a mobile wind machine [J]. *Plant Methods*, 2019, 15(1): 96.
- [49] Ferguson J N, Fernandes S B, Monier B, et al. Machine learning-enabled phenotyping for GWAS and TWAS of WUE traits in 869 field-grown sorghum accessions [J]. *Plant Physiology*, 2021, 187(3): 1481–1500.
- [50] Wu X, Feng H, Wu D, et al. Using high-throughput multiple optical phenotyping to decipher the genetic architecture of maize drought tolerance [J]. *Genome Biology*, 2021, 22(1): 185.
- [51] Wang W, Guo W, 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: 1–20.
- [52] Le L, Guo W J, Du D Y, et al. A spatiotemporal transcriptomic network dynamically modulates stalk development in maize [J]. *Plant Biotechnology Journal*, 2022, 20(12): 2313–2331.
- [53] Ren W, Zhao L F, Liang J X, et al. Genome-wide dissection of changes in maize root system architecture during modern breeding [J]. *Nature Plants*, 2022, 8(12): 1408.
- [54] 张颖, 廖生进, 王璟璐, 等. 信息技术与智能装备助力智能设计育种 [J]. *吉林农业大学学报*, 2021, 43(2): 119–129.
Zhang Y, Liao S J, Wang J L, et al. Information technology and in-

- telligent equipment facilitating smart breeding [J]. Journal of Jilin Agricultural University, 2021, 43(2): 119–129.
- [55] Xu Y, Zhang X, Li H, et al. Smart breeding driven by big data, artificial intelligence, and integrated genomic-enviromic prediction [J]. Molecular Plant, 2022, 15(11): 1664–1695.
- [56] Ninomiya S. High-throughput field crop phenotyping: Current status and challenges [J]. Breeding Science, 2022, 72(1): 3–18.
- [57] 顾生浩, 温维亮, 卢宪菊, 等. 作物智慧栽培学——信息-农艺-农机深度融合的新农科 [J]. 农学学报, 2023, 13(2): 67–76.
- Gu S H, Wen W L, Lu X J, et al. Smart crop cultivation: A new agricultural science toward deep integration of information, agronomy and machinery [J]. Journal of Agriculture, 2023, 13(2): 67–76.
- [58] Li Y, Wen W, Fan J, et al. Multi-source data fusion improves time-series phenotype accuracy in maize under a field high-throughput phenotyping platform [J]. Plant Phenomics, 2023, 5: 0043.