

马仁罡, 孙健英, 李宗芸. 基于生物信息学的甘薯基因组学等研究进展[J]. 江苏农业学报, 2021, 37(2): 531-538.
doi: 10.3969/j.issn.1000-4440.2021.02.032

基于生物信息学的甘薯基因组学等研究进展

马仁罡, 孙健英, 李宗芸

(江苏师范大学生命科学学院, 江苏 徐州 221116)

摘要: 甘薯是重要的粮食、工业原料和新型能源作物,同时具有较高的营养价值。近年来,随着生物信息学分析手段、二代测序技术的发展,甘薯的基因组学、转录组学、蛋白质组学、代谢组学研究取得了较大进展。本文主要综述了近年来基于生物信息学技术,甘薯及其近缘野生种在基因组学、转录组学、蛋白质组学和代谢组学等方面研究中的进展,为甘薯品种改良提供参考和借鉴,并展望了甘薯未来的研究方向。

关键词: 甘薯; 生物信息学; 基因组学; 转录组学; 蛋白质组学; 代谢组学

中图分类号: Q78; S531

文献标识码: A

文章编号: 1000-4440(2021)02-0531-08

Research progress of sweet potato genomics and other omics based on bioinformatics

MA Ren-gang, SUN Jian-ying, LI Zong-yun

(School of Life Sciences, Jiangsu Normal University, Xuzhou 221116, China)

Abstract: Sweet potato is an important food, industrial raw material and new energy crop with high nutritional value. In recent years, with the development of bioinformatics and the next generation sequencing technology, great progress had been got in the study of genomics, transcriptomics, proteomics and metabonomics of sweet potato. This paper mainly reviewed the research progress of genomics, transcriptomics, proteomics and metabonomics of sweet potato and its wild relatives based on bioinformatics in recent years. It can be used as a reference for the variety improvement of sweet potato. Finally, the future research direction of sweet potato is prospected.

Key words: sweet potato; bioinformatics; genomics; transcriptomics; proteomics; metabonomics

甘薯 [*Ipomoea batatas* (L.) Lam] 为旋花科 (Convolvulaceae) 番薯属 (*Ipomoea*) 植物,是世界重要的粮食、饲料、新型能源和工业原料作物,同时甘薯富含淀粉、花青素和硒等营养元素,是世界卫生组织推荐的最佳食品。因此,发展甘薯产业具有重要

的战略意义。甘薯具有 90 条染色体 ($2n=6x=90$), 基因组大小为 2 508 Mb^[1]。与其他作物相比,甘薯基因组学相关研究难度较大,进展较为缓慢^[2]。近几年来,生物信息学技术发展迅速,特别是新一代测序技术具有低成本和高效率的特点,促进了甘薯基因组、转录组、蛋白质组和代谢组的生物信息学分析(图 1),并建立了一系列甘薯基因组的数据库(表 1)。本文对近年来甘薯的基因组、转录组、蛋白质组和代谢组的研究进展进行了分析和总结,并以此为基础对甘薯研究进行了展望。

1 甘薯基因组研究进展

基因组学是一门交叉学科,一般分为以全基因

收稿日期: 2020-07-27

基金项目: 国家自然科学基金项目(31771367); 江苏高校优势学科建设工程项目(PAPD); 现代农业产业技术体系建设专项(CARS-10-B3); 江苏师范大学研究生科研与实践创新计划项目(2020XKT480)

作者简介: 马仁罡(1995-),男,江苏徐州人,硕士研究生,从事物种进化方向的研究。(E-mail) 13101956091@163.com

通讯作者: 李宗芸, (Tel) 0516-83500083; (E-mail) zongyunli@xsnu.edu.cn。孙健英, (E-mail) jianyingsun@jsnu.edu.cn

测序和系统表征的结构基因组学^[3],对基因功能研究的基因组学^[4]和对基因组分析比较的比较基因组学^[5]。

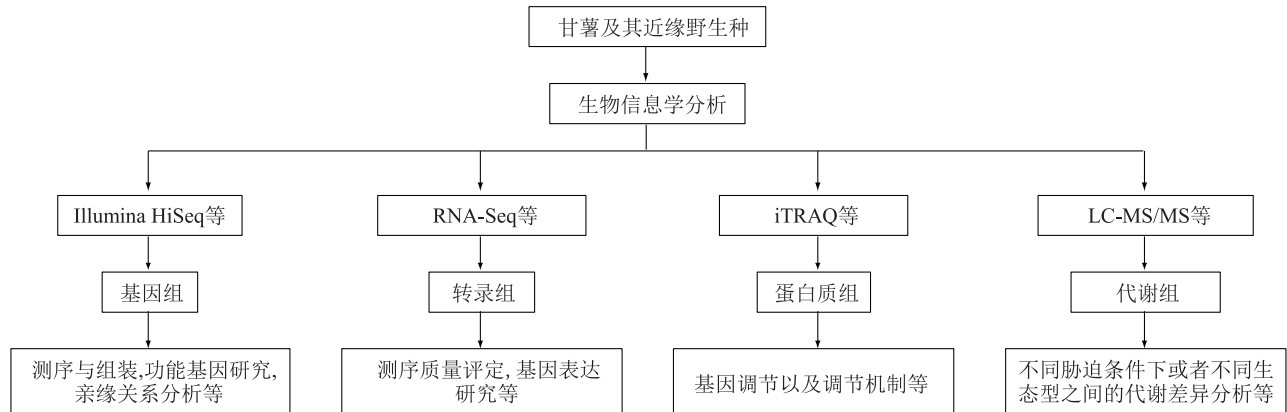


图 1 基因组、转录组、蛋白质组和代谢组的生物信息学分析

Fig.1 Bioinformatics analysis on genomics, transcriptomics, proteomics and metabonomics

表 1 甘薯基因组常用数据库

Table 1 Common genomic databases of sweet potato

名 称	网址
甘薯基因组资源(Sweetpotato genomics resource)	http://sweetpotato.plantbiology.msu.edu/index.shtml
甘薯基因组浏览器(<i>Ipomoea batatas</i> Genome Browser)	http://public-genomes-ngs.molgen.mpg.de/cgi-bin/hgGateway?db=ipoBat4
NCBI 甘薯基因组	https://www.ncbi.nlm.nih.gov/genome/?term=sweet+potato
<i>I. nil</i> 基因组	https://www.ncbi.nlm.nih.gov/genome/46552
<i>I. trifida</i> 基因组	https://www.ncbi.nlm.nih.gov/genome/37016
<i>I. triloba</i> 基因组	https://www.ncbi.nlm.nih.gov/genome/70682
<i>Ipomoea</i> 基因组	https://ipomoea-genome.org/blast.html

1.1 甘薯基因组测序和组装

近年来随着测序技术的不断发展,为了从近缘野生种中获取与生长发育和抗性相关的大量优异基因,提高甘薯育种技术,改善甘薯品质,并了解甘薯的起源和进化历程,研究者们对甘薯及其近缘野生种进行了基因组测序,包括核基因组测序以及质体基因组测序。

1.1.1 核基因组测序和组装 *Ipomoea trifida* (2x) 是六倍体甘薯最有可能的二倍体祖先,为了辅助甘薯基因组的分析,Hirakawa 等^[6]使用 Illumina HiSeq 平台对 *I. trifida* 的 2 个品系 (Mx23Hm 和 0431-1) 进行了全基因组从头测序,也是第一次进行二倍体甘薯野生种全基因组从头测序。根据 k-mer 分布估算 Mx23Hm 和 0431-1 基因组大小分别为 515.8 Mb 和 539.9 Mb, 组装获得的核心序列为 240.0 Mb 和

353.0 Mb,共鉴定出62 407个和109 449个可能的基因,大小分别为 62.4 Mb 和 87.2 Mb,但此次组装并未达到染色体水平。直到 2018 年,报道了第一个关于 *I. trifida* 以及异源多倍体假说中甘薯祖先 *I. triloba* 的染色体水平的基因组测序^[7],*I. trifida* 株系 NCNSP0306 基因组大小为 526.4 Mb,杂合度较高 (0.24%),注释得到 32 301 个基因,*I. triloba* 株系 NCNSP0323 的基因组大小为 495.9 Mb,杂合度相对较低,注释得到31 426个基因。2019 年,Li 等^[8]通过使用 Illumina PE125 对 *I. trifida* var. Y22 的基因组进行测序,生成了一个染色体水平基因组序列,基因组大小约为 476.4 Mb,杂合度为 2.2%,预测了 30 227个可能的基因,其中 79.76%被转录组数据所验证。

除了 *I. trifida* 和 *I. triloba*,同属 *I. nil* 的基因组

测序组装也已完成。2016年,Hoshino等^[9]使用二代和三代测序技术对*I. nil*进行了测序,利用PacBio reads从头组装后基因组大小约为736.4 Mb,contig N50为1.87 Mb,scaffold N50为2.88 Mb,91.42%的组装序列被锚定在15条假染色体上,是旋花科中第一个组装到染色体水平的基因组。高质量*I. nil*基因组的公布也促进了甘薯基因组学的研究,Yang等^[1]开创了一种基于二代测序序列的单倍型组装策略,对泰中6号进行全基因组从头组装,组装后1个染色体组大小约为836.3 Mb,scaffolds N50约为200.7 Kb,以*I. nil*基因组为参考基因组,有75.5%的组装序列锚定在15个假染色体上。

此外,一些甘薯近缘野生种的基因组特征信息也被报道,为这些物种全基因组深度测序提供了参考信息。马鞍藤(*I. pes-caprae*)基因组^[10]初步测序分析预估其基因组大小为1 041.65 Mb,重复序列所占比率为74.52%,杂合度为0.99%;*I. littoralis*的基因组^[11]大小预估为676.27 Mb,重复序列占比为60.98%,杂合度为0.81%;*I. cordatotriloba*基因组^[12]大小约为539.69 MB,重复序列比例57.93%,杂合度0.40%。

基因文库的构建对甘薯基因组研究起到重要作用,颜朗等^[13]为徐薯18构建180 bp、500 bp和2 Kb文库,进行从头组合式组装,预测甘薯基因组大小为2.7 Gb,重复序列比例为88%,挖掘出93 162条蛋白质编码基因序列,建立了具有创新性的从头组合式组装策略,为甘薯育种和基因组研究奠定了基础。第一个甘薯细菌人工染色体(Bacterial artificial chromosome,BAC)文库于2016年构建^[14],此BAC文库包含240 384个克隆,平均插入大小为101 kb,基因组覆盖度7.93~10.82 X。

1.1.2 质体基因组测序和组装 甘薯及其野生种的叶绿体基因组结构为典型的四分体结构,包括2个反向重复区(Inverted repeats,IRs),一个大的单拷贝区(Large single copy,LSC)和一个小的单拷贝区(Small single copy,SSC)。Eserman等^[15]对33个番薯亚族物种进行了叶绿体基因组测序,并基于叶绿体基因组序列对番薯亚族物种进行了详细的系统发育分析。Yan等^[16]对甘薯叶绿体全基因组进行了分析,结果表明甘薯叶绿体基因组包含145个基因,其中蛋白质编码基因为94个(单拷贝基因72个,双拷贝基因11个)。Sun等^[17]对8个甘薯近缘野生

种的叶绿体基因组进行了测序和组装,结果显示,这些近缘野生种叶绿体基因组长度为161 225~161 721 bp,含有80个蛋白质编码基因、4个rRNA和37个tRNA基因。Park等^[18]利用NextSeq平台对6个甘薯野生近缘种进行了叶绿体全基因组测序,基因组长度范围为161 354~161 750 bp,共鉴定出112个基因,含有78个蛋白质编码区,30个tRNA基因和4个rRNA基因。随着*I. nil*基因组的发表,研究者也公布了*I. nil*的线粒体和叶绿体基因组^[9],它们的基因组大小分别为0.27 Mb和0.16 Mb,G和C含量分别为44.45%和37.47%。

以上研究结果不仅促进了我们对甘薯及其野生近缘种基因组特征的认识,为进行甘薯遗传改良提升育种品质提供了重要的遗传信息,还加速了甘薯系统发生和起源进化的研究。如Wu等^[7]确定了贮藏根中与类胡萝卜素生物合成相关的基因和等位基因,这使高维生素A含量品种的高效育种成为可能;Li等^[8]则发现了*BMV11*与甘薯贮藏根发育有关,可促进淀粉的积累和贮藏根的膨胀,为后期分子育种提供了研究基础。甘薯泰中6号、甘薯近缘野生种*I. trifida*和*I. triloba*基因组的测序和组装为甘薯不同起源和进化假说提供了基因组方面的证据。

1.2 遗传图谱构建

遗传图谱是甘薯分子育种的重要工具之一,构建甘薯遗传图谱在遗传演化分析和数量性状位点(Quantitative trait locus,QTL)基因定位研究中起到重要作用^[19]。随着测序技术的发展,遗传图谱构建所使用的技术从相对落后的基于聚合酶链式反应(Polymerase chain reaction,PCR)的分子标记技术发展到现在基于高通量测序的分子标记技术。

1.2.1 基于PCR技术的遗传图谱构建 李爱贤等^[20]以漯徐薯8号和郑薯20杂交产生的后代为研究对象,利用序列相关扩增多态性(Sequence related amplified polymorphism,SRAP)技术和JoinMap 3.0软件构建了分子连锁图谱,漯徐薯8号连锁图谱有473个SRAP标记组成了81个连锁群,总图距为5 802.46 cM,平均标记间距为10.16 cM。郑薯20连锁图谱由328个SRAP标记组成了66个连锁群,总图距为3 967.90 cM,平均标记间距为12.02 cM。揭琴等^[21]也利用JoinMap 3.0软件,结合扩增片段长度多态性(Amplified fragment length polymorphism,AFLP)标记,构建了徐781和徐薯18的分子

连锁图谱,均含有 90 个连锁群,分别含有 1 878 和 1 868 个 AFLP 标记,总图距为 8 214.0 cM 和 8 319.0 cM,平均标记间距为 4.4 cM 和 4.5 cM。Kim 等^[22]使用表达序列标签简单重复序列 (Expressed sequence tag simple sequence repeat, EST-SSR) 标记技术,以 Yeseumi 和 Annobeney 杂交产生的后代为研究对象,共开发了 245 个 EST-SSR 标记,使用 Mapmaker 3.0 软件进行标记分组,并使用 Mapchart 2.2 软件利用其中 210 个标记进行甘薯遗传图谱的构建,构建出的图谱总长为 1 508.1 cM,平均距离为 7.2 cM。

1.2.2 基于高通量测序的遗传图谱构建 随着测序技术的发展,产生了基于高通量测序的遗传图谱构建方法,与基于 PCR 的分子标记技术相比,具有工作量更小,成本低,准确性高以及可重复等优点。Shirasawa 等^[23]通过徐薯 18 的自花授粉产生了 S1 定位群体,利用限制性双酶切位点关联 DNA 测序技术 (Double digest restriction-site associated DNA sequence, ddRAD-Seq) 对此群体文库构建和测序分析,共得到 28 087 个单核苷酸多态性 (Single nucleotide polymorphism, SNP) 位点,可映射到 96 个连锁群上,总距离为 33 020.4 cM,分析得出,国内品种间遗传距离较近,国内品种与国外品种遗传距离较远。同样, Su 等^[24]选用 264~314 bp 的 DNA 片段作为特异性位点扩增片段 (Specific-locus amplified fragment, SLAF),使用 Illumina HiSeq 2500 进行测序,利用 62 363 个 SNP 将 197 个来自全球各地的样本分为 3 组,各组材料间遗传距离值为 0.290~0.311,平均多态性信息含量均为 0.232~0.251,最小等位基因频率为 0.207~0.222。李慧峰等^[25]通过简化基因组测序技术对 122 份甘薯种质资源进行 SNP 位点开发并对种质资源群体结构和遗传进化进行分析,共获得基因组大小为 563.18 Mb, G+C 含量平均为 38.17%,共开发出 2 388 759 个高质量的 SLAF 标签,平均深度为 17.45 X,其中含有 3.26% 的多态性 SLAF 标签,共 77 761 个,含有 129 063 个群体 SNP 位点。Mollinari 等^[26]利用 MAPpoly 软件,构建了第一个六倍体甘薯的多位点整合遗传图谱,最终图谱包含 30 684 个 SNP 标记,全长 2 708.4 cM。此外, Sasai 等^[27]将几种标记技术组合起来,对 J-Red 和 Choshu 种间杂交产生的 F_1 后代结合 SSR 标记, SNP 标记以及逆转录转座子插入多态性分析,建立了高密度连锁图谱, J-Red 连锁图谱总长为 13 247 cM,各连锁群

平均长度为 147.2 cM,标记间平均距离为 2.09 cM,平均有 70.5 个标记。Choshu 连锁图谱总长为 12 241.8 cM,各连锁群平均长度 136.0 cM,标记间平均距离 2.0 cM,平均有 66.8 个标记。

以上研究结果不仅绘制了遗传图谱,还对甘薯的群体结构和遗传多样性进行了分析,为甘薯基因改良和甘薯育种提供了帮助。

2 甘薯转录组研究进展

转录组是指特定细胞或组织中全部转录产物,包括信使 RNA,核糖体 RNA、转运 RNA 以及非编码 RNA^[28],对转录组分析能够提供测序质量评定、信息定位、基因表达和可变剪接位点等信息^[29]。

2.1 非生物胁迫相关转录组分析

转录组测序方面, Ding 等^[30]利用单分子实时测序和第二代测序技术,对甘薯和 *I. trifida* 的全长转录序列进行了测定,分别获得 53 864 个和 51 184 个高质量的长阅读转录本,覆盖了基因组中约 10 439 个和 10 452 个基因座,能够预测 96.83% 和 96.82% 的转录本的开放阅读框架,并识别出 34 963 个和 33 637 个全长 cDNA 序列, 1 401 个和 1 457 个转录因子, 25 315 个和 27 090 个简单重复序列, 1 656 个和 1 389 个长的非编码 RNA, 5 251 个和 8 901 个可变剪接事件,促进了六倍体甘薯的比较和功能基因组研究。

在转录组测序过程中,研究者们发现了一些病毒和非生物胁迫相关信息。Jo 等^[31]使用 RNA 测序技术 (RNA Sequencing, RNA-Seq) 对韩国广泛栽培的 2 个甘薯品种 Beni Haruka 和 Hogammi 的 10 个不同文库进行测序分析,发现 2 种新病毒 [Sweet potato virus E (SPVE) 和 Sweet potato virus F (SPVF)] 感染,结合前人的研究,提供了一个完整的韩国甘薯可能会感染的病毒清单。Kuo 等^[32]通过 Illumina MiSeq 平台对转录组文库进行测序,利用 trinity 软件重新组装了 contigs RNA 序列,进行转录组分析,研究了甘薯中与创伤相关的 miR408 及其靶基因。Yang 等^[33]共鉴定出 475 个已知的 miRNA 和 175 个新的 miRNA,并通过降解组测序验证了 314 个 miRNA 的 636 个靶基因,发现大部分 miRNA 在盐胁迫下的表达与其靶基因的表达呈负相关。Weng 等^[34]基于甘薯转录组数据库,鉴定了 miR2111 的靶基因 *lbfBK*,通过试验发现 miR2111 的抑制作用导致

IbFBK 表达量的增加,并可能调节 *IbCNR8* 在损伤时的蛋白质降解。

在非生物胁迫下进行转录组分析有利于帮助我们筛选出与胁迫相关的基因,如 Ji 等^[35]对低温胁迫下和恢复过程中的叶片进行了转录组从头测序,分别鉴定出 2 461 个和 1 017 个差异表达基因,同时研究了抗氧化酶途径介导的活性氧(Reactive oxygen species, ROS)响应的相关基因,进一步研究了甘薯对冷胁迫的响应机制。并在之后的研究中从头组装冷藏能力较强的徐薯 15-1 和冷藏能力较弱的徐薯 15-4 2 个甘薯株系的 RNA-Seq 数据^[36],产生了 27 636 个基因, N_{50} 值为 1 204 bp,共有 525 个差异表达基因,并筛选出抗寒的候选基因。干旱胁迫下, Lau 等^[37]使用 RNA-Seq 技术,从干旱处理下的甘薯中,筛选出 122 个耐旱候选基因。在盐胁迫下, Arisha 等^[38]建立了 5 个 cDNA 文库,进行转录组测序和分析,筛选出了新的耐盐胁迫候选基因。吴燕等^[39]从甘薯的转录组数据中筛选出 238 个与耐盐、耐旱相关的基因。Sung 等^[40]通过分析甘薯在根结线虫感染过程中转录组的变化,确定了可能有助于预防甘薯块根根结线虫感染的候选基因。

2.2 发育相关转录组分析

转录组测序分析不仅可以发现与非生物胁迫相关的基因,还可以发现与发育相关的基因。块根发育方面, Dong 等^[41]为探讨甘薯贮藏根形成和发育的分子机制,以 4 个不同时期的须根中段和贮藏根为材料,制备了 5 个 cDNA 文库,通过 Illumina HiSeq 2000 测序平台,共鉴定出 26 273 个差异表达基因,根据对贮藏根形成和发育过程中表现出相似表达谱的基因进行聚类分析,确定了 4 个显著的基因亚群。Li 等^[42]则在不同甘氨酸处理的贮藏根中共鉴定出 4 836 个差异表达基因,其中涉及碳水化合物代谢的基因最多,为 1 830 个,并通过实时荧光定量 PCR (qRT-PCR)证实外源甘氨酸通过加强光合作用和增加植物激素促进贮藏根的生长,通过加速碳水化合物的代谢和调控淀粉相关基因的表达,促进贮藏根淀粉的生物合成。He 等^[43]从徐薯 18 和徐紫薯 3 的块根中构建了小分子 RNA 和降解组文库,共鉴定出 191 个已知的 miRNA、33 个新的 miRNA、180 个靶基因和 5 个新的 ib-miRNA,并筛选出与花青素相关的 miRNA 和相应的靶基因。Ponniah 等^[44]将甘薯与 *I. trifida* 转录组数据相比较,发现了一些甘薯

特异的转录因子、蛋白激酶和家族蛋白质,提示这些基因可能在贮藏根的形成中起到重要作用,并筛选出一些分子标记,为进一步研究奠定了基础。

花发育方面, Tao 等^[45]分析比较甘薯花的转录组数据,筛选出了调控开花的相关基因。Wei 等^[46]对嫁接后的 *I. nil* 进行转录组测序分析,结合 PacBioIso-Seq 和 IlluminaRNA-seq,分析了花期嫁接过程中的转录组变化,筛选出与花青素生物合成、光合作用和乙烯信号转导通路相关基因,帮助研究者们更好地认识了嫁接作用的分子机制,为甘薯育种奠定了基础。

3 甘薯蛋白质组研究进展

蛋白质组的研究在甘薯领域内相对较少, Almo-hanna 等^[47]采用 2 种互补的蛋白质提取方法和自动化的蛋白质组平台,分析了甘薯组织特异性的蛋白质组,成功鉴定出与 4 321 个非冗余蛋白质相对应的 74 255 个多肽, 39 916 个多肽定位于叶中 3 143 个特异的蛋白质上, 34 339 个多肽定位于根中 2 928 个独特的蛋白质上, 总共预测了 741 个新的蛋白质编码基因, 鉴定了甘薯叶片和贮藏根蛋白质组的组成和功能特征, 为甘薯基因组功能注释研究做出了贡献。Dong 等^[41]利用同位素标记相对和绝对定量(Isobaric tags for relative and absolute quantification, iTRAQ)对不同时期的须根和贮藏根进行了比较蛋白质组分析,从 5 个文库中共鉴定 7 727 个蛋白质,其中,在基因本体(Gene ontology, GO)数据库的注释率为 98.81%,蛋白质家族、域和作用位点信息整合数据资源(InterPro, IPR)数据库的注释率为 88.08%,京都基因与基因组百科全书(Kyoto encyclopedia of genes and genomes, KEGG)数据库的注释率为 87.03%,蛋白质相邻类的聚簇(Cluster of orthologous groups of proteins, COG)数据库的注释率为 56.59%,与转录组研究结合,证明贮藏根的发育存在转录后调控,揭示了关于贮藏根形成的途径。

4 甘薯代谢组研究进展

代谢组学是系统研究生物体或生物体内小分子代谢产物的科学^[48],在植物中的研究主要是对植物在受到外界影响前后代谢成分的定性、定量分析^[49],也可以通过对代谢组数据的分析来区分不同表型内在代谢差异^[50]。在研究中,代谢组多与基因

组、转录组以及蛋白质组联合进行分析^[51]。

在热处理过程中,甘薯的代谢产物会减少或变化,如 4 种在巴西生长的橙肉红薯经过热处理后,类胡萝卜素、酚类化合物的含量以及抗氧化能力都显著降低^[52]。Rautenbach 等^[53]的研究结果也证明了这一点,热处理后的 4 个甘薯品种的类胡萝卜素和维生素 C 含量降低,且相比于橙肉甘薯,白肉甘薯降低的更多。Kim 等^[54]在热处理前后对 12 个甘薯品种的酚类化合物含量和抗氧化活性进行了比对,发现热处理后红薯的水杨酸、香草酸、五倍子酸和咖啡酸含量增加,原儿茶酸和绿原酸含量下降。

对不同薯肉颜色甘薯进行代谢组研究,可以分析不同薯肉颜色甘薯的代谢差异和营养成分。Teow 等^[55]通过比较 19 种不同薯肉颜色甘薯的抗氧化活性、酚类含量和 β -胡萝卜素含量,总结出总酚含量可以作为甘薯抗氧化活性的指标,且试验中紫肉甘薯的抗氧化活性比白肉甘薯高。根据研究卢旺达种植的 2 个白肉和 2 个橙肉甘薯的水分、蛋白质、粗纤维、胡萝卜素和还原糖含量,推测橙肉甘薯更有营养,胡萝卜素只在橙肉甘薯中存在^[56]。此外,对不同薯肉颜色甘薯的 14 个初级代谢产物和 18 个次生代谢产物分析结果显示,花青素只存在于紫肉甘薯中,且酚类和类黄酮含量高于其他肉色甘薯,橙肉甘薯的类胡萝卜素高于其他甘薯^[57]。但在之后的研究中发现,紫肉甘薯 P40 的花青素却低于白肉甘薯 Bonita 和橙肉甘薯 Beauregard,这一出乎意料的结果刷新了对甘薯中花青素含量的认知^[58]。对不同薯肉颜色甘薯的类黄酮也开展了更深研究,共鉴定了 213 种代谢物,其中黄酮类化合物 29 种,酚酸类化合物 27 种,发现代谢差异的原因是薯肉颜色的不同,且与苯丙和黄酮的生物合成相关^[59]。对薯肉颜色相同的甘薯进行研究,有利于选取适合种植的高营养物质含量的甘薯品种。如对 5 个紫肉品种甘薯的酚类化合物进行研究,可以区分各品种所富含的酚类化合物^[60]。

5 展望

生物信息学和测序技术的发展^[61-64]极大地促进了对甘薯的研究,目前已对部分甘薯品种及 2 个近缘野生种完成了测序和较高水平的组装。此外,通过对基因组和转录组的测序分析,筛选得到一些基因,并对一些已知的基因功能进行了验证。目前,

对蛋白质组和代谢组的研究也日益增多,对甘薯品种的选择提供了帮助。综上所述,近几年来生物信息学与基因组学等研究的结合,使得甘薯及其近缘野生种的基因组、转录组、蛋白质组和代谢组研究都取得了长足的进展。在未来的研究中,需进行深度测序和高质量的组装,加强对甘薯及其野生近缘种基因组特征的认识,绘制出更精细的遗传图谱,分析筛选出更全面的功能基因,为甘薯系统发生和进化研究提出更多证据,并通过对蛋白质组和代谢组的进一步研究,选择培育更有营养的甘薯品种,为甘薯育种奠定更为坚实的基础,为中国的甘薯产业做出贡献。

参考文献:

- [1] YANG J, MOEINZADEH M, KUHL H, et al. Haplotype-resolved sweet potato genome traces back its hexaploidization history[J]. *Nature Plants*, 2017, 3(9): 696-703.
- [2] 李强,刘庆昌,马代夫,等. 甘薯遗传转化研究现状、问题及展望[J]. *分子植物育种*, 2005(1): 99-106.
- [3] 薛方方,王义聪,杜美,等. 高通量全基因组测序应用于淋球菌耐药的研究进展[J]. *中国艾滋病性病*, 2020, 26(6): 671-672.
- [4] WANG D, LI F, CAO S, et al. Genomic and functional genomics analyses of gluten proteins and prospect for simultaneous improvement of end-use and health-related traits in wheat[J]. *Theor Appl Genet*, 2020, 133(5): 1521-1539.
- [5] TORRES-CORRAL Y, SANTOS Y. Comparative genomics of *Streptococcus parauberis*: new target for molecular identification of serotype III[J]. *Applied Microbiology and Biotechnology*, 2020, 104(14): 6211-6222.
- [6] HIRAKAWA H, OKADA Y, TABUCHI H, et al. Survey of genome sequences in a wild sweet potato, *Ipomoea trifida* (H.B.K) G. Don[J]. *DNA Research*, 2015, 22(2): 171-179.
- [7] WU S, LAU K H, CAO Q, et al. Genome sequences of two diploid wild relatives of cultivated sweetpotato reveal targets for genetic improvement[J]. *Nature Communications*, 2018, 9(1): 4512-4580.
- [8] LI M, YANG S, XU W, et al. The wild sweetpotato (*Ipomoea trifida*) genome provides insights into storage root development[J]. *BMC Plant Biology*, 2019, 19(1): 119.
- [9] HOSHINO A, JAYAKUMAR V, NITASAKA E, et al. Genome sequence and analysis of the Japanese morning glory *Ipomoea nil* [J]. *Nature Communications*, 2016, 7(1): 13295.
- [10] 霍恺森,赵冬兰,陈艳丽,等. 甘薯属耐盐植物马鞍藤基因组大小及特征分析[J]. *植物遗传资源学报*, 2019, 20(3): 728-735.
- [11] 霍恺森,曹清河,王珅,等. 甘薯近缘野生种 *Ipomoea littoralis*

- 全基因组 Survey 分析[J]. 热带作物学报, 2019, 40(10): 2001-2005.
- [12] 王 珏,邓逸桐,戴习彬,等. 甘薯近缘种 *Ipomoea cordatotriloba* 基因组大小测定及高通量调查测序[J]. 热带作物学报, 2020, 41(6): 1154-1159.
- [13] 颜 朗,李雪丹,吴 燕,等. 甘薯基因组概览分析及重要功能基因挖掘[C].//中国遗传学会. 2015 中国遗传学会大会论文摘要汇编: 北京:科学出版社, 2015: 42.
- [14] SI Z, DU B, HUO J, et al. A genome-wide BAC-end sequence survey provides first insights into sweetpotato [*Ipomoea batatas* (L.) Lam.] genome composition[J]. BMC Genomics, 2016, 17(1): 945.
- [15] ESERMAN L A, TILEY G P, JARRET R L, et al. Phylogenetics and diversification of morning glories (tribe *Ipomoeae*, Convolvulaceae) based on whole plastome sequences[J]. American Journal of Botany, 2014, 101(1): 92-103.
- [16] YAN L, LAI X, LI X, et al. Analyses of the complete genome and gene expression of chloroplast of sweet potato (*Ipomoea batata*) [J]. PLoS One, 2015, 10(4): e124083.
- [17] SUN J, DONG X, CAO Q, et al. A systematic comparison of eight new plastome sequences from *Ipomoea* L.[J]. Peer J, 2019, 7: e6563.
- [18] PARK I, YANG S, KIM W J, et al. The complete chloroplast genomes of six *Ipomoea* species and indel marker development for the discrimination of authentic pharbitidis semen (Seeds of *I. nil* or *I. purpurea*) [J]. Frontiers in Plant Science, 2018, 9: 965.
- [19] 胡志程,周梦迪,吕建春,等. 甜瓜遗传图谱与基因定位研究进展[J]. 分子植物育种, 2020, 18(7): 2290-2295.
- [20] 李爱贤,刘庆昌,王庆美,等. 利用 SRAP 标记构建甘薯分子连锁图谱[J]. 作物学报, 2010, 36(8): 1286-1295.
- [21] 揭 琴,李 华,翟 红,等. 甘薯抗茎线虫病基因 AFLP 标记的开发[J]. 农业生物技术学报, 2008, 16(5): 837-841.
- [22] KIM J, CHUNG I K, KIM K. Construction of a genetic map using EST-SSR markers and QTL analysis of major agronomic characters in hexaploid sweet potato [*Ipomoea batatas* (L.) Lam.] [J]. PLoS One, 2017, 12(10): e185073.
- [23] SHIRASAWA K, TANAKA M, TAKAHATA Y, et al. A high-density SNP genetic map consisting of a complete set of homologous groups in autohexaploid sweetpotato (*Ipomoea batatas*) [J]. Scientific Reports, 2017, 7(1): 44207.
- [24] SU W, WANG L, LEI J, et al. Genome-wide assessment of population structure and genetic diversity and development of a core germplasm set for sweet potato based on specific length amplified fragment (SLAF) sequencing [J]. PLoS One, 2017, 12(2): e172066.
- [25] 李慧峰,黄咏梅,李彦青,等. 基于 SLAF-seq 技术的甘薯种质资源群体遗传进化分析[J]. 热带作物学报, 2019, 40(12): 2390-2396.
- [26] MOLLINARI M, OLUKOLU B A, PEREIRA G D S, et al. Unravelling the hexaploid sweetpotato inheritance using ultra-dense multi-locus mapping[J]. G3-Genes Genomes Genetics, 2020, 10(1): 281-292.
- [27] SASAI R, TABUCHI H, SHIRASAWA K, et al. Development of molecular markers associated with resistance to *Meloidogyne incognita* by performing quantitative trait locus analysis and genome-wide association study in sweetpotato [J]. DNA Research, 2019, 26(5): 399-409.
- [28] 崔 凯,吴伟伟,刁其玉. 转录组测序技术的研究和应用进展[J]. 生物技术通报, 2019, 35(7): 1-9.
- [29] 宋尚桥,马围围,张超龙,等. 基于转录组测序生物信息学分析的研究进展[J]. 中国畜牧兽医, 2020, 47(2): 392-398.
- [30] DING N, CUI H, MIAO Y, et al. Single-molecule real-time sequencing identifies massive full-length cDNAs and alternative-splicing events that facilitate comparative and functional genomics study in the hexaploid crop sweet potato [J]. Peer J, 2019, 7: e7933.
- [31] JO Y, KIM S, CHOI H, et al. Sweet potato viromes in eight different geographical regions in Korea and two different cultivars [J]. Scientific Reports, 2020, 10(1): 2588.
- [32] KUO Y, LIN Y, LI Y, et al. MicroR408 regulates defense response upon wounding in sweet potato [J]. Journal of Experimental Botany, 2019, 70(2): 469-483.
- [33] YANG Z, ZHU P, KANG H, et al. High-throughput deep sequencing reveals the important role that microRNAs play in the salt response in sweet potato (*Ipomoea batatas* L.) [J]. BMC Genomics, 2020, 21(1): 116-164.
- [34] WENG S, KUO Y, KING Y, et al. Regulation of microRNA2111 and its target IbFBK in sweet potato on wounding [J]. Plant Science, 2020, 292: 110391.
- [35] JI C Y, BIAN X, LEE C, et al. De novo transcriptome sequencing and gene expression profiling of sweet potato leaves during low temperature stress and recovery [J]. Gene, 2019, 700: 23-30.
- [36] JI C Y, HO S K, LEE C, et al. Comparative transcriptome profiling of tuberous roots of two sweetpotato lines with contrasting low temperature tolerance during storage [J]. Gene, 2020, 727: 144244.
- [37] LAU K H, ROSARIO HERRERA M, CRISOVAN E, et al. Transcriptomic analysis of sweet potato under dehydration stress identifies candidate genes for drought tolerance [J]. Plant Direct, 2018, 2(10): e92.
- [38] ARISHA M H, ABOELNASR H, AHMAD M Q, et al. Transcriptome sequencing and whole genome expression profiling of hexaploid sweetpotato under salt stress [J]. BMC Genomics, 2020, 21(1): 118-197.
- [39] 吴 燕,颜 朗,李雪丹,等. 甘薯耐旱和耐盐基因的挖掘和表达分析[J]. 四川大学学报, 2016, 53(5): 1147-1154.
- [40] SUNG Y W, LEE I H, SHIM D, et al. Transcriptomic changes in sweetpotato peroxidases in response to infection with the root-knot nematode *Meloidogyne incognita* [J]. Molecular Biology Reports, 2019, 46(4): 4555-4564.

- [41] DONG T, ZHU M, YU J, et al. RNA-Seq and iTRAQ reveal multiple pathways involved in storage root formation and development in sweet potato (*Ipomoea batatas* L.) [J]. BMC Plant Biology, 2019, 19(1): 136.
- [42] LI C, YAO W, WANG J, et al. A novel effect of glycine on the growth and starch biosynthesis of storage root in sweetpotato (*Ipomoea batatas* Lam.) [J]. Plant Physiology and Biochemistry, 2019, 144: 395-403.
- [43] HE L, TANG R, SHI X, et al. Uncovering anthocyanin biosynthesis related microRNAs and their target genes by small RNA and degradome sequencing in tuberous roots of sweetpotato [J]. BMC Plant Biology, 2019, 19(1): 232.
- [44] PONNIAH S K, THIMMAPURAM J, BHIDE K, et al. Comparative analysis of the root transcriptomes of cultivated sweetpotato [*Ipomoea batatas* (L.) Lam] and its wild ancestor [*Ipomoea trifida* (Kunth) G. Don] [J]. BMC Plant Biology, 2017, 17(1): 9.
- [45] TAO X, GU Y H, JIANG Y Z, et al. Transcriptome analysis to identify putative floral-specific genes and flowering regulatory-related genes of sweet potato [J]. Bioscience, Biotechnology, and Biochemistry, 2013, 77(11): 2169-2174.
- [46] WEI C, LI M, QIN J, et al. Transcriptome analysis reveals the effects of grafting on sweetpotato scions during the full blooming stages [J]. Genes & Genomics, 2019, 41(8): 895-907.
- [47] ALMOHANNA T, AHSAN N, BOKROS N T, et al. Proteomics and proteogenomics analysis of sweetpotato (*Ipomoea batatas*) leaf and root [J]. Journal of Proteome Research, 2019, 18(7): 2719-2734.
- [48] 杨倩春, 李思宁, 陈 硕, 等. 代谢组学的运用及其研究进展 [J]. 临床合理用药杂志, 2020, 13(2): 176-178.
- [49] 王佳钰. 重金属胁迫下植物代谢组学研究进展 [J]. 绿色科技, 2020(1): 33-34.
- [50] LEE W, YEO Y, OH S, et al. Compositional analyses of diverse phytochemicals and polar metabolites from different-colored potato (*Solanum tuberosum* L.) tubers [J]. Food Science and Biotechnology, 2017, 26(5): 1379-1389.
- [51] 杨慧菊, 兰玉倩, 王石华. 植物响应低温胁迫组学研究进展 [J]. 山东农业科学, 2020, 52(5): 142-148.
- [52] DONADO-PESTANA C M, SALGADO J M, DE OLIVEIRA RIOS A, et al. Stability of carotenoids, total phenolics and *in vitro* antioxidant capacity in the thermal processing of orange-fleshed sweet potato (*Ipomoea batatas* Lam.) cultivars grown in Brazil [J]. Plant Foods for Human Nutrition, 2012, 67(3): 262-270.
- [53] RAUTENBACH F, FABER M, LAURIE S, et al. Antioxidant capacity and antioxidant content in roots of 4 sweetpotato varieties [J]. Journal of Food Science, 2010, 75(5): C400-C405.
- [54] KIM M Y, LEE B W, LEE H U, et al. Phenolic compounds and antioxidant activity in sweet potato after heat treatment [J]. Journal of the Science of Food and Agriculture, 2019, 99(15): 6833-6840.
- [55] TEOW C C, TRUONG V, MCFEETERS R F, et al. Antioxidant activities, phenolic and β -carotene contents of sweet potato genotypes with varying flesh colours [J]. Food Chemistry, 2006, 103: 829-838.
- [56] ROSE I M, VASANTHAKAALAM H. Comparison of the Nutrient composition of four sweet potato varieties cultivated in Rwanda [J]. American Journal of Food and Nutrition, 2011, 1(1): 34-38.
- [57] PARK S, LEE S Y, YANG J W, et al. Comparative analysis of phytochemicals and polar metabolites from colored sweet potato (*Ipomoea batatas* L.) tubers [J]. Food Science and Biotechnology, 2016, 25(1): 283-291.
- [58] SU X, GRIFFIN J, XU J, et al. Identification and quantification of anthocyanins in purple-fleshed sweet potato leaves [J]. Heliyon, 2019, 5(6): e1964.
- [59] WANG A, LI R, REN L, et al. A comparative metabolomics study of flavonoids in sweet potato with different flesh colors [*Ipomoea batatas* (L.) Lam] [J]. Food Chemistry, 2018, 260: 124-134.
- [60] OKI T, MASUDA M, FURUTA S, et al. Involvement of anthocyanins and other phenolic compounds in radical-scavenging activity of purple-fleshed sweet potato cultivars [J]. Journal of Food Science, 2002, 67(5): 1752-1756.
- [61] 杨妍梅, 李 玉, 覃 圣, 等. 静宁鸡 PPAR α 基因克隆与生物信息学分析 [J]. 江苏农业学报, 2019, 35(2): 370-377.
- [62] 庞宁宁, 樊怀福, 王 哲, 等. 黄瓜 PP2-A1 蛋白的生物信息学分析 [J]. 江苏农业科学, 2019, 47(2): 46-49.
- [63] 祖盘玉, 李 维, 林家栋, 等. 赤水乌骨鸡 TYR 基因多态性及生物信息学分析 [J]. 南方农业学报, 2019, 50(12): 2806-2811.
- [64] 冯 磊, 石元豹, 汪贵斌, 等. 银杏 bHLH 家族转录因子生物信息学及表达分析 [J]. 江苏农业学报, 2019, 35(2): 400-411.

(责任编辑: 陈海霞)