

高通量测序技术在植物内生菌领域研究中的应用

庞煜¹, 马达², 王波², 蔡燕雪², 王际辉², 肖珊^{1,2*}

1 大连工业大学 生物工程学院, 辽宁 大连 116034

2 东莞理工学院 生命健康技术学院, 广东 东莞 523808

庞煜, 马达, 王波, 蔡燕雪, 王际辉, 肖珊. 高通量测序技术在植物内生菌领域研究中的应用[J]. 生物工程学报, 2024, 40(10): 3395-3406.

PANG Yu, MA Da, WANG Bo, CAI Yanxue, WANG Jihui, XIAO Shan. Application of high-throughput sequencing in research on plant endophytes[J]. Chinese Journal of Biotechnology, 2024, 40(10): 3395-3406.

摘要: 植物内生菌是一类能够在植物体内度过部分生命周期而不会引起宿主病害的微生物, 是种类和功能十分丰富的微生物资源。随着测序技术的进步, 在微生物领域中关于植物内生菌的研究愈发深入。Sanger 测序的定向验证性以及较低的测序成本使得该测序技术沿用至今, 但其测序的通量较低, 不适用于更深层次的内生菌基因组研究。本文简要概述了植物内生菌的研究历程, 并综述了以高通量为技术特点的下一代测序(next-generation sequencing, NGS)和以单分子实时测序技术(single-molecule real-time, SMRT)为代表的第三代测序技术在植物内生菌研究领域的应用, 总结了不同测序技术在实际研究中的优势和局限性, 并探讨了如何根据研究需求选择合适的测序技术, 希望为研究者进一步发掘植物内生菌的潜在价值提供参考。

关键词: 植物内生菌; Sanger 测序; 高通量测序; 单分子实时测序技术

Application of high-throughput sequencing in research on plant endophytes

PANG Yu¹, MA Da², WANG Bo², CAI Yanxue², WANG Jihui², XIAO Shan^{1,2*}

1 School of Biological Engineering, Dalian Polytechnic University, Dalian 116034, Liaoning, China

2 School of Life and Health Technology, Dongguan University of Technology, Dongguan 523808, Guangdong, China

Abstract: Plant endophytes spend at least part of their life cycle in plants without causing diseases in the hosts, being the microbial resources with rich species and diverse functions. With

资助项目: 东莞理工学院高层次人才科研启动项目(GC300501-139)

This work was supported by the Dongguan University of Technology High-level Talent Research Initiation Project (GC300501-139).

*Corresponding author. E-mail: xiaoshan@dgut.edu.cn

Received: 2023-12-18; Accepted: 2024-03-25

the advancement in sequencing technology, the microbiological study of endophytes has become increasingly intensive. Being praised for the targeted validation and low cost, Sanger sequencing has been preferred by researchers. However, Sanger sequencing is no longer suitable for deeper genomic study of endophytes due to the low throughput. In this paper, we briefly summarize the research history of endophytes, review the applications of next-generation sequencing characterized by high throughput and third-generation sequencing (single-molecule real-time sequencing) in the research on endophytes, and summarize the research results of different sequencing technologies. Furthermore, we summarize the advantages and limitations of different sequencing technologies and discuss how to choose the appropriate sequencing technology according to the research needs. This review provides a reference for researchers to further explore the potential value of plant endophytes.

Keywords: plant endophytes; Sanger sequencing; high-throughput sequencing; single-molecule real-time sequencing

植物内生菌(endophyte)是指在植物体内度过部分或全部生命周期且不会对宿主产生病害的一类微生物,主要包括细菌、真菌和古细菌。植物内生菌研究初期(19 世纪初),研究者们无法准确鉴定植物内微生物种类,只能依靠形态学方法对微生物进行分类,这限制了研究者进一步对内生菌生理功能及其和宿主之间共生关系的探索,使得针对植物内生菌的研究停留于表面。

1977 年,第一代 DNA 测序技术 Sanger 测序和化学测序技术的问世标志着人类掌握了解析内生菌基因组信息的工具,并在近 40 年的时间里为植物内生菌相关的研究积累了可靠的植物内生菌序列信息,同时为植物内生菌资源和物种信息数据库增添了资料。随着研究的不断深入, Sanger 测序的低通量不再能够支持解析大规模的基因组数据^[1]。近年来高通量测序技术不断取得突破,下一代测序(next-generation sequencing, NGS)技术以及第三代测序技术使得研究者能够更加深入全面地探索植物内生菌领域,新的植物内生菌种类、菌群功能和全基因组信息得到了扩充,更进一步揭示了植物与其内生微生物的相互关系。

植物内生菌广泛存在于各种植物中,开发

植物内生菌资源将是未来微生物领域研究的热门方向。本文简要回顾了植物内生菌的研究历程,综述了近年来高通量测序技术在内生菌研究中的具体应用,对目前研究中存在的问题和难点进行了分析总结并进行了展望,旨在为相关内生菌研究者提供参考,进一步推动植物内生菌领域的发展。

1 植物内生菌的主要研究进程

植物内部存在着不同的微生物群落,包括细菌、真菌和古细菌。内生菌最早的概念来自于 de Bary,他在 1866 年将栖息于植物内部的致病微生物称为内生菌^[2]。直到 1887 年,Compant 等^[3]提出了蔬菜植物内部存在真菌和细菌的理论,并认为这些微生物来源于土壤且对植物本身有益。1888 年,Beijerinck 发现从豆科植物根瘤中分离出的细菌可以固定大气中的氮,首次揭示了内生菌对宿主植物的有益反馈^[4]。但受限于当时的微生物鉴定平台,并不能确定此种微生物的种类,直到 1977 年 Sanger^[5]测序的发明,才确定了在豆科植物根瘤中起固氮作用的微生物为豌豆根瘤菌(*Rhizobium leguminosarum*)。

近代对植物内生菌的系统研究始于 20 世纪 60 年代至 70 年代。Bacon 和 Strobel 对寄生在植物内部组织中的真菌开展研究^[6]；1993 年 Stierle 等^[6]从短叶红豆杉(*Taxusbrevifolia* Nutt.) 中分离出能产生抗癌活性成分紫杉醇的内生菌，这为发掘植物内生菌在药物医疗领域的潜在价值奠定了基础。

20 世纪 90 年代之后，由于认识到了植物内生菌在农业和医药领域的潜在价值，研究人员针对具有生理活性的内生菌开展了大量筛选工作，如产抗生素^[7]和抗癌活性成分^[8]的植物内生菌，自此迎来了植物内生菌的研究热潮。

到 21 世纪初，得益于 DNA 测序技术和分子生物信息学的发展，植物内生菌研究领域的重点转移到了基因组学和分子研究，研究人员能够更深入地了解内生菌的多样性和功能性。表 1 对植物内生菌的历史主要研究历程进行了总结。

2 高通量测序技术在植物内生菌领域的应用

Sanger 测序技术自发明以来得到了广泛应用，但随着植物内生菌研究领域的不断扩大，其测序通量和效率低下的缺点被暴露出来。过去

20 年里，在对高通量测序和高效率测序的强烈需求下，高通量测序技术迅速发展并日渐成熟，包括下一代测序技术以及第三代测序技术，其测序的高通量以及超强的解读基因组信息能力使其成为在植物内生菌研究领域应用最广泛的测序技术。

2.1 基于 NGS 的植物内生菌群落时空多样性研究

NGS 技术揭示了关于寄生在植物组织中的各种微生物的丰富信息，现已广泛地应用于植物内生菌时空多样性研究。植物在不同的生长阶段具有不同的内生菌多样性，一项关于茶树内生菌群落随时间动态变化的研究发现老叶比新叶具有更丰富的内生真菌多样性，推测可能与植物自身生产发育周期所需的营养物质和抗病抗虫能力有关^[10]。另一项针对白花前胡(*Peucedanum praeruptorum* Dunn)抽薹前、抽薹后及开花期内生菌的多样性与差异性研究发现，开花期的内生细菌多样性指数显著高于抽薹前，而内生真菌的多样性指数无显著差异；在属水平上，假单胞菌属(*Pseudomonas*)和伯克霍尔德菌(*Burkholderia*)为主要优势细菌属，表明了抽薹可能是影响白花前胡内生菌多样性与差异性的因素^[11]。

表 1 植物内生菌研究进程

Table 1 Processes of research on plant endophytes

Crucial moment	Research priorities	Research representatives	Research methods	Reference
1809–1887	Discovery of plant endophytes	Link, Galippe	Microbial <i>in vitro</i> culture	[3]
19 centuries	Plant growth-promoting effects of endophytes	Beijerinck	Microbial <i>in vivo</i> and <i>in vitro</i> culture	[4]
1970s	Medicinal value of metabolites of plant endophytes	Strobel	First-generation sequencing technology	[6]
1990s	Bioexploitation of plant endophytes	Strobel	High-throughput sequencing technology	[7-8]
Early 21st century	Genomics and molecular studies of plant endophytes	Sessitsch	Third-generation sequencing technology	[9]

此外,植物内生菌可以帮助宿主植物在环境条件变化时作出适应性反应。植物组织的表面消毒是植物内生菌研究中十分重要的一个环节,不彻底的表面消毒无法排除植物样本表面微生物的干扰,Yu等^[12]聚焦于植物内生菌筛选的样品前处理过程,探讨了次氯酸钠(NaClO)表面消毒处理对茶树叶片和茎内生细菌和真菌多样性变化的影响,研究发现,NaClO浓度和预处理时间对细菌内生菌的多样性变化有显著影响,对内生真菌的多样性变化影响不大,并给出了一种针对菌种的样品表面消毒的最佳方案。除了外界的人为因素,植物与动物之间的自然生态关系也会对内生菌的多样性产生影响,近期针对两种苔原草(*Anthoxanthum nipponicum*和*Calamagrostis phragmitoides*)与食草动物的相互作用的研究揭示了放牧对苔原草中真菌内生菌多样性的影响,发现当苔原草暴露于食草动物时具有更高的内生真菌多样性,即适当地放牧可以提高苔原草内生真菌多样性^[13]。

NGS可以在不同时间和空间尺度上产生大量内生菌序列数据。这使得研究人员可以深入解析内生菌组成随宿主发育时期和组织类型的变化规律,揭示微生物组成与宿主生长发育和适应性的内在联系,为功能微生物的筛选提供理论指导。此外,NGS为深入解析植物内生菌的时空动态变化规律和环境响应奠定了基础,使功能型和宿主适应性内生菌的发现与应用成为可能。

2.2 基于 NGS 的植物内生菌群落功能特性研究

使用下一代测序技术可以全面分析植物内生菌基因功能,探索内生菌对宿主植物的正向反馈作用,主要表现在促进宿主植物生长方面。如Nivedita等^[14]采用高通量RNA-Seq方法研究了一种植物根系内生菌印度梨形孢

(*Piriformospora indica*)对高盐胁迫的转录反应,在对照样本和处理样本之间共检索到661个差异表达基因(differentially expressed genes, DEGs)、基因本体(gene ontology, GO)和真核直系同源群(eukaryotic orthologous homologous group, KOG);富集分析结果表明,DEGs特异性参与了对盐胁迫的反应,此外,参与细胞壁完整性、甾醇生物合成和氧化应激的单基因如Rho型GTP酶、羟甲基戊二酰辅酶A合酶和硫氧还蛋白过氧化物酶在盐胁迫下的水稻中上调。此研究通过盐响应DEG表明该内生菌可能在高盐胁迫的环境下对植物生长发育进行调节。

近年来使用NGS测序技术研究植物内生菌群落功能特性的方法已在多种农作物上应用,如水稻(*Oryza sativa* L.)和苹果等常见农作物。研究发现在水稻种子内生菌垂直传播的过程中,从亲本和后代种子中分离出的泛菌属(*Pantoea*)和*Xanthomanas*属内生菌基因组序列显示出高度的平均核苷酸和核心蛋白同一性,其基因组内含有长度短、代谢复杂度低的流线性基因组,这些基因组具有多种促进植物生长的功能^[15]。此外,有研究人员在对不同品种苹果树内生细菌群落多样性及功能性研究中发现,新疆维吾尔自治区本地品种微生物群落的功能信息多于吉尔吉斯斯坦品种,主要体现在化能异养、需氧化能异养、尿素分解、磷酸盐解毒和异化砷还原等功能方面^[16]。程欢等^[17]则对用于嫁接苹果的砧木T337中不同组织内生菌群落结构进行了分析并进行功能预测,经过16S rRNA和内部转录间隔区(internally transcribed spacer, ITS)测序以及通过重建未观测状态对群落进行系统发育研究(phylogenetic investigation of communities by reconstruction of unobserved states, PICRUSt)功能预测,从内生细菌和真菌中分别预测到417个和73个代谢通路,内生菌群

的预测功能分别属于生物合成, 降解、利用、同化, 前体代谢产物和能量产生, 聚糖途径以及其他代谢共 5 个方面。

NGS 使得鉴定内生菌响应环境胁迫的关键基因和代谢途径成为可能, 如抵抗高盐胁迫相关基因的差异表达, 这有助于阐明内生菌减轻植物受高盐胁迫的分子机制。此外, NGS 揭示了一些内生菌具有流线型基因组, 即具有较低 G+C 含量、低非编码 DNA 百分比以及编码细胞质膜蛋白、周质蛋白、相关转录蛋白和信号转导通路基因的基因组^[18], 拥有较小基因组的生物体能够缩小细胞的体积, 从而增加表面积与体积比, 并增加养分的吸收^[19]。拥有更小基因组的优点还包括细胞分裂时基因组复制速度更快和更容易共调控多个相关基因, 且具有流线型基因组的内生菌一般具有更高的适应性^[20]。NGS 还可以通过比较不同植物内生菌群落的功能预测, 解析微生物组成与宿主适应性的关系。NGS 为深入挖掘和利用植物内生菌的功能提供了高效途径。充分运用 NGS 技术有望加速功能型内生菌的发现, 实现其在改善植物生长环境和提高农作物产量等方面的应用。

2.3 基于 NGS 的植物内生菌生物活性物质研究

许多植物内生菌在其生长过程中往往能产生具有各种生物活性的物质, 这是内生菌成为研究热点的重要原因。NGS 拓宽了研究人员发现由植物内生菌产生的新型生物活性化合物的途径。2012 年一项关于印楝(*Azadirachta indica* A. Juss.) 内生真菌分离和表征的研究表明, 通过 ITS 分析鉴定出的内生真菌正青霉属(*Eupenicillium*), 在其培养过程中可以产生天然杀虫剂印楝素及其结构类似物^[21]。孙碧琪等^[22]通过高通量测序技术筛选了 4 种诺丽果(*Morinda citrifolia* L.) 的内生菌并探讨了其生物活性物质, 通过抑菌实

验和抗氧化实验发现具有抑菌性的诺丽果内生细菌均表现出抗氧化性, 其发酵液的总抗氧化值在 3.07–20.78 U/mL 之间, 并且内生细菌的抗氧化性和抑菌活性呈正相关性, 表明内生菌可以产生具有抑菌性和抗氧化性的活性物质, 且二者之间相互关联。在促植物生长活性成分研究领域, 黄小茜等^[23]采用生理生化方法测定发现了所分离菌株的固氮活性、溶磷特性、产生嗜铁素(siderophore)、分泌吲哚乙酸(indole-3-acetic acid, IAA)等促生特性, 该菌株可作为植物促生长剂使用。此外, 2021 年的一项研究首次利用宏基因组测序对乌拉尔甘草(*Glycyrrhiza uralensis* Fisch.) 内生菌群落结构与功能基因组进行了测定, 通过 COG 数据库、KEGG 数据库和 CAZy 数据库对其功能基因进行注释, 发现在甘草内生菌中含有 123 个细胞色素 P450 (cytochrome P450) 及 520 个 UDP-糖基转移酶(UDP-glycosyltransferase, UGT)编码基因, 为后续对甘草内生菌的生物学功能和挖掘其药用价值, 以及它们如何转变为甘草酸生物合成资源奠定了理论基础^[24]。

新一代测序技术的应用极大地促进了植物内生菌生物活性物质的探索研究。与传统方法相比, 新一代测序能够在更短的时间内产生更丰富的基因组数据, 从而提高对内生菌种类及其代谢潜能的检测效率。此外, 丰富的基因组信息不仅有助于预测编码次生代谢相关酶类的基因, 为新化合物的筛选和结构解析提供参考, 还可以揭示代谢途径和调控机制, 为后续的代谢工程奠定基础。因此, 新一代测序技术的应用极大地拓宽了植物内生菌次生代谢产物挖掘的广度与深度, 是发现高价值内生菌的有力工具。充分利用新一代测序的优势, 开展植物内生菌基因组测序与生物活性关联研究, 可以推动微生物资源的深入开发利用。

3 新一代高通量测序技术

NGS 技术相较于一代测序技术在通量、成本以及测序速率方面都取得了较大的突破,但由于 NGS 在测序过程中需要 PCR 扩增,导致该测序技术在对一些复杂基因进行测序时的准确率和读长都不及第一代测序技术。高通量测序技术依然在不断发展,如今已步入新一代高通量测序技术的时代,即第三代测序时代。三代测序技术解决了 NGS 读长短的问题。目前,比较成熟的第三代测序技术是 PacBio 公司的单分子实时测序技术(single-molecule real-time, SMRT)。

3.1 基于 SMRT 测序的植物内生菌基因组研究

SMRT 测序技术在解析植物内生菌基因组的复杂性方面起到了关键作用,它为研究人员提供了深入理解这些微生物的结构、功能、进化历程、定位以及基因编辑的可能性。许多研究都使用了 SMRT 测序对植物内生菌的基因组信息进行建库或扩充,主要集中于促生植物内生菌的全基因组结构组成研究,这些促生作用表现在固氮作用与抵抗外界胁迫上。近期,利用 PacBio 技术对寡干腐霉菌(*Pythium oligandrum*)进行长序列基因组测序的研究,构建了新的寡干腐霉菌基因组数据库^[25],可以用于对植物有益或致病、寄生于真菌和卵菌的腐霉菌进行基因组比较,从而确定其不同生活方式的关键遗传决定因素。同样使用 PacBio 技术, Li 等^[26]获得了一株麻风树(*Jatropha curcas* L.)茎组织的内生菌 L9-754T 的基因组信息,发现了一个与抗氧化酶相关的基因簇 GQR91_18 700-GQR91_18 715,揭示了该菌株在生物防治应用的潜在功能。此外,还有研究使用 SMRT 测序完整地测得了从广西壮族自治区种植的甘

蔗(*Saccharum officinarum* L.)中分离出的一种内生固氮细菌变异克雷伯氏菌(*Klebsiella variicola* DX120E)的基因组信息^[27]。在抵抗外界胁迫功能方面, Zhang 等^[28]从健康的白杨(*Populus davidiana*×*P. alba* var. *pyramidalis*, PdPap)木茎中分离出了可以抑制白杨木中病原菌生长的芽孢杆菌(*Bacillus velezensis*) BY6, SMRT 技术对菌株 BY6 进行的全基因组测序结果表明其中 9 个次生代谢物的基因簇具有抗真菌活性和促生功能,主要参与合成表面活性剂、细菌素、积累铁离子和相关抗生素。还有研究者使用 SMRT 测序技术阐明了 LK11 鞘氨醇单胞菌(*Sphingomonas* sp.)基因组的结构和组成^[29]。表 2 对 SMRT 测序在植物内生菌基因组研究中的应用进行了总结。

3.2 基于 SMRT 测序的 DNA 甲基化分析

DNA 甲基化是广泛存在于微生物中重要的基因表观修饰方式之一。在细菌中,可以调控基因表达及改变细胞循环周期,同时还可以增强基因的抗突变能力,并有助于维持基因组的稳定性。SMRT 测序技术可以很好地识别内生菌中的 DNA 甲基化,其基本原理是依据 DNA 聚合酶不同的作用时间而产生不同的脉冲间隔来判断甲基化位点。如 Bianco 等^[34]在探究非洲光叶水稻(*Oryza glaberrima* L.)的内生菌能否在亚洲水稻(*Oryza sativa* L.)中有效定殖并促进宿主生长的研究中,对菌株 BDA62-3、BDA134-6 和 BDA59-3 进行了 SMRT 测序,得出了完整的基因组序列,并检测到 BDA62-3 和 BDA59-3 的甲基化 DNA 基序主要为 m6A 修饰(6-甲基腺嘌呤),涉及经典的 GATC 基序,与固氮作用相关,得出了光叶水稻的内生菌能够相当有效地定殖驯化亚洲水稻的结论。此外 DNA 甲基化分析还可以帮助研究人员了解一些内生菌的致病性,如对从感染枯萎病的核桃(*Juglans regia* L.)枝中

表 2 SMRT 测序在促生内生菌基因组研究中的应用

Table 2 SMRT sequencing in the genome of endophyte-promoting bacteria

Host plant	Strain name	Genetic structure composition (genome size; GC content)	Functional activity	GenBank login ID	Reference
<i>Jatropha curcas</i> L.	<i>Sphingomonas carotini</i> <i>faciens</i> L9-754T	Chromosome: 4 094 655 bp; 66.90%	Enhanced antioxidant activity; anti-fungal diseases	WSUT00000000	[26]
<i>Saccharum officinarum</i> L.	<i>Klebsiella variicola</i> DX120E	Chromosome: 5 501 013 bp; 57.30% Plasmids 1: 162 706 bp; 50.70% Plasmids 2: 54 715 bp; 53.10%	Nitrogen fixation	CP009274; CP009275; CP009276	[27]
<i>Populus davidiana</i> × <i>P. alba</i> var. <i>pyramidalis</i> , PdPap	<i>Bacillus velezensis</i> BY6	Chromosome: 3 898 273 bp; 47.33% Plasmids: 7 256 bp; 37.53%	Anti-fungal activity; promotes iron ion absorption; nitrogen fixation	CP051011; CP051012	[28]
<i>Alhagi sparsifolia</i>	<i>Klebsiella</i> sp. LTGPAF-6F	Chromosome: 6 174 401 bp; 56.00% Plasmids 1: 526 446 bp; 49.00% Plasmids 2: 89 552 bp; 51.00%	Drought resistance	CP017450; CP017451; CP017452	[30]
<i>Zygophyllum simplex</i>	<i>Paenibacillus</i> sp. JZ16	Chromosome: 7 421 843 bp; 49.25%	Promotes the synthesis of phospholipids; plant degrader	CP017659	[31]
<i>Triticum aestivum</i> L.	<i>Bacillus</i> sp. WR11	Chromosome: 4 202 080 bp; 43.53%	Resistance to high salt stress; resistance to low phosphorus stress; promotion of plant colonization	CP033064	[32]
<i>Solanum lycopersicum</i> L.	<i>Rhodanobacter glycinis</i> T01E-68	Chromosome: 4 172 240 bp; 64.57%	Resistance to high salt stress; enhanced antioxidant activity; resistance to bacterial wilt	CP042807	[33]

分离出来的内生真菌 *Diaporthe capsici* 进行了基因组资源的整合, 并使用 SMRT 测序技术平台对 *Diaporthe capsici* 的基因组组装结果进行甲基化位点检测, 检测到的 m4C、m6A 和未知类型的修饰碱基分别为 9 057 646 个和 233 045 个修饰位点, *Diaporthe capsici* 的致病性在基因组层面得到了解释^[35]。

SMRT 测序技术可以进行 DNA 甲基化分析, 为深入解析植物内生菌的表观遗传学调控和宿主适应性进化机制提供了有力工具, 通过识别植物内生菌甲基化位点为解析其表观遗传

学调控机制奠定了基础, 这有助于深入挖掘其宿主适应性形成机制和功能分化过程, 为筛选适应性强的功能型植物内生菌提供新策略。表 3 对不同测序技术在植物内生菌领域的应用范围和测序特点进行了总结。

4 总结与展望

Sanger 测序作为第一代测序技术为早期内生菌研究奠定了基础, 但由于测序技术不断进步, 其使用范围逐渐缩小。NGS 通过提供大规模基因组数据, 显著扩展了研究人员对植物内

表3 植物内生菌领域不同测序技术平台的应用

Table 3 Comparison of applicability of different sequencing technology platforms

	Technical name/platform	Year of birth	Field of application	Advantages and limitations
First-generation sequencing technology	Sanger sequencing, chemical sequencing	1997	Species identification, characterization of new isolates of plant endophytes	High accuracy in small-scale sequencing with easy-to-understand data; short sequencing read lengths, low data throughput, and high sequencing costs;
Second-generation sequencing (high-throughput sequencing)	454 pyrophosphate sequencing platform, Solexa sequencing platform, HiSeq sequencing instrument, SOLid technology platform, Ion PGM sequencer	2005	Diversity and structural analysis of plant endophytes, genomic study of endophytes, study of biologically produced active substances of endophytes	High-throughput sequencing, also known as high-throughput sequencing, is fast and widely used; higher error rates in homopolymer regions, more difficult to analyze data, high cost of large-scale sequencing
Third-generation sequencing (high-throughput sequencing)	SMRT sequencing technology	2009	Plant endophyte gene assembly and characterization, growth-promoting mechanism studies, DNA methylation analysis	Longer sequencing reads, lower error rates relative to second-generation sequencing, and the ability to perform DNA methylation analysis; high cost of sequencing and increased difficulty in data processing

生菌与宿主植物相互作用的理解。SMRT 测序技术在解决复杂基因组和研究内生菌中的表观遗传修饰方面产生了深远的影响。NGS 和 SMRT 测序技术通过全面分析植物基因组、转录组和微生物群,极大地推动了该领域的革新。

在植物内生菌研究领域,测序方式的选择是对具体的研究目标、研究预算和可用资源的综合考量。Sanger 测序可能适用于有针对性的研究,NGS 在各种应用和大型项目中具有多样性,而 SMRT 测序可以有利于解决复杂的基因组区域和表观遗传学见解。针对鉴定特定基因或分析少量样本,Sanger 测序可以提供准确的结果,比如用于内生细菌鉴定的 16S rRNA 测序,本课题组对从菠萝(*Ananas comosus* L.)中筛选出的细菌进行了测序,鉴定为乳酸菌和酵母菌,并成功地将其作为一种复配发酵剂用于菠萝副产物的加工^[36],在植物内生菌的应用方面

提出了新的见解。此外 Sanger 测序还比较适合当研究预算有限且需要对少量样本进行特定基因标记的筛选。当研究涉及组装内生菌的完整基因组或转录组和内生菌群多样性分析时,NGS 是理想的选择,因此本课题组在涉及药用植物土沉香[*Aquilaria sinensis* (Lour.) Spreng.]中内生乳酸菌基因组学功能性研究中使用了 NGS 测序技术,试图寻找其与宿主植物相关联的基因组信息。NGS 非常适合宏基因组学和微生物标记物的引物测序,其测序的高通量和多样的应用使其成为大型项目的首选。SMRT 测序适用于研究复杂基因组区域和表观遗传学研究以及基因甲基化分析,在处理复杂的基因组区域时,例如重复序列或结构变异,SMRT 测序的长读长可以提供有价值的见解,包括宏基因组学和微生物组分析,特别适用于一些具有大型和复杂基因组的植物。此外,NGS 和 SMRT 测

序都会产生大量数据,需要生物信息学专业知
识进行分析。技术进步影响着测序方法的能力
和成本,了解最新的测序技术发展,并在不确
定如何处理特定研究目标时尝试综合使用多种
方法十分重要。

测序技术的进步推动着植物内生菌领域的
深入研究,近期迅速发展的第三代测序技术正
在掀起一场新的革命,特别是近期开始进入市
场的纳米孔测序技术(nanopore sequencing)。不
同于 SMRT 测序,纳米孔测序技术的测序读长
只受限于待测 DNA 分子长度,而非技术本身^[37],
超长的测序读长使内生菌领域的研究得到了
前所未有的进步,主要体现在宏基因组学研究
中的基因组装和表征方面。然而,由于植物内
生菌研究的重复性较差,高通量测序技术类型
较多,数据处理平台较为多样,且新兴的测序
技术如纳米孔测序技术并不像 SMRT 测序一
样对同一条链进行多次测序,导致测序错误率
仍相对较高,因此,高通量测序技术在植物内
生菌领域的未来研究方向可以集中在以下 5 个
方面。

(1) 未培养植物内生菌基因组解析。培养后
鉴定是目前研究植物内生菌的主要方式,由于
许多内生菌无法进行体外培养,致使这种研究
方法具有较大的局限性。来自植物内部未培养
和未鉴定的微生物可能具有新的遗传属性和基
因组信息,利用高通量测序对其进行解析可用
于农业和生物技术领域。未来可通过进一步探
索宏基因组学在未知植物内生菌领域的应用或
与其他同位素探测技术共同使用来实现^[38]。

(2) 复杂的植物内生菌网络。植物内生菌所
生存的环境是一个微生物与周围环境构成的复
杂网络,包括微生物与微生物以及微生物与环
境因素相互影响的因素,而这些影响每时每刻
都在发生。内生菌研究的另一大难点在于很难

完美复现植物内生菌所处的环境,所以,在实
验室所获得的测序数据与实际的植物内生菌种
类和基因组信息仍存在着差距。

(3) 非细菌 16S rRNA 的干扰。16S rRNA
通常被用于一个物种系统发育的研究,因为它
在不同种类的细菌和古细菌之间高度保守^[39]。
水平基因转移和转座因子理论诞生后,人们意
识到基因信息并不能完全准确地定性一个物
种,且这一理论已被证实^[40]。水平基因转移指
遗传物质在生物体之间发生移动的现象,而非
由亲本传向子代;转座因子是指基因组中可以
改变位置的核酸序列,水平基因转移现象会干
扰植物内生菌的鉴定。内生菌研究的困难之处
还在于如何排除非细菌 16S rRNA 的干扰,目
前针对该问题并没有较好的方法,是未来植物
内生菌领域的一个研究方向。

(4) 较高的学习成本。高通量测序技术的一
大特点就是数据产出量大且数据处理难度较
高,需要学习一定的生物信息学知识,有时还
需有计算机领域的基础,且目前每个测序平台
都有各自的数据处理方法,不像一般的生物学
实验数据一般简单易懂,这使得高通量测序的
使用成本一直居高不下。建立统一的高通量数
据处理平台能大大降低使用成本,进而使得高
通量测序技术成为更普遍的生物分析工具。

(5) 较高的测序错误率。目前,高通量测序
技术相比于一代测序技术在测序通量与测序读
长方面都取得了较大的突破,但由于测序数据
量较大,错误率也相对提升,这将导致在对某
些促生内生菌基因组信息进行解析时发生错
误,或在检测和鉴定其他有价值的基因组信息
上发生遗漏。针对测序错误率的问题可以从测
序机器的原理入手或开发一种新的数据处理方
法对测序数据进行修正,这需要较强的生物信
息学基础以及强大的计算机水平。

REFERENCES

- [1] KU CS, PAWITAN Y, WU M, ROUKOS DH, COOPER DN. The evolution of high-throughput sequencing technologies: from sanger to single-molecule sequencing[J]. *Next Generation Sequencing in Cancer Research: Volume 1: Decoding the Cancer Genome*, 2013: 1-30.
- [2] HARRISON JG, GRIFFIN EA. The diversity and distribution of endophytes across biomes, plant phylogeny and host tissues: how far have we come and where do we go from here?[J]. *Environmental microbiology*, 2020, 22(6): 2107-2123.
- [3] COMPANT S, SESSITSCH A, MATHIEU F. The 125th anniversary of the first postulation of the soil origin of endophytic bacteria—a tribute to M.L.V. Galippe[J]. *Plant and Soil*, 2012, 356(1): 299-301.
- [4] HARDOIM PR, van OVERBEEK LS, BERG G, PIRTILÄ AM, COMPANT S, CAMPISANO A, DÖRING M, SESSITSCH A. The hidden world within plants: ecological and evolutionary considerations for defining functioning of microbial endophytes[J]. *Microbiology and Molecular Biology Reviews: MMBR*, 2015, 79(3): 293-320.
- [5] SANGER F, NICKLEN S, COULSON AR. DNA sequencing with chain-terminating inhibitors[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 1977, 74(12): 5463-5467.
- [6] STIERLE A, STROBEL G, STIERLE D. Taxol and taxane production by *Taxomyces andreanae*, an endophytic fungus of Pacific yew[J]. *Science*, 1993, 260(5105): 214-216.
- [7] LI JY, HARPER JK, GRANT DM, TOMBE BO, BASHYAL B, HESS WM, STROBEL GA. Ambuic acid, a highly functionalized cyclohexenone with antifungal activity from *Pestalotiopsis* spp. and *Monochaetia* sp.[J]. *Phytochemistry*, 2001, 56(5): 463-468.
- [8] STROBEL G, LI JY, SUGAWARA F, KOSHINO H, HARPER J, HESS WM. Oocydin A, a chlorinated macrocyclic lactone with potent anti-oomycete activity from *Serratia marcescens*[J]. *Microbiology*, 1999, 145(Pt 12): 3557-3564.
- [9] SESSITSCH A, HARDARSON G, de VOS WM, WILSON KJ. Use of marker genes in competition studies of *Rhizobium*[M]//Hardarson G, Broughton WJ. *Molecular Microbial Ecology of the Soil*. Dordrecht: Springer, 1998: 35-45.
- [10] LIN HY, LIU CW, PENG Z, TAN B, WANG KB, LIU ZH. Distribution pattern of endophytic bacteria and fungi in tea plants[J]. *Frontiers in Microbiology*, 2022, 13: 872034.
- [11] 罗书岚, 刘丽, 李肖莉, 宋程, 贾彬, 欧金梅, 韩邦兴. 不同时期白花前胡内生菌的多样性和差异性[J]. *中国野生植物资源*, 2022, 41(10): 25-29.
- LUO SL, LIU L, LI XL, SONG C, JIA B, OU JM, HAN BX. Diversity and difference of endophytes in *Peucedanum praeruptorum* dunn at different stages[J]. *Chinese Wild Plant Resources*, 2022, 41(10): 25-29 (in Chinese).
- [12] YU YE, CHEN ZM, XIE HT, FENG XX, WANG YF, XU P. Overhauling the effect of surface sterilization on analysis of endophytes in tea plants[J]. *Frontiers in Plant Science*, 2022, 13: 849658.
- [13] BRÅTHEN KA, JAHIRI X, JUSDADO JGH, SOININEN EM, JENSEN JB. Fungal endophyte diversity in tundra grasses increases by grazing[J]. *Fungal Ecology*, 2015, 17: 41-51.
- [14] NIVEDITA, RAWOOF A, RAMCHIARY N, ABDIN MZ. A high-throughput RNA-Seq approach to elucidate the transcriptional response of *Piriformospora indica* to high salt stress[J]. *Scientific Reports*, 2021, 11: 4129.
- [15] ZHANG XX, MA YN, WANG X, LIAO KJ, HE SW, ZHAO X, GUO HB, ZHAO DF, WEI HL. Dynamics of rice microbiomes reveal core vertically transmitted seed endophytes[J]. *Microbiome*, 2022, 10(1): 216.
- [16] 古丽尼沙·沙依木, 张志东, 杨波, 章世奎, 唐琦勇, 宋素琴, 朱静, 郭春苗, 顾美. 不同品种苹果树内生细菌群落多样性及功能[J]. *微生物学通报*, 2020, 47(2): 500-511.
- SHAYIMU G, ZHANG ZD, YANG B, ZHANG SK, TANG QY, SONG SQ, ZHU J, GUO CM, GU M. Endophytic bacterial community diversity and function in apple trees of different varieties[J]. *Microbiology China*, 2020, 47(2): 500-511 (in Chinese).
- [17] 程欢, 张东华, 张俊忠, 刘丽. 苹果砧木 T337 不同组织内生菌群落及其功能预测[J]. *江苏农业科学*, 2022, 50(14): 144-154.
- CHENG H, ZHANG DH, ZHANG JZ, LIU L. Endophytic microbial community analysis and function prediction in different tissue parts of apple rootstock T337[J]. *Jiangsu Agricultural Sciences*, 2022, 50(14): 144-154 (in Chinese).

- [18] LAURO FM, McDOUGALD D, THOMAS T, WILLIAMS TJ, EGAN S, RICE S, DeMAERE MZ, TING L, ERTAN H, JOHNSON J, FERRIERA S, LAPIDUS A, ANDERSON I, KYRPIDES N, MUNK AC, DETTER C, HAN CS, BROWN MV, ROBB FT, KJELLEBERG S, et al. The genomic basis of trophic strategy in marine bacteria[J]. Proceedings of the National Academy of Sciences of the United States of America, 2009, 106(37): 15527-15533.
- [19] CHEN BZ, LIU HB. Relationships between phytoplankton growth and cell size in surface oceans: interactive effects of temperature, nutrients, and grazing[J]. Limnology and Oceanography, 2010, 55(3): 965-972.
- [20] SELA I, WOLF YI, KOONIN EV. Theory of prokaryotic genome evolution[J]. Proceedings of the National Academy of Sciences of the United States of America, 2016, 113(41): 11399-11407.
- [21] KUSARI S, VERMA VC, LAMSHOEFT M, SPITELLER M. An endophytic fungus from *Azadirachta indica* A. Juss. that produces azadirachtin[J]. World Journal of Microbiology and Biotechnology, 2012, 28(3): 1287-1294.
- [22] 孙碧琪, 史涵博, 田亮, 王伟, 刘洋. 4种海南诺丽果内生细菌多样性及其关键生物活性研究[J]. 食品科学技术学报, 2023, 41(3): 85-97.
SUN BQ, SHI HB, TIAN L, WANG W, LIU Y. Diversity and key biological activities of endophytic bacteria in 4 varieties of Hainan noni fruit[J]. Journal of Food Science and Technology, 2023, 41(3): 85-97 (in Chinese).
- [23] 黄小茜, 陈翰, 李发活, 梁任繁, 肖冬, 何龙飞, 王爱勤. 广西葛根内生细菌的分离鉴定及其促生特性[J]. 微生物学通报, 2023, 50(5): 2017-2028.
HUANG XX, CHEN H, LI FH, LIANG RF, XIAO D, HE LF, WANG AQ. Endophytic bacteria from *Pueraria montana* var. *lobata* in Guangxi: isolation, identification, and characterization of plant growth-promoting effect[J]. Microbiology China, 2023, 50(5): 2017-2028 (in Chinese).
- [24] 陈亚超, 李楠楠, 刘子迪, 胡冰, 李春. 源于甘草内生菌的甘草酸合成相关功能基因的宏基因组挖掘[J]. 中国生物工程杂志, 2021, 41(9): 37-47.
CHEN YC, LI NN, LIU ZD, HU B, LI C. Metagenomic mining of functional genes related to glycyrrhizin synthesis from endophytes of licorice[J]. China Biotechnology, 2021, 41(9): 37-47 (in Chinese).
- [25] FAURE C, VEYSSIÈRE M, BOËLLE B, SAN CLEMENTE H, BOUCHEZ O, LOPEZ-ROQUES C, CHAUBET A, MARTINEZ Y, BEZOUŠKA K, SUCHÁNEK M, GAULIN E, REY T, DUMAS B. Long-read genome sequence of the sugar beet rhizosphere mycoparasite *Pythium oligandrum*[J]. G3 Genes Genomes Genetics, 2020, 10(2): 431-436.
- [26] LI GM, QIN J, ZHANG T, YANG L, YANG QH, CAO YJ, TANG B, LI ZX, LUO J, YOU SM, WAN XQ, LIU Y, GUO LB, JIANG KF, ZHENG JK. Genome resource of *Sphingomonas carotinifaciens* L9-754^T, an endophyte isolated from leaf tissues of *Jatropha curcas*[J]. Plant Disease, 2021, 105(1): 205-206.
- [27] LIN L, WEI CY, CHEN MY, WANG HC, LI YY, LI YR, YANG LT, AN QL. Complete genome sequence of endophytic nitrogen-fixing *Klebsiella variicola* strain DX120E[J]. Standards in Genomic Sciences, 2015, 10(1): 22.
- [28] ZHANG P, DIAO J, XIE GQ, MA L, WANG LH. A complete genome sequence of the wood stem endophyte *Bacillus velezensis* BY6 strain possessing plant growth-promoting and antifungal activities[J]. BioMed Research International, 2021, 2021: 3904120.
- [29] ASAF S, KHAN AL, KHAN MA, AL-HARRASI A, LEE IJ. Complete genome sequencing and analysis of endophytic *Sphingomonas* sp. LK11 and its potential in plant growth[J]. 3 Biotech, 2018, 8(9): 389.
- [30] ZHANG L, ZHONG J, LIU H, XIN KY, CHEN CQ, LI QQ, WEI YH, WANG Y, CHEN F, SHEN XH. Complete genome sequence of the drought resistance-promoting endophyte *Klebsiella* sp. LTGPAF-6F[J]. Journal of Biotechnology, 2017, 246: 36-39.
- [31] EIDA AA, BOUGOUFFA S, ALAM I, HIRT H, SAAD MM. Complete genome sequence of *Paenibacillus* sp. JZ16, a plant growth promoting root endophytic bacterium of the desert halophyte *Zygophyllum simplex*[J]. Current Microbiology, 2020, 77(6): 1097-1103.
- [32] CHEN C, YUE ZH, CHU CW, MA KS, LI LL, SUN ZK. Complete genome sequence of *Bacillus* sp. strain WR11, an endophyte isolated from wheat root providing genomic insights into its plant growth-promoting Effects[J]. Molecular Plant-Microbe Interactions: MPMI, 2020, 33(7): 876-879.
- [33] LEE S, KANTH B, KIM HS, KIM TW, SANG M, SONG J, WEON AHY. Complete genome sequence of

- the plant growth-promoting endophytic bacterium *Rhodanobacter glycinis* T01E-68 isolated from tomato (*Solanum lycopersicum* L.) plant roots[J]. The Microbiological Society of Korea, 55(4): 422-412.
- [34] BIANCO C, ANDREOZZI A, ROMANO S, FAGORZI C, CANGIOLI L, PRIETO P, CISCHE F, NIANGADO O, SIDIBÉ A, PIANEZZE S, PERINI M, MENGONI A, DEFEZ R. Endophytes from African rice (*Oryza glaberrima* L.) efficiently colonize Asian rice (*Oryza sativa* L.) stimulating the activity of its antioxidant enzymes and increasing the content of nitrogen, carbon, and chlorophyll[J]. Microorganisms, 2021, 9(8): 1714.
- [35] FANG XM, QIN K, LI SJ, HAN S, ZHU TH, FANG XM, QIN K. Whole genome sequence of *Diaporthe capsici*, a new pathogen of walnut blight[J]. Genomics, 2020, 112(5): 3751-3761.
- [36] LUO JW, XIAO S, WANG JH, WANG B, CAI YX, HU WF. The metabolite profiling and microbial community dynamics during pineapple by-product fermentation using co-inoculation of lactic acid bacteria and yeast[J]. Fermentation, 2023, 9(2): 79.
- [37] van DIJK EL, JASZCZYSSZY Y, NAQUIN D, THERMES C. The third revolution in sequencing technology[J]. Trends in Genetics: TIG, 2018, 34(9): 666-681.
- [38] ADELEKE BS, MULLER D, BABALOLA OO. A metagenomic lens into endosphere microbial communities, promises, and discoveries[J]. Letters in Applied Microbiology, 2023, 76(2): ovac030.
- [39] COENYE T, VANDAMME P. Intragenomic heterogeneity between multiple 16S ribosomal RNA operons in sequenced bacterial genomes[J]. FEMS Microbiology Letters, 2003, 228(1): 45-49.
- [40] ARMIJOS JARAMILLO VD, VARGAS WA, SUKNO SA, THON MR. New insights into the evolution and structure of *Colletotrichum* plant-like subtilisins (CPLSs)[J]. Communicative & Integrative Biology, 2013, 6(6): e25727.

(本文责编 陈宏宇)