# 纳米孔测序技术在实蝇类害虫检疫鉴定中的应用

姜 帆<sup>1,2</sup>, 丁新华<sup>3</sup>, 梁 亮<sup>4</sup>, 付开赟<sup>3</sup>, 郭文超<sup>3</sup>, 朱水芳<sup>2\*</sup> <sup>1</sup>三亚中国检科院生物安全中心,海南 三亚 572025; <sup>2</sup>中国检验检疫科学研究院,北京 100176; <sup>3</sup>新疆农业科学院植物保护研究所,农业农村部西北荒漠绿洲作物有害生物综合治理 重点实验室,新疆 乌鲁木齐 830091; <sup>4</sup>农业农村部规划设计研究院,北京 100125

摘要:【目的】纳米孔测序技术是单分子实时测序技术的一种,正在被广泛应用于临床快速诊断及微生物 检测等领域。本研究以实蝇这一类重要的检疫性有害生物为例,探究该技术在昆虫检疫鉴定中应用的可 行性,为昆虫检疫鉴定提供新方法。【方法】分别采用一代测序技术和纳米孔测序技术,对 14 种经形态 学鉴定的实蝇成虫进行 DNA 条形码测序,通过 BOLD 和 NCBI 数据库对测序结果进行比对分析,并比较 2 种测序技术所得序列结果准确性的差异。【结果】通过纳米孔测序,14 个实蝇样品在 44 min 内获得 181 Mb bases.每个样品平均得到 11280 条 reads.单个 reads 的准确度为 92.10%~94.53%:经一致性序列校正.



开放科学标识码 (OSID 码)

所有实蝇样品均可得到与一代测序结果完全一致的序列,序列分析结果与形态学鉴定结果完全一致。【结论】采用本研究 的实验流程和数据分析方法,纳米孔测序技术可以应用于实蝇类害虫的检疫鉴定,测序结果准确、高效;本研究提供的实验 方案同样适用于基于扩增子测序的物种鉴定,满足大规模样品的高通量精准鉴定需求。

关键词:纳米孔测序;实蝇;昆虫;鉴定;扩增子

# Preliminary investigation of quarantine and identification of Tephritidae based on nanopore sequencing technology

JIANG Fan<sup>1,2</sup>, DING Xinhua<sup>3</sup>, LIANG Liang<sup>4</sup>, FU Kaiyun<sup>3</sup>, GUO Wenchao<sup>3</sup>, ZHU Shuifang<sup>2\*</sup>

<sup>1</sup>CAIQ Center for Biosafety, Sanya, Hainan 572025, China; <sup>2</sup>Chinese Academy of Inspection and Quarantine, Beijing 100176,

China; <sup>3</sup>Key Laboratory of Integrated Pest Management on Crops in Northwestern Oasis, Institute of Plant Protection,

Xinjiang Academy of Agricultural Sciences, Urumqi, Xinjiang 830091, China; <sup>4</sup>Academy of Agricultural Planning

and Engineering, Ministry of Agriculture and Rural Affairs, Beijing 100125, China

Abstract: [Aim] Nanopore sequencing technology, a single-molecule real-time sequencing approach, plays a crucial role in rapid clinical diagnosis and microbiological detection. Using Tephritidae as a crucial example of quarantine pests, the applicability of nanopore sequencing technology was assessed for insect quarantine identification, offering a new assessment approach. [Method] We employed Sanger sequencing technology and nanopore sequencing to obtain the DNA barcodes of 14 Tephritidae species previously identified based on morphology. Each sequencing result was cross-referenced in the NCBI and BOLD databases to confirm species accuracy. We then compared the accuracy of sequences obtained through these two distinct sequencing technologies for each species. [Result] Nanopore sequencing generated 181-megabyte bases in just 44 min across 14 samples, with an average of 11280 reads per sample and individual read accuracy of 92.10%-94.53 %. After sequence correction for consistency, the results of nanopore sequencing were consistent with those of Sanger sequencing, and the sequencing analysis findings were consistent with the outcome of morphological identification. [Conclusion] Our study demonstrates the full applicability of nanopore sequencing results proved to be both accurate and efficient. Additionally, the experimental protocol provided in this study is suitable for species identification based on amplicon sequencing, fulfilling the requirements for high-throughput and accurate identification of large-scale samples. Key words; nanopore sequencing; Tephritidae; insects; identification; amplicon

纳米孔测序技术是英国牛津纳米孔技术公司 (Oxford Nanopore Technologies, ONT)研发的一种 单分子测序技术,它依赖一个纳米级的蛋白质孔 (称为"纳米孔")作为生物传感器,嵌入一个浸没 在电生理溶液的薄膜中,施加恒定的电压,通过纳 米孔产生离子电流,在马达蛋白的控制下使单个核 苷酸分子逐一通过纳米孔,根据不同类型核苷酸碱 基带电性质的不同,通过电信号的差异识别碱基类 别,从而实现测序(Deamer et al., 2016)。与其他测 序平台相比,纳米孔测序具有超长读长、试验周期 短、实时测序、操作便捷等优势,正在被大量应用于 基因组测序组装、全长转录本测序、基因突变检测、 临床快速诊断以及疫情监测等领域(Wang et al., 2021)。MinION 纳米孔测序仪是 2014 年 ONT 提供 的第一个便携式纳米孔测序仪(Jain et al., 2016), 只需连接一台电脑,使用者便可在任何时间、任何 地点给任何生物测序。由于其小巧易携带、操作简 单,特别适用于在现场检测及病虫害监测。

目前,基于纳米孔测序技术的宏基因组测序在 植物病原微生物检测方面开展了一些初步研究()陈 大嵩等,2021; 卢慧林等,2021; Chalupowicz et al., 2019; Filloux et al., 2018; Naito et al., 2019), 而未 见该技术应用于害虫检疫鉴定的报道。实蝇(双翅 目 Diptera 实蝇科 Tephritidae) 是具有重要经济意义 的有害生物类群、在世界植物检疫中占有重要的地 位,其在我国的入侵危害也日益严重,多个种属均 被列为我国进境植物检疫对象和我国重点管理的 外来物种(李志红等,2013)。基于线粒体 CO I 基 因的 DNA 条形码技术已被广泛应用于除复合体以 外的实蝇物种鉴定(Jiang et al., 2014),在以往的研 究中,研究人员通常将扩增产物送至测序公司,最 快也需1d时间得到测序结果,然后才能进行种类 鉴定分析。本研究尝试利用 MinION 纳米孔测序 仪,以实蝇类害虫为研究对象,探究纳米孔测序技 术在昆虫检疫鉴定中应用的可能性:特别是与一代 测序技术相比,评估纳米孔测序技术对扩增子的测 序效率和准确性。

#### 1 材料与方法

#### 1.1 供试样品

本研究所用实蝇成虫样品包括南美按实蝇 Anastrepha fraterculus (Wiedemann)、墨西哥按实蝇 A. ludens (Loew)、西印度按实蝇 A. obliqua (Macquart)、番石榴实蝇 Bactrocera correcta (Bezzi)、橘小 实蝇 B. dorsalis Hendel、橘大实蝇 B. minax (Enderlein)、蜜柑大实蝇 B. tsuneonis (Miyake)、昆士兰实 蝇 B. tryoni (Froggatt)、桃实蝇 B. zonata (Saunders)、枣实蝇 Carpomya vesuviana Costa、地中海实 蝇 Ceratitis capitata (Wiedemann)、樱桃绕实蝇 Rhagoletis cerasi Loew、瓜实蝇 Zeugodacus cucurbitae (Coquillett)和南亚果实蝇 Z. tau (Walker),共计14 种。均来自田间采集、口岸截获以及国际原子能机 构赠予。全部样品均经过形态学鉴定确认种类。

#### 1.2 DNA 提取

采用天根生化科技(北京)有限公司"血液/细胞/组织基因组 DNA 提取试剂盒"(DP304)对实蝇样品进行 DNA 提取。取实蝇单足于 1.5 mL 离心管中,向离心管中加入液氮充分研磨,再根据试剂盒说明书对样品进行 DNA 提取,得到的 DNA 于-20 ℃冰箱中保存备用。

### 1.3 PCR 扩增

采用 DNA 条形码所用通用引物 LCO1490 (5'-GGTCAACAAATCATAAAGATATTGG-3')/HCO2198 (5'-TAAACTTCAGGGTGACCAAAAAATCA-3')(Fol mer *et al.*,1994) 对样品 DNA 进行扩增,目的片段 为线粒体 DNA 中序列长度为 658 bp 的 CO I 基因 片段。PCR 反应的总体系为 50  $\mu$ L:2×Taq PCR MasterMix [天根生化科技(北京)有限公司, KT201] 25  $\mu$ L, LCO1490/HCO2198 (10  $\mu$ mol·L<sup>-1</sup>)各2  $\mu$ L, ddH<sub>2</sub>O 19  $\mu$ L, DNA 模板 2  $\mu$ L。PCR 反应条件:94 ℃ 3 min;94 ℃ 30 min、50 ℃ 30 s、72 ℃ 30 s, 共35 个循环;72 ℃ 5 min。取5  $\mu$ L PCR 产物在 1.5%的琼脂糖凝胶上进行电泳检测,将有目的条带的 20  $\mu$ L 扩增产物送至北京擎科生物科技 有限公司进行一代双向测序。

#### 1.4 纳米孔测序文库构建

使用 ONT 公司 LSK109 连接及 EXP-NBD104/ 114 建库试剂盒,根据北京金果壳公司的纳米孔测 序 DNA 多样本辅助建库试剂盒(JC-EXP024)及其 说明书进行测序文库构建。首先,采用磁珠法对剩 余的 PCR 产物进行纯化,并取 1 μL 纯化产物通过 Qubit 2.0 荧光定量仪(美国 Life Qubit)进行定量。 为使每个样品测序得到相对一致的 reads 数,根据 PCR 纯化产物定量结果和说明书推荐的 300 fmol PCR 产物用量,分别对每个样品进行末端修复;每 个样品配制 54.5 µL 的混合体系于 PCR 管中进行 末端修复,反应体系为3.5 µL 末端修复缓冲液、3 μL 末端修复酶、PCR 纯化产物、加无核酸酶水补足 至 54.5 µL;末端修复反应条件:20 ℃ 10 min;65 ℃ 5 min:20 ℃保存:对反应产物进行磁珠纯化并进行 Qubit 定量。向每个样本添加对应条形码并记录, 配制 30 μL 混合体系于 PCR 管中,反应体系为 12.5 μL 末端修复后产物、2.5 μL 条形码 Native Barcode、 15 µL TA 连接预混液;反应条件:25 ℃ 15 min;65 ℃ 10 min;20 ℃保存。混合所有已添加条形码的 样本进行磁珠纯化,用 50 µL 无核酸酶水洗脱。最 后,配制 71 μL 混合体系于 PCR 管中,反应体系为 45 μL 混合样本、5 μL 测序接头 II (AMII)、14 μL 快速连接缓冲液、7 μL T4 快速连接酶;连接反应条 件:20 ℃ 15 min;21 ℃保存;对连接产物进行磁珠 纯化,15 μL EB 洗脱,取 1 μL 纯化产物进行 Qubit 定量,其余制备好的文库用于上机测序。

#### 1.5 纳米孔测序与分析

使用 Flow Cell R9.4.1 的测序芯片。首先,将测 序芯片放入 MinION 纳米孔测序仪并连接电脑,对测 序仪和芯片进行质检,保证测序芯片有 800 个以上的 可用纳米孔。然后,准备将构建好的测序文库滴入 芯片用来测序:从芯片的 Priming 孔吸出 20~30 µL 缓冲液除去气泡;吸取 30 µL Flush Tether 加入到一 管 Flush Buffer 中混匀配制成 Priming Mix,向 Priming 孔中匀速加入 800 µL 混匀的 Priming Mix,前置 5 min;按如下体系配制总体积为 75 µL 的 Loading Mix: 37.5 µL Sequencing Buffer、25.5 µL Loading Beads、12 µL 测序文库;再由 Priming 孔匀速推入 200 µL 混匀的 Priming Mix,此时需看到 Sample 孔有大 液滴向上涌出;向 Sample 孔滴加混好的 75 µL 待测 文库 Loading Mix,控制好滴加速度以保证文库充满 芯片区域;关闭 Priming 孔和 Sample 孔,开始测序。

MinION 纳米孔测序仪运行 44 min。用 seqtk 软件将测序所得的 fastq 文件转换为 fasta 文件,下 载 BOLD 数据库中实蝇科物种的基因序列作为候 选比对数据库,用 blast 将 fasta 数据与候选比对数 据库进行比对,过滤掉相似度低于 90%且覆盖度低 于 90%的比对结果,统计比对到候选序列上的 fasta 条数,挑选比对数量最多的序列作为参考序列;使 用 minimap2 将原始 fastq 比对到挑选出的参考序列 上,过滤掉比对质量值小于 60 的序列,并对 bam 文 件排序,使用 medaka 软件检测突变,获得变异 vcf 文件,使用 samtools 软件统计每个位点上的深度, 使用自定义脚本将深度低于 100 的位点标记为 N, 使用 bcftools consensus 获得一致性序列。利用 DNAMAN 软件,将获得的每个样品的一致性序列与 一代测序数据进行差异比较分析。

#### 2 结果与分析

#### 2.1 一代测序结果

电泳结果显示,经 PCR 扩增所有样品均产生目的扩增条带(图1),PCR 产物的一代测序结果经与 BOLD 和 NCBI 数据库比对,也显示与形态学鉴定结果一致。



图 1 PCR 产物电泳检测结果图 Fig.1 The result of PCR products

#### 2.2 纳米孔测序结果

纳米孔测序文库构建过程中,PCR 产物纯化后 浓度如表 1 所示。为使后续每个样品测序得到相 对一致的 reads 数,本研究在末端修复的步骤对每 个样品浓度进行了均一化处理,从末端修复后的产 物浓度可知,后续建库过程中每个样品的浓度相差 不大。经 Qubit 检测,最终用于上级测序的文库浓 度为 13.9 ng · μL<sup>-1</sup>。

MinION 纳米孔测序仪运行 44 min 后,共有 3.77 GB 数据产出,获得 181 Mb bases。序列长度分 布如图 2 所示,序列长度集中分布在 700~800 bp, 与目的片段的长度较一致。每个样品产生的 reads 情况如图 3 所示,14 个样品的 reads 数分布在 8863 ~13758,每个样品的平均 reads 数为 11280。

从每个样品的测序结果中随机挑选 10 条单条 序列,经比对分析可知,与该物种的一代测序结果 相比,纳米孔测序的单条序列与一代测序的序列相 似度为 92.10%-94.53%。经一致性序列校正后,每 个样品的一致性序列与一代测序结果完全一致。 以本研究中产生 reads 数最少的枣实蝇为例,展示 纳米孔测序文库构建各阶段结果

e 1\*1

单条序列和一致性校正后序列与一代测序所得序 列的比对结果(图4)。

表1

Table 1 Results of each stage of indrary construction for nanopore sequencing					
序号 Number	样品种类 Species	PCR 产物纯化后浓度 Concentration of purified PCR products/(ng・µL <sup>-1</sup> )	末端修复用 PCR 纯化产物体积 Volume of purified PCR products for terminal repair/µL	末端修复后产物 浓度 Concentration of terminal repair products/(ng・μL <sup>-1</sup> )	文库浓度 Concentration of library ∕(ng・μL <sup>-1</sup> )
1	南美按实蝇 Anastrepha fraterculus	25.6	4.69	4.88	13.9
2	墨西哥按实蝇 Anastrepha ludens	14.5	8.28	4.44	
3	西印度按实蝇 Anastrepha obliqua	29.4	4.08	4.86	
4	番石榴实蝇 Bactrocera correcta	14.5	8.28	4.92	
5	橘小实蝇 Bactrocera dorsalis	14.7	8.16	4.88	
6	橘大实蝇 Bactrocera minax	3.7	32.26	4.82	
7	蜜柑大实蝇 Bactrocera tsuneonis	19.4	6.19	5.20	
8	昆士兰实蝇 Bactrocera tryoni	18.1	6.63	5.14	
9	桃实蝇 Bactrocera zonata	21.6	5.56	4.38	
10	枣实蝇 Carpomya vesuviana	14.7	8.16	5.00	
11	地中海实蝇 Ceratitis capitata	11.3	10.62	5.38	
12	樱桃绕实蝇 Rhagoletis cerasi	13.5	8.89	4.74	
13	瓜实蝇 Zeugodacus cucurbitae	20.8	5.77	4.52	
14	南亚果实蝇 Zeugodacus tau	3.7	32.43	4.67	



序列长度 Sequence length/k





南美按实蝇 Anastrepha fraterculus 墨西哥按实蝇 Anastrepha ludens 西印度按实蝇 Anastrepha obliqua 番石榴实蝇 Bactrocera correcta 橘小实蝇 Bactrocera dorsalis 橘大实蝇 Bactrocera tsuneonis 昆士兰实蝇 Bactrocera tsuneonis 昆士兰实蝇 Bactrocera tryoni 桃实蝇 Bactrocera zonata 枣实蝇 Carpomya vesuviana 地中海实蝇 Ceratitis capitata 樱桃绕实蝇 Rhagoletis cerasi 瓜实蝇 Zeugodacus cucurbitae 南亚果实蝇 Zeugodacus tau



#### 图 3 各样品序列数量统计图

Fig.3 The total number of reads for each sample generated by MinION nanopore sequencing



图 4 纳米孔测序与一代测序比对结果图(以枣实蝇为例)

Fig.4 Sequence alignment between nanopore sequencing and sanger sequencing (example of C. vesuviana)
Sanger\_Cv: 枣实蝇一代测序序列; Nanosingle\_Cv: 枣实蝇纳米孔测序单条序列; Nanoconsensus\_Cv: 枣实蝇纳米孔测序一致性序列。
Sanger\_Cv: Sanger sequencing result of C. vesuviana; Nanosingle\_Cv: Single sequence of C. vesuviana based on nanopore sequencing;
Nanoconsensus\_Cv: Consensus sequence of C. vesuviana based on nanopore sequencing.

## 3 讨论

本研究利用纳米孔测序技术,对14个实蝇样 品进行了一次性高通量测序,测序结果显示,经一 致性序列校正,所有实蝇样品均可得到与一代测序 结果一致的序列。说明采用本研究的试验流程和 数据分析方法,纳米孔测序技术可以应用于基于 DNA条形码技术的实蝇类害虫的检疫鉴定。与一 代测序相比,本研究只需44 min即可获得181 Mb bases,每个样品可得到平均11280条 reads,且测序 结果足以保证了序列的准确性,大大提高了基因序 列测定的效率,缩短了基于序列分析进行检疫鉴定 所需的时间。

纳米孔测序技术自问世以来,其本身的测序精 度一直是限制其应用的一个主要因素。本研究结 果显示,采用 R9.4.1 单个 reads 的准确度为 92.10% ~94.53%,但通过一致性序列校正,测序结果可达 到与一代测序结果完全一致。迄今为止,ONT 发布 了 8 个系统版本,2022 年发布的最新版 R10.4.1 官 方公布的准确度可达 99.6%,且双链数据的准确度 达到 99.92%。R10.4.1 的穿孔速度有 260、400 及 520 bp · s<sup>-1</sup>3 种模式,可根据准确度及测序通量要 求,选择最合适的模式(Oxford,2022)。除了测序精 度,测序通量也是检测应用特别是大规模检测中考虑的另一个重要因素。本研究采用的 MinION 测序 平台,数据通量取决于活性纳米孔数量以及核苷酸 分子通过纳米孔的速度和测序运行时间。本研究 采用的 R9.4.1 测序芯片纳米孔每秒通过 400 个碱 基,在今后的应用中可进一步缩短测序时间,即可 得到准确的测序结果。

除纳米孔测序本身的技术优势外,合理的检测 方案的制定也是提高检测准确性和时效性、大幅降 低检测成本的重要手段。基于核酸序列分析的检 测方法是除形态学鉴定外唯一有效的非靶向分子 检测方法。在病原微生物高通量非靶向检测中,宏 基因组测序是目前较为常用的手段(路惠馨等, 2021)。但与微生物相比,昆虫基因组较大,若采用 检测样品基因组作为物种鉴定的依据,无论从测序 成本还是检测效率来说,都不是一个切实可行的方 案。此外,对于昆虫来说,DNA 条形码对于大部分 类群具有较好的鉴定效果(Hebert *et al.*,2003)。因 此,扩增子测序更适用于昆虫的物种鉴定。本研究 提供了一种基于纳米孔测序技术适用于扩增子测 序进行物种鉴定的实验流程和数据分析方法,且能 够满足大规模样品的高通量精准鉴定需求。 纳米孔测序技术以其长读长、通量高、测序精 度高、测序成本低的优势,在检疫鉴定领域中有着 广泛的应用前景,特别是对于一代和二代测序无法 一次性测通的长片段扩增子,纳米孔测序技术显示 出了独一无二的优势。在今后的研究中,仍需进一 步测试完善可以满足检疫鉴定工作需求的样品浓 度、测序时间和数据量;特别是低浓度样品和混合 样品中可实现有效检测鉴定的数据量;对于不同类 群生物、不同检测目的制定不同的测序检测方案, 使纳米孔测序技术在检疫鉴定领域应用中发挥实 质性作用。

#### 参考文献

- 陈大嵩, 黄鸿, 吴华, 李健雄, 戴建青, 2021. 利用纳米孔 测序鉴定蓖麻蚕病害. 环境昆虫学报, 43(1): 260-264.
- 李志红,姜帆,马兴莉,方焱,孙壮志,秦誉嘉,王巧铃, 2013. 实蝇科害虫入侵防控技术研究进展. 植物检疫, 27 (2):1-10.
- 卢慧林,陈大嵩,陈逢浩,叶静文,欧阳革成,2021.基于 纳米孔测序技术的柑橘黄龙病检测方法建立.环境昆虫 学报,43(6):1596-1600.
- 路惠馨, 孙凯, 尹传林, 黄昱栋, 俞晓平, 2021. 纳米孔测 序技术在植物病原检测中的应用与展望. 农业生物技术 学报, 29(9): 1817-1824.
- CHALUPOWICZ L, DOMBROVSKY A, GABA V, LURIA N, REUVEN M, BEERMAN A, LACHMAN O, DROR O, NISSAN G, MANULIS-SASSON S, 2019. Diagnosis of plant diseases using the nanopore sequencing platform. *Plant Pathology*, 68: 229–238.
- DEAMER D, AKESON M, BRANTON D, 2016. Three decades of nanopore sequencing. *Nature Biotechnology*, 34(5): 518-524.

FILLOUX D, FERNANDEZ E, LOIRE E, CLAUDE L, GALZI

S, CANDRESSE T, WINTER S, JEEVA M L, MAKESH-KUMAR T, MARTIN D P, ROUMAGNAC P, 2018. Nanopore-based detection and characterization of yam viruses. *Scientific Reports*, 8: 17879.

- FOLMER O, BLACK M, HOEH W, LUTZ R, VRIJENHOEK R, 1994. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology*, 3: 294–299.
- HEBERT P D N, CYWINSKA A, BALL S L, DEWAARD J R, 2003. Biological identifications through DNA barcodes. *Proceedings the Royal Society B*, 270: 313–321.
- JAIN M, OLSEN H E, PATEN B, AKESON M, 2016. The Oxford nanopore MinION: delivery of nanopore sequencing to the genomics community. *Genome Biology*, 17: 239.
- JIANG F, JIN Q, LIANG L, ZHANG A B, LI Z H, 2014. Existence of species complex largely reduced barcoding success for invasive species of Tephritidae: a case study in *Bactrocera* spp. *Molecular Ecology Resources*, 14: 1114–1128.
- NAITO F Y B, MELO F L, FONSECA M E N, SANTOS C A
  F, CHANES C R, RIBEIRO B M, GILBERTSON R L,
  BOITEUX L S, PEREIRA-CARVALHO R D C, 2019. Nanopore sequencing of a novel bipartite new world begomovirus infecting cowpe. Archives of Virology, 164: 1907-1910.
- OXFORD, 2022. Oxford nanopore delivers technology update at annual London calling conference: bringing together years of innovation to showcase one sensing platform for all biological analyses. [2022-05-20]. https:// nanoporetech.com/about-us/news/oxford-nanopore-delivers-technology-updateannual-london-calling-conference-bringing.
- WANG Y, ZHAO Y, BOLLAS A, WANG Y, AU K F, 2021. Nanopore sequencing technology, bioinformatics and applications. *Nature Biotechnology*, 39 (11): 1348–1365.

(责任编辑:郭莹)