

基于大数据分析的农作物智慧育种信息平台建设初探

郑媛¹ 冉永迪² 潘德斌²

1. 甘肃省种子总站 甘肃兰州 730020

2. 甘肃新空间信息技术有限公司 甘肃兰州 730020

摘要: 本文介绍了农作物育种信息化发展现状及存在的问题,探讨了“物联网+大数据”育种的新模式,初步构建了农作物智慧育种信息平台,提供了农作物育种信息数字化、育种过程自动化、育种结果智能化解决方案,为农业育种科研提供了智能辅助决策工具,支撑企业育种知识积累和育种创新,缩短育种周期,提升育种单位科研管理和核心竞争力。

关键词: 农作物; 育种; 大数据; 信息平台

我国是农业生产大国和用种大国,种业是国家战略性、基础性核心产业,是促进农业长期稳定发展、保障国家粮食安全的根本。近年来,随着国内外作物科学研究和种业的不断发展,育种试验规模不断扩大,大量的表型数据和基因组数据不断产生,育种工作信息化趋势不断加强,数据管理、挖掘和共享的数字化管理技术需求日益迫切。

一、农作物育种信息化发展现状

进入21世纪,随着自动化采样和数据获取技术的突破,从分子水平的基因表达数据到宏观性状表现型数据都实现了质的飞跃。育种数据的存储、分析、利用成为现代数字育种技术的关键。伴随着当代育种技术进步的加快,特别是生物技术的广泛应用,在作物育种中产生了信息爆炸,数据不再是单一的田间性状调查结果,还有气候、土壤等动态环境数据、分子标记等基因型数据、基因表达和代谢物动态数据、生产管理数据。这些数据类型复杂多变,组成一个复杂的大数据集,为确保这些数据不断积累和丰富,需要使用先进的数据库管理系统进行管理维护,清晰地建立各种数据之间的相互关系,便于对这些数据进行交叉分析,从中找出规律,提取有用的知识,为后代的选择和新品种的定位提供可靠的依据。

国际上欧美发达国家持续加强种业基础理论研究与技术集成创新,高通量、高效率,规模化、智能化的商业化育种体系不断强化,全方位、全天候、全产业链的数字种业加快构建,种业的高科技优势进一步显现。例如,孟山都、杜邦先锋、先正达等各大种企都建有自成体系的私有数据库和管理系统,且功能十分先进与完善,存储了整个产业链条从研究部门到销售部门的各种相关

数据资料,这些私有数据库系统,其结构、功能及内涵的商业机密是保密的,但可以肯定都具备海量数据的超大存储能力、复杂数据的高效分析能力、庞大系统的科学管理能力,能够为研究、管理人员提供简捷、高效、精准的服务,更好地完成相关应的育种研究,如数据的自动采集、分类、存储、分析、建模等。从国内看,虽然我国种业发展虽然取得了很大进步,但总体上仍处于初级阶段,基础研究弱,企业实力不强,还没有形成种业科技创新的整体合力,整体水平与发达国家形成了代际差距。^[1-2]

二、目前农作物育种信息化存在的问题

1. 育种信息化程度偏低

随着农业机械化及各种现代栽培技术的发展,农业生产技术水平有了显著的提高,而在育种过程中,虽然现代生物技术的发展带动了育种方式的革命性变化,但性状参数的获取大多还停留在人工操作的阶段。新品种育成时间长、工作量大、性状参数获取的配套技术设备落后等瓶颈问题,限制了育种的规模化发展。

2. 育种数据分散,利用率不高

目前,国内已有一些组织和部门构建了一些数据库和共享平台,虽然育种相关数据量很大,但分散,未有效组织。目前育种者在育种过程中利用的数据主要为自身内部数据,而公开的文献和基因组相关数据等其他数据很少利用或无法利用。导致大量内部数据成为“数据孤岛”,无从下手。上述问题极大地限制了育种相关数据的利用和育种效率的提高。如何有效利用科学家多年来产生的育种数据,整合处理成有效资源,并反哺于农业发展,大数据育种系统的发展,将会成为解决悬而未决育种难题的有效手段。

3. 田间信息采集强度大, 成本高

在实际生产中, 农作物的性状参数依靠人工测量并记录, 在作物不同的生长周期, 育种工作人员需要频繁往返各个试验田采集作物性状数据, 工作强度高, 投入成本高, 效率低、准确度差。

三、农作物智慧育种信息平台构建

该平台旨在为玉米、大豆、高粱、棉花、油菜、水稻、小麦等作物进行育种试验提供种质资源管理, 田间的信息采集、数据管理、数据分析、试验设计、试验预警以及田间远程物联和农作物考种性状的硬件、数据汇总为一体的、全方位的综合信息平台。通过大数据分析为农作物新品种研发提供决策支持, 最终实现农作物育种信息数字化、育种过程自动化、育种结果智能化。

1. 技术框架

根据作物育种相关的专业知识和涉及到的具体试验, 以及试验中涉及的具体内容, 经过市场调研和用户需求分析, 构建了育种众源信息共享与服务系统、智慧育种系统、企业决策和信息公开四大功能模块。

(1) 育种众源信息共享与服务系统

现代育种(生物育种)技术体系的建立需要常规育种、分子辅助育种、转基因育种及抗逆性检测多项关键技术共同支撑, 涉及多个交叉学科, 需要强大的信息平台进行有效的数据组织和管理。例如, 分子设计育种可以实现育种的定向变异、准确选择的目标, 但是这一技术只有借助大数据强大的信息处理能力、共享机制及计算模拟集成技术才能更好地发挥作用。利用大数据挖掘计算技术, 以育种目标为纽带, 充分发挥信息技术的提升和串联带动作用, 促进多种常规育种技术与生物育种技术在不同产业链环节的协同发展, 形成合力, 建立有明确育种目标的高效率的现代育种体系, 共同促进现代种业的发展。育种众源信息共享与服务系统的主要功能是怎样处理和分析挖掘通过市场反馈、网络爬取和合作中获得数据, 并将处理结果展示出来。

(2) 智慧育种系统

该模块主要涉及系统的软件平台研发, 通过对育种行业的需求调研, 依据育种业务具体的工作方式和业务数据的存储关系, 将该模块总体划分为种质资源库管理、试验设计、试验管理、数据分析四个部分。种质资源库是可共享的资源, 主要功能是种质性状可查询、种质数量可实时更新, 存储位置固定, 种质出入库审批及复核等。试验设计分为选择圃和鉴定圃两种, 选种圃的功能是可以快速建立杂交育种和自交选种计划, 列出

父母本清单、田间布局、杂交方法, 杂交数量, 授粉计划、性状调查要求等信息供育种家选择; 鉴定圃主要涉及鉴定试验和品比试验, 包括试验设计和试验管理两部分, 具体有试验材料选择、试验方法选择、田间布局设定、田间记载本以及材料决选等功能; 鉴定圃也包括区域试验, 区域试验分为一年多点和多年多点试验, 是品比试验的升级, 也是产量测试试验中的重要试验。智慧育种系统主要设计的硬件部分为数据采集系统和产量测试系统。

(3) 企业决策

企业决策包括科研定位、成本预算、市场定位和生产管理。平台通过集成利用大数据、数据挖掘、决策支持系统等技术和分析工具, 实现跨企业数据的综合分析, 从海量数据中挖掘知识、发现规律, 为农作物育种提供决策辅助。

(4) 信息公开

信息公开包括资源状态、品种成果和科研团队三方面的信息功能公开和共享, 资源状态公开可以使目前的种质资源得到最大化的利用; 品种成果公开可避免各单位间无效、低水平重复工作, 优化资源配置, 实现资源有效利用率最大化的同时, 达到各方共赢的目的; 科研团队公开可实现人力资源的最大化利用, 集思广益, 使企业利益最大化。

2. 主要技术路线

技术软件平台类型为标准的MIS系统, 网络应用环境以Internet技术为核心, 系统采用C/S的网络结构。软件开发工程方法采用目前被广泛使用的RUP(Rational Unified Process)方法来进行分析、设计和开发。

农作物智慧育种信息平台基于WEB、微信公众号、微信小程序、手机APP为一体化解决方案, 使用SQLSERVER关系型数据库和REDIS非关系型数据库共同存储数据。

(1) WEB使用ASP.NET MVC4.5实现, 为整个平台的基础结构, 承载整个平台的数据运算以及提供统一的数据接口, 同时要完成处理各种业务流程的功能模块供用户在PC端使用。

(2) 微信公众号提供育种众源信息系统的资讯信息。

(3) 微信小程序提供育种众源信息系统所需的反馈数据的入口, 以及处理育种企业关于审核流程相关的业务流程。

(4) 手机APP需要实现对田间数据的采集以及企业所需的周边业务, 如考勤等。

四、社会效益分析

1. 农作物智慧育种信息平台贯穿在整个育种流程之中, 以强大的种质资源库为基础, 在自交系选育、产量测试等阶段提供大量的数据支撑和统计分析, 能给育种家提供丰富的材料和便捷高效的分析手段。

2. 平台开发完成后可解决育种周期长、成果效率转换低、成果缺乏保障、智能化程度低等问题; 同时, 还能解决当前育种企业信息化程度偏低、育种人员不能实时掌握田间数据信息和育种相关设备格式不统一的问题。该平台开发完成后为农作物新品种选育提供育种设计、数据集成、智能分析、品种优化、大数据分析管理与决策等综合服务; 为实现农作物育种信息数字化、育种过程自动化、育种结果智能化提供解决方案。

3. 平台上线后, 可为农业供种和用种安全提供保障。随着异常天气频繁发生和市场变化难测, 平台拥有快速、准确、完善地监测和检测信息系统, 可实现及时调度发

布主要农作物种子供需, 积极应对天气、病虫等灾害对制种产量和质量的不利影响, 有效监督市场上农作物种子质量, 确保农业生产供种和用种的数量安全、质量安全、结构平衡。

4. 平台可实现种苗信息共享。让种业企业及时了解种业市场发展情况和作物市场细分状况, 掌握市场的竞争态势, 根据市场实时变化情况, 调整自己的经营策略与经营活动。

5. 对于政府管理, 平台可及时提供信息服务, 为政府决策提供参考, 引导种业行业有序前行, 为我国种业健康发展提供重要手段和保障措施。

参考文献:

[1]董春水, 才卓. 现代数字育种技术的研究进展. 玉米科学, 2013, 21(1): 1-8.

[2]王冰冰. 大数据: 植物育种的加速器. 高科技与产业化, 2015, 5(228): 50-52.