

DOI:10.11686/cyxb2025155

http://cyxb.magtech.com.cn

郭一获, 刘凌云, 范希峰, 等. 狼尾草属植物基因组学研究进展. 草业学报, 2026, 35(4): 221—232.

GUO Yi-di, LIU Ling-yun, FAN Xi-feng, et al. Progress in genomic research on *Pennisetum*. Acta Prataculturae Sinica, 2026, 35(4): 221—232.

狼尾草属植物基因组学研究进展

郭一获^{1,2}, 刘凌云², 范希峰², 岳跃森², 穆娜², 常智慧^{1*}, 滕珂^{2*}

(1. 北京林业大学草业与草原学院, 北京 100083; 2. 北京市农林科学院草业花卉与生态景观研究所, 北京 100097)

摘要: 狼尾草属植物兼具牧草、观赏草、能源草和生态修复草的重要价值。自 2017 年美洲狼尾草基因组首个高质量参考序列发布, 狼尾草属植物研究进入了基因组学时代, 为其属资源开发与分子育种提供了重要基础。开展基因组学研究为解析狼尾草属植物的遗传进化机制、挖掘重要性状功能基因及设计精准育种策略等提供了关键支撑。本研究系统梳理了狼尾草属植物基因组学的研究进展, 涵盖基因组测序与组装技术, 比较基因组学与进化分析、功能基因组学研究、群体遗传学及分子育种应用, 旨在为狼尾草属植物的基因组学及其相关应用研究提供理论参考。

关键词: 狼尾草属; 基因组测序; 比较基因组学; 功能基因组学; 遗传进化

Progress in genomic research on *Pennisetum*

GUO Yi-di^{1,2}, LIU Ling-yun², FAN Xi-feng², YUE Yue-sen², MU Na², CHANG Zhi-hui^{1*}, TENG Ke^{2*}

1. School of Grassland Science, Beijing Forestry University, Beijing 100083, China; 2. Institute of Grassland, Flowers and Landscape Ecology, Beijing Academy of Agriculture and Forestry Sciences, Beijing 100097, China

Abstract: Members of the *Pennisetum* genus (recently reclassified as *Cenchrus*) are vital resources for forage production, ornamental horticulture, bioenergy feedstocks, and ecological restoration. The release of the first high-quality reference genome for *Pennisetum americanus* (also called *Cenchrus americanus*) in 2017 heralded a new genomics era in *Pennisetum* research, laying the foundation for resource development and molecular breeding. Genomic studies have since offered critical insights into the mechanisms of genetic evolution, the functional genes underlying key traits, and precision breeding strategies. This review systematically summarizes progress in *Pennisetum* genomic research, encompassing genome sequencing and assembly technologies, comparative genomics and evolutionary analyses, functional genomics, population genetics, and molecular breeding applications. It aims to provide a reference for future innovations in genomics-driven research on this genus.

Key words: *Pennisetum*; genome sequencing; comparative genomics; functional genomics; genetic evolution

1 狼尾草属物种及全基因组测序简介

1.1 狼尾草属物种简介

狼尾草属 (*Pennisetum*) 为禾本科 (Poaceae) 黍亚科 (Panicoideae) 植物, 包含约 140 个种, 具有重要的生态与经济价值^[1]。在我国有 11 种狼尾草分布, 主要包括狼尾草 (*P. alopecuroides*)、美洲狼尾草 (*P. americanus*)、长序狼尾草 (*P. longissimum*)、象草 (*P. purpureum*)、陕西狼尾草 (*P. shaanxiense*) 等 (图 1), 多为优良牧草, 主要分布于

收稿日期: 2025-04-24; 改回日期: 2025-07-10

基金项目: 北京市农林科学院创新能力建设专项 (KJCX20251208; KJCX20250917) 和云南省科技计划项目 (202403AP140045) 资助。

作者简介: 郭一获 (1999—), 女, 山东菏泽人, 在读博士。E-mail: guoyidi99@163.com

* 通信作者 Corresponding author. E-mail: changzh@bjfu.edu.cn; tengke@baafs.net.cn

西南、西北地区,部分种类分布广泛^[2-3]。目前有4种狼尾草属植物进行了基因组测序研究,包括美洲狼尾草、象草、狼尾草、巨菌草(*Pennisetum giganteum*)。

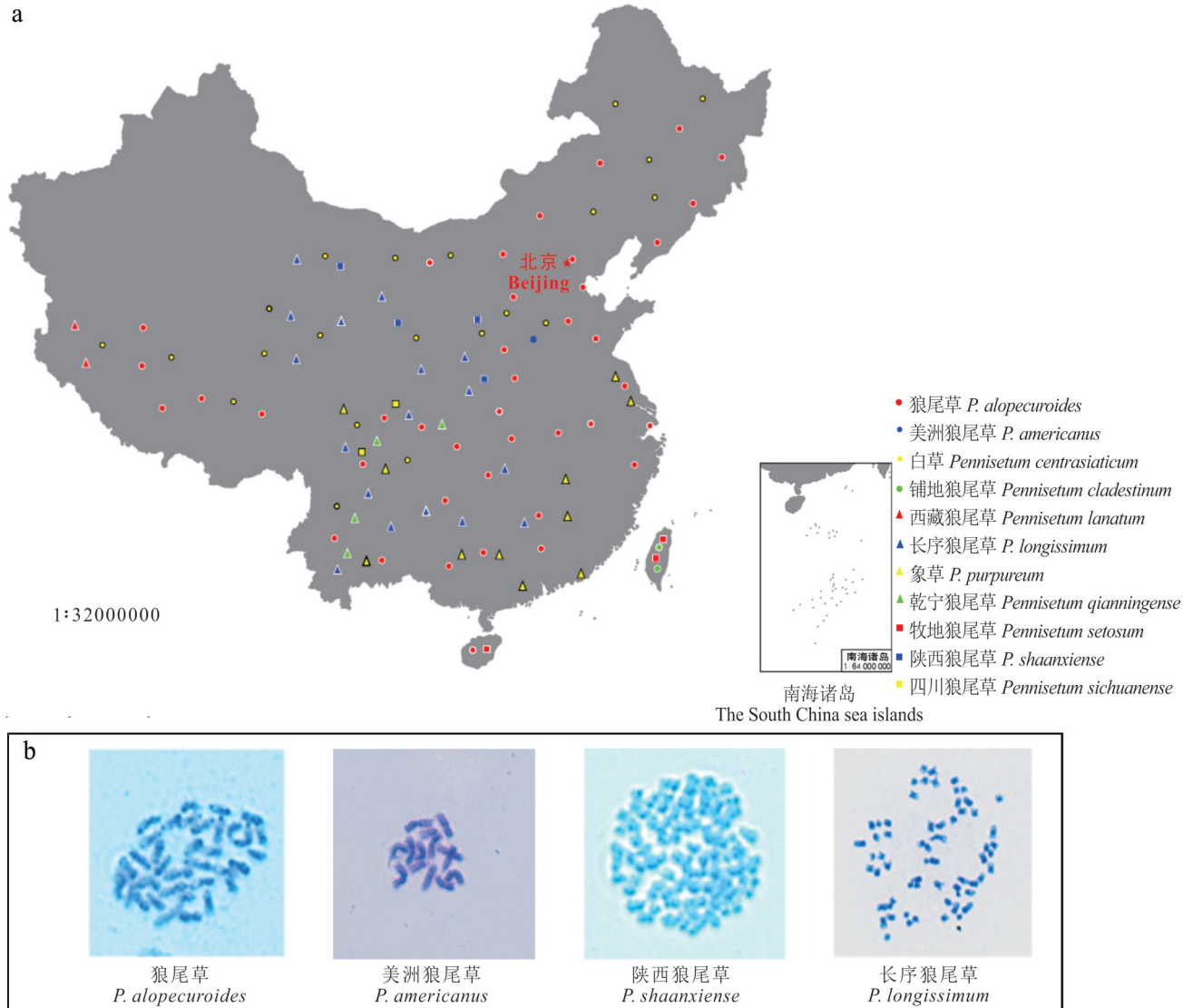


图1 中国狼尾草属植物分布情况

Fig. 1 The geographical distribution of *Pennisetum* in China

a: 中国狼尾草属植物分布情况(根据中国植物志绘制),基于自然资源部标准地图服务网站GS(2016)1571号标准地图制作,底图边界无修改 The geographical distribution of *Pennisetum* in China (based on the Flora Reipublicae Popularis Sinicae), based on the standard map service website GS (2016)1571 of the Ministry of Natural Resources, the boundary of the base map is not modified; b: 部分狼尾草属代表性物种的核型 Karyotypes of representative species of *Pennisetum*.

美洲狼尾草作为亚洲、非洲干旱和半干旱热带地区旱作农业的支柱作物,营养价值与玉米(*Zea mays*)相当,是全球干旱地区的主粮之一,是低投入的作物-畜牧系统重要的饲料来源(图2a)^[4]。作为C₄植物,美洲狼尾草具有较高的光合效率,强大的耐高温、干旱和耐贫瘠能力^[5]。象草广泛分布于热带地区,也是一种C₄植物,它在高温和干旱环境下仍能保持较高的生长速率,在不同生长阶段都能保持较高的生物量,还具有较高的粗蛋白含量和干物质消化率,是一种高质量的饲料作物(图2b)^[6]。象草的纤维素含量较高(12.3%),木质素含量较低(10.73%),使其成为造纸工业中重要的非木材纤维来源,并且象草较高的生物量和纤维素含量使其成为生物能源开发的明星物种^[7-8]。狼尾草是一种C₃植物,作为优良的观赏草、牧草、生态草(图2c),其根系发达,能够有效防止土壤侵

蚀和水土流失,对生态修复和植被恢复具有重要作用^[9]。狼尾草独特的花序和优雅的外观使其成为城市绿化、公园景观和庭院装饰的理想草种,被广泛应用于园林景观设计^[10]。狼尾草还具有较高的生物量和良好的营养价值,干物质产量高、粗蛋白含量丰富、适口性好,是一种优质的饲草^[11]。巨菌草作为一种新发现的禾本科植物,具有典型的C₄植物特征及很强的环境适应性,其生物产量高,是已知栽培植物中生物量较高的草种之一^[12]。

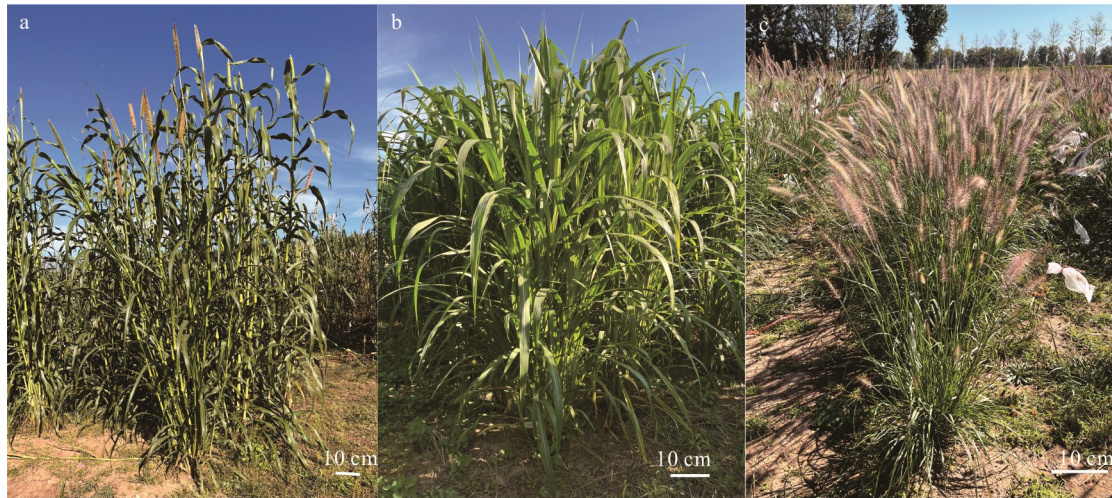


图 2 部分狼尾草属基因组测序植物

Fig. 2 Genome-sequenced plants of certain species within the genus *Pennisetum*

a: 美洲狼尾草 *P. americanum*; b: 象草 *P. purpureum*; c: 狼尾草 *P. alopecuroides*.

1.2 全基因组研究进展

全基因组测序经历了第一代测序到第三代测序,已成为现代生物学的重要研究方法,大量的研究成果被用于医学、农业、生物技术等领域。第一代测序也称为 Sanger 测序,是 20 世纪 70 年代由 Sanger 等^[13]开创的 DNA 双脱氧链终止法测序。Sanger 测序操作简单、误差小,但是成本高、通量低,测序长度有限,限制了其在基因组大规模测序的进展。第二代测序(next-generation sequencing, NGS)也称为高通量测序,测序平台主要包括 Roche 的 454FLX 和 Illumina 的 Miseq/Hiseq 等^[14]。第二代测序基于半导体芯片或毛细管电泳技术在短时间内对大量 DNA 片段进行测序,极大地提高了测序的通量和效率,但其读长较短,并且在进行复杂基因组测序时会产生较高的重复序列和杂合率。第三代测序包括 SMRT 和 Nanopore 测序技术,是通过纳米孔单分子成像技术捕捉 DNA 序列中的每个碱基,并对其进行解码和分析,提高了基因组测序的准确性与可重复性^[15]。但第三代测序仍存在一些挑战和局限性,包括读长较长,需要更准确地组装算法和技术来解析基因组,错误率仍较高,需要进一步改进和优化测序技术以提高准确性。基因组组装的目的是将测序产生的短序列(reads)组装成为一条连续的基因组序列,为后续的基因组研究提供基础数据^[16]。基因组组装技术包括高通量染色体构象捕获(Hi-C)、BioNano、10X Genomics 等^[17-18]。目前,基因组的组装采用多种技术组合的方法进行,以提高组装的准确度和整体质量。

随着测序技术的迅速发展,对植物基因组的深入探究已成为揭示植物性状形成机制、优化植物育种以及推动农业可持续发展的重要手段^[19]。狼尾草属植物自 2017 年开始有基因组数据公布,并且随着长读长测序技术和多组学整合分析的突破,狼尾草属基因组研究取得显著进展,为耐逆机制解析、分子育种及生物质能开发奠定了基础^[20-21]。基因组数据对研究植物的群体遗传学、物种演化和抗性机制有重要意义^[22]。狼尾草属植物基因组复杂度高,具有丰富的遗传多样性,不同种差异显著。同时,狼尾草属植物存在有性生殖与无融合生殖两种生殖方式,倍性水平包括二倍体至八倍体,生命周期为一年生或多年生。当前我国饲草育种面临品种创新能力不足、种源依赖进口等问题^[23]。狼尾草属作为重要的饲草资源,其基因组研究亟须结合育种技术,以加速高产、抗逆品种的培育。对比其他禾本科的基因组研究进展,如玉米、水稻(*Oryza sativa*)等,在基因组测序及应用方面狼尾草属植物

具有滞后性^[24-26]。综述国内外狼尾草属植物基因组研究的最新进展(包括测序技术、组装策略、组装质量以及应用),并展望其未来发展的趋势和前景,对发掘利用狼尾草属植物资源和开展分子育种具有重要的现实意义。

2 狼尾草基因组的测序与组装

随着基因组测序技术的发展,越来越多的狼尾草属植物完成了基因组测序与组装(图3和表1)。目前有22个狼尾草属材料的基因组数据已公布,包括17种美洲狼尾草、1种狼尾草、2种象草、2种巨菌草。第一个狼尾草属植物基因组发表于2017年,是美洲狼尾草的 Tift 23D2B1-P1-P5^[20]。2023年发表了6个狼尾草属植物基因组,包括3个美洲狼尾草、1个狼尾草和2个巨菌草基因组(表1),2023年也被联合国宣布为2023国际小米年(The International Year of Millets 2023)^[27]。美洲狼尾草、狼尾草、象草、巨菌草的基因组测序时间、测序技术与组装策略以及组装指标如表1所示。图3为狼尾草属植物基因组测序的时间分布以及来源机构,2017-2024年共有10篇关于狼尾草属植物基因组测序的相关研究,其中6项来自中国。

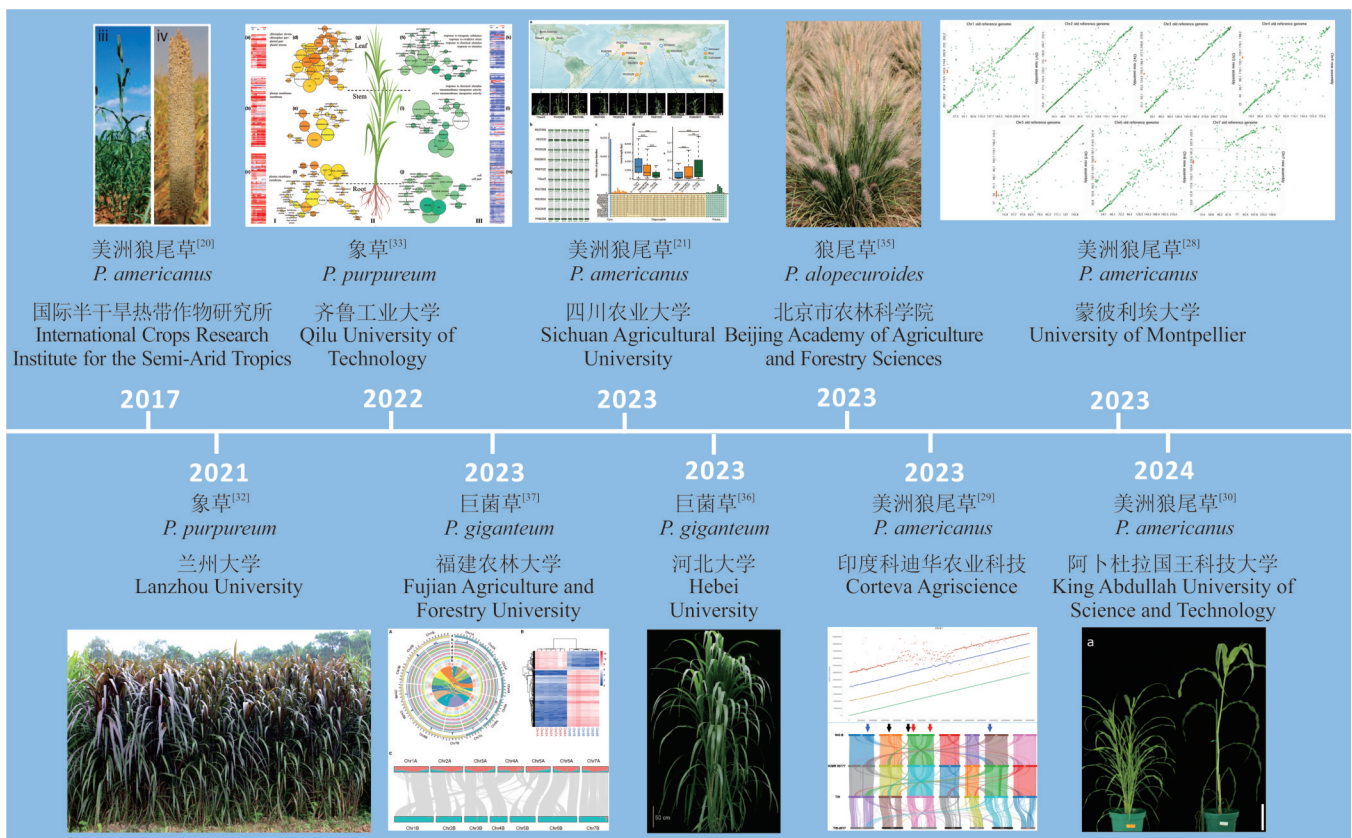


图3 狼尾草属植物基因组测序进展

Fig. 3 Progress in genome sequencing of the genus *Pennisetum*

2.1 美洲狼尾草基因组

Varshney等^[20]在2017年通过WGS(whole genome shotgun sequencing)和BAC(bacterial artificial chromosome)测序技术对美洲狼尾草的 Tift 23D2B1-P1-P5品种进行全基因组测序,这是国际半干旱热带作物研究所对狼尾草属植物基因组研究的开创性成果,是狼尾草属基因组研究的开端。该研究使用SOAPdenovo 53软件组装了约1.79 Gb的参考基因组,此版本基因组的BUSCO完整性为95.4%,CEGMA完整性为97%,但受限于测序技术,其重叠群N50(Contig N50)仅为18180 bp,支架N50(Scaffold N50)仅为884945 bp。该版本基因组鉴定了约38579个基因,基因组中重复序列占比高达77.2%,显示出较高的复杂性。

表 1 狼尾草属植物基因组测序研究进展

Table 1 Research progress of genome sequencing of *Pennisetum*

材料 Materials	物种 Species	测序和组装方法 Sequencing and assembly methods	支架 N50 Scaffold N50	重叠群 N50 Contig N50	基因组大小 Genome size	基因数量 Gene No.	文献 Reference
Tift 23D2B1-P1-P5	PA	Illumina WGS/BAC+Multiple	0.88 Mb	0.02 Mb	1.76 Gb	38579	[20]
PI537069	PA	PacBio HiFi+Bionano+Hi-C+Illumina	266.84 Mb	61.62 Mb	—	35486	[21]
PI521612	PA	PacBio HiFi+Bionano+Hi-C+Illumina	278.46 Mb	5.40 Mb	—	37906	[21]
PI587025	PA	PacBio HiFi+Bionano+Hi-C+Illumina	257.50 Mb	5.15 Mb	—	38076	[21]
PI583800	PA	PacBio HiFi+Bionano+Hi-C+Illumina	261.45 Mb	3.10 Mb	—	35826	[21]
Tifleaf3	PA	PacBio HiFi+Bionano+Hi-C+Illumina	279.17 Mb	25.57 Mb	—	37280	[21]
PI526529	PA	PacBio HiFi+Bionano+Hi-C+Illumina	286.98 Mb	79.18 Mb	—	36451	[21]
PI186338	PA	PacBio HiFi+Bionano+Hi-C+Illumina	284.64 Mb	3.80 Mb	—	36343	[21]
PI343841	PA	PacBio HiFi+Bionano+Hi-C+Illumina	263.66 Mb	5.10 Mb	—	36312	[21]
PI527388	PA	PacBio HiFi+Bionano+Hi-C+Illumina	193.80 Mb	3.10 Mb	—	37866	[21]
PI250656	PA	PacBio HiFi+Bionano+Hi-C+Illumina	276.63 Mb	4.20 Mb	—	36923	[21]
Tift23D2B1-P1-P5	PA	PacBio HiFi+hifasm+BioNano	279.00 Mb	129.40 Mb	—	50339	[29]
ICMR 06777	PA	PacBio HiFi+hifasm+BioNano	144.60 Mb	71.70 Mb	—	57822	[29]
843 B	PA	PacBio HiFi+hifasm+BioNano	144.00 Mb	126.50 Mb	—	54176	[29]
Tift23D2B1-P1-P5	PA	Oxford nanopore+Bionano	81.82 Mb	1.15 Mb	1.94 Gb	36898	[30]
Aw	PA	PacBio HiFi+Omni-C	—	284.00 Mb	1.92 Gb	38920	[30]
P10	PA	PacBio HiFi+Omni-C	—	244.00 Mb	1.93 Gb	40869	[30]
紫色象草 Purple	PP	Oxford nanopore technologies+Hi-C	455.00 bp	1.83 Mb	2.01 Gb	65927	[32]
CIAT6263	PP	Nanopore+HiC	8.47 Mb	2.90 Mb	2.07 Gb	77139	[33]
丽秋 Liqiu	PAL	PacBio+Illumina+Hi-C	91.33 Mb	84.83 Mb	845.71 Mb	34312	[35]
<i>P. giganteum</i>	PF	PacBio+Hi-C	—	88.47 Mb	2.03 Gb	—	[36]
JUJUNCAO	PF	PacBio HiFi+Hi-C	—	134.10 Mb	1.99 Gb	68526	[37]

PA: 美洲狼尾草 *P. americanus*; PP: 象草 *P. purpureum*; PAL: 狼尾草 *P. alopecuroides*; PF: 巨菌草 *P. giganteum*.

随着长读长测序技术升级,Salson 等^[28]聚焦于提升基因组组装质量,通过结合 Oxford nanopore 长读长测序技术和 Bionano 光学图谱的方法,对 Tift 23D2B1-P1-P5 美洲狼尾草基因组进行改进组装,显著提高了美洲狼尾草基因组的组装质量。新的组装版本在染色体级别的连续性上有了显著提升,与 2017 年 Varshney 等^[20]的版本相比增加了约 200 Mb 的序列,N50 值提升了两个数量级,达到 86 Mb。新组装的基因组在 BUSCO 评估中达到了 98.4%,基因组组装的完整性显著提高。

Yan 等^[21]采用 PacBio HiFi 长读长测序、Bionano 光学图谱、Hi-C 数据以及 Illumina 短读长测序技术的多技术联合测序策略,完成了 10 个美洲狼尾草品种 (PI537069、PI521612、PI587025、PI583800、Tifleaf3、PI526529、PI186338、PI343841、PI527388、PI250656) 的泛基因组组装。该研究组装的 10 个品种美洲狼尾草基因组大小为 1.89~2.00 Gb,Scaffold N50 为 193.80~286.98 Mb,揭示了重复序列中占比 71.58% 的保守特征。

Ramu 等^[29]通过 PacBio HiFi 测序与 Bionano 光学图谱技术构建的 3 个美洲狼尾草系的高连续性、完整性和准确性的高质量基因组,实现了染色体水平突破。其中 Tift 23D2B1-P1-P5 品种的 contig N50 达到了 129.4 Mb,较 2017 版本提高了 7000 倍。使用 *de novo* 基因预测和证据基础方法,对基因组进行注释,识别了 35000~38000 个基因,其中 500~700 个是 tRNA 基因,其余为蛋白编码基因。Kuijer 等^[30]创新性地结合 PacBio HiFi 长读长测序和

Omni-C 三维基因组技术构建了易感品系 P10 和抗性品系 29Aw 美洲狼尾草的染色体级别基因组,解析了抗性品系的 4 条无间隙染色体结构(Contig N50>240 Mb),为功能基因组学的研究提供了新范式。

2.2 象草基因组

相较于美洲狼尾草,象草的基因组学研究起点较晚。2018 年,Paudel 等^[31]首次对象草基因组进行了全面调查,发现其与美洲狼尾草基因组具有高度同源性,但存在染色体重排现象,这为象草的基因组结构和进化研究提供了重要线索。2021 年,Yan 等^[32]首次采用短读测序、单分子长读测序(oxford nanopore technologies)和 Hi-C 技术组装了象草的高质量染色体级别的基因组,总大小为 2.01 Gb,Contig N50 为 1.83 Mb,染色体 N50 (Chromosome N50)为 150 Mb,BUSCO 完整性为 97.8%,CEGMA 完整性为 93.9%,杂合度为 1.5%,包含 66.32% 的重复序列,共鉴定出 65927 个蛋白质编码基因。

2022 年,Zhang 等^[33]通过 Nanopore 测序和 Hi-C 技术完成了象草的高质量染色体水平基因组组装,基因组大小为 2.07 Gb,该版本象草基因组在保持同等 BUSCO 完整性(97.8%)的情况下,将 Contig N50 提升至 2.90 Mb。基因组的 Scaffold N50 为 8.47 Mb,Chromosome N50 为 146.84 Mb,重复序列比例为 60.74%,杂合率为 1.3%,发现了 77139 个基因,将其分为 A' 和 B 两个亚基因组并进行亚基因组分化的系统解析,为理解该多倍体物种进化机制提供了新视角。

2.3 狼尾草基因组

狼尾草作为我国一种重要的乡土牧草与观赏草资源,具有巨大的开发潜力。北京市农林科学院通过根尖压片进行了 9 种观赏型狼尾草的核型分析,并通过流式细胞术测定其基因组大小,发现不同狼尾草的染色体数目为 18~45,基本染色体数 X 为 7 或 9,核 DNA 含量为 $1.84 \times 10^9 \sim 4.61 \times 10^9$ bp。并结合形态学指标系统解析了 9 种狼尾草的遗传关系与进化规律,为后续基因组测序提供了研究基础^[34]。北京市农林科学院于 2023 年成功组装了‘丽秋’狼尾草的染色体水平参考基因组,为该领域的研究提供了新的视角。研究团队综合运用 PacBio、Illumina 和 Hi-C 技术,组装出一个大小约为 845.71 Mb 的基因组,其中 833.41 Mb 的序列被锚定到 9 条染色体上,基因组完整性为 99.13%、Contig N50 为 84.83 Mb,Scaffold N50 为 91.33 Mb。经分析,狼尾草基因组中重复序列占 60.66%,共预测到 34312 个基因^[35]。

2.4 巨菌草基因组

针对异源多倍体巨菌草的基因组研究,目前取得了两项重要进展。Xing 等^[36]研究了异源四倍体巨菌草的基因组,使用 PacBio 测序技术和 Hi-C 技术生成了巨菌草的高质量基因组组装,基因组大小为 2.03 Gb,Contig N50 为 88.47 Mb,分为 A 和 B 两个亚基因组。而 Zheng 等^[37]利用 PacBio HiFi 测序技术,结合 Hi-C 技术辅助基因组组装,达到了“端粒到端粒”(T2T)水平,总大小约为 1.99 Gb,其中 99% 的序列被锚定到 14 条染色体上。

3 狼尾草的比较基因组学与进化分析

3.1 狼尾草属内的物种进化

狼尾草属多倍体物种在进化中呈现出显著的功能亚基因组分化模式,其核心特征在象草和巨菌草中得到系统解析。Zhang 等^[33]通过系统进化树构建与基因家族动态分析,首次揭示了象草的四倍体起源可以追溯到约 1500 万年前,其两次全基因组复制(whole genome duplication, WGD)事件引发的基因家族扩张(尤其是生长调控通路)可能响应非洲草食动物的选择压力。阐明了 A' 和 B 亚基因组在功能上的分化,A' 亚基因组主要负责植物生长、发育和光合作用,而 B 亚基因组则主要负责物质运输和抗逆性,这一“功能模块化”分化为解析多倍体适应性优势提供了范式。Xing 等^[36]在巨菌草中的研究也进行了亚基因组的分析,发现巨菌草的 A 亚基因组与美洲狼尾草关系更近,而 B 亚基因组与狗尾草(*Setaria viridis*)关系更近,表明两个亚基因组在进化上存在差异。对巨菌草的 6 个组织进行转录组测序,分析基因表达模式,转录组分析显示两个亚基因组在组织特异性表达中呈现显著分化。但 C₄ 光合作用相关基因在两个亚基因组中高度保守且表达水平相似,表明两个亚基因组对巨菌草的高效

C₄光合作用贡献几乎相等。揭示了巨菌草的两个亚基因组在功能上的分化情况,为理解多倍体植物的高生物量积累提供了新见解^[36]。

值得关注的是,Teng等^[35]基于狼尾草基因组数据,通过比较基因组学方法,分析了狼尾草与其他物种的基因家族扩展和收缩,揭示了其特有基因家族的功能富集。分析结果将关键 WGD 事件定位于约 69.67 百万与 8.20 百万年前,其 615 个特有基因家族在 RNA 聚合酶通路中的特异性富集,可能驱动了该物种独特的生态适应性演化。

3.2 近缘物种比较

跨物种比较基因组学为解析狼尾草属植物的进化历程提供了重要依据。Yan等^[32]将象草基因组与其他 9 种植物进行基因组学比较分析,研究了象草的基因家族扩展和收缩以及全基因组复制事件。Teng等^[35]发现狼尾草与狗尾草属(*Setaria*)的分化时间为 7.53 百万~13.80 百万年前,蒺藜草属(*Cenchrus*)的分化时间为 5.33 百万~8.99 百万年前,狼尾草的基因组数据为理解禾本科植物的进化历史提供了新的视角。Zheng等^[37]通过与其他黍亚科植物的比较,重建了巨菌草的进化历程,并提出了其祖先核型的模型,推测其两个亚基因组可能独立起源于同一个祖先核型。巨菌草的基因组中发现了 3 种中心粒重复序列(*CEN137*、*CEN148*和*CEN156*),它们在两个亚基因组中独立进化,并可能通过基因流在亚基因组之间传播。

4 狼尾草功能基因组学的机制解析

狼尾草属功能基因组研究聚焦于生物量形成、抗逆机制与特殊代谢三大领域,通过多组学整合揭示其关键农艺性状的分子机制。

4.1 生物量形成机制

Yan等^[32]通过比较基因组学和表达分析,揭示了象草快速生长的分子机制,特别是 C₄光合作用和植物激素信号转导途径的基因扩展和高表达。通过分析两个亚基因组之间的表达差异,揭示了象草中亚基因组的互补功能,对于理解多倍体植物的适应性和进化具有重要意义。Zheng等^[37]研究发现巨菌草具有典型的 C₄植物 Kranz 解剖结构,其光合作用效率极高,最大光合速率达到 40 mmol·m⁻²·s⁻¹。巨菌草的 C₄光合作用主要依赖于烟酰胺腺嘌呤二核苷酸苹果酸酶(NADP-Malic enzyme, NADP-ME)和磷酸烯醇式丙酮酸羧化酶(phosphoenolpyruvate carboxykinase, PEPC)两种脱羧酶的协同作用,这与其他 C₄植物的单一脱羧酶途径不同,可能是其高光合效率的关键因素。

象草中纤维素合成相关基因的扩张和表达模式对于其作为能源物种的研究至关重要,Zhang等^[33]通过基因组分析得出象草中与纤维素合成相关的基因家族显著扩张,并且在茎中高表达,这可能是其高纤维素含量的原因,纤维素合成机制的分析为生物燃料开发提供了重要的基因资源。

4.2 抗逆机制解析

Varshney等^[20]研究发现,与蜡质合成和次生代谢物转运相关的基因家族在美洲狼尾草中显著扩张,这可能与耐热和耐旱能力有关,为耐逆境作物的改良提供了新的靶点。Yan等^[21]通过不同品种美洲狼尾草的基因组,构建了基于图谱的泛基因组,识别了多个与耐热性相关的 SVs,并结合美洲狼尾草在高温胁迫下的转录组数据进行分析,识别了差异表达基因(differentially expressed genes, DEGs),发现 RWP-RK 转录因子家族在美洲狼尾草中显著扩张,并可能通过调控内质网(endoplasmic reticulum, ER)相关基因参与耐热性响应。通过转基因试验和双荧光素酶报告基因试验,验证了 RWP-RK 转录因子的功能。结合基因组学、转录组学和功能验证试验,多组学整合分析全面解析了美洲狼尾草耐热性的遗传基础,提出了其耐热的分子机制,即通过 RWP-RK 转录因子与 ER 相关基因的协同调控,快速响应高温胁迫,维持细胞内蛋白质的稳态。

Kuijter等^[30]解析了美洲狼尾草对寄生杂草独脚金(*Striga hermonthica*)抗性的遗传机制,聚焦于植物激素独脚金内酯(strigolactones, SL)的合成调控基因,首次揭示了美洲狼尾草独脚金内酯合成路径的遗传调控机制,填补

了C₄作物抗寄生杂草的分子基础空白,为非洲及亚洲干旱区美洲狼尾草抗独脚金杂草育种提供了直接靶点。Yan等^[32]通过基因组分析揭示了象草紫色叶片的形成机制,发现其关键基因家族(如*ANS*、*UFGT*)的扩展和高表达是花青素积累的主要原因。

5 狼尾草群体遗传学与分子育种应用

5.1 狼尾草群体遗传学

Varshney等^[20]对994个美洲狼尾草品系进行了重测序,分析其种群结构和遗传多样性,研究人员发现美洲狼尾草的野生种群具有较高的遗传多样性,而栽培种群的多样性较低,表明驯化过程中可能发生了瓶颈效应,并且结合考古证据发现美洲狼尾草的驯化可能起源于西非的中心地带。利用重测序数据建立基因组选择模型,用以预测杂交后代的表现,为杂交育种提供了新的工具。Burgarella等^[38]将美洲狼尾草的全基因组测序结果^[20]与考古学结合推断了美洲狼尾草的驯化起源。研究结果表明,美洲狼尾草的驯化起源于西撒哈拉地区的陶德尼盆地,并且其扩散时间可追溯至4892年前。野生种群与栽培种群之间的基因流增加了栽培种群的遗传多样性,并在西非和东非形成了多样性热点。此外,研究揭示了15个基因组区域的适应性渗入事件,这些事件可能促进了栽培品种的本地适应性。这项研究将遗传学和考古学证据相结合,为理解其作为非洲最古老的谷物之一的驯化历史提供了新的视角。

5.2 分子标记辅助育种

Paudel等^[31]构建了狼尾草的首个高密度遗传图谱,覆盖14个连锁群,平均标记间隔为0.73 cM,开发了5339个简单序列重复(simple sequence repeats, SSR)标记,为狼尾草的分子育种提供了重要的遗传工具。Muktar等^[39]通过DARtSeq平台对105个象草样本进行了基因分型测序(genotyping by sequencing, GBS),构建了高密度的全基因组标记,获得116190个SilicoDARt和85452个单核苷酸多态性(single nucleotide polymorphism, SNP)标记,包括980个高多态性SNP标记,填补了象草分子标记的资源空白。并且利用近缘物种美洲狼尾草的参考基因组进行标记定位及比较基因组分析,发现17% SilicoDARt和33% SNP标记成功定位至美洲狼尾草基因组,标记在染色体末端富集,鉴定出22个候选基因(如类受体激酶、热激蛋白),与生物量、抗旱性相关。研究还揭示了象草不同收集样本间的连续不平衡(linkage disequilibrium, LD)模式差异,全群体LD平均衰减距离为2.54 kbp,但ILRI种质库衰减较慢(15.69 kbp vs. EMBRAPA种质库10.24 kbp),反映其较少经历减数分裂,遗传重组受限。EMBRAPA精英系LD衰减最快(16.56 kbp),表明育种过程有效打破了连锁区块。STRUCTURE分析揭示了7个亚群,ILRI种质(野生及历史育种材料)与EMBRAPA种质(现代选育系)遗传分化显著(AMOVA组间变异为13.28%~14.42%)。构建了核心种质库,筛选出14个代表性种质(覆盖7个亚群),遗传多样性指数($H_e=0.46$)高于全群体($H_e=0.20$),适用于田间抗逆表型评估及异交优势利用。Wang等^[40]通过全基因组测序开发出83706个高质量的基因组SSR标记。

2024年,Yan等^[41]通过分析242份美洲狼尾草种质的基因组数据,包括145万个SNP标记和12.4万个结构变异(structure variant, SV)标记,研究发现,SV标记能捕捉SNP无法检测的遗传多样性,尤其在连锁不平衡分析中,超过1/3的SV与附近SNP关联性较低,表明SV携带独特的遗传信息。研究测试了9种基因组预测模型(包括传统统计模型和人工智能算法)。结果显示,基于SV的模型预测准确性更高,其中基因组最佳线性无偏预测(genomic BLUP, gBLUP)模型表现最佳,预测精度接近99%,相比之下,SNP模型更适合贝叶斯方法。通过模型交叉验证,研究鉴定出8个基于SNP的高潜力品种(涉及穗数、粒重等性状),以及4个基于SV的候选品种(如“P12”在15个穗粒数相关性状中表现突出)。尤为重要的是,“P23”品种在SNP和SV模型中均被预测为穗数性状的优质资源,显示出其育种潜力。

5.3 分子育种平台

Sun等^[27]开发了一个数据全面、功能多样、以狼尾草为主的多组学数据库平台(Milletdb, <http://milletdb>).

novogene.com/)。Milletdb 包含谷子一类 7 个属的基因组和 1800 份不同的组学数据, Milletdb 包含来自 18 个基因组的 824198 个基因条目, 其中 18 个基因组中有 11 个美洲狼尾草基因组以及 2 个象草基因组, 涵盖 7184 条生物通路(其中 939 条与抗逆性相关), 以及基于图谱的泛基因组学数据, 包括 30050483 个单核苷酸变异(SNPs)和 424085 个结构变异(SVs)。Milletdb 是迄今为止最全面的狼尾草多组学数据库, 它包含并可视化了大量的多组学数据, 使得用户可以从单个基因的研究扩展到遗传网络的水平, 并开发了多种生物信息学工具, 用于基因鉴定、表达分析、基因网络构建等, 可有效助推狼尾草类的基因功能研究和分子育种进程。通过全基因组关联分析, 研究人员成功鉴定出产量和抗逆性相关的基因, *PMA1G03779.1* 基因与生长指数(growth index, GI)相关, 该基因在不同环境胁迫下均表现出高表达。以 *PLATZ* 转录因子家族为例, 通过 Milletdb 工具分析发现, 该家族在谷子中有显著扩张, 并在 87.5% 的谷子品种中参与植物生长和抗逆性响应。

6 结论与展望

6.1 结论

狼尾草属植物基因组学研究已实现从基础测序向功能解析与育种应用的跨越。技术突破方面, 长读长测序与三维基因组技术的结合, 实现了美洲狼尾草、象草等高杂合度多倍体物种的染色体级别组装(Contig N50 > 100 Mb), 为解析复杂基因组结构奠定了基础^[20,33]。在功能解析方面, 耐逆基因(如 *DREB2A*、*CLAMT1b*)和生物质合成通路(如纤维素 *CesA* 家族)的挖掘, 揭示了狼尾草属适应极端环境与高效碳固定的分子机制^[30,32]。育种应用中, 全基因组复制(WGD)事件与亚基因组功能分化的解析(如象草 A/B 亚基因组互补), 为多倍体优势利用提供了理论依据^[36]。此外, 群体遗传学研究通过重测序与基因组选择模型, 加速了耐旱、抗病等性状的标记开发, 推动分子育种从“经验驱动”转向“数据驱动”^[21,38]。

6.2 展望

尽管成果丰硕, 但是狼尾草属基因组研究仍面临多重挑战, 需要在以下方面重点突破。目前狼尾草属植物基因组研究的对象主要集中在美洲狼尾草、象草以及巨菌草^[20,32,37]。其他种的研究较少, 其中 80% 的物种尚未开展基因组研究, 尤其是具有独特生态适应性与应用潜力的野生种及地方品种, 难以全面解析狼尾草属植物的适应性进化机制^[35]。未来需通过技术革新与资源整合, 实现从模式物种主导到全属覆盖的转变, 指导优异种质的鉴定与利用, 释放狼尾草属植物在牧草饲料、生态修复、生物能源与观赏利用方面的多维价值。

通过比较不同生态型狼尾草的基因组, 可解析 C_4 光合、耐逆等性状的趋同进化机制。拓展狼尾草属基因组的研究对象, 不仅有助于揭示其全属遗传多样性规律, 更能为分子设计育种提供基础。不同研究中产生的狼尾草组学数据缺乏统一标准, 难以跨物种比较。建立覆盖狼尾草属 80% 以上物种的泛基因组以及多组学数据库将加速狼尾草关键性状基因的挖掘和分子育种进展。

Hi-C 与单倍型分型技术已部分实现染色体水平组装, 但亚基因组间的高度同源性(>95%)导致同源染色体错误锚定率仍在 12%~18%。针对错误锚定问题, 亟须开发和应用更有效的策略, 将同源染色体区分错误率降低^[33,36], 可以通过开发单细胞 Hi-C 技术(scHi-C)分辨单倍型, 以及整合表观标记辅助组装来降低错误率。三维基因组技术的发展为植物基因组结构的研究提供了新的视角, 植物染色质在多个尺度上具有层次性组织, RNA 与染色质的相互作用以及液-液相分离现象在染色质动态变化中扮演了关键角色^[42]。在水稻中首次建立了植物单细胞 Hi-C 体系, 揭示水稻精细胞、卵细胞及合子的染色质三维结构动态, 阐明了亲本基因组空间互作驱动合子基因组激活的机制, 为多倍体植物杂种优势研究提供范式^[43]。狼尾草多倍体特有的亚基因组互作尚未在三维染色质构象层面系统解析, 未来研究可以通过进一步整合三维基因组与表观多组学数据, 揭示其调控网络, 为分子育种提供新的理论依据^[44]。未来狼尾草的研究需要在多个方面深入展开, 包括结合智能技术育种、单倍体育种以及转基因体系和基因编辑体系的构建, 由传统育种方式转向设计育种^[23]。

参考文献 References:

- [1] Wu J Y, Teng W J, Wang Q H. Basic botanic characters, adaptabilities and applying in landscape architecture of *Pennisetum alopecuroides*. Chinese Landscape Architecture, 2005, 21(12): 57–59.
武菊英, 滕文军, 王庆海. 狼尾草的生物学特性及在园林中的应用. 中国园林, 2005, 21(12): 57–59.
- [2] Hou X C, Teng K, Guo Q, *et al.* Research advances in forage *Pennisetum* resource. Chinese Bulletin of Botany, 2022, 57(6): 814–825.
侯新村, 滕珂, 郭强, 等. 狼尾草属牧草研究进展. 植物学报, 2022, 57(6): 814–825.
- [3] Chen C, Yuan X H, Teng W J, *et al.* Biological characteristics, ecological adaptability, ornamentality, and invasion risk of *Pennisetum* species. Chinese Journal of Ecology, 2017, 36(2): 374–381.
陈超, 袁小环, 滕文军, 等. 狼尾草属植物生物学特性、生态适应性、观赏性和入侵风险关系的探讨. 生态学杂志, 2017, 36(2): 374–381.
- [4] Dudhate A, Shinde H, Tsugama D, *et al.* Transcriptomic analysis reveals the differentially expressed genes and pathways involved in drought tolerance in pearl millet [*Pennisetum glaucum* (L.) R. Br]. PLoS One, 2018, 13(4): e0195908.
- [5] Vadez V, Hash T, Bidinger F R, *et al.* Phenotyping pearl millet for adaptation to drought. Frontiers in Physiology, 2012, 3: 386.
- [6] Kebede G, Feyissa F, Assefa G, *et al.* Agronomic performance, dry matter yield stability and herbage quality of Napier grass [*Pennisetum purpureum* (L.) Schumach] accessions in different agro-ecological zones of Ethiopia. Journal of Agricultural & Crop Research, 2017, 4(5): 49–65.
- [7] Daud Z, Mohd H M Z, Mohd K A S, *et al.* Analysis of napier grass (*Pennisetum purpureum*) as a potential alternative fibre in paper industry. Materials Research Innovations, 2014, 18(6): 18–20.
- [8] Eliana C, Jorge R, Juan P, *et al.* Effects of the pretreatment method on enzymatic hydrolysis and ethanol fermentability of the cellulosic fraction from elephant grass. Fuel, 2014, 118: 41–47.
- [9] Kumar S, Saxena S, Rai A, *et al.* Ecological, genetic, and reproductive features of *Cenchrus* species indicate evolutionary superiority of apomixis under environmental stresses. Ecological Indicators, 2019, 105: 126–136.
- [10] Guo Y D, Liu L Y, Yue Y S, *et al.* Development of SSR markers based on transcriptome sequencing and verification of their conservation across species of ornamental *Pennisetum* Rich. (Poaceae). Agronomy, 2022, 12(7): 1683.
- [11] Xu J, Liu C, Song Y, *et al.* Comparative analysis of the chloroplast genome for four *Pennisetum* species: Molecular structure and phylogenetic relationships. Frontiers in Genetics, 2021, 12: 687844.
- [12] Lin Z X, Lin D M, Liu Z J, *et al.* *Cenchrus fungigraminus* Z. X. Lin & D. M. Lin & S. R. Lan sp. nov., a new species of Panicoideae (Poaceae): Evidence from morphological, nuclear and plastid genome data. Journal of Forest and Environment, 2022, 42(5): 514–520.
林占熿, 林冬梅, 刘仲健, 等. 基于形态和基因组证据的禾本科新种——巨菌草. 森林与环境学报, 2022, 42(5): 514–520.
- [13] Sanger F, Coulson A R. A rapid method for determining sequences in DNA. Journal of Molecular Biology, 1975, 94(3): 441–448.
- [14] Ronaghi M, Uhlen M, Nyren P. A sequencing method based on real-time pyrophosphate. Science, 1998, 5375(281): 363–365.
- [15] Eid J, Fehr A, Gray J, *et al.* Real-time DNA sequencing from single polymerase molecules. Science, 2009, 5910(323): 133–138.
- [16] Tang D, Zhou Q. Research advances in plant genome assembly. Biotechnology Bulletin, 2021, 37(6): 1–12.
唐蝶, 周倩. 植物基因组组装技术研究进展. 生物技术通报, 2021, 37(6): 1–12.
- [17] Pendleton M, Sebra R, Pang A W C, *et al.* Assembly and diploid architecture of an individual human genome via single-molecule technologies. Nature Methods, 2015, 12(8): 780–786.
- [18] Dudchenko O, Batra S S, Omer A D, *et al.* *De novo* assembly of the *Aedes aegypti* genome using Hi-C yields chromosome-length scaffolds. Science, 2017, 356(6333): 92–95.
- [19] Hu Y W, Chen L. Research advance and applications in maize wild relatives genomes. Biotechnology Bulletin, 2024, 40(3): 14–24.

- 胡伊娃, 陈露. 玉米野生种基因组研究进展及应用. 生物技术通报, 2024, 40(3): 14–24.
- [20] Varshney R K, Shi C, Thudi M, *et al.* Pearl millet genome sequence provides a resource to improve agronomic traits in arid environments. *Nature Biotechnology*, 2017, 35(10): 969–976.
- [21] Yan H D, Sun M, Zhang Z R, *et al.* Pangenomic analysis identifies structural variation associated with heat tolerance in pearl millet. *Nature Genetics*, 2023, 55(3): 507–518.
- [22] Tang J, Zhou H L, Wang W Q, *et al.* Advances in breeding and molecular biology research of *Pennisetum*. *Chinese Journal of Tropical Crops*, 2018, 39(11): 2313–2320.
唐军, 周汉林, 王文强, 等. 狼尾草属牧草育种及分子生物学研究进展. 热带作物学报, 2018, 39(11): 2313–2320.
- [23] Jing H C, Hu W J, Jin J B, *et al.* Accelerate innovation of forage intelligent breeding technology: Reflection and suggestions. *Bulletin of Chinese Academy of Sciences*, 2025, 40(2): 310–319.
景海春, 胡伟娟, 金京波, 等. 加快饲草智能育种科技创新的思考与建议. 中国科学院院刊, 2025, 40(2): 310–319.
- [24] Song S H, Tian D M, Zhang Z, *et al.* Rice genomics: Over the past two decades and into the future. *Genomics, Proteomics & Bioinformatics*, 2018, 16(6): 397–404.
- [25] Schnable J C. Genome evolution in maize: From genomes back to genes. *Annual Review Plant Biology*, 2015, 66(1): 329–343.
- [26] Wang X, Xu Y Y, Xu Y, *et al.* Research progress in genomic selection breeding technology for crops. *Biotechnology Bulletin*, 2024, 40(3): 1–13.
王欣, 徐一亿, 徐扬, 等. 作物全基因组选择育种技术研究进展. 生物技术通报, 2024, 40(3): 1–13.
- [27] Sun M, Yan H D, Zhang A L, *et al.* Milletdb: A multi-omics database to accelerate the research of functional genomics and molecular breeding of millets. *Plant Biotechnology Journal*, 2023, 21(11): 2348–2357.
- [28] Salson M, Orjuela J, Mariac C, *et al.* An improved assembly of the pearl millet reference genome using oxford nanopore long reads and optical mapping. *G3-Genes, Genomes, Genetics*, 2023, 13(5): jkad051.
- [29] Ramu P, Srivastava R K, Sanyal A, *et al.* Improved pearl millet genomes representing the global heterotic pool offer a framework for molecular breeding applications. *Communications Biology*, 2023, 6(1): 902.
- [30] Kuijjer H N J, Wang J Y, Bougouffa S, *et al.* Chromosome-scale pearl millet genomes reveal *CLAMT1b* as key determinant of strigolactone pattern and striga susceptibility. *Nature Communications*, 2024, 15(1): 1–12.
- [31] Paudel D, Kannan B, Yang X P, *et al.* Surveying the genome and constructing a high-density genetic map of napiergrass (*Cenchrus purpureus* Schumach). *Scientific Reports*, 2018, 8(1): 14419.
- [32] Yan Q, Wu F, Xu P, *et al.* The elephant grass (*Cenchrus purpureus*) genome provides insights into anthocyanidin accumulation and fast growth. *Molecular Ecology Resources*, 2021, 21(2): 526–542.
- [33] Zhang S K, Xia Z Q, Li C, *et al.* Chromosome-scale genome assembly provides insights into speciation of allotetraploid and massive biomass accumulation of elephant grass (*Pennisetum purpureum* Schum.). *Molecular Ecology Resources*, 2022, 22(6): 2363–2378.
- [34] Zhang Y, Yuan X H, Teng W J, *et al.* Karyotype diversity analysis and nuclear genome size estimation for *Pennisetum* Rich. (Poaceae) ornamental grasses reveal genetic relationship and chromosomal evolution. *Scientia Horticulturae*, 2015, 193: 22–31.
- [35] Teng K, Guo Q, Liu L Y, *et al.* Chromosome-level reference genome assembly provides insights into the evolution of *Pennisetum alopecuroides*. *Frontiers in Plant Science*, 2023, 14: 1195479.
- [36] Xing L S, Wang M J, He Q, *et al.* Differential subgenome expression underlies biomass accumulation in allotetraploid *Pennisetum giganteum*. *BMC Biology*, 2023, 21(1): 161.
- [37] Zheng H K, Wang B Y, Hua X T, *et al.* A near-complete genome assembly of the allotetraployploid *Cenchrus fungigraminus* (jujucao) provides insights into its evolution and C₄ photosynthesis. *Plant Communications*, 2023, 4(5): 100633.
- [38] Burgarella C, Cubry P, Kane N A, *et al.* A western Sahara centre of domestication inferred from pearl millet genomes. *Nature Ecology & Evolution*, 2018, 2(9): 1377–1380.
- [39] Muktar M S, Teshome A, Hanson J, *et al.* Genotyping by sequencing provides new insights into the diversity of napier grass (*Cenchrus purpureus*) and reveals variation in genome-wide LD patterns between collections. *Scientific Reports*, 2019, 9(1): 6936.

- [40] Wang C R, Yan H D, Li J, *et al.* Genome survey sequencing of purple elephant grass (*Pennisetum purpureum* Schum 'Zise') and identification of its SSR markers. *Molecular Breeding*, 2018, 38(7): 795–800.
- [41] Yan H D, Jin Y R, Yu H P, *et al.* Genomic selection for agronomical phenotypes using genome-wide SNPs and SVs in pearl millet. *Theoretical and Applied Genetics*, 2024, 137(10): 185–192.
- [42] Ouyang W Z, Xiong D, Li G L, *et al.* Unraveling the 3D genome architecture in plants: present and future. *Molecular Plant*, 2020, 13(12): 1676–1693.
- [43] Zhou S L, Jiang W, Zhao Y, *et al.* Single-cell three-dimensional genome structures of rice gametes and unicellular zygotes. *Nature Plants*, 2019, 5(8): 795–800.
- [44] Falk M, Feodorova Y, Naumova N, *et al.* Heterochromatin drives compartmentalization of inverted and conventional nuclei. *Nature*, 2019, 570(7761): 395–399.