

从「经验选种」到「数据算种」
怎样缩小搜索范围

“传统育种靠什么？很大程度上靠育种家的眼睛、尺子和长期经验。在三亚崖州区的试验田里，黎亮穿行玉米地，能说出一株株玉米的“父亲”“母亲”和主要特点，甚至不剥开苞叶也能判断籽粒是否饱满。几十年积累的种质资源和田间观察，像一册写在育种家心里的“记录本”。但经验难以复制和规模化，育种的迭代速度仍受土地、人力和自然节律制约。

如今，测序和高通量表型设备正把这本“记录本”转化为可计算的数据。基因分型成本持续下降，使大规模检测成为可能。三亚智数生物科技有限公司CRO团队负责人隋耀介绍，团队利用编码微芯片技术开展基因分型，每日可完成5000个样本提取和20万个SNP位点检测，为大规模基因数据采集提速。

在中国农科院国家南繁研究院，5.7米高的轨道式高通量表型平台可全天候监测3800株玉米。设备搭载可见光、三维激光、热红外、高光谱等模块，自动采集作物形态、生理、产量和品质信息，效率比人工提升4至5倍。

数据采集快了，新的瓶颈随之出现：海量数据如何处理，模型如何被普通育种家使用。

“玉米有数万个基因、20多亿个碱基。数据量过大时，一些软件会运行缓慢甚至宕机；高光谱图像中，一个叶片就包含数百个波段，再叠加时间和空间信息，传统统计方法难以处理。”海南大学三亚南繁研究院副院长夏志强说。多数育种家并非计算机专业出身，“白天下地、晚上跑数据”成为现实压力。

AI的价值，首先不是替代育种家，而是帮助他们从巨量组合中更快找到值得下田验证的少数材料。

海南具备训练和验证农业模型的独特条件。常年数十万份育种材料在这里种植观测，持续产生基因组、表型、气象环境等数据；南繁科研从季节性“候鸟式”向全年“驻留式”转变，为模型迭代提供连续场景；自贸港制度则为联合研发和数据合作打开空间。

围绕这些条件，海南已形成一批技术平台：崖州湾国家实验室与华为发布“繁—未来农业智能枢纽”；国内首个种业大模型“丰登”可“读基因、编序列、算组合”；中国农科院国家南繁研究院联合作物科学研究所搭建了国内首个一体化南繁智慧育种平台；海南大学开发植物基因组语义大模型XiaLLM/Creator 1.0；三亚华大生命科学研究院构建DeepGxE模型；中国种子集团研发玉米智能育种工具GASEL。平台正在增多，下一步更关键的是走出实验室。

5000份父本玉米，搭配5000份母本玉米，理论上可以形成2500万个杂交组合。

对中国农科院育种专家黎亮来说，这道算术题藏着传统育种最大的困境：受试验田规模、人力和时间限制，他每年最多实地测评约2万个杂交材料。面对2500万个可能组合，人工筛选如同“大海捞针”。

漫长的田间种植和逐季表型观测，是育种家的工作常态。一个常规玉米品种从选系到审定通常需要8至10年，优质高产玉米品种“农大108”更是历经了18年才问世。“你得种下去，等它长出来，看它表现好不好，才知道这个组合行不行。一年一季，每一轮都要淘汰大批材料。”黎亮说。

如何让筛选更快、更准，成为新品种创制的核心命题。传统南繁让科研人员每年多种一季，育种能多迭代一次。为此，大批科研人员向海南“借”时间；如今，基因测序、高通量表型和人工智能又让育种家向技术“争”时间。

当前，各类农业数据成为农业新质生产力核心生产要素，育种5.0时代——AI驱动的模式化智能育种走向前台。拥有南繁硅谷国家战略平台和自贸港制度优势的海南，能否把“南繁育种基地”升级为智能育种服务枢纽？答案不只取决于模型有多先进，更取决于技术能否走进田间，获得育种家信任并形成可持续的产业服务。

当南繁育种用上「新农具」

种业CRO降低落地门槛
从「有模型」到「用起来」

夏日炎炎，三亚崖州湾的田间南繁季已经结束，海南大学南繁学院机房里的服务器仍在运行。

“我们的‘田’就在这里。”夏志强指着数据机房说。

该团队研发的AI生物信息科学家智能体，可以在虚拟空间分析基因功能和作物性状。打开助手对话框输入指令，一份基因功能鉴定报告几分钟即可生成；过去1个人3天的工作量，AI助手约5小时完成。

今年5月，团队发布首个水稻泛基因组大模型OryzaG3。模型通过学习149份高质量水稻基因组，能够识别DNA序列规律，精准捕捉水稻特有的遗传密码。7亿参数的轻量化设计，使推理速度比同类模型快4倍以上。

“这片‘田’，不求大水漫灌，而是精准滴灌。”夏志强说，OryzaG3可服务基因功能预测、亲本组合推荐、品种适应性评估和抗逆性状挖掘等场景。

模型具备能力，并不等于一线育种家愿意使用。

“转型成

本是真实的。”三亚智数生物科技有限公司副总经理吴丹丹说，许多育种家数十年习惯用Excel管理数据、用纸质记录田间观察。如今切换到AI辅助决策系统，不只需要培训、改变工作习惯，还需要能印证AI优于原有方法的案例。

海南正在用升级版种业CRO（合同研究组织）模式跨过这道门槛：服务不再停留在“帮客户做实验”，而是延伸到种质数字化、杂交组合模拟、性状预测和育种方案优化。

广西一家科研机构委托夏志强团队开发米粉稻直链淀粉研究模型。团队以417份水稻种质为材料，一次性完成基因型数据产出，定位调控直链淀粉、胶稠度的关键位点，并用7种模型交叉验证，直链淀粉含量预测准确度突破0.83。这套“低成本测序+自动化实验+AI算法”方案还可迁移到玉米、大豆和蔬菜。

“育种家不需要自己搭建AI模型，也不需要自己写代码。”夏志强说，轻量化、低成本、按需订阅的数据和算法服务，可降低中小种企和基层科研团队的使用门槛。对海南而言，种业CRO的价值正在从承接实验，扩展为连接“种质引进、数据采集、AI建模、联合育种、成果输出”的服务链。

自贸港制度则为联合研发和数据合作打开空间
为模型迭代提供连续场景
南繁科研从季节性“候鸟式”向全年“驻留式”转变
持续产生基因组、表型、气象环境等数据
常年数十万份育种材料在这里种植观测
海南优势

海南日报全媒体记者 黄媛艳

2024年，黎亮在田间种下300个特别组合：AI预测了100个“上品”、100个“中品”和100个“下品”。

“田间验证结果还不错，已经有一些组合陆续出来了，今年秋天再看品种表现。”黎亮说。2024年起，他与中国农科院李慧慧团队合作，用南繁智慧育种平台初筛玉米材料，目前已生产3000多份纯系材料，配成约3000个杂交组合，并开展多年、多点、多环境评价。

李慧慧期待，智慧育种能把8至10年的育种周期压缩到3至5年。平台通过虚拟筛选减少无效田间工作，并已服务国内科研机构、种业企业及荷兰瓦赫宁根大学等院校。

传统育种依靠大面积播种、长期观测和逐步淘汰；AI则先在计算机中筛掉大量低潜力组合，把有限的土地、人力集中到更有希望的材料上。它改变的是试验规模和顺序，而不是取消试验。

难点在于，作物性状由基因、年份气候和土壤环境共同决定。许多模型能给出结论，却难以解释哪些基因发挥作用、基因如何与环境互动。“黑箱”如果打不开，AI就很难成为育种家的核心决策伙伴。

三亚华大生命科学研究院构建的DeepGxE模型，尝试融合温度、降水、土壤等环境因子，预测不同区域、不同气候年份的品种表现，并追溯碱基变异与性状的关联。团队从4000多个玉米候选组合中推选前100个和后100个组合，已在3个地点、2个重复的田间试验中验证。

市场需求正在提供更多验证场景。各地水土、气候和种养对象差异巨大，近日，黑龙江省农科院将寒地水稻品质提升项目委托给华大团队。

“我们致力于打造‘样本海南测、数据海南算、方案海南出’的服务闭环。”三亚华大生命科学研究院智慧育

种方向负责人刘光宇介绍，客户交付水稻种质材料后，可获得测序数据、田间鉴定报告、AI模拟杂交方案、优良亲本推荐和区域适应性预测。团队还把蛋白质从头设计平台用于绿色生物农药、抗病蛋白分子设计，拓展智能育种技术的应用边界。

“AI不会取代育种家。再先进的模型仍需田间试验验证，再强的算法也离不开高质量种质资源和长期积累的育种经验。”吴丹丹说。AI先行虚拟筛选，让有限土地和人力用于高价值试验，但最终判断仍要回到泥土里。

要把单个案例变成产业能力，海南还需建设高质量数据底座。当前，海南正打造“种业数据一张网”，国家南繁生物育种服务平台汇聚种质资源、育种试验和表型信息等数据，并探索可信数据空间和科研数据流通机制。

放眼全球，海南的优势并非没有短板：高端算法人才不足，本土龙头企业量有限，跨机构数据共享机制尚未成熟，商业化产品和国际合作网络仍待完善。能否形成统一的数据标准、可信的流通规则和可复制的付费服务，将决定技术热度能否转化为产业竞争力。

夏志强认为，海南可以南繁硅谷科创平台为底座，继续做强自动化实验和种业CRO服务，把零散试验数据治理为可训练的数据集；以真实育种场景牵引模型迭代，吸引国内外团队将多环境试验、种质数字化和模型训练业务落地；同时依托自贸港政策，稳妥探索国际种质数据合作。

从“大海捞针”到“计算设计”，AI正在改变育种家寻找答案的方式。但一粒种子能否走出模型、长进田野，仍取决于高质量数据、可信算法和一季又一季的田间验证。海南要争取的，不只是多发布几个模型，而是把南繁积累转化为一套育种家用得起、信得过、能够持续迭代的智能育种服务体系。

（本报三亚9月6日电）



育种实验室。



工作人员在操作自动化基因编辑继代仪。本组图片均由海南日报全媒体记者 黄媛艳 摄

深读

