Evaluation of NGS DNA barcoding for biosecurity diagnostic applications

Case study from banana freckle incursion in Australia

Galaihalage, Kalpani; Patel, Shreya; Yadav, Sonu

Published in:
Australasian Plant Pathology

DOI:
10.1007/s13313-024-00978-4

Published: 01/05/2024

Document Version
Publisher's PDF, also known as Version of record

Australasian Plant Pathology (2024) 53:291–295
https://doi.org/10.1007/s13313-024-00978-4

RESEARCH NOTE

Evaluation of NGS DNA barcoding for biosecurity diagnostic applications: case study from banana freckle incursion in Australia

Kalpani Galaihalage[1] · Shreya Patel[1] · Sonu Yadav[1,2]

Received: 5 October 2023 / Accepted: 7 April 2024 / Published online: 15 April 2024

Abstract

Molecular diagnostics in combination with morphological identification is the method of choice for several cryptic microbial plant pathogens. For some diagnostic applications, traditional sequencing techniques can be time consuming, making them ill-suited for biosecurity incursion responses, where accurate results are needed in real time. More rapid next generation sequencing tools must be tested and compared with traditional methods to assess their utility in biosecurity applications. Here utilizing 95 samples infected with fungal pathogen *Phyllosticta cavendishii*, from a recent incursion in Australia, we compare species identification success using Internal Transcribed Spacer (ITS) gene barcode on conventional Sanger and Oxford Nanopore MinION sequencing platforms. For Sanger sequencing, the average pairwise identity percentage score between generated consensus sequences and *P. cavendishii* sequence from holotype material on NCBI database was 99.9% ± SE 0.0 whereas for MinION sequencing the average pairwise identity percentage was 99.1% ± SE 0.1. Relatively larger consensus sequences (mean 486 bp ± SE 2.4) were generated by Sanger sequencing compared to MinION sequencing (mean 435 bp ± SE 4.6). Our results confirm that both sequencing methods can reliably identify *P. cavendishii*. MinION sequencing, provided quicker results compared to Sanger sequencing and demonstrated diagnostic competence, with the added advantage of being portable, for front-line "point of incursion" biosecurity applications.

Keywords Biosecurity · Banana freckle · Incursion · MinION · Sanger sequencing · Fungal plant pathogen

The intensification of global trade, human assisted movement of species and climate change is facilitating introduction of invasive species, agronomic pests and plant pathogens into new areas (Capinha et al. 2015; Seebens et al. 2017; Santini et al. 2018). The scale of global plant movements has seen an increase in the frequency of exotic plant pest and disease incursions (Hu et al. 2020), leading to increased cost of biosecurity management worldwide. Biosecurity response, containment and management success largely depends on the timely and accurate diagnosis of the invasive species in the early stages of infestation, before establishment.

Traditionally fungal pathogen identification was based upon examination of morphological and physiological characteristics. However, identification of species lacking unique morphological characters and limited well-trained experts pose challenges in morphological species identification (Fajarningsih 2016). This limitation has been to some extent alleviated by the application of DNA sequencing based molecular methods (Armstrong and Ball 2005). Internal transcribed spacer (ITS) region has been accepted as a universal barcode for identification of fungi (Schoch et al. 2012). Sanger sequencing is a robust method to generate sequence data and has been widely used in species identification (Taylor and Harris 2012) but it can lead to incorrect identification in case of mixed infection (Langsiri et al. 2023) and requires an expensive sequencer. Outsourcing sequencing to a commercial company offering this service is a common practice, but can take time to turn-around the diagnostic outcome. Species identification using Sanger sequencing can take up to 4–5 days and even longer for

✉ Sonu Yadav
sonu.yadav@nt.gov.au

[1] Biosecurity and Animal Welfare, Department of Industry, Tourism and Trade, Northern Territory Government, Berrimah Farm Science Precinct, Darwin, Northern Territory 0810, Australia

[2] Research Institute for the Environment and Livelihoods, Faculty of Science and Technology, Charles Darwin University, Darwin, Northern Territory 0909, Australia

laboratories located in remote areas due to increased sample shipping time. The rapid spread of exotic species demands locally sourced sequencing techniques to enable "point-of-care" diagnostic capability.

Relatively new sequencing platforms, such as the Oxford Nanopore Technologies' MinION (ONT, Oxford, UK) can reportedly perform amplicon sequencing, long-read and short-read sequencing, targeted and whole genome sequencing in a timely and cost-effective manner (Chen et al. 2023; Zheng et al. 2023). Although Nanopore-based sequencing has been applied to pathogen diagnostics, surveillance and post-entry quarantine (McCormack et al. 2013; Bronzato Badial et al. 2018; Loit et al. 2019; Maina et al. 2022; Delmiglio et al. 2023), case studies showing its utility during an active biosecurity incursion are limited. Species identification during biosecurity incursion requires a robust methodology, high confidence in sequencing output and analysis. Nanopore's MinION sequencing has been shown to produce high base-calling errors and technical biases (Loit et al. 2019; Chen et al. 2023) especially for long read sequencing and poses challenges for accurate identification. MinION sequencing is yet to be established as a reliable tool for the detection of microbial plant pathogens during biosecurity incursion response.

Banana freckle is a plant disease caused by the fungal species *Phyllosticta cavendishii* on Cavendish and Lady Finger banana cultivars. *P. cavendishii* is classified as a high priority pest of banana in Australia. It was detected on Cavendish bananas in the Northern Territory, Australia in 2013 and was eradicated in 2019 (McMaster et al. 2020). A second outbreak of the disease on Dwarf Cavendish cultivar was reported in May 2022 in the Northern Territory (Banana freckle | NT.GOV.AU). In this study, utilizing 95 *P. cavendishii* infected samples from the 2022 incursion, we assess species identification success using Sanger and MinION amplicon sequencing to examine MinION's utility as an alternative to Sanger sequencing in a biosecurity incursion setting.

Materials and methods

Leaf cuttings and/or infected bananas were collected from symptomatic plants in remote and rural-communities as part of active biosecurity incursion surveillance. Pycnidia were scraped from the infected plant material and placed into 2 ml Eppendorf tube containing lysis buffer and a 5 mm stainless steel bead (QIAGEN cat no. 69989). Total genomic DNA was extracted from 95 samples using BIOLINE Isolate II Plant DNA Kit (BIO-52070) following the manufacturer's instructions. Samples were extracted in small batches in a dedicated biocontainment laboratory. We amplified *Phyllosticta* spp. partial internal transcribed spacer (ITS) gene including ITS1 region, 5.8S rRNA gene and the ITS2 region using forward primer GmF1 (5'-GTG CGCACCTCTCGAAC-3') and reverse primer GmR2 (5'-ACCTGATCCGAGGTCAAC-3') (Wong et al. 2012). To monitor potential contamination extraction controls (kit reagents except plant material) and PCR negative (no template) were used in each run (further information is in Text S1). Purified products were Sanger sequenced bi-directionally at the Australian Genome Research Facility (www.agrf.org.au), Brisbane, Australia.

Library preparation for amplicon sequencing using the same primers on MinION was completed using the Oxford Nanopore rapid barcoding kit (SQK-RBK110.96) following the manufacturer's instructions (Oxford Nanopore Technologies, Oxford, U.K.; see Text S1 for further information). Sequencing run was initiated on MinKNOW (v21.11.7) for 72 h with basecalling mode turned to the "ON" setting.

Sanger sequencing data was analysed on Geneious Prime® software (v2022.1.1; https://www.geneious.com/; Biomatters, New Zealand). Sequencing reads obtained from MinION sequencing were filtered using NanoFilt software (De Coster et al. 2018) and consensus sequences were generated from the filtered reads using NGSpeciesID (Sahlin et al. 2021). See Text S1 for further information. For both sequencing approaches, species identity was confirmed by analysing percentage nucleotide similarity, query cover and multiple sequence alignments between generated consensus sequences and sequences from *Phyllosticta* spp. type material including *P. cavendishii* holotype (Bft26/CBS H-20,918/BRIP 55,419; NCBI accession: JQ743562.1 (Wong et al. 2012)) available on NCBI BLASTn using a megablast algorithm to obtain only highly similar hit sequences. Percentage identity score with the holotype was recorded for each sample. Sequences were also examined against closely related species epitype– *Phyllosticta maculata* (Bf002/CBS H-20,919/BRIP 55,425; NCBI accession: JQ743570.1) and *Phyllosticta musarum* (Bfd34/CBS H-20,922/BRIP 55,434; NCBI accession: JQ743584.1).

Results and discussion

For Sanger sequencing, 94 samples were successfully identified as *P. cavendishii* with the pairwise percentage identity on NCBI BLASTn between generated consensus sequence and holotype sequence ranging from 99.7 to 100% (average 99.9% ± SE 0.0). One sample (sample18) failed to produce good quality sequence and was excluded from the analysis. The consensus sequence length ranged from 407 bp to 503 bp (average 486 bp ± SE 2.4; Table S1). 42 samples showed a single nucleotide transition from C to T.

The single nucleotide variation is also present in the holotype and other *P. cavendishii* accessions: JQ743565 has a "C" whereas JQ743562.1 and JQ743566 has a "T". Sanger sequencing took an average of 5–7 days for completion including 3–4 days for sample shipping.

For MinION sequencing, a total of 19,937 reads were generated with mean read length of 919 bp and mean read quality of 7.2. All samples were successfully sequenced with a mean read score of 135 reads/sample (Figure S1). After filtering for maximum and minimum length and average quality score the total number of reads were reduced to 5,284 with mean read length of 365.9 bp and mean read quality of 10.4 (Figure S2a, b). No reads were removed by mapping reads to human genome and 64 reads were removed by mapping to banana genome. The majority of the reads were filtered out and no consensus sequence was obtained for one sample (sample 58). Two samples (sample 68 and sample 86) failed to generate high quality consensus sequences using NGSpeciesID. Sample 68 contained 16.7% ambiguous bases and sample 86 contained 8.9% ambiguous bases and were excluded from further analysis. *P. cavendishii* was identified with identity score ranging from 96.8 to 100% (average 99.1% ± SE 0.1) (Fig. 1; Table S1). Consensus sequence length ranged from 310 bp to 535 bp (average 435 bp ± SE 4.6). The MinION sequencer generated data within 24–72 h (Figure S3 a, b).

We found that both techniques successfully detected *P. cavendishii* in samples with species identity score percentage ranging from 99 to 100% for Sanger sequencing and 97–100% for MinION between the holotype and generated barcodes. Both sequencing technologies were able to discriminate the target pathogen from two closely related species: *P. musarum* and *P. maculata* with a minimum of 1–3% difference in the pairwise identity score. False negative detection due to nanopore's high error rate can be minimized by running technical replicates of samples at a higher sequencing depth. The low quality noisy read for sample 18 in Sanger sequencing compared to MinION sequencing could be due to the co-amplification of other components in the PCR product.

In our study, species identification and discrimination between closely related species using BLAST was valuable due to comprehensive ITS region reference database for *Phyllosticta* spp., however precautions must be taken while using BLAST in the absence of high-quality reference database. Our findings agree with recent studies comparing the accuracy and speed of traditional sequencing and MinION sequencing for species identification at border checkpoint;

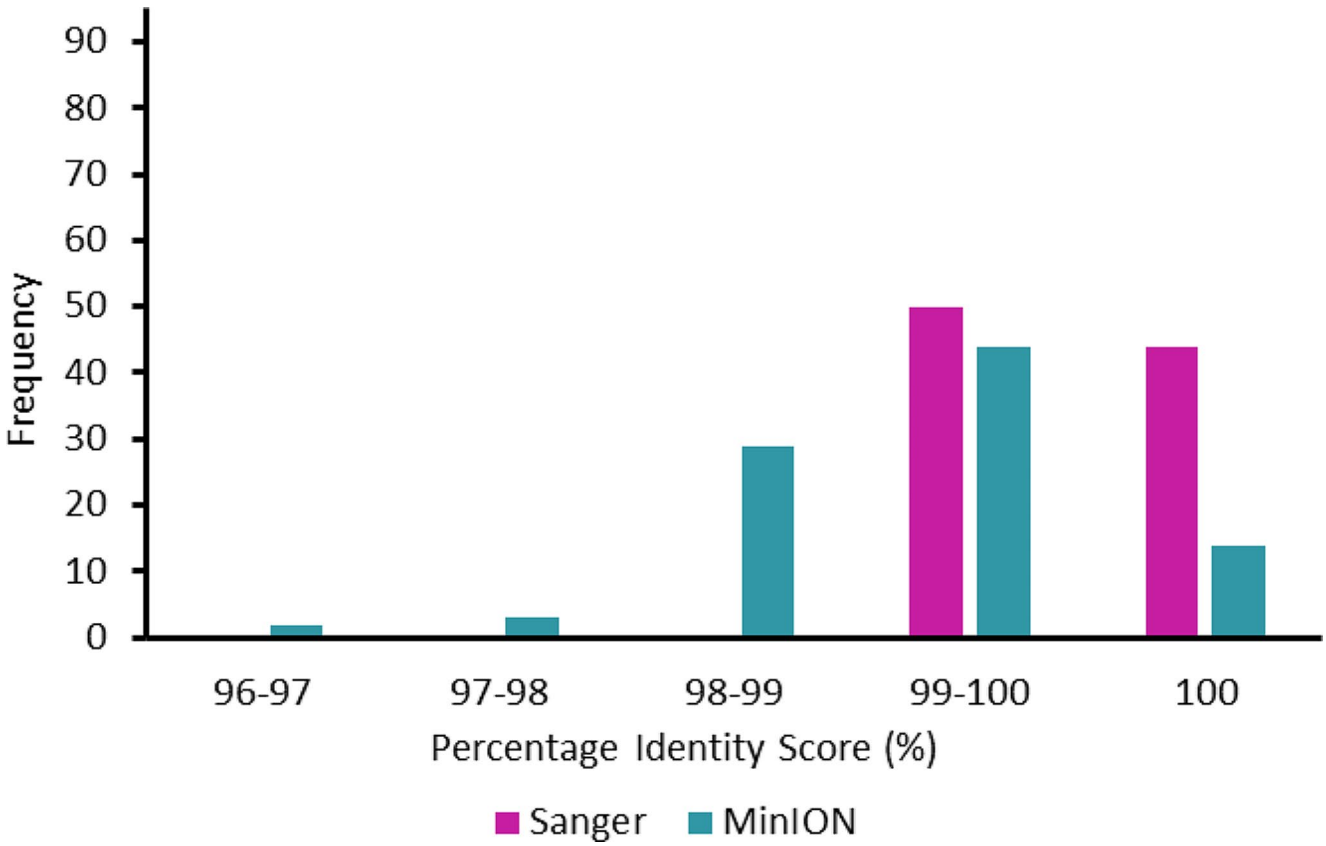

Fig. 1 Comparison of the frequency of percentage identity score between Sanger sequencing and Nanopore-based amplicon sequencing of the ITS gene region the Y-axis in the submitted manuscript was cropped to around 60 as most of the data values are below 60

and differentiating closely related species using MinION and NGSpeciesID (Abeynayake et al. 2021; Vasiljevic et al. 2021).

Our study examined the utility of Nanopore sequencing for the rapid and accurate diagnosis of *P. cavendishii* in a biosecurity emergency response, that requires a rapid (< 48-hour) turn-around. Nanopore sequencing provided results faster than Sanger sequencing and can be used for the molecular diagnosis of high risk pathogen in the event of a suspected incursion. Three samples failed to generate a high-quality consensus sequence, in such cases re-sampling is recommended to minimize false negative detection. In our study, a small number of unsuccessful identifications did not make a substantial difference as several replicates were tested from infected locations. MinION amplicon sequencing combined with improved bioinformatics workflows is particularly useful for laboratories located in remote geographic locations, with restricted access to Sanger sequencing. Other quicker molecular assays such as quantitative PCR and high resolution melting assay targeting shorter amplicons (Wong et al. 2013) can be tested in future for their utility in a suspected incursion.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s13313-024-00978-4.

Acknowledgements We would like to thank surveillance staff for collecting samples as a part of the Northern Territory Banana Freckle Emergency Response and Stanley E. Bellgard for proof-reading the manuscript. KG was funded under Activity ID 4-IF99YJT; Immediate Assistance Funding, administered through the Community Grants Hub on behalf of the Commonwealth of Australia.

Author contributions SY conceptualised, designed and coordinated the study. KG, SP and SY performed laboratory work and data analysis. SY wrote the initial draft. All authors edited and approved the final manuscript.

Data availability 12 sequences from different collection sites have been uploaded to NCBI GenBank with accession numbers OR793007-OR793018.

Declarations

Conflict of interest The authors have no conflict of interest to declare that are relevant to this article.

References

Abeynayake SW, Fiorito S, Dinsdale A, Whattam M, Crowe B, Sparks K, Campbell PR, Gambley C (2021) A rapid and cost-effective identification of invertebrate pests at the borders using MinION sequencing of DNA barcodes. Genes 12(8): 1138

Armstrong K, Ball S (2005) DNA barcodes for biosecurity: invasive species identification. Philosophical Trans Royal Soc B: Biol Sci 360(1462):1813–1823

Bronzato Badial A, Sherman D, Stone A, Gopakumar A, Wilson V, Schneider W, King J (2018) Nanopore sequencing as a surveillance tool for plant pathogens in plant and insect tissues. Plant Dis 102(8):1648–1652

Capinha C, Essl F, Seebens H, Moser D, Pereira HM (2015) Dispersal Alien Species Redefines Biogeogr Anthropocene Sci 348(6240):1248–1251

Chen P, Sun Z, Wang J, Liu X, Bai Y, Chen J, Liu A, Qiao F, Chen Y, Yuan C (2023) Portable nanopore-sequencing technology: Trends in development and applications. Front Microbiol 14:1043967

De Coster W, D'hert S, Schultz DT, Cruts M, Van Broeckhoven C (2018) NanoPack: visualizing and processing long-read sequencing data. Bioinformatics 34(15):2666–2669

Delmiglio C, Waite DW, Lilly ST, Yan J, Elliott CE, Pattemore J, Guy PL, Thompson JR (2023) New Virus Diagnostic approaches to ensuring the Ongoing Plant Biosecurity of Aotearoa New Zealand. Viruses 15(2):418

Fajarningsih ND (2016) Internal transcribed spacer (ITS) as DNA barcoding to identify fungal species: a review. Squalen Bull Mar Fisheries Postharvest Biotechnol 11(2):37–44

Hu Y, Wilson S, Schwessinger B, Rathjen JP (2020) Blurred lines: integrating emerging technologies to advance plant biosecurity. Curr Opin Plant Biol 56:127–134

Langsiri N, Worasilchai N, Irinyi L, Jenjaroenpun P, Wongsurawat T, Luangsa-Ard JJ, Meyer W, Chindamporn A (2023) Targeted sequencing analysis pipeline for species identification of human pathogenic fungi using long-read nanopore sequencing. IMA Fungus 14(1):18

Loit K, Adamson K, Bahram M, Puusepp R, Anslan S, Kiiker R, Drenkhan R, Tedersoo L (2019) Relative performance of MinION (Oxford Nanopore Technologies) versus sequel (Pacific Biosciences) third-generation sequencing instruments in identification of agricultural and forest fungal pathogens. Appl Environ Microbiol 85(21):e01368–e01319

Maina S, Norton S, McQueen V, King S, Rodoni B (2022) Oxford Nanopore sequencing reveals an exotic broad bean mottle virus genome within Australian grains Post-entry Quarantine. Microbiol Resource Announcements 11(7):e00211–00222

McCormack JE, Hird SM, Zellmer AJ, Carstens BC, Brumfield RT (2013) Applications of next-generation sequencing to phylogeography and phylogenetics. Mol Phylogenet Evol 66(2):526–538

McMaster C, Tran-Nguyen L, Voutsinos M, Cook S, Condé B, West S, Nguyen T, Liberato J (2020) Outbreak of freckle disease (Phyllosticta Cavendishii) on Cavendish bananas in Australia and the initial biosecurity response. Plant Pathol 69(1):38–49

Sahlin K, Lim MC, Prost S (2021) NGSpeciesID: DNA barcode and amplicon consensus generation from long-read sequencing data. Ecol Evol 11(3):1392–1398

Santini A, Liebhold A, Migliorini D, Woodward S (2018) Tracing the role of human civilization in the globalization of plant pathogens. ISME J 12(3):647–652

Schoch CL, Seifert KA, Huhndorf S, Robert V, Spouge JL, Levesque CA, Chen W, Consortium FB, C. A. List FB, Bolchacova E (2012) Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. Proc Natl Acad Sci 109(16):6241–6246

Seebens H, Blackburn TM, Dyer EE, Genovesi P, Hulme PE, Jeschke JM, Pagad S, Pyšek P, Winter M, Arianoutsou M (2017) No saturation in the accumulation of alien species worldwide. Nat Commun 8(1):14435

Taylor H, Harris W (2012) An emergent science on the brink of irrelevance: a review of the past 8 years of DNA barcoding. Mol Ecol Resour 12(3):377–388

Vasiljevic N, Lim M, Humble E, Seah A, Kratzer A, Morf NV, Prost S, Ogden R (2021) Developmental validation of Oxford Nanopore Technology MinION sequence data and the NGSpeciesID bioinformatic pipeline for forensic genetic species identification. Forensic Sci International: Genet 53:102493

Wong M-H, Crous PW, Henderson J, Groenewald JZ, Drenth A (2012) Phyllosticta species associated with freckle disease of banana. Fungal Divers 56:173–187

Wong MH, Henderson J, Drenth A (2013) Identification and differentiation of Phyllosticta species causing freckle disease of banana using high resolution melting (HRM) analysis. Plant Pathol 62(6):1285–1293

Zheng P, Zhou C, Ding Y, Liu B, Lu L, Zhu F, Duan S (2023) Nanopore Sequencing Technol its Appl MedComm 4(4):e316