

潜在位点难预测 人工智能来帮忙

■本报记者 李晨

基因组包括两类遗传信息——DNA序列遗传信息及表观遗传学信息。后者涉及何时、何地以何种方式应用遗传学信息,是作物农艺性状多样性的基础。

近年来,表观遗传学成为生命科学热点领域,检测技术和高通量测序的发展推动了作物表观组学的发展。然而,受到多种因素制约,仍有大量表观修饰位点没有得到发掘和研究。

近日,中国农业科学院生物技术研究所谷晓峰课题组、田健课题组和普利课题组在《新植物学家》上发表了最新研究成果。他们构建了植物表观遗传修饰智能预测在线工具 SMEP。

该项工作利用人工智能,深度学习植物 DNA 甲基化、RNA 甲基化、组蛋白修饰等序列信息,系统实现了水稻、玉米等物种中表观修饰位点的预测,为作物功能基因组研究和智能设计育种提供工具和数据库支撑。

真核生物基因转录的主要驱动力

谷晓峰告诉《中国科学报》,表观遗传修饰是指在基因的 DNA 序列没有发生改变的情况下,基因功能发生了可遗传的变化,并最终导致了表型的变化。这些可遗传的表观修饰包括 DNA 甲基化修饰、RNA 甲基化修饰、组蛋白修饰、非编码 RNA、染色质重塑等。

“表观遗传修饰是调控真核生物基因转录、RNA 代谢以及其他生物过程的主要驱动因素。”谷晓峰说,植物、动物等真核生物细胞正常功能的维持是 DNA 序列遗传、表观遗传这两种信息相互作用、保持平衡的结果。“如果这两种因素任何一种表达失衡,都有可能导出不正常的基因表达,不能形成正常的生长发育过程。”

表观遗传在调控基因表达、RNA 稳定性、蛋白质翻译等方面起着关键作用,参与作物产量、品质、抗逆等多个重要的生物学过程。

普利介绍,作为主要表观修饰类型之一,甲基化可以发生在 DNA、RNA 或组蛋白尾部序列上,且甲基化类型具有多样性,不同位置、不同类型的甲基化修饰对基因的表达具有激活或抑制作用。

目前,以表观遗传学理论和生物技术为中心的精准表观育种打开了培育农作物的全新路径。普利解释说,表观遗传修饰位点是指在 DNA、RNA 或者组蛋白序列上的甲基化、乙酰化等修饰位点。

“解析作物的品质产量、广适性



“考虑到人工智能在处理大数据方面的优势,以及表观遗传动态可逆的特点,应用人工智能的技术预测修饰位点可以弥补由检测技术、实验成本、取材组织造成的表观修饰位点鉴定数量的不足。”

正常生长条件和 45℃ 热激处理 36 小时的水稻苗期表型。
中国农科院生物所供图

和耐逆性等重要农艺性状的调控网络,鉴定重要的表观修饰位点,尤其是新的未知的修饰位点,可以为高产优质的作物新品种提供表观遗传基因资源、修饰位点等理论支持和技术支持。”谷晓峰说。

然而,“由于表观遗传修饰在不同发育时期、不同生长环境等条件下具有动态可逆的调控特性,受到检测技术、实验成本、取材组织的限制,目前仍有大量重要的参与调控基因表达、生长发育等的表观修饰位点没有得到发掘和研究。因此,如何利用现有表观遗传学数据深入探索潜在的表观修饰位点是表观遗传研究面临的关键问题。”田健说。

人工智能生物技术成为关键

目前,科学家已经获得了一些作物的表观遗传学大数据。用什么方法解析这些数据呢?

谷晓峰团队将目光投向了人工智能等信息科学的创新和应用。人工智能驱动基因编辑、合成生物、全基因组选择等前沿生物技术快速发展,促进了人工智能和生物技术的深度融合。这种深度融合产生的人工智能生物技术成为生物农业领域的关键技术。

在此背景下,以深度学习为核心的

数据、组学数据)方面具有显著优势。

“考虑到人工智能在处理大数据方面的优势,以及表观遗传动态可逆的特点,应用人工智能的技术预测修饰位点可以弥补由检测技术、实验成本、取材组织造成的表观修饰位点鉴定数量的不足。利用深度学习算法和模型优势,我们可以训练 AI 分析表观遗传多组学大数据,从而获得有重要应用价值的潜在表观修饰位点。”谷晓峰说。

谷晓峰介绍,通过深度学习和训练,该团队基于卷积神经网络(CNN)方法,利用课题组前期绘制的粳稻日本晴的多种表观修饰图谱,构建得到了具有高精度的智能预测模型(SMEP)。

“人工智能通过收集与整理已发表的重要表观遗传学数据和修饰位点,利用深度学习的算法从中进行特征序列提取、特征序列学习训练,构建智能预测模型,并对基因组序列进行扫描,发现潜在的表观修饰位点。”田健解释道。

普利介绍,为了验证该模型的预测是否成功,他们采用多个准确率评估指标来评估预测准确率;利用两个不同水稻品种(籼稻和粳稻)的表观数据进行相互验证;将预测结果与已发表的传统机器学习方法进行对比,进一步验证该方法的准确率;利用表观遗传检测手段,验证预测位点的表观修饰水平变化。

古茶树保护有了科学依据

■本报记者 张晴丹

“世界之茶看中国,中国之茶看云贵”。中国是最早发现 and 栽培茶树的国家之一,也是茶文化的起源地。而在云南和贵州地区,分布着非常丰富的古茶树群。这些古茶树经历过千百年岁月的洗礼,依然保留着古老的气息,研究古茶树对于茶产业发展具有重要意义。

近日,贵州大学山地植物资源保护与种质创新教育部重点实验室教授赵德刚团队采用群体遗传学研究方法,从基因组层面,结合原产地茶树特征和特性,探讨了云贵地区 8 个古茶树群体的起源进化关系,并确定了茶树重要农艺性状的相关候选基因。

相关成果在线发表于《园艺研究》。该研究为中国西南地区古茶树的系统发育关系、种群结构和遗传多样性研究提供了重要补充,为茶树亲缘关系及演化路径研究提供了重要参考。

宝贵的天然基因库

茶是世界上最受欢迎的非酒精饮料,具有诱人的香气和宜人的口感。茶饮料含有丰富的人体有益的化合物,比如多酚、咖啡因、茶氨酸、维生素、多糖、挥发油和矿物质等。茶产业发展空间巨大。

云贵地区独特的地形地貌和生态环境,蕴藏大量的古茶树资源。“这些古茶树具有重要的生态价值、科学价值、经济价值和社会文化价值,是研究茶树起源与进化的重要遗传资源,也是现代茶树育种的天然基因库。”该论文通讯作者赵德刚在接受《中国科学报》采访时说。

而贵州这个地区更是有其独特之处。作为茶树向江北、江南、华南茶区迁徙的过渡地带,这里的气候条件和生态环境孕育了丰富的古茶树资源,至少有 50 多个县发现有古茶树存在。普安、沿河、花溪永安、毕节七星关等地被命名为“中国古茶树之乡”,是除云南省外我国野生乔木大茶树和灌木茶树保存最多的省份。古茶树已经

是贵州一张绚丽的名片,成就了当地别具一格的茶文化。

据统计,贵州全省古茶树总量达 65 万余株,1000 亩以上的连片古茶园有 10 余处,并存有多种珍稀野生及人工选择的具有地理标志的地方特色群体品种。

“不仅如此,贵州还是世界唯一的茶籽化石发现地,这些都足以体现出贵州古茶树重要的地位。”该论文第一作者、贵州大学山地植物资源保护与种质创新教育部重点实验室教授吕立堂在接受《中国科学报》采访时说。

宝贵的古茶树资源需要传承和保护,随着《贵州省古茶树保护条例》的颁布与施行,目前古茶树保护得到加强。“科学研究古茶树是为了最大程度地挖掘价值,发挥资源优势,为更好地保护古茶树提供有力的依据。”赵德刚表示。

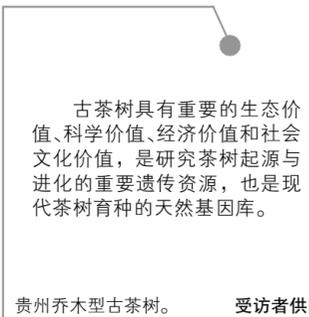
揭开起源之谜

关于茶树的起源研究,早期有起源于东南亚、印度、我国西藏和云贵高原等地之说。研究者在云贵地区发现丰富的古茶树种质资源,权威教科书认为茶树起源于云贵高原。

“茶生思州、播州、贵州、夷州……往往得之,其味极佳。”“茶圣”陆羽在《茶经》中记载的这句话里提到的地区,主要属于贵州,许多学者以此作为贵州是茶原产地的证据。

一直以来,茶树的具体起源地存在争议。“云贵地区丰富的古茶树资源到底起源于哪里,目前尚不清楚。因此,非常有必要应用形态特征结合现代分子生物学技术研究古茶树的分布、结构、起源与进化。”赵德刚指出。

近年来,依托贵州大学建设的山地植物资源保护与种质创新教育部重点实验室以及贵州大学茶学院、贵州省农科院茶科所等单位的种质资源研究人员分别对贵州习水、普安、沿河、石阡、都匀和三都以及云南大理和贺开收集



贵州乔木型古茶树。受访者供图

了 120 份具有不同形态特征的古茶树。他们利用基因组测序技术,以舒茶早基因组为参考基因组,共识别了 8082370 个单核苷酸多态性位点(SNPs),其中 1404774 个 SNPs 分布在非编码区域,175824 个 SNPs 分布在编码区域。

系统发育进化树分析表明,8 个群体 120 份古茶树聚集成 3 个组共 5 个单支,单支结构表明古茶树之间存在基因交换。科研人员还推断出古茶树群体在进化过程中存在相对较弱的选择和较长的世代时间,古茶树相比人工选育的栽培茶树具有较高的遗传多样性。群体遗传结构分析结果显示,120 株古茶树聚集成 7 个亚组,来自相同地域的株系形成单系进化支,具有共同的地理起源。

“我们发现,与栽培型茶树比较,来自贵州习水县的古茶树是最为原始的古茶树,比云南贺开古茶树和临沧古茶树更为原始。习水县也是贵州省古茶树数量最多的一个县,原始森林里随处可见。”吕立堂表示,这些发现为进一步探究茶树的起源、进化提供了重要线索。

找到候选基因,研究有待深入

古茶树有着丰富的形态特征。按照茶树树型,古茶树可分为灌木型和乔木

型;按照叶片大小特征,则可分为小叶种、中叶种、大叶种;而且叶子的颜色也深浅不一,有绿色、白色、黄色和紫色等;叶缘引起的锯齿也各不相同。这些表型性状引起了科研人员的注意。

免费在线智能预测

“这个模型从数据库中预测到了水稻、玉米等作物中大量潜在的表观修饰位点,这些修饰位点是常规技术尚未检测到的,参与了调控基因表达、环境胁迫响应等重要过程。”谷晓峰说。

水稻、玉米等是重要的粮食作物,其产量和品质受到全球变暖 and 极端高温的严重威胁而显著降低。研究显示,平均气温每上升 1 摄氏度,水稻产量就可能下降 10%,玉米产量下降 5.8%以上。

谷晓峰说,应对挑战的关键是系统阐明作物响应高温胁迫的生物学基础,鉴定和优化响应高温胁迫的关键基因和调控位点,培育耐受高温胁迫的新品种。

他们研究发现,水稻中 DNA 腺嘌呤甲基化能够影响基因表达,从而响应高温胁迫的环境信号。在热胁迫响应中,主要通过动态调节 DNA 腺嘌呤甲基化水平和影响热胁迫响应基因的表达两种方式发挥作用。

为了验证预测位点的准确性,他们利用水稻幼苗期经高温胁迫(45℃)处理 36 小时的材料,进行 DNA 腺嘌呤甲基化特异抗体实验。他们选取三个热响应转录调节基因作为候选基因,根据预测的潜在的表观修饰位点所在区域设计引物。

“结果表明,预测位点和实验结果吻合度很高,显示了智能预测位点的准确性和实用性。”普利说,预测的表观遗传修饰位点可以为水稻、玉米的耐高温育种提供有用的基因资源和位点,加速培育耐高温的作物新品种。

基因组数据非常庞大,其中还有更多表观遗传修饰位点待发现。为了让更多科学家关注和参与相关研究,该团队构建了 SMEP 免费在线智能预测工具,可向用户免费提供检索表观遗传修饰位点和基因表达数据的可视化界面。

“科研人员可以快速、准确地查询水稻、玉米等作物基因组任何区域或任何基因已知的和潜在的表观修饰位点,为作物功能基因组研究、重要基因克隆和功能分析、作物设计育种提供数据支撑。”田健说,他们还分享了所有数据和代码,从事人工智能研究的工作者也可以应用这些代码和模型。

相关论文信息: <https://doi.org/10.1111/nph.17630>

“2021 年,戴庄村越光有机再生稻头季开镰收割!”近日,随着全国脱贫攻坚楷模赵亚夫的一声令下,两台轻盈小巧的收获机驶入江苏戴庄南庄自然村的稻田,在田间灵活敏捷地前行转弯。

这两台价值 30 万元的“小铁牛”,全名叫作“4LZ-6.0 低碾率再生稻联合收获机”,是江苏大学为戴庄越光有机再生稻“量身定制”的。随着再生稻头季在机械收获难题的攻克,我国再生稻发展有望迎来新的春天。

解决两大关键问题

“所谓再生稻,就是种一茬收两回,这种水稻头季收割后,利用稻桩重新发苗、长穗,再收一季。”江苏大学农业工程学院教授李耀明介绍,再生稻不仅可以“一种双收”,还具有省种、省工、省肥、省药、品种好等优点。

随着品种改进和技术推广,再生稻迎来了示范种植的好时机,然而,传统的收获机械却成为了再生稻大面积种植的绊脚石。“传统机器在收获再生稻头季时的碾压率达到 40%~50%,导致再生稻严重减产。”多年来,李耀明一直深耕于联合收获机关键技术及装备领域,从 2017 年开始,他带领团队开始了再生稻收获机的研发工作。

李耀明认为,再生稻联合收获机关键要解决两大问题,第一是要降低碾压率,第二是要结合再生稻生长特性,保证 40~50 厘米的留茬高度。

和双季稻明显不同的是,再生稻头季收割的时候,田块里的水是放掉的,田块是硬地。硬地收获对李耀明团队来说无疑是个利好消息,他们设计将收获机的履带变窄、割幅加宽,来降低碾压率,保证高效率。

我国大面积种植的再生稻品种头季收割时,一般要求 40~50 厘米的留茬高度,一旦割多了会直接影响第二茬的产量。为此,李耀明在割台部分安装了两大“法宝”,通过智能控制图像处理技术和位置位移传感器,来实时控制割台的留茬高度,并且确保了低草谷比喂入。

历经四年有余,李耀明团队与沃得农机联合研发出了低碾率再生稻联合收获机,并在湖南、海南、四川等多地进行了试验。

适用生产实际的才是好机器

越光稻是从日本引进的一种水稻品种。在多年种植越光稻的基础上,句容市戴庄村从 2019 年开始试种越光稻再生稻。

当听说江苏大学新研发的再生稻联合收获机在湖南试验反响热烈时,赵亚夫立即联系了李耀明团队。

进展

研究揭示过热蒸汽对小麦淀粉性质的影响

本报近日,江南大学食品学院教授徐学明团队在《食品凝胶》在线发表研究论文,报道了过热蒸汽对小麦淀粉水分迁移、结构特性和流变学性质的影响。

过热蒸汽处理能够改善食物基食品的品质。然而,在应用过热蒸汽的情况下,淀粉颗粒和蒸汽环境之间的水分迁移相关信息较少。而且,过热蒸汽处理的小麦淀粉的分子量分布和流变学性质也需要阐明。

在这项研究中,相比天然淀粉,过

根据戴庄越光稻硬地收割、留茬高度的实际情况,李耀明团队又进一步加宽割幅到 2.8 米、缩窄履带到 35 厘米,让机器的嘴更“宽”,脚更“细”,身更“轻”,“这样就保证了碾压率仅有 25%,比传统机型碾压率降低了一半,增产效果更好”。

受台风影响,今年的越光稻有一点倒伏,赵亚夫一直担心会对产量造成影响,但是看到“小铁牛”现场作业情况,又有了信心,“第一茬产量估计在 650 斤左右,第二茬在去年 200 多斤基础上,今年有可能提高到 300 斤以上”。

“一巴掌按下去,田里遗留的稻粒不超过 5 粒,比普通收割机会余下 10~20 粒左右有了长足的提高,收割尽可能减损就是在为农民增收。”李耀明说。

“这两个季稻的经济效益,农民一亩地净收益可以达到 3000 元。”一小时 10 亩的收获效率让赵亚夫十分满意,他表示,有了这样的收获帮手,明年戴庄将把越光稻种植面积从 500 亩提升到 1000 亩,并用三到五年时间,在句容市发展越光有机稻 1.5 万亩。

李耀明表示,有了再生稻收获机的助力,再生稻的产量并不会低于双季稻产量,同时再生稻还具有自身独具的经济优势、口感优势,必将迎来大面积推广种植的良好时机。“将来加强农机农艺结合,通过宽窄行栽培方式,最终实现直行零碾压,同时继续提高智能化水平,帮助农民实现再生稻种植增产增收。”

隆隆声中,金黄的稻谷被“小铁牛”吞进肚中,田埂边拖拉机正等待着卸粮,一吞一吐预示着戴庄今年又迎来了一个丰收年。



4LZ-6.0 低碾率再生稻联合收获机。江苏大学供图

『小铁牛』助力再生稻丰产增收

■本报记者 温才妃 通讯员 吴奕

热蒸汽处理能够减少结晶层的平均厚度和引起淀粉分子的解聚,这归因于有限的水分迁移导致淀粉颗粒的部分糊化;过热蒸汽处理的淀粉样品的粒径和降解温度增加,而且,淀粉结构的变化能够引起淀粉样品的流变学性质的变化,包括稳态剪切特性、蠕变性和黏弹性;过热蒸汽处理的淀粉样品具有更弱的假塑性特性,但能够增加凝胶系统的可变形性和流动性。(王方)

相关论文信息: <https://doi.org/10.1016/j.foodhyd.2021.107089>

250 份大豆种质材料完成全基因组重测序

本报近日,吉林省农业科学院郝东云团队联合中国农业大学王向峰团队在《基因组蛋白组学与生物信息学报》在线发表研究论文,对 250 份大豆种质材料进行了深度全基因组重测序和生物信息学分析,构建了大豆性状耦合的遗传关联网络。

该研究解析了全球 250 个代表性大豆种质的群体结构和遗传分类,鉴定了 370 个与大豆改良相关的强选择信号,这些选择信号大部分位于与产量和品质性状相关的基因组区域。研究支持大豆遗传水平的亚群划分与其地理分布的强相关性,支持大

植物病毒靠介体昆虫进入中肠细胞囊泡高效传毒

本报日前,《分子植物病理学》在线发表了中国农业科学院植物保护研究所作物病毒病流行与控制创新团队的研究论文。该研究发现了南方水稻黑条矮缩病毒(SRBSDV)通过水稻外壳蛋白 P10 与介体白背飞虱 SNARE 家族的膜融合相关蛋白(VAMP7、Vt1a)互作,而进入到中肠上皮细胞的囊泡,并融合形成一个大的囊泡,最终突破中肠细胞释放大量病毒粒子,以实现病毒高效传播的过程。

自然界中近 70% 的植物病毒传播需要依靠介体昆虫。介体昆虫在取食植物韧皮部汁液时,植物病毒尤其

是持久型病毒会随着汁液被摄入到昆虫的肠道中,通过跨越上皮细胞进入血淋巴,借助昆虫的循环系统到达唾液腺。再次取食时,病毒会随着唾液进入到健康植物体。介体昆虫的高效传毒是导致病害暴流行的关键,而突破中肠是病毒进入到昆虫体内第一个也是最为重要的屏障。

该研究揭示了植物呼吸囊泡病毒劫持囊泡以克服媒介昆虫中肠逃逸障碍的关键机制,为囊泡运输在病毒传播中的作用提供了新的见解,同时为阻断病毒病的传播提供靶标。(李晨)

相关论文信息: <https://doi.org/10.1111/mpp.13109>