Final Report

Industry preparedness for exotic root knot nematode (Meloidogyne enterolobii)

Project leader:

Mike Hodda

Report authors:

Mike Hodda and Dan Huston

Delivery partner:

Commonwealth Scientific and Industrial Research Organization

Project code:

MT22012

Project:

Industry preparedness for exotic root knot nematode (Meloidogyne enterolobii) (MT22012)

Disclaimer:

Funding statement:

This project has been funded by Hort Innovation, using the fresh potato, processing potato, sweetpotato and vegetable research and development levies and contributions from the Australian Government. Hort Innovation is the grower-owned, not-for-profit research and development corporation for Australian horticulture.

Publishing details:

Published and distributed by: Hort Innovation

Level 7
141 Walker Street
North Sydney NSW 2060

Telephone: (02) 8295 2300

www.horticulture.com.au

Contents

Public summary

In 2022, the exotic guava root knot nematode (*Meloidogyne enterolobii;* herein referred to as GRKN) was detected in Australia for the first time from the greater Darwin Area of the Northern Territory. This nematode is rated as a High/Extreme risk in biosecurity plans for the Onion, Potato, Sweet potato & Vegetable Industries. *Meloidogyne enterolobii* is highly pathogenic on multiple crops with genetic resistance to other root-knot nematodes, including sweet potato, tomato, capsicum, and soybean. The wide host range of this species, including many invasive weeds, several native Australian plants, and a lack of resistant crop varieties, presents ample opportunity for this pest to spread if left unchecked.

The recent and limited number of detections of GRKN in Australia around the time of first detection posed an important question: Had this nematode just recently arrived, or had it occurred in Australia for a long time and was detected because diagnosticians are now using molecular sequencing more regularly? GRKN may truly be a recent arrival, but with five other species of root-knot nematode (RKN) occurring widely in Australia, it is possible that GRKN has long been present but misidentified and confused with other RKN species. Considering the significant cost and logistical challenges associated with eradication, answering the above question was considered important to help guide management decisions regarding this incursion.

Some information regarding historic occurrence may be gleaned from Australian collections, however, the specialised expertise and time intensive nature associated with morphological identification of root-knot nematodes mean this would be long and costly process. Furthermore, because most nematode specimens lodged in collections will have been processed using formaldehyde (formalin), obtaining diagnostic molecular data from such material poses a significant challenge. This project aimed to solve both these issues so that in future incursions scenarios specimen collections could rapidly be screened to determine potential historic occurrence.

This project successfully developed and tested two complementary methods for identifying RKNs from historic samples. First, we developed an AI-driven image-analysis method that enables fast, accurate, and cost-effective species-level RKN identification. Second, we developed a new method for acquiring genetic information from formalin-preserved plant pathology material, potentially applicable to other small plant pests and pathogens.

Applying these methods to historic samples has enhanced understanding of RKN species in Australia, identified potential misidentifications for future study, and suggested that *M. enterolobii* was not present in Australia before its 2022 detection. Further development of the AI diagnostic method into a public-facing interface could benefit Australian diagnosticians, growers, and have positive impacts on trade and biosecurity through enhanced species identification and area freedom designations. Further investigation is proposed to fully utilise these methods to improve knowledge of RKN and their management.

Technical summary

The arrival of guava root-knot nematode (GRKN, *Meloidogyne enterolobii*) in Australia presented a biosecurity dilemma: did this globally emerging plant pest represent a new arrival requiring an eradication effort, or had it been in Australia for a long time, but had been previously mistaken for another root-knot nematode (RKN) species? To answer this a two-pronged project was executed using innovative techniques to retrospectively screen historic plant pathology collections. The approach involved developing a novel molecular protocol to recover DNA from formalin-fixed specimens and creating an Artificial Intelligence (AI) model for rapid identification based on species-specific morphological patterns.

Development of a Protocol for obtaining DNA from formalin-fixed root knot nematodes

Formalin-fixed plant pathology material poses a unique challenge for molecular analyses. Formaldehyde preservation leads to extensive DNA-protein cross-linking. Furthermore, the gradual conversion of formaldehyde into formic acid causes progressive acidification. Long-term storage in formalin solutions degrades and fragments nucleic acids inhibiting standard DNA extraction and amplification methods.

To address this challenge, we developed a protocol which de-crosslinks and then amplifies highly fragmented DNA. Standard DNA barcodes for nematodes, such as the *cox1* gene fragment amplified by the JB3/JB5 primers (approx. 390 bp), were unsuitable, as formalin preservation typically limits amplifiable fragments to under 150 bp. Therefore, a new primer pair was designed to amplify a short, diagnostic 120 base-pair region of the *cox1* gene.

Specimens for protocol development were sourced from the Northern Territory Plant Pathology Herbarium (DNAP) and the NSW Plant Pathology Herbarium (DAR), along with experimental material held by CSIRO fixed in unbuffered formalin for one to two years. Initial screening of historic samples revealed most samples were highly acidic. Multiple DNA extraction protocols were evaluated along with various custom formulations incorporating different proteinase K concentrations and lysis buffers. Ultimately, a hot alkaline lysis method, followed by Proteinase K digestion and purification using silica spin columns, was found to be the most effective for de-crosslinking and DNA recovery.

Proof-of-concept tests confirmed that silica spin columns can effectively retain highly fragmented DNA. However, successful PCR amplification was reliably achieved only from recently fixed experimental material (≤ 2 years). Evaluation of samples older than two years through random-hexamer amplification showed severe fragmentation, demonstrating that progressive acidification in formalin-preserved root material renders DNA irretrievable in most cases. This finding is critical for future collection management. We recommend that where formalin fixation is required, high-capacity buffering agents should be used to neutralize acidic degradation and extend the molecular shelf life of specimens.

AI image classification model for the identification of root knot nematodes via the perineal pattern

A complementary AI-driven diagnostic system was developed to identify RKN species based on their perineal patterns. Perineal patterns were prepared for five RKN species (*M. arenaria, M. enterolobii, M. hapla, M. incognita, M. javanica*) and imaged using a compound microscope.

A dataset of 598 images was curated, cropped to the region of interest (ROI), and split 70/15/15% for training, validation, and testing. Three AI software model backbones (ResNet, ConvNeXt, and Swin Transformer) were evaluated. The final models utilized the Swin V2-B architecture. Images were optimized with moderate augmentation, simple temperature scaling and Test-Time Augmentation. The AI system delivered robust performance, with the best model achieving an identification accuracy of 97.56% on the test set.

Critically, *M. enterolobii* was never confused with any other species in any tested scenario, achieving a perfect score of across all model classification metrics. Saliency mapping confirmed the model was relying on diagnostically relevant features, specifically the dorsal arch and lateral striae, consistent with human expert identification.

We screened perineal patterns from the CSIRO and Agriculture Victoria collections as well as newly prepared perineal patterns made from historic formalin-fixed material from the Northern Territory. The AI model did not recognise any of these perineal patterns as representing *M. enterolobii*.

Conclusion

The project successfully delivered two novel diagnostic tools. The molecular protocol informs best practices for future sample preservation. With further development the AI model is likely to become an accessible, and high-speed diagnostic tool for industry. Crucially, none of our investigations indicated that *M. enterolobii* has occurred in Australia for any length of time. Thus, we conclude that *M. enterolobii* is a relatively recent arrival to Australia.

Keywords

Root-knot nematode; plant parasitic nematode; *Meloidogynidae*; *Meloidogyne*; pests; AI, diagnostics, computer vision; formalin DNA.

Introduction

Guava root knot nematode (*Meloidogyne enterolobii*; herein abbreviated as GRKN) was reported in Australia for the first time in late 2022 from the Darwin area of the Northern Territory. Since this initial discovery, further detections have been made in Queensland. GRKN is a devastating plant parasite causing significant crop losses on a global scale. This nematode has a broad host range spanning 30 plant families, including many important food crops. Notably, GRKN is a resistance-breaking species, causing severe damage to crops with genetic resistance to other root-knot nematode (RKN) species. Consequently, this nematode is a significant biosecurity risk for multiple crops in Australia and is rated in the high to extreme risk categories on the biosecurity plans for multiple industries, e.g., onions, potatoes, sweet potatoes, and other vegetables. The wide host range and resistance-breaking characteristics of GRKN facilitates rapid spread, and this species is very difficult to control. Preventing establishment of this species in production areas where it was previously not present is thus critically important.

The very recent, and thus far limited, number of detections of GRKN in Australia pose an important question: Has this nematode just recently arrived, or has it occurred in Australia for a long time and has now been detected because diagnosticians are now looking for GRKN using DNA sequencing which is more sensitive than perineal fingerprints which the available human expertise often did not interpret to species? GRKN may be a recent arrival in Australia, but with five other species of RKN occurring in Australia, it is possible that GRKN has long been present but misidentified and confused with other RKN species. Considering the significant cost and logistical challenges associated with eradication, answering the above question is needed to justify either an eradication program or a move to management.

Some information regarding historic occurrence may be gleaned from Australian collections, however, the specialised expertise and time intensive nature associated with morphological identification of root-knot nematodes mean this would be long and costly process. Furthermore, because most nematode specimens lodged in collections will have been processed using formaldehyde, obtaining diagnostic molecular data from such material poses a significant challenge.

We proposed combining two promising lines of research to solve the above problems, and which would provide the ability to quickly screen historic material from Australian state and territory collections for *M. enterolobii*.

First, CSIRO has been developing AI-based image recognition software that can be used to identify pests (see https://blog.csiro.au/ai-stink-bug/; https://vimeo.com/676128183/fd9e5fe74a). Root-knot nematodes have a distinctive pattern like a fingerprint on part of their outside termed a 'perineal pattern'. These patterns are species-specific but are difficult for humans to interpret. We aimed to use AI to identify species quicker and better than a person could. This would be particularly useful as lots of historic museum specimens are prepared microscope slides of these perineal patterns – useful for AI, but which acquiring DNA is not feasible.

Second, it is generally presumed that you cannot obtain DNA from formalin-preserved material, but various methods exist and a modern method has recently been pioneered by CSIRO that allows sequencing of formalin-fixed specimens (see https://bit.ly/3Utr9MA). Many collections hold plant pathology specimens, including root material that was identified as being infested with Root-Knot Nematodes. In many cases such material was identified no further than '*Meloidogyne* sp.', and in most cases a formaldehyde-based solution was used to preserve the root specimens. We aimed to use new techniques to facilitate molecular sequence data to screen this historic formalin preserved material for *M. enterolobii*.

Methodology

Development of a protocol for obtaining DNA from formalin-fixed root knot nematodes

Specimen acquisition : We obtained historic, formalin-fixed, root-knot nematode infested root material from the most relevant collection for the purposes of our project, the Northern Territory Plant Pathology Herbarium (DNAP). We also obtained a large collection of infested root material from the New South Wales Plant Pathology Herbarium (DAR). Although we contacted a total of 19 collections at the onset of this project (Table A1); only the two above institutions had formalin-fixed infested root material in sufficient quantities such that loanable material could be destructively sampled. Additionally, prior to the onset of the project we fixed subsets of root-knot nematode infested roots obtained from the Queensland Department of Primary industries and the study of Bond et al. (2022) in 10% unbuffered formalin, which, across the course of the project, made samples fixed in formalin from between one and two years available for study.

Development and validation of primers: Because long-term formalin preservation fragments DNA, the standard DNA barcodes used for root knot nematodes are not suitable for the study of nematode material which has preserved in formalin for a long period of time. For example, the standard primer pair used to amplify the cytochrome oxidase subunit 1 gene (cox1), JB3/JB5 (Bowles et al. 1992; Derycke et al. 2005) amplifies a fragment approximately 390 base pairs long. However, in animal tissues which have been preserved in formalin for a long period of time, only DNA fragments which are less than 150 base pairs are generally amplified successfully. Thus, new primers needed to be developed which would amplify a shorter section of root-knot nematode DNA, and which could discriminate *Meloidogyne enterolobii* from other root-knot nematode species. We screened three different mitochondrial genetic barcodes which have been shown to have the diagnostic power to distinguish *M. enterolobii* from other common root-knot nematodes (Janssen et al. 2016): cox1, cox2 and NAD5. We found that both the cox2 and NAD5 genes included too many indels (large gaps in multi-species alignment) for the successful design of suitable primers. However, a region of the cox1 gene was found to be suitable for our purposes and we designed a set of primers which amplify an approximately 120 base-pair fragment: Forward Primer 'Formelo-F' (GGG CGA AAA GGA ATA ATT TTA GCA ATT) and reverse primer 'Formelo-R' (GTC GCT GAA CTA AAA TAA GCA CGA).

The Formelo-F / Formelo-R primer pair was validated by amplifying previously extracted genomic DNA of *M. enterolobii* obtained during the study of Bond et al. (2022) and from genomic DNA of specimens of *M. arenaria*, M. *hapla*, *M. incognita* and *M. javanica* obtained from the Queensland Department of Primary Industries. A PCR cycling optimization revealed that the primer set was robust under a wide variety of amplification conditions. The following PCR cycling conditions were used for downstream protocol development: an initial denaturation step at 94°C for four minutes followed by 35 cycles of 30 seconds at 94°C, 30 seconds at 52°C, then 30 seconds at 72°C, then concluding with a five minute extension step at 72°C. Amplification was successful for all species using the above PCR cycling conditions. We further validated the approach by Sanger Sequencing amplicons of *M. enterolobii* to ensure diagnostic differences in the sequence were obtainable. The primers and sequencing performed as expected. Thus, this primer set was selected for downstream protocol development.

Initial screening: Acquired samples were initially screened for pH and formaldehyde content to ensure they were suitable for the development of the protocol. The pH of samples was measured with pH-Fix pH Indicator Strips (Macherey-Nagel) on the original storage medium or residual liquid in vials. Readings recorded to nearest 0.5 pH unit. In total, 341 samples (20 from NT and 321 from NSW collections) were pH tested. Free formaldehyde was measured with Supelco HCHO test strips. 112 specimens were tested (20 from the Northern Territory and 92 from the New South Wales Collection to verify formaldehyde preservation. Results were recorded following manufacturer band equivalents.

Tissue preparation: To recover nematodes for testing, female root knot nematodes were dissected from roots under a stereomicroscope using fine tweezers and a scalpel. Nematodes were rehydrated and washed in 1X phosphate-buffered saline (PBS) for 48–72 h at room temperature with at least two buffer changes over this period. Early trials used cutting or crushing to aid downstream tissue lysis; later refinements omitted mechanical disruption after chemical lysis and heat were optimized.

Testing DNA extraction methods: We evaluated multiple DNA extraction protocols that differed in tissue lysis and cross-link reversal steps. Most extractions used 1–50 nematodes; a subset used >100 individuals.

- The method described by Appleyard et al. (2021) involved overnight proteinase K digestion (4 mg/ml total conc.), followed by incubation at 80 °C for two hours for reverse crosslinking. DNA extraction was then performed using the Promega ReliaPrep FFPE gDNA Miniprep System kit.

- A custom protocol was developed, incorporating heat treatment with a retrieval solution and tissue lysis. Samples were Incubated at 80-100 °C for up to three hours in citric acid and Tris-based solutions. Lysis was carried out using a modified "worm lysis buffer" (100 mM Tris-HCl, pH 8.0; 50 mM EDTA, 100 mM NaCl, 1% SDS, Proteinase K at 0.3 mg/mL) and Qiagen ATL lysis buffer. Various lysis durations were tested, ranging from a few hours to overnight digestion, along with different proteinase K concentrations (0.3 – 4 mg/ml) and SDS concentrations (up to 2%). DNA extraction was conducted using the Promega ReliaPrep FFPE gDNA Miniprep System kit and the Qiagen DNeasy Blood & Tissue Kits with Zymo-Spin IC binding columns. To enhance lysis efficiency, slight modifications were introduced to the Qiagen extraction protocol such as a 15-minute incubation in buffer AL at 70 °C.

- The HiTE (highly concentrated tris-mediated DNA extraction) method, following Oba et al. (2022), was also tested. This method featured a tissue lysis phase, including proteinase K digestion (3.6 mg/ml total conc.), followed by a 24h incubation at 80 °C in the HiTE buffer. The proteinase K treatment was extended (overnight digestion) to allow for complete lysis. DNA extraction was performed using Qiagen DNeasy Blood & Tissue Kits with Zymo-Spin IC binding columns.

- The hot alkaline lysis method, as described by Campos and Gilbert (2012) and Hahn et al. (2022) was also evaluated. Samples were placed in PCR strip tubes containing alkaline lysis buffer (0.1 M NaOH with 1% SDS) and heated to 100 °C for up to two hours (in a thermocycler). Different NaOH concentrations (0.1-0.5 M) and SDS concentrations (up to 2%) were tested. To further improve tissue lysis, after heat treatment, samples were neutralized using 1M Tris-HCl pH 8 and subjected to a 3-hour to overnight proteinase K digestion (up to 2.5 mg/ml total conc.). DNA was extracted using the Qiagen DNeasy Blood & Tissue Kits with Zymo-Spin IC binding columns.

Proof of concept and control: The use of spin columns streamlines extraction and reduces handling losses. To confirm that Zymo-Spin IC columns retain highly fragmented DNA, extracts were subjected to random-hexamer priming and the amplified products were visualised on agarose gel. Successful DNA recovery was verified by PCR of ~60 bp and/or ~120 bp targets using custom primers, followed by gel electrophoresis. Extraction blanks and no-template PCR controls were included to monitor contamination.

Methods – AI model

Specimen acquisition and preparation: The Queensland Department of Primary Industries (QDPI) maintains cultures of four species of root-knot nematode (RKN) on tomatoes: *Meloidogyne arenaria*, *M. hapla*, *M. incognita* and *M. javanica*. Infested tomato root material was obtained from QDPI for each of these four species. Root-knot nematode females were dissected from the roots under a stereomicroscope using micro-forceps. Perineal patterns were prepared for the four species sourced from QDPI following Hartman & Sasser (1985). Perineal patterns of *M. enterolobii* from the Northern Territory, prepared as part of the study of Bond et al. (2022), were also obtained from the CSIRO Australian National Insect Collection.

Specimen imaging, curation and inclusion: Perineal patterns were imaged on a Leica DM6 compound microscope under brightfield conditions. For each specimen, images were taken at multiple depths to acquire a z-stack and then focus-fused

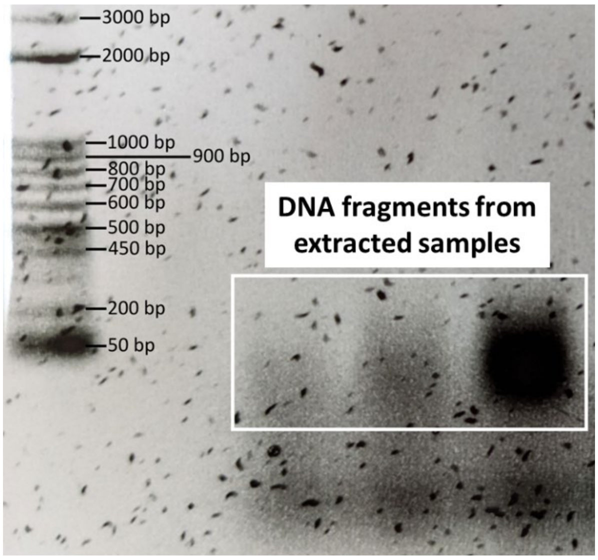

Figure 1. Random-hexamer amplification of DNA extracted from female root knot nematodes dissected from formalin-fixed plant roots visualised through gel-electrophoresis.

to a single image with the open source software PICOLAY (www.picolay.de). Two stacking regimes (views) were used: a shallow view with a narrow slice window to emphasise dorsal-arch and striae edges (comprising 5–15 z-slices) and a deep view with a wider slice window to retain additional peripheral and contextual detail (comprising 15–30 z-slices). We used manual screening to remove images with poor focus, glare, debris or damage. An RGB colour profile was used for compatibility with standard downstream processing pipelines.

Regions of interest (ROI) cropping, Pre-processing, layout, and split: Perineal areas were annotated in YOLO format with open source Labellmg/Label studio software and used to crop each image to the ROI (the perineal pattern). ROI annotations were retained for potential future automation (e.g., auto-stacking/auto-cropping). Cropped images were centre-padded to square and resized to 384 x 384 pixels. A stratified 70/15/15% split by species was used for train/validation/test (fixed random seed). 275 total images were obtained of perineal patterns before being further split and comprised 56 images of *M. Arenaria*, 18 of *M. Enterolobii*, 60 of *M. Hapla*, 73 of *M. Incognita* and 68 of *M. Javanica*.

Model family choice ("backbones") and training approach: A model Backbone is the part of the model that learns general visual patterns (edges, curves, textures, shapes). A small classification head on top converts those patterns into one of the five species. For the purposes of our work, we evaluated the performance of three different model backbones: ConvNeXt, ResNet, and Swin Transformer.

Our Chosen training approach involved Final per-view models used Swin V2-B initialised with ImageNet pre-trained weights (starting from general image knowledge to speed learning and improve accuracy). We then compared learning rate, augmentation strength, degree of backbone fine-tuning, and class-imbalance handling, selecting the combination that maximised validation performance while limiting overfitting. We kept Inference and calibration simple using Test-time augmentation (TTA) and Probability calibration. For TTA we averaged predictions from the original image, a left–right flip, and small ±5° rotations to improve stability. For probability calculation we used a simple temperature scaling fitted on the validation set was applied so predicted probabilities reflect true likelihoods (shallow T≈1.195, deep T≈1.424).

Development environment : We moved from an initial trial using AutoML (Microsoft Azure) to a code-first workflow in Cursor to fully control curation and splits, implement and adjust our own training code, keep exact versions of scripts and settings, and run locally or offline when needed.

Results and discussion

Results - Formalin DNA extraction protocol

Tissue screening: Most historic samples that we tested were acidic (pH 3.5–5.0) whereas recently prepared experimental material fixed in 10% formalin for ≤2 years typically measured ~pH 5 (Fig. A1). Of the 112 specimens tested with HCHO strips (20 NT, 92 NSW), 101 were positive for formaldehyde (18 NT, 83 NSW) and 11 showed no detectable formaldehyde, indicating formaldehyde preservation was common in these collections.

Spin-column proof-of-concept: Random-hexamer amplification of extracts followed by agarose gel electrophoresis confirmed that Zymo-Spin IC columns retain highly fragmented DNA. Strikingly, specimens held in unbuffered formalin for as little as ~2 years already showed severe fragmentation, with the bulk of DNA fragments being less than 200 base pairs in length (Fig. 1). This result underscores how rapidly damage accrues under unbuffered formalin preservation.

Prolonged exposure to formaldehyde leads to extensive DNA-protein crosslinking, making DNA extraction increasingly difficult. Additionally, long-term storage in formaldehyde-based solutions results in acidification, as formaldehyde is gradually converted into formic acid, creating an increasingly unstable environment for DNA preservation. This issue is particularly pronounced in specimens preserved using FAA (formaldehyde, ethanol, acetic acid), a fixative historically employed in nematode research, resulting in extremely low pH levels that further compromise DNA integrity. A further complication arises from the common practice of preserving whole roots rather than isolating nematodes and preserving them on their own, as plant roots can release organic acids, contributing to acidification. pH measurements of our samples confirm that this is not merely a theoretical concern. All tested specimens had pH values below 6, with some as low as 2–3, demonstrating the severity of these preservation challenges. Additionally, many samples have uncertain treatment histories, further complicating DNA extraction. Traces of formaldehyde were detected in specimens recorded as ethanol-preserved, making it impossible to determine their true preservation conditions over time. This adds another layer of difficulty in selecting suitable samples for DNA analysis.

In recently fixed test material (*M. incognita* and *M. javanica* in unbuffered 10% formalin ≤ 2 years), hot alkaline lysis followed by Proteinase K was the most effective and relatively simple approach, yielding amplifiable DNA from single nematodes (Fig. 2).

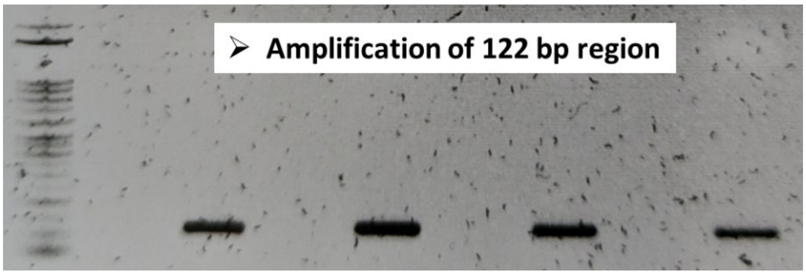

Figure 2. Successful amplification of DNA from root-knot nematodes dissected from infested roots preserved in 10% unbuffered formalin for 1–2 years.

In contrast, applying the same and related protocols to 28 historic, formalin-preserved specimens (1–200 nematodes per extraction) produced no recoverable DNA. Together with our random-hexamer gels, these results indicate that unbuffered formalin rapidly degrades DNA, with severe fragmentation evident by ~2 years, and progressive acidification likely rendering older material irretrievable.Our finalized protocol was provided in previous project milestone reports and is herein provided again as Appendix 2.

Results of development of AI model

We found that of the three backbones evaluated here (ResNet, ConvNeXt, and Swin families) Swin V2-B consistently gave the best validation and test performance (especially for shallow view), so it was chosen for the final models. It was also determined that heavier augmentation reduced performance; moderate augmentation plus simple calibration resulted in the best performance. Temperature scaling (shallow T≈1.195, deep T≈1.424) and Test-time augmentation (original + H-flip ±5°) yielded a small but consistent gain in probability calibration. It was also determined that a class-balanced sampler generally outperformed loss re-weighting, improving minority-class stability without harming major classes (species IDs).

When evaluating a shallow vs deep view model, it was found that the shallow model performed better. The shallow view model (Fig. A2) using Swin V2-B yielded an identification accuracy of 97.56% (40/41 diagnoses correct), whereas the deep view model (Fig. A3) yielded an identification accuracy of 90.24% (37/41 diagnoses correct). The shallow model showed only a single error, where a specimen of *Meloidogyne hapla* was mistaken for *Meloidogyne arenaria*. In the deep model errors were mostly specimens of *M. hapla* being confused for *M. javanica* and vice versa, with more rare *M. arenaria* and *M. javanica* mix ups. The AI models evaluated rely on three standard classification metrics: Precision, Recall and F1. Precision refers to how many predictions were correct. Recall means of all true images of a species, how many were found. F1 is a single score that balances precision and recall. The shallow model showed higher overall precision and recall and associated F1 (Table A2).

Notably, *Meloidogyne enterolobii* was never confused for any other species, nor were any other species confused for *M. enterolobii*, in any of our tested scenarios. Furthermore, in all three standard classification metrics (precision, recall, F1), *M. enterolobii* scored perfectly (1.0). Thus, the model has a robust ability to diagnosis this species in particular.

To better understand which elements of the perineal patterns were being used by the AI model to yield an identification probability (saliency), we used RISE (*Randomised Input Sampling for Explanation*) (Petsiuk et al. 2018) saliency maps to show which regions supported each prediction (Fig. 3). From these we are building class-level “atlases” by averaging per-image saliency and consistency (frequency of top-k salient pixels) to visualise species-defining regions used by the AI.

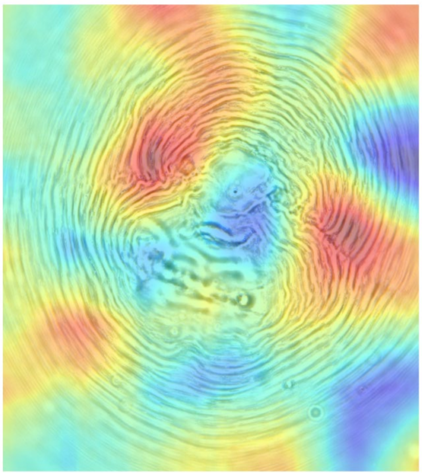

Figure 3. Correctly classified perineal pattern of *Meloidogyne arenaria* using AI diagnostic model. Heat map emphasizes the dorsal arch and lateral striae which are features used by human experts for identification. Red regions are those the model relies on most for identification (high saliency); Blue regions are those that contribute little or nothing to an identification probability (low saliency).

Discussion.

Methodology for extraction of root-knot nematode DNA from formalin-fixed root material

Formalin fixation remains the standard method for preserving nematodes and nematology-material (i.e., infested roots), however, such fixation renders DNA inaccessible due to protein-DNA cross-linking. Our development of the alkaline lysis method for extracting viable nematode DNA from formalin-fixed root-knot nematodes embedded in roots offers a useful tool for addressing a unique challenge in plant pathology. The methods developed here are not merely for nematology; they provide the groundwork for a broad strategy for obtaining DNA from historic plant pathogens.

We identified a sharp decline in DNA extraction efficiency beyond two years. This issue seems not to be due to the formaldehyde itself, rather it appears due to increased acidification over time due to the release of organic acids from the cells of the fixed plant roots. This acidification likely inhibits the alkaline lysis reagents from effectively de-crosslinking the nucleic acids, thereby inhibiting successful PCR. This acidity also further degrades and fragments DNA over time. In many cases, highly acidified material will have no obtainable DNA left.

Our findings that formalin-fixed plant root material leads to acidification of solution over time suggests that pests and pathogens associated with formalin fixed plant root material may have a short molecular shelf life. Hence, we have developed and promoted changes to standard protocols aimed at minimizing this phenomenon where formalin must be used as a fixative. Where formalin is required, the solution should be adequately buffered to prevent acidification using a high-capacity buffering agent such as Tris. This will extend the period that useful DNA is available from nematode samples fixed using standard methods and mean more samples will be available for molecular identification in the future.

The alkaline lysis protocol has clear positive implications for horticultural and agricultural industries. This method provides a valuable molecular tool which can unlock previously inaccessible formalin archives. This capability can enable retrospective studies across multiple pathogen types (e.g., nematodes, insects, fungi) facilitating mapping of the historical spread of economically important pests. Although we found that formalin-fixed root material becomes acidified over time inhibiting molecular work, this finding better informs stakeholders on how to best preserve samples for long term use. Furthermore, plant pathology materials of other plant tissue types may still have viable DNA after long periods of time. Where formalin fixation is required, using adequate buffering can significantly extend the molecular shelf life of plant pathology material.

Development of an AI model for identification of root-knot nematodes

As a complementary approach to our protocol for extracting DNA from formalin-fixed material, we also sought to develop an AI-driven diagnostic model for the identification of root knot nematodes via the perineal pattern. This model has proven effective on morphological specimens of all ages tested, overcoming the acidification-related age limitations observed in the formalin-fixed molecular samples. Importantly, the library of perineal pattern slides in Australian collections is far more extensive than formalin-preserved root material, providing enhanced opportunity for rapid and iterative re-evaluation of an extensive set of historic root knot nematode specimen records. Using this system we have found no evidence in available perineal pattern slides that *M. enterolobii* arrived in Australia prior to the 2022 report (Bond et al. 2022). Notably, this includes perineal patterns we produced based on material from the Northern Territory preserved about 15 years in the past (2010–2011); see Singh et al. 2012) – while this material was too acidified for molecular analyses, we were able to produce perineal patterns and, importantly, we have found the root knot nematodes in question from these collections were not *M. enterolobii*. This suggests limited initial spread and provides a baseline for future management actions and efforts.

Although we initially developed the AI system to screen Australian collections for historic presence of *M. enterolobii*, we quickly found that the AI model is likely to be a highly valuable diagnostic tool for all root knot nematode species. The primary benefit of this system for horticultural industries lies in the speed in which an identification can be reached, accessibility to a very broad user base, and cost efficiency. Our AI system delivers rapid identification results with an accuracy of nearly 100% in our tests samples, suggesting it will be a potentially more reliable and less costly alternative to standard molecular methods (molecular diagnostic methods often produce a small percentage of indefinite results for field samples, so seldom achieve 100% accuracy either). Presently, the AI method requires only a basic laboratory setup including a microscope and a camera which will cater to a wide range of users beyond well-equipped diagnostic labs. Furthermore, the automation of key steps, such as diagnostic region identification and image stacking, makes the system available to non-experts. Future works will explore further increasing the accessibility of the system by developing an identification app that could be used on a smartphone with a microscope attachment to the phone camera.

Using the current version of our AI model we have screened a large number of historic perineal pattern slides sourced from the Australian National Insect Collection and Agriculture Victoria. This process has highlighted a number of potential misidentifications in historic samples which enhances our understanding of what species of root knot nematodes occurred on what crops at what times and where across a deeper time scale. The AI model works well on the five major species of root knot nematodes that we have trained the model on and tested. However, we do presently know that the model can, in a sense, recognise species for which it has not been trained. We had the model attempt identification on slides of *Meloidogyne fallax* from Victoria and found the model returned very low identification confidence—the system is recognizing that this is NOT one of the species it was trained on. This further demonstrates the robustness of the AI diagnostics approach.

It would be ideal if we could incorporate additional species into the model. There are presently two additional species reported from Australia (*M. fallax* and *M. triopholiophila*). Incorporation of these species, as well as additional exotic species, into the model will help enhance the model's reliability. The majority of the specimens used to train the model came from just a single population each. All our specimens of *M. arenaria*, *M. hapla*, *M. incognita* and *M. javanica* came from cultures that have been held over a long period of time by the Queensland Department of Primary Industries. These RKN populations have been grown on tomatoes as the only host for many years. Our specimens of *M. enterolobii* came from the initial detection in the Northern Territory (Bond et al. 2022) and are all from a single cucumber plant. There is some evidence that differences in host and environmental condition can have an impact on elements of the perineal pattern morphology in root knot nematodes. Therefore, to build a truly robust diagnostic tool which can identify root knot nematodes across crop types and different regions of Australia, we need to incorporate far more specimens into the training of our model. We require specimens of all species from multiple crop types and regions. Such work is timely as a survey of root knot nematodes across Australia is presently underway. The next step in the development of the model is to obtain as many populations of root knot nematodes as possible and use them to further refine our AI model.

Development of this AI-driven diagnostic model represents a substantial advancement for nematode diagnostics and should prove of significant benefit to Australian horticultural industries. By rapidly analysing historic perineal patterns, the model has already helped confirm that *M. enterolobii* is indeed a recent arrival. More importantly, the AI model can deliver diagnostic results with high speed and accuracy, and, coupled with its accessibility to non-experts through basic lab equipment, may offer a cost-effective alternative to molecular methods. Incorporating additional RKN species which occur in Australia such as *M. fallax* and *M. trifoliophila* and expanding the model's training data across diverse host populations and regions are the necessary next steps. However, the current AI model provides a strong foundation for future development. This method is likely to become an invaluable diagnostic tool, enabling better and more rapid management actions to protect Australian horticultural industries from root knot nematodes.

Outputs

We have authored a procedure for obtaining DNA from formalin-preserved Root-Knot Nematodes (Appendix 1). We have also now produced a large collection of female root knot nematode perineal patterns which have been mounted on microscope slides covering 5 different species (*M. enterolobii*, *M. arenaria*, *M. hapla*, *M. javanica*, *M. incognita*; see Figure 1 for example). These are held in the Australian National Insect Collection and have also been imaged. These images will be disseminated online open access through the CSIRO data access portal and/or other appropriate data depositories upon publication of a scientific paper describing the AI model and its use, together with the release of the model (free and open access). The perineal pattern slides will be accessioned and reported on in the same publication. We have also imaged 386 root-knot nematode perineal pattern specimens from the collection of Agriculture Victoria which include the species *M. arenaria*, *M. fallax*, *M. hapla*, *M. javanica* and *M. incognita*. A presentation on this work was given at the 25th Annual Australasian Plant Pathology Meeting. An extension article was published in the spring 2024 issue of Australian Grower.

Table 1. Output summary

Output	**Description**	**Detail**
White paper on	A written procedure for	The white paper detailing the procedure for extraction of DNA

procedure for extraction of DNA from formalin preserved root knot nematodes	extraction of DNA from formalin-preserved Root-Knot Nematode material, followed by amplification and sequencing of extracted DNA. Procedure has been validated.	from formalin-preserved root-knot nematodes is attached to this report as an appendix and has been uploaded to the delivery partner portal where it can be accessed by any interested stakeholders.
Collection of Perineal Pattern Slides produced, and images obtained	A large collection of root knot nematode perineal patterns has been produced and imaged.	Slides will be lodged in Australian National Insect Collection and images lodged in the publicly-available Data Access Portal after publication of paper on AI model. (This is to prevent anyone else using the images commercially without permission.)
Presentation at 25th Australasian Plant Pathology Society (APPS) Conference	A presentation was given by the project leader describing both the work on the AI model and the formalin DNA work.	A presentation was given at the Australasian Plant Pathology Society conference, May 2025. The presentation was well received and a number of insightful questions were asked.
AI model which can identify root knot nematodes to species based on perineal patterns (computer code).	Open source software was used to develop the model which consists of two parts: A YOLOv5 model which successfully extracts the 'area of interest' (the perineal pattern) from a source image and a combination of the AI software models ResNet101 and Swin Transformer which was trained to make the root knot nematode identifications based on the area of interest extracted by YOLOv5. The resulting model can take an input image and produces an identification.	The data and software code is first intended to be published in the form of a scientific paper under an open access license agreement, at which point the computer code and all images used to train the model will be openly available.
Extension article on the project in Spring 2024 issue of Australian Grower	An article detailing the project prepared in collaboration with AusVeg. The intended audience was growers and agronomists.	Article in Spring 2024 issue of Australian Grower. (https://ausveg.com.au/knowledge-hub/ai-and-dna-come-together-to-identify-nematodes/).

Photos/images/other audio-visual material

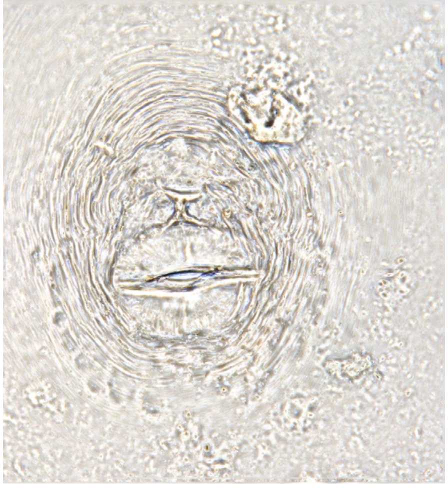

Figure 4. Example perineal pattern image generated for identification through automated identification of critical areas, and stacking multiple focal planes. This image identified with 100% certainty as *Meloidogyne hapla* (confirmed by conventional DNA sequencing). CSIRO image from specimen prepared from material sourced from QDAF cultures.

Outcomes

The outcomes of this project directly support the biosecurity, trade, and on-farm profitability of the potato, sweet potato, onion, and leviable vegetable sectors through the development of new methods for interrogating historic specimen collections to better guide management and policy actions during novel pest incursion events. This project was primarily to develop and test methods for identifying root-knot nematodes from historic samples, so a primary outcome is demonstrating that this is possible using the two complementary methods developed during the project. A by-product of the project is the image-analysis method, which could be used for species-level identification of Root-Knot Nematodes that is faster and cheaper than other methods.

1. ***Enhanced understanding of the species of Root-Knot Nematodes occurring in Australia from a historic perspective.***

   Understanding has been increased, but also some potential misidentifications have been recognized. We propose further investigating these potential misidentifications and their consequences. We have identified many more slides available for this. Using the efficient AI system for identification on these slides will further increase the understanding of Root-Knot Nematodes in Australia. Better host and occurrence records will enable better characterization of the Root-Knot Nematodes present and better management options such as planting resistant varieties, crop rotations, fallow periods required, and timing.

2. ***Development of methods for acquiring genetic information from formalin-preserved plant pathology material which can be applied to other species of nematodes and plant pests/pathogens.*** As a result of this project we have a method which is potentially applicable for identifying small eukaryotic plant pathogens from formalin-fixed material, including other nematodes, protists, fungi and oomycetes. Further testing on these groups is required to confirm this.

3. ***Enhanced understanding regarding the timing of arrival of* M. enterolobii *into Australia.*** Applying both methods developed during the project—but especially the AI image analysis method—to the historic samples located and processed so far has not produced any evidence that *M. enterolobii* was present in Australia before the first detection in 2022. This suggests the spread is likely limited, but more samples have been located recently, which are being tested currently. Further searching for historic slides and recent samples will increase confidence, as well as clarifying what might be required to manage the pest in the future, such as minimizing spread.

4. ***More accurate and rapid method of species-level identification of Root-Knot Nematodes via morphological methods. This method would also be less costly per unit than molecular methods.*** Method developed with nearly 100% accuracy for test samples. (For every method, including standard molecular and morphological identification methods, there are a small percentage of ambiguous or anomalous results: these are being investigated as part of a separate HIA project PW22000.) With validation for a wider geographic and host data set, this could be turned into a public-facing interface available free to Australian diagnosticians and growers. Where older historic samples are unsuitable for identification using DNA because of fragmentation, there is the opportunity to use these samples for identification using the imaging system, even though the specimens were not originally collected or processed for morphological identification methods. The current work has also opened the prospect of identification of RKN species from completely unprocessed specimens: that is, whole females still in the roots. At the moment, all species identification methods require extraction of adult females from plant roots, which is time-consuming, costly and painstaking. The success of AI recognition of the area of the nematode critical for identification, plus the automated focal-plane stacking suggests that completely automated ID of whole nematodes from the roots may be possible. This would represent a quantum advance in identification allowing fast and accurate identification of RKN species directly from roots by growers with only a basic magnification and imaging system (mobile phone with clip-on lens?).

5. ***Improved understanding of the distribution of root-knot nematode species, including* M. enterolobii*, enabling better methods of management.*** Our understanding of the distribution of Root-Knot Nematodes in Australia has increased

to the extent that we now understand that there is considerable complexity in the host and geographic distributions of Root-Knot Nematodes in Australia. Consequently, rapid, cheap methods for accurate identification to species such as those developed in the present project will need to be deployed to determine what is likely to occur where (on a farm or paddock scale), on what hosts, and where it is likely to move.

6. ***Positive impacts on trade; both import and export. The proposed technologies would be useful for area freedom designations, and if provided to trade partners, could assist in pre-identification of species before import, thus enhancing Australian biosecurity.*** This outcome will be realized with additional validation and development of a public-facing interface and deployment. Value will be added to the currently-planned extensive national collections of Root-Knot Nematodes by validation and analysis using these methods.

Table 2. Outcome summary

Outcome	**Alignment to fund outcome, strategy and KPI**	**Description**	**Evidence**
Development of methods for acquiring genetic information from formalin-preserved plant pathology material which can be applied to other species of nematodes and plant pests/pathogens	Provides a tool for vegetable industries (potatoes, onions, sweet potatoes) to track past distribution and timing of key pest incursions such as RKNs, providing the historical context required for pest management strategies, trade, and area freedom designations.	The method has high transferability across various preserved plant pests/pathogens. This method enhances diagnostic capability and strengthens national biosecurity frameworks by enabling faster identification of emerging threats.	Draft of procedure for extraction DNA from formalin preserved root knot nematode material authored (appendix 1)
Shorter diagnostic barcode sequence	Provides a critical tool for identification of *M. enterolobii* from formalin fixed material.	These primers were developed to amplify a short barcode of the cox1 gene which is required when dealing with highly fragmented formalin fixed material.	Primer set provided in results and as part of Appendix 2.
More accurate and rapid method of species-level identification of Root-Knot Nematodes via morphological methods.	AI-driven RKN identification method will enable faster management decisions (e.g., resistant varieties, rotation), boosting productivity. It will reduce diagnostic costs and strengthen biosecurity and trade integrity.	AI method is low cost and accurate, facilitating RKN diagnostics essential for maximizing on-farm productivity, minimizing grower costs, and enhancing biosecurity and export confidence.	AI image recognition model developed and validated; this model achieves high accuracy for species level identification on our test images
Enhanced understanding regarding the timing of arrival of *M. enterolobii* into Australia; Enhanced understanding of the species of Root-Knot Nematodes occurring in Australia from a historic perspective	We have found no evidence that M. *enterolobii* has occurred in Australia for any length of time suggesting recent arrival. This knowledge is vital for targeted containment and protecting levy-crop productivity.	Lack of data suggesting long term establishment of *M. enterolobii* indicates that the species is not yet widespread and that efforts should be made to contain the species	Testing of historic material stored in formalin as well as perineal patterns produced from specimens collected in the Northern Territory 15 years ago (2010-2011) indicated the RKNs occurring at that time were not *M. enterolobii*.

Monitoring and evaluation

Table 3. Key Evaluation Questions

Key Evaluation Question	Project performance	Continuous improvement opportunities
Effectiveness **- To what extent has the project achieved its expected outcomes?** Has the project provided an improved understanding of the timing of arrival of the guava root knot nematode into Australia?	Yes. We evaluated historic material, including slides of perineal patterns and formalin-preserved roots showing signs of RKN infestation from the Northern Territory and other Australian states and found no evidence that *M. enterolobii* has occurred in Australia prior to the initial report in 2022.	The methods developed to evaluate this question have significant future applicability to future questions of a similar nature. The AI method can be used to rapidly screen collections for previously unrecognized species so that in future incursions we can quickly answer questions regarding timing of arrival. The method developed for extracting DNA from formalin will be applicable to other nematode species and pests/pathogens so formalin preserved specimens can also be evaluated in incursion situations.
Effectiveness **- To what extent has the project achieved its expected outcomes?** Has the project provided a better understanding of the historic distribution of root knot nematodes in Australia?	Yes. We have found no evidence of *M. enterolobii* in Australia prior to the 2022 report and we have also highlighted a number of misidentifications among historic collections of root knot nematodes providing a better understanding of what species occurred where and when.	As we continue to improve upon the AI model we are likely to further enhance our understanding of what species occur where and when in time. Notably, once imaged, perineal pattern slides can be quickly re-evaluated at any time allowing continuous refinement of our understanding of distribution across time as we further refine the model.
Effectiveness **- To what extent has the project achieved its expected outcomes?** Has the project validated a method for extracting DNA from formalin-fixed root knot nematode material?	Yes. We have developed a protocol that is successful at extracting DNA sufficient for diagnosis of RKNs from formalin-preserved material.	This method should have wide applicability for other pests and pathogens and should be evaluated against other kinds of formalin preserved material such as nematodes on slides. This will require some additional R&D. Development of methods to extract DNA from dried material (such as herbarium specimens) would be advantageous.
Effectiveness **- To what extent has the project achieved its expected outcomes?** Has the project determined the suitability of AI for the identification of root knot nematodes?	Yes. We have developed an AI model that can successfully identify five different species of RKN based on the perineal pattern demonstrating the promise of this method.	Our AI model is strikingly successful – it shows significant promise as a fast and cheap tool for rapid RKN diagnosis. Further development is needed to transform the current model into a truly robust tool that can identify all species of RKN occurring in Australia (and potentially exotics) – to do this we must incorporate material from as many species/locality/crop type combinations as possible.

***Relevance* - How relevant was the project to the needs of intended beneficiaries?** To what extent has the project met the needs of industry levy payers?	We see our project results as addressing two important needs of the agricultural industries impacted by root knot nematodes. 1) farm biosecurity, i.e., keeping out new pests and diseases that impact productivity and increase trade barriers and 2) low-cost and rapid pest diagnostics which will assist growers to make more proactive integrated pest management decisions. We think our project has contributed to both these needs. Our methods will help us better understand the historic context of pest incursions thus helping to determine high level policy management decisions (attempt eradication, move to management etc.). Our AI method was initially designed to contribute to this but also has the future side benefit of being a lower cost and faster diagnostic method for RKN that anyone should be able to use, given some more R&D.	There is a timely opportunity to continue to improve upon the AI model as a diagnostic tool – we must integrate as much material from across Australia into the model so it can be a tool with a broad reach. The ability to extract DNA from formalin preserved plant pathology material also presents further opportunities to obtain a better historic context for other pests and diseases where relevant.
***Process appropriateness* - How well have intended beneficiaries been engaged in the project?** To what extent were the target engagement levels of industry levy payers achieved?	The present project was primarily a scoping project designed to evaluate the feasibility of the methods described here for answering questions about historic occurrence of plant pests, specifically root knot nematodes. Thus, the target audience for this work was primarily diagnosticians and researchers. We have engaged with most plant pathology diagnostic labs across Australia during the course of this study, requesting specimen which we could test with this method. We have also communicated the research associated with this project at several relevant meetings such as the Australasian Plant Pathology society meeting which is attended by many agricultural industry professionals. We did engage with levy payers directly through publication of a short article on this in the spring 2024 issue of Australian Grower.	Further development of the AI diagnostic model into a public facing interface that can be used by anyone, including diagnosticians, agronomists and even growers themselves, will ultimately directly engage with the levy payers.
***Efficiency* - What efforts did the project make to improve efficiency?** What efforts did the project make to improve efficiency?	We increased Team efficiency by having informal meetings about the project every few days to discuss project progress and how various members could contribute to progress. For development of the	In refinements to the AI model, obtain further assistance from IT and AI experts in CSIRO to help increase efficiency of computer code and experimentation.

	formalin-DNA extraction method, we continuously refined our protocols and used parallel experiments to increase the efficiency of our experiments during lab-use hours. For the AI model we relied extensively on expert expertise from other CSIRO research units (who graciously donated their time free of charge) to help select appropriate AI software and to help increase the efficiency of the underlying computing code process.	

Recommendations

Recommendation 1. Enhance collections and sample preservation.

Australian specimen collections have the potential to provide critical insights into the timing of arrival and historic distributions of pests and pathogens. However, when dealing with putatively known and established pests, diagnosticians, researchers and other stakeholders rarely voucher specimens plant pests and pathogens in Australian collections. Typically, only new records or significant range extensions result in the vouchering of pest and pathogen specimens. Gaps in Australian collections due to insufficient vouchering inhibits our ability to perform retrospective studies of pests and pathogens such as the one described here. We recommend that pest and pathogen specimens are submitted to Australian collections more regularly, even when the organisms in question are considered common or established. For example, plant pathology material can be submitted to the NSW Plant Pathology Herbarium and nematology material can be submitted to the CSIRO Australian National Insect Collection. This will enhance our ability to understand changes in pest and pathogen distributions and hosts over time. One of the core findings of our study was that the use of formalin to preserve root-knot nematode infested root material led to acidification of the solution over time, leading to DNA degradation. For future collections we recommend root-knot nematode infested roots be preserved in ethanol. However, if formalin must be used then it should be highly buffered to enhance the molecular shelf life of the sample.

Recommendation 2. Invest in AI for root-knot nematode diagnostics.

The AI model we developed shows very high accuracy for the diagnosis of common root-knot nematode species. To make this AI method a robust and reliable tool for diagnostics of root knot nematode we need to broaden the AI model's diagnostic power through Phase 2 training. We need root-knot nematodes from across Australia, including as many species, crop types, and climatic regions, as possible. Each population of nematodes will be grown in the lab so that we can use an integrated taxonomic approach to ensure we know the taxonomic identity of each population with high confidence. These nematodes can then be grown on different crop types to ensure we capture variation within species and host-induced variation. From there we can produce an extensive set of perineal patterns that will cover all morphological variation within and among species and use it to re-train the model. Expanding the training data will make the AI model highly accurate across all species no matter the host plant or location.

Recommendation 3. Take the AI model to the Next Generation.

One of the most valuable and exciting elements of AI-based pest diagnostics is providing low-cost and rapid diagnostics without the need for a specialised laboratory set up. The present AI model requires only a very basic laboratory set up, but it is possible that these requirements could be even further simplified. Future work should explore the development of a smartphone-based App that could use a simple microscope attachment to allow root-knot nematode diagnostics in the field.

Refereed scientific publications

None as yet to report.

References

Appleyard SA, Maher S, Pogonoski JJ, Bent SJ, Chua XY, McGrath A. 2021. Assessing DNA for fish identifications from reference collections: the good, bad and ugly shed light on formalin fixation and sequencing approaches. *Journal of Fish Biology* 98: 1421-1432.

Bond SJ, Huston DC, Patel S, Hodda M, Yadav S, Bellgard SE. 2024. Scientific data and background behind the first detection of *Meloidogyne enterolobii* in Australia. *Australasian Plant Disease Notes* 19: 16.

Bowles J, Blair D, McManus DP. 1992. Genetic variants within the genus *Echinococcus* identified by mitochondrial DNA sequencing. *Molecular and Biochemical Parasitology* 54: 165-173.

Campos PF, Gilbert TM. 2011. DNA extraction from formalin-fixed material *Ancient DNA: methods and protocols*: Springer. 81-85.

Derycke S, Remerie T, Vierstraete A, Backeljau T, Vanfleteren J, Vincx M, Moens T. 2005. Mitochondrial DNA variation and cryptic speciation within the free-living marine nematode *Pellioditis marina*. *Marine Ecology Progress Series* **300:** 91-103.

Hahn EE, Alexander MR, Grealy A, Stiller J, Gardiner DM, Holleley CE. 2022. Unlocking inaccessible historical genomes preserved in formalin. *Molecular Ecology Resources* 22: 2130-2147.

Hartman KM, Sasser JN. 1985. Identification of *Meloidogyne* species on the basis of differential host test and perineal pattern morphology *An Advanced Treatise on Meloidogyne. Volume II. Methodology.* 69-77.

Janssen T, Karssen G, Verhaeven M, Coyne D, Bert W. 2016. Mitochondrial coding genome analysis of tropical root-knot nematodes (*Meloidogyne*) supports haplotype based diagnostics and reveals evidence of recent reticulate evolution. *Scientific Reports* **6:** 22591.

Oba U, Kohashi K, Sangatsuda Y, Oda Y, Sonoda K-H, Ohga S, Yoshimoto K, Arai Y, Yachida S, Shibata T. 2022. An efficient procedure for the recovery of DNA from formalin-fixed paraffin-embedded tissue sections. *Biology Methods and Protocols* **7:** bpac014.

Petsiuk V, Das A, Saenko K. 2018. Rise: Randomized input sampling for explanation of black-box models. *arXiv preprint arXiv:1806.07421*.

Singh SK, Conde B, Hodda M. 2012. Root-knot nematode (*Meloidogyne incognita*) on bitter melon (*Momordica charantia*) near Darwin, Australia. *Australasian Plant Disease Notes* **7:** 75-78.

Intellectual property

Project IP includes images of prepared perineal patterns mounted on microscope slides, the computer code used for training of the AI model and the protocol for obtaining DNA from formalin fixed material.

The following slides and images were created: *M. arenaria* – 120 images of 59 specimens; *M. enterolobii* – 45 images of 20 specimens; *M. hapla* – 163 images of 66 specimens; *M. incognita* – 164 images of 64 specimens; *M. javanica* – 147 images of 61 specimens. In total we obtained 639 high-quality stacked images of root knot nematode perineal patterns derived from the above slides (598 used for model development, 41 used for verification) and 386 high-quality stacked images of root knot nematode perineal patterns obtained from slides loaned by Agriculture Victoria.

Acknowledgements

We are grateful to the Queensland Department of Primary Industries Team (Dylan Corner, Tim Shuey, Wayne O'Neil) who maintain cultures of root knot nematodes and who graciously provided this material to us for the purposes of this work. We thank John Wainer of Agriculture Victoria for lending a substantial collection of perineal pattern slides to us and Jordan Bailey of the New South Wales Plant Pathology Herbarium for facilitating a collection visit and permitting us to subsample a significant collection of formalin-preserved root material. We thank Samantha Bond (at the time with Northern Territory Department of Industry, Tourism and Trade) for providing us with material from the Northern Territory Plant Pathology Herbarium (DNAP), and Melanie Ford and Sharl Mintoff (Northern Territory Department of Industry, Tourism and Trade) for some follow-up specimens. We thank Erin Hahn (Australian National Wildlife Collection) for assistance and advice regarding development of the formalin-DNA extraction protocol.

Appendices

Appendix 1. Additional Data from Methods and Results of project.

Table A1. Australian collections and other institutions contacted at the onset of the project in search of material for the present study.

New South Wales
NSW Plant Pathology Herbarium (DAR) - Orange
Australian Museum
Queensland
Biological Crop Protection Pty. Ltd.
Plant Pathology Herbarium (BRIP)
Queensland Museum
QDAF
South Australia
SARDI
South Australian Museum
Western Australia
Department of Agriculture Western Australia Plant Pathogen Collection (WAC)
Western Australian Museum
AGWEST Plant Laboratories, Perth (Now called DDLS - Plant pathology services)
Victoria
Plant Pathology Herbarium (VPRI)
Museum Victoria
Northern Territory
Northern Territory Plant Pathology Herbarium (DNAP)
Museum and Art Gallery of the Northern Territory
Tasmania
Tasmanian Agricultural Insect Collection
Tasmanian Institute of Agricultural Research
Tasmanian Museum and Art Gallery, Hobart (TMAG)
Queen Victoria Museum and Art Gallery, Launceston (QVMAG)

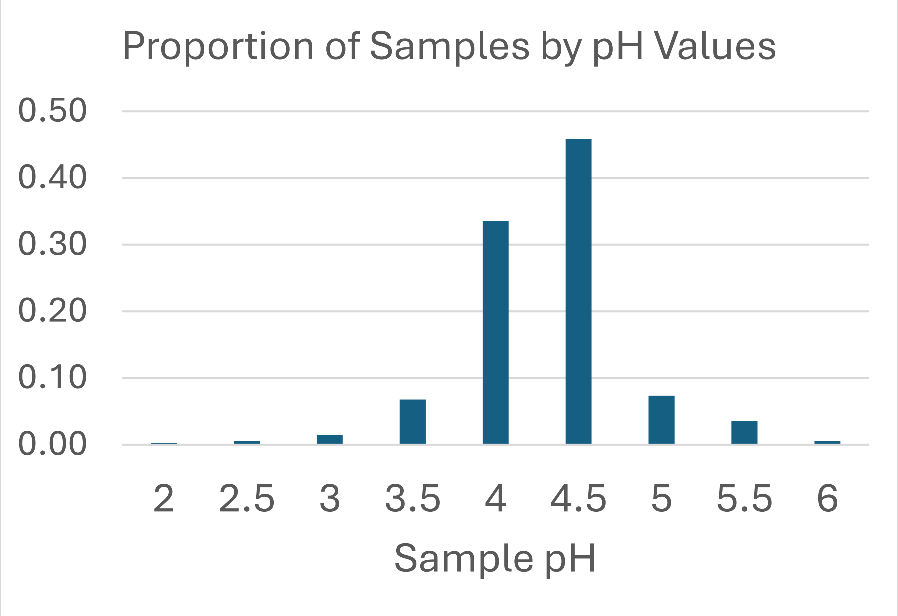

Figure A1. Summary of pH of samples tested across life of project. The majority of the historic samples tested were acidic.

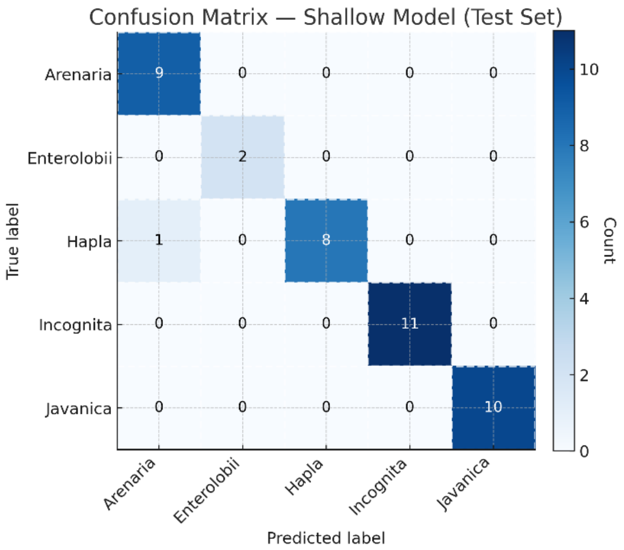

Figure A2. Confusion matrix detailing identity diagnoses of five species of root knot nematode based on the ‘shallow’ model using the Swin V2-B backbone (rows = true, cols = predicted; Arenaria, Enterolobii, Hapla, Incognita, Javanica). Values on the diagonal represent agreement of the AI model with confirmed identifications. Note that there is only a single non-zero value off the diagonal.

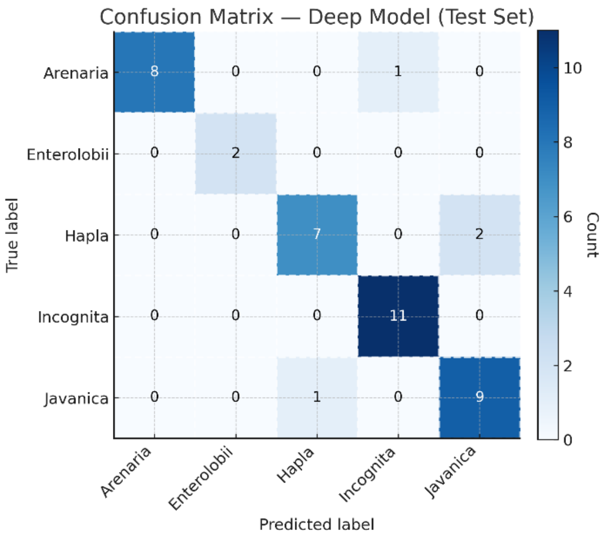

Figure A3. Confusion matrix detailing identity diagnoses of five species of root knot nematode based on the ‘deep’ model using the Swin V2-B backbone (rows = true, cols = predicted; Arenaria, Enterolobii, Hapla, Incognita, Javanica). Note more non-zero values off the diagonal. (This is not the final recommended model.)

Table A2. Performance of the shallow and deep AI Swin V2-B models at diagnosing five species of root knot nematodes using in terms of precision, recall and F1.

Species	Precision	Recall	F1	Precision	Recall	F1

	(Shallow)	(Shallow)	(Shallow)	(Deep)	(Deep)	(Deep)
Arenaria	0.9	1	0.9474	1	0.8889	0.9412
Enterolobii	1	1	1	1	1	1
Hapla	1	0.8889	0.9412	0.875	0.7778	0.8235
Incognita	1	1	1	0.9167	1	0.9565
Javanica	1	1	1	0.8182	0.9	0.8571
Macro avg	0.98	0.9778	0.9777	0.922	0.9133	0.9157

Appendix 2. Protocol: DNA Extraction from Formalin-Fixed Nematodes Using Alkaline Lysis and Proteinase K Treatment.

Introduction

There is significant data of biosecurity relevance locked inside historic specimens stored in collections. Much of this material was preserved using the chemical preservative formalin. It is generally presumed that you cannot obtain DNA from formalin-preserved material, but several methods exist, including some developed by CSIRO (see Hahn et al. 2022). Many collections hold plant pathology specimens, including root material that was identified as being infested with Root-Knot Nematodes. In many cases such material was identified no further than '*Meloidogyne* sp.', and in most cases a formaldehyde-based solution was used to preserve the root specimens. The following protocol was designed to extract DNA from adult female root knot nematodes preserved in this way.

Procedure

Materials & Reagents Required

- Lysis Buffer (1 mL)
    - 100 µL of 1 M NaOH
    - 50 µL of 20% SDS
    - 850 µL of nuclease-free water
- Neutralization Buffer (1 mL, 100 mM Tris-HCl, pH 8.0)
    - 100 µL of 1 M Tris-HCl (pH 8.0)
    - 900 µL of nuclease-free water
- Specimens
    - Dissect roots to obtain adult female nematodes

Protocol Steps

Step 1: Sample Preparation

1. Place 1 nematode per 200 µL PCR tube.
2. Add 70 µL of lysis buffer (0.1 M NaOH + 1% SDS) to each tube.

Step 2: Lysis & Crosslink Reversal

1. Incubate at 100°C for 2 hours in a thermocycler with a heated lid.
2. Allow the samples to cool to room temperature.

Step 3: Neutralization

1. Add 70 µL of neutralization buffer (100 mM Tris-HCl, pH 8.0) to each tube.
2. Mix gently by pipetting up and down.

Step 4: Proteinase K Digestion (for selected replicates)

1. Add 20 µL of Proteinase K (20 mg/mL stock).
2. Incubate at 56°C for 3 hours (up to overnight if required).

Step 5: DNA Purification (Qiagen Kit Extraction)

a. Spin samples down quickly. Add 330 µL Buffer AL to the sample, mix thoroughly and incubate at 70 °C for 15 minutes. (Use this time to prepare the columns).

b. Add 330 µL ethanol (96–100%) to the sample and mix thoroughly.

c. Pipette the mixture from step 1 (including any precipitate) into the Zymo micro spin column placed in a 2 mL collection tube. Centrifuge at 10000 x g for 1 min. Discard flow-through and collection tube.

d. Place the Zymo micro spin column in a new 2 mL collection tube, add 500 µL Buffer AW1, and centrifuge for 1 min at 10000 x g. Discard flow-through and collection tube.

e. Place the Zymo micro spin column in a new 2 ml collection tube, add 500 µL Buffer AW2, and centrifuge for 1 min at 10,000 x g. Discard flow-through.

f. Repeat step e). Discard flow-through and collection tube.

g. Place the Zymo micro spin column in a new 2 ml collection tube, open the lid and centrifuge for 5 min at 20,000 x g to dry the membrane. Discard flow-through and collection tube.

h. Place the Zymo micro spin column in a clean 1.5 ml microcentrifuge tube, and pipet 10 µL nuclease-free water directly onto the Zymo column membrane. Incubate at room temperature for 10 min, and then centrifuge for 1 min at 10000 x g to elute.

i. Pipette additional 10 µL nuclease-free water directly onto the Zymo column membrane. Incubate at room temperature for 5 min, and then centrifuge for 1 min at 10000 x g to elute.

Amplification and sequencing

DNA obtained using the above protocol is suitable for standard TAQ based PCR amplification using common commercially available reagents. However, formalin preservation results in excessive fragmentation of the DNA. Thus, specialised primers are required for during amplification and sequencing. The following primers, targeting the mitochondrial cox1 gene fragment, were designed. These primers amplify an approximately 90 bp region that is diagnostic for *M. enterolobii*.

Forward Primer (Formelo-F)

GGGCGAAAAGGAATAATTTTAGCAATT

Reverse Primer (Formelo-R)

GTCGCTGAACTAAAATAAGCACGA