

现阶段,我国进入口粮饲料并重新时期,饲草作为动物的“口粮”,是饲料的核心组成部分,其产业的发展对于保障国家大粮食安全具有深远的战略意义。文章分析了我国饲草种业的现状,综述了作物基础科学前沿成果及人工智能(AI)在智能育种的应用,剖析了饲草特化性状与育种关键问题,指出饲草智能育种有望加速饲草新品种的培育,带动产业发展。文章最后给出了助力饲草智能育种发展的建议,并希望引起政府相关部门和广大科技工作者对饲草育种与产业的关注。

加快饲草智能育种科技创新的思考与建议

□景海春 胡伟娟 金京波 张景昱 周姚 巩玥 姚刚 王雷 康康

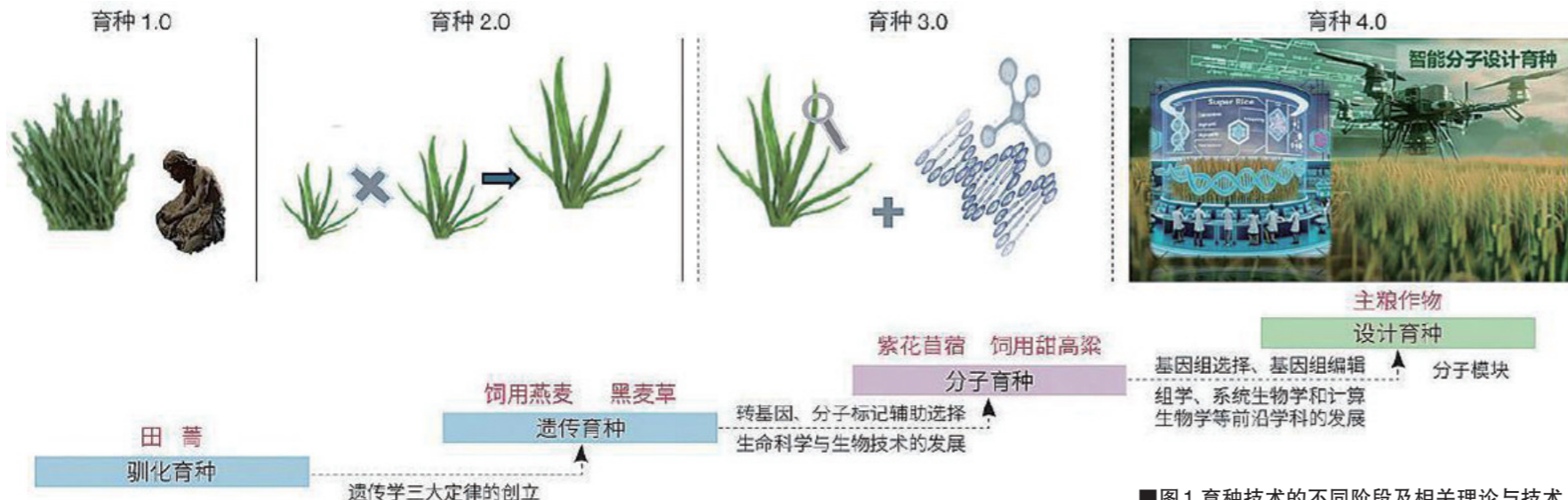


图1 育种技术的不同阶段及相关理论与技术。

随着我国畜牧业的发展,饲草和饲草种子的需求日益增加,直接关乎肉乳供给和国家粮食安全。为此,我国“十四五”期间,在重点研发计划、种业“卡脖子”攻关、农业种质资源专项、生物育种专项等部分饲草育种布局;党的第二十届中央委员会将饲草饲料确定为基本农作物写入了第三次全体会议公报;2024年,《国务院办公厅关于践行大食物观构建多元化食物供给体系的意见》也明确提出“大力发展饲草产业,增加草食畜产品供给”。饲草种子是饲草产业的芯片,育种科技水平直接决定了一个国家种源保障、产业发展和世界饲草种子贸易能力。

全球饲草育种战略与科技水平概况

欧美发达国家长期重视饲草育种。在美国,饲草被誉为“绿色黄金产业”。美国农业部2013年启动《面向21世纪的首蓿研究路线图》(国家奶牛草技术路线图);2019年启动5年计划“天然草地,人工草地和农牧耦合系统”。欧盟自2014年启动“LIFE-Viva-Grass计划”资助全欧草地畜牧业,2020年投入1000万欧元启动“Smart-Proteins水平线计划”系统开始饲草蛋白研究。澳大利亚在2018年启动“面向2030年农业创新研究计划”,侧重草—畜种、环境监测。

美国是世界饲草种业大国和强国,我国则是世界饲草种子的进口大国。美国在20世纪50年代就将紫花苜蓿列入战略物资名录,草产业已成为美国农业中的重要支柱产业,年产值约110亿美元,仅次于玉米和大豆。《草种研究动态报告(2022)》显示,2021年,世界饲草种子贸易量为87万吨,主要以黑麦草、羊茅、紫花苜蓿、三叶草和早熟禾等种子为主。美国2021年饲草种子出口量世界第一,市场份额达到27%。全球进口饲草种子的国家主要有荷兰、德国、中国、法国、意大利、加拿大、土耳其、比利时、英国、巴基斯坦、美国等。我国2021年饲草种子进口份额为8%,位居全球第三。可见,我国是世界饲草种子的进口大国。

相对于粮食作物,饲草虽然也有着万年的驯化历史,但其育种水平落后。作物育种技术随着生命科学基础理论的发展经历了4个不同的阶段(图1),而饲草育种还处在人工表型选育等早期阶段,依赖“老把式”经验育种。全球饲草育种水平有着如下特征:1.基于表型选择的常规育种为饲草育种的主要路径。选择育种、诱变育种、杂交育种是目前育成品种的主要技术手段,广泛应用于新种质新品种(品系)的创制,优异生产性状(单产量、品质)和区域适应性(耐逆、抗病等)的育种材料主要通过人工田间观察和表型筛选获得。2.注重饲草种质资源的收集保存与发掘利用。各国将饲草视为国家战略生物资源,开展了种质资源库建设,广泛收集鉴定饲草种质资源,保护力度不断增加。在资源评价方面,结合表型、核型、分子遗传等技术,鉴定饲草种质及其近缘种的农艺性状(如高产优质、环境抗性、病虫害抗性)等。3.逐步开展饲草重要性状分子遗传机制解析与分子标记等生物育种技术在育种中的应用。获得了主要饲草高质量参考基因组,组学技术与分子遗传工具应用到重要基因的鉴定与功能分析。全基因组分子标记技术、基因编辑技术也在饲草育种中得到应用,加速了性状关联位点的聚合和育种效率。

我国饲草育种战略布局晚,起点低,短板突出。我国在饲草种质资源发掘与育种技术方面,同其他发达国家没

有大的区别,加之长期得不到重视,表现出如下3个突出问题。1.育成品种少,性状不突出。截至2024年,我国共有720个饲草新品种通过国家审定。选育的饲草品种品质、生产能力和抗逆性都无法超越引进品种,且有的品种出现了严重退化现象。与此相反,美国每年用于生产的饲草品种中豆科达4000多种,禾本科约1500种,西方发达国家经贸易成员国互认的登记饲草品种达到5000多个。2.主栽品种以进口品种为主体。商业用种对外依存度高,2022年草种进口6.84万吨,紫花苜蓿用种80%以上进口。3.丰富的牧草资源未得到充分利用。仅草地饲用植物就达246科1545属6704种,但无论是国家种质资源库的收集保存量还是育成草类品种都低于30%总量,宝贵的草资源尚未得到完全认识和保护。

总之,全球范围内,饲草育种的基本面为基础生物学研究不系统,对基因组变异了解不够,功能基因解析不足,遗传转化与基因编辑等高效生物育种技术不成熟。因此,迫切需要强化饲草智能育种,从根本上破解饲草产业与种源问题。

智能育种技术在作物上的应用实践及其发展态势

2000年以来数据驱动科研表现为3种形式:数据驱动科研(data-driven science)、科学智能(artificial intelligence)和智能科学(artificial intelligence scientist)。在作物育种领域,人工智能(AI)的应用也成为热点。近期,李家洋等提出了“未来育种5.0时代”的概念,将其定义为“智能作物育种”,并详细阐述了其两大基本特征:1.“智能品种”,指能够自主应对环境变化的作物品种;2.“智能培育”,指在品种培育过程中发展与利用前沿生物技术及信息技术,实现生物技术(BT)与AI的深度融合。具体而言,作物智能育种是指利用AI、大数据、基因组学、表型组学等前沿技术,结合传统育种方法,实现作物品种的高效、精准改良。它通过整合多维度数据,优化育种流程,提高育种效率和精准度,以满足现代农业对高产、优质、抗逆性强作物品种的需求。这一过程不仅依赖传统的育种经验,更通过数据的深度分析,实现了对育种过程的全面优化。

作物智能育种有以下4个特征:1.数据驱动。其往往利用大数据分析和机器学习算法,从海量的基因组和表型数据中挖掘有价值的信息,指导育种决策。通过深度学习模型预测基因型与表型之间的关系,提高育种的准确性和效率。本文基于大数据构建知识图谱和复杂网络理论,构建了包含60年以来的中国水稻品种的系谱关系网络,发现中国水稻天然区分了亚种的交流亲疏程度。2.多学科融合解析。综合利用基因组学、表型组学、生物信息学、计算机科学等多学科技术,实现从基因到表型的全面解析。3.智能化决策。通过AI算法和模型,实现对育种过程的智能化管理和决策支持。例如,利用深度学习模型预测作物的生长趋势和病害发生概率,提前采取措施。目前在作物育种方面已有常用的AI模型。4.高效精准。通过精准基因编辑和分子标记辅助选择,提高育种效率和精准度。例如,利用CRISPR/Cas9技术对目标基因进行编辑,快速培育出具有优良性状的作物品种。近期,许操团队通过基因编辑技术将热响应元件(HSE)精准融入番茄细胞壁蔗糖转化酶(CWIN)基因的启动子中,使番茄能够感应温度变化并自动调节光合产物分配。

作物智能育种实施要素。区别于传统育种,作物智能育种需要以下4个方面要素。1.高通量的表型组、基因组及环境组数据的采集与管理。图2归纳了目前流行的作物表型获取的传感技术,例如:无人机成像、高光谱成像、激光雷达等用于实时监测获取作物的生长和生理状态;快速高效的基因组测序技术用于获取作物的基因信息,构建基因组数据库;精准高效的环境参数监测系统获取和管理不同生态区的光、温、水等各项环境参数。2.数据分析与建模。需要研发各种机器学习和深度学习算法,实现从海量数据中挖掘有价值的信息,构建预测模型。例如,利用卷积神经网络(CNN)和循环神经网络(RNN)对基因型和表型数据进行分析,预测作物的产量和抗逆性。3.高效精准的育种技术与工具。如利用CRISPR/Cas9基因编辑技术精准改良作物的遗传特性;分子标记辅助选择技术实现快速筛选优良性状的个体。4.智能决策系统。应用该系统实现对育种过程的智能化管理和决策支持。例如,通过机器学习模型预测作物的生长趋势和病害发生概率,提前采取措施。

AI在作物育种方面的应用进展。作物智能育种处于兴起阶段,近年来,多有AI育种的理论内涵、方法体系和应用场景的观点与综述文章,涉及算法模型、表型获取、传感技术、过程检测和系统集成等各个侧面。目前,智能育种仅在有限的主粮作物中开展,进展可归纳为4个方面:

1. AI助力理解作物遗传学基础。中心法则的各个环节通过大数据驱动来助力物种个体发育的新科学发现。CNN鉴定了更多高质量单核苷酸变异并实现基因组变异的精准预测。以3000多万单细胞测序数据作为学习语料,单细胞基底大模型优化了基因表达模式和分子机制的预测,如细胞类型注释、基因共表达网络和调控网络推断等。轰动世界的AlphaFold模型利用蛋白质结构数据库开展深度学习和算法优化,从而高精度预测了对未知蛋白复杂空间结构和分子互作的解析。

2. AI助力高通量表型组学研究。我国在表型预测方面开展了有益的探索,例如:通过深度学习大样本基因型与表型的非线性关系从而提高准确性,应用无人机遥感数据估算玉米地上生物量,基于高光谱图像估算小麦产量和地上生物量;使用生成式对抗网络预测水稻籽粒蛋白含量,采用单模态或多模态深度学习监测小麦条锈病、番茄叶病;高光谱成像技术在作物表型上具有较大应用潜力,也有开发出多功能的无监督学习框架。

3. AI助力优化作物编辑新工具。高彩霞团队等使用RNN开发了PRE-DICT的深度学习模型,高通量筛选了92423个pegRNA的主要编辑结果。通过对30多万引导RNA的高通量分析鉴定了最佳引导RNA,DeepPrime预测引导编辑效率,并优化出适用于特定细胞类型的DeepPrime-FT和预测脱靶效应的DeepPrime-Off。DeepCas9variants预测了9种Cas9变体的效率,DeepBE预测了63个碱基编辑器的效率。

4. AI助力田间集约化高效管理。借助机器学习或深度学习可实现杂草精准管理、土壤墒情、土壤肥力评估、土壤污染和土壤生物多样性评估等。总体来讲,智能育种技术尚处于兴起阶段,鉴于前期知识积累、数据充沛量、功能机制解析的深度等原因,目前智能育种仅在有限的主粮作物中开展。饲草智能育种尚未形成体系,仅限于少数表型高通量获取方法探索与平台建设,DNN和CNN等方法的尝试应用等,其现阶段水平离实质性的智能育种技术要求相去甚远。

饲草智能育种的 关键科学问题及初步尝试

饲草智能育种的 关键科学问题 通过借鉴智能育种技术在作物上的应用经验,应当从饲草基础生物学的角度,针对以下科学问题及饲草特化性状进行布局研究。

1. 饲草种质多样性与驯化性状。在37种开花植物中,有1000—2000个物种被驯化。饲草和粮食作物一样,驯化改良利用始于万年前,如苜蓿等。然而,同粮食作物相比,其育种技术的发展水平与前沿基础研究相差甚远,可见的是目前仅有6—7种不同饲草得到利用为人类供给能量和蛋白,绝大多数资源多样性丢失或等待发掘利用。驯化性状及驯化基因的鉴定和利用是作物遗传改良的核心,但饲草因收获器官和利用方式显著不同于籽粒为经济产量的粮食作物,如何定义饲草的驯化性状,发展驯化育种基础理论和开发驯化技术等成为需要首要思考的问题。

2. 饲草再生与生物量生产性状基因模块及其网络。饲草作物不同于粮食、油料作物的最大区别是地上生物量的全部收获和利用,而其具有的刈割再生和多年生等特征显著影响生物量的形成。应研究生物量构成要素与产量函数,利用群体遗传学、基因组学、基因编辑等手段,解析饲草刈割再生、多年生等特化性状的遗传基础,发掘重要基因模块的功能及其调控机制,创制高生物量优异种质。

3. 饲草蛋白和能量总量与积累过程的生长发育规律。饲草为畜牧养殖提供蛋白质和能量。应通过转录组学、蛋白质组学与代谢组学等现代全景组学技术方法,阐明饲草地上部分蛋白与能量代谢、分布、积累的生长发育规律,解析饲草蛋白与能量积累的遗传基础、基因模块的功能及其调控机制,创制高蛋白或高能量积累的优异种质。

4. 饲草特化生长繁育性状调控的基因模块。饲草特化生长繁育特性决定着生产方式和经济效益。应解析器官分化、营养生长、开花期、自交不亲和、近交衰退等形成的分子调控机制,创制生长发育优异、繁育障碍消减的新种质。

5. 饲草逆境韧性与生物量耦合的遗传规律。我国饲草产业发展必须善加利用边际土地,并适应南北气候差异大的特点;同时需要探索逆境韧性生长与高产的耦合机制。应发展高通量无损表型组学等手段,解析饲草耐非生物胁迫与生物胁迫的基因模块,探索逆境韧性生长与生物量形成的耦合机制,创制逆境稳产优异种质。

中国科学院等相关机构饲草智能育种的初步尝试

近年来,中国科学院等相关机构关注到饲草的重要性,布局了相关的科技创新战略,围绕AI辅助饲草育种系统开展工作,在以下方面进行了实践和布局。

1. 饲草基因组学与基因编辑技术方面。国内科研人员先后成功获得了紫花苜蓿、羊茅、燕麦、黑麦草、狼尾草、田菁等饲草的全基因组序列;建立了紫花苜蓿、羊茅、老芒麦、柳枝稷、甜高粱、饲用燕麦和田菁策略系统的遗传转化与基因编辑体系;发现紫花苜蓿、高粱、羊茅等重要基因的功能,相关育种技术得到发展。如在甜高粱方面,通过泛基因组和群体基因组策略系统分析了高粱不同育种目标对基因组变异的影响,解析了驯化基因的不同单倍型变化与利用方向,尤其克隆到调控甜高粱秆秆含糖量的重要节点基因,并开展基因组选择育种,贯通了从基础研究到产业化育种的链条。通过解析紫花苜蓿调控重要性状的11个分子元件,开发了分子标记10个,选育出苜蓿新品系4个,形成了紫花苜蓿基因组设计育种技术。

2. 基于传感技术的饲草获取表型应用方面。传感技术发挥着至关重要的作用。配备RGB色彩模式和NDVI(归一化差异植被指数)成像的无人机技术尤为突出,它能够提供饲草作物的生长状况、光合效率与叶绿素含量等多维度表型数据,这为精准农业和作物表型分析开辟了新的方向。通过多时相遥感图像结合RGB植被指数(NDVI),能够有效监测草地的生物量和叶片覆盖度等关键性状,为草地生产管理和质量控制提供数据支持。此外,基于传感器对土壤水分、温度、pH等环境因素的实时监测,可以有效反映饲草作物对环境变化的响应。多模态传感器技术在紫花苜蓿(Medicagosativa)中实现了实时监控其在不同环境条件下的生长状态。这些传感器不仅能精确测量作物的物理特征(如植物高度、叶面积、根系分布等),还可以实时监测作物的生理状态(如水分状况、氮素含量等重要生理指标)。例如,红外传感器技术在实时监测作物水分状况方面具有显著优势,通过探测作物叶片的温度变化来评估其水分状况,从而为研究作物的耐旱性提供重要依据;激光扫描技术能够精确测量作物

的三维结构,利用高精度的点云数据为研究根系分布、叶片结构及植物整体生长提供详细信息;近红外光谱传感器则可实时监测作物的氮素含量、水分水平及其他关键营养元素,进而优化作物的施肥策略和水分管理。

3. 表型组学数据分析与知识图谱的构建方面。种康团队发展了表型组、代谢组双组学表型鉴定方法,采用目标数据特定型数据模型,无需大量数据实现目标表型精准鉴定,应用于饲草育种将成为新品种创制强有力的工具。已有团队开始基于大数据和AI算法构建农业物种的表型知识图谱,并结合基因组数据进行联合分析,以推动育种效率和精准化发展。例如,AgroLD知识图谱平台已将表型数据、基因型数据与环境数据相结合,提供关于植物科学的知识图谱,助力作物育种。类似的概念被引入饲草领域,逐步推动了饲草育种的智能化进程。例如,通过对苜蓿在铅污染下的耐性机制,显著提升了其产量和抗逆性。首蓿的GWAS研究揭示了在盐碱胁迫和Phomamedicaginis病害感染下,影响生长与生物量恢复的关键基因。也有耦合高光谱、代谢组学分析与特定型数据模型开展苜蓿耐盐突变体筛选的工作。

的三维结构,利用高精度的点云数据为研究根系分布、叶片结构及植物整体生长提供详细信息;近红外光谱传感器则可实时监测作物的氮素含量、水分水平及其他关键营养元素,进而优化作物的施肥策略和水分管理。

3. 表型组学数据分析与知识图谱的构建方面。种康团队发展了表型组、代谢组双组学表型鉴定方法,采用目标数据特定型数据模型,无需大量数据实现目标表型精准鉴定,应用于饲草育种将成为新品种创制强有力的工具。已有团队开始基于大数据和AI算法构建农业物种的表型知识图谱,并结合基因组数据进行联合分析,以推动育种效率和精准化发展。例如,AgroLD知识图谱平台已将表型数据、基因型数据与环境数据相结合,提供关于植物科学的知识图谱,助力作物育种。类似的概念被引入饲草领域,逐步推动了饲草育种的智能化进程。例如,通过对苜蓿在铅污染下的耐性机制,显著提升了其产量和抗逆性。首蓿的GWAS研究揭示了在盐碱胁迫和Phomamedicaginis病害感染下,影响生长与生物量恢复的关键基因。也有耦合高光谱、代谢组学分析与特定型数据模型开展苜蓿耐盐突变体筛选的工作。

相关建议

1. 系统布局我国饲草智能育种BT+IT底座,开辟基础科研新赛道。在饲草饲料被确定为基本农作物,以及出台《国务院办公厅关于践行大食物观构建多元化食物供给体系的背景》的背景下,国家发展和改革委员会、农业农村部和国家林业和草原局联合发表了饲草产业高质量发展意见,为饲草产业的未来发展提供了明确的行动方案。饲草智能育种涉及种质资源发掘、复杂基因组解析、基因组/表型组大数据与知识图谱构建,以及基因组智能选择设计等多个方面,对BT和IT资源有着巨大的科技创新需求。因此,建议结合国家战略,发展基于BT+IT的饲草智能育种体系。

2. 加强全国饲草智能育种基地网络建设。我国自然资源禀赋差异大,适宜饲草产业发展的土地资源为盐碱荒地、酸性贫瘠等障碍性土壤和草山草坡等。根据上述情况,建议充分发挥举国体制的优势,按照生态区划,系统布局主要饲草作物的智能育种基地网络,全国一盘棋,在传感器、表型获取、数据分析、育种模型等多方面实现归一化、标准化,以缩短育种周期,加速饲草品种的产业化。例如,由于开展DUS和VCU测试具有一定的复杂性,很多饲草(如紫花苜蓿)为自交不亲和,一个品种的最小单株群体应该多大才能代表一个品种符合DUS和VCU的测试,智能育种联网测试体系的建立有利于系统解决上述问题。

3. 发展饲草AI育种与数字孪生。发展饲草育种的数字孪生虚拟表达体系,模拟、分析和优化育种场景的现实过程,结合传感器数据、机器学习算法、高级建模技术和合成育种环境创建,准确地反映饲草育种现实的对应场景,从而实现“虚拟育种”。建议加速二者的融合发展,以实现更复杂、更精准的饲草育种表达和建模,促进饲草数字生命超越现实生命得以保存和发展,从而改善饲草育种决策并提高整体育种效率。

	成像技术	图像数据示例	元数据	波长范围	表型参数
二维成像	可见光成像		灰度或彩色图像 RGB 通道反射值	400—700 nm	形态学结构相关: 株高、株型、叶型、叶片弯曲指数、穗型、种子形态等; 产量相关: 生物量、叶面积、生长率、开花率等; 颜色相关: 绿度、叶片衰老指数等
	热成像		灰度图像 IR 反射值	8 000—14 000 nm	叶片或冠层温度; 冠层温度下降差
	荧光成像		颜色图像 荧光反射值	400—700 nm	色素含量: 叶绿素指数、花青素指数; 光合效率; 衰老指数; 胁迫指数
	高光谱成像		灰度或彩色图像 光谱吸收曲线	400—2 500 nm 连续波长	色素含量: 叶绿素含量、花青素含量; 生化组分: 蛋白质、可溶性固形物、淀粉、脂肪、叶片 N、P 元素含量, 水分含量 植物光谱反射指数: 归一化植被指数 (NDVI)、RVI、GVI
三维成像	计算机断层扫描成像 (CT 成像)		连续灰度图像	100 μm 或更低	形态学结构相关: 茎秆强度、植株内部结构、分蘖角度等; 产量相关: 分蘖数、穗粒数等; 各部位密度分布
	激光雷达成像		点云图像	532 nm	形态学结构相关: 株型、叶型、叶倾角、种子形态等; 产量相关: 生物量
	磁共振成像		连续灰度图像	200—500 μm	形态学结构相关: 根系构型; 生化组分: 脂肪含量、水分含量; 各部位电磁分布
					细胞 组织 器官 个体 群体

图2 适用于不同尺度的作物表型获取的传感技术。